

The responsible governance of hate speech in app stores: The case of Parler

Thesis submitted in partial fulfilment of the requirements for the degree of
DPhil in Information, Communication and the Social Sciences
in the Oxford Internet Institute
at the University of Oxford

Word count: 79,957

Josh Cowls
Oxford Internet Institute
St. Hugh's College

Supervisors:
Prof Luciano Floridi
and Prof Mariarosaria Taddeo

Hilary Term 2023

Abstract

This thesis is anchored in an empirical case study of the fringe social media platform Parler, with a focus on how the platform was governed by Apple and Google as operators of the two largest smartphone app stores. Parler was launched in 2018 and gained popularity during 2020 following the "deplatforming" of conservatives from mainstream platforms. It sought to discursively legitimate a platform governance regime that it claimed was grounded in "constitutional free speech". Parler's removal from app stores following the US Capitol riot in January 2021 casts light on the quiet yet significant governance role that app store operators like Apple and Google play with respect to the prevalence of problematic material on social media platforms that their app stores host. The aim of this thesis project is to critically assess the role that app store operators purport to play, the role that they actually play, and the role that they ought to play with respect to hate speech on social media platforms. The project is designed to generate novel theoretical, empirical, and policy contributions. First, I examine the responsibility to govern hate speech that app store operators accord themselves and identify several ambiguities related to the enforcement of their rules. Then, in the empirical case study, I find that Parler's failure to manage hate speech on the platform in 2020 would have warranted Apple and Google to sanction it earlier than January 2021, based on analysis of hate speech on Parler and the two companies' guidelines. Finally, I assess the extent to which forthcoming legislation will address challenges in app store governance. The central finding of this thesis is that, although Apple and Google accord themselves the responsibility to govern hate speech in their app stores, this responsibility is ambiguous and is not always adequately met in practice.

Acknowledgements

If it takes a village, as the proverb has it, to raise a child, then it takes a metaphorical megapolis to sustain a DPhil student. I am indebted to an extraordinary number of friends, colleagues, family members—and, at times of particular desperation, the occasional allotment plant—without whose support it would have been impossible for me to make it this far. I will inevitably fail to name them all here.

I must begin by expressing gratitude to my supervisors, Luciano Floridi and Rosaria Taddeo. Both of you have been a source of immense support, wisdom, patience, and encouragement throughout this process. I appreciate your faith in me even—especially—in those moments when I've lacked it in myself.

I am also very appreciative to my examiners, especially Vicki Nash, who has been an invaluable source of guidance at each stage of the examination process. Many thanks are also due to Ralph Schroeder, Rebecca Eynon, and Matthias Kettemann.

I was the grateful recipient of a doctoral studentship from the Alan Turing Institute, and it was a pleasure to remain involved with the Institute's Ethics Advisory Group and Public Policy Programme through much of the DPhil—so thanks must go to Kirstie and David for keeping me around. I'm also indebted to Paul for helpful conversations and API access.

I also owe thanks to the intellectual environment provided by the Oxford Internet Institute, and above all to the community of doctoral students, past and present, who offered support and solidarity in abundance. Special thanks must go to my research group peers, including Corinne, David, and Carl, who showed us all how it's done, as well as to Nikita, Vincent, Jakob, Jess, Marta, Andreas, Huw, and Emmie, with many whom I've been fortunate enough to collaborate on work in the field of digital ethics. I also had fortunate opportunities to collaborate with colleagues within my institute, including Yung Au and Cindy Ma, and beyond, including Philip Darius, Dominiquo Santistevan, Moritz Schramm, and Bertie Vidgen. I belonged to a small but mighty cohort, with Cindy, Fabian, Julian, Marie-Therese, Tomas, and Yung. It's hard to imagine

a more talented and empathic group of scholars, and I'm excited to find out where you go and what you'll do next. I'm also indebted to the brain trust: my formidable co-podcaster Nayana, co-punster Patrick and co-plotter Yung—as a source of ever-illuminating insights. Finally, I most would certainly not have made it this far without the friendship and support of Jess.

One of the great privileges of graduate study at the University of Oxford is the college system, which more often than not yields friendships that last even longer than the length of a DPhil. I was lucky enough to get to know Alex, Alice, Amartya, André, Anwar, Caroline, Ed, Florence, Jake, JT, and Tom at St Hugh's College. I'm grateful that Sam, Izzie, Evan, Kate, and Serena were close by in Oxford too, and that Dan, Ed, Hannah, Jonny, Justin, Kate, Liam, Lizzie, Oli, Otis, Ronnie, and Siamak weren't all that far away either. Special thanks must go to Campbell, with whom I navigated two years of on-off lockdowns, variable harvests and questionable brews.

I'm indebted to the support and encouragement of my family, not least my grandmother. I'm thankful for the park runs and occasional pints with Stephen and Louisa. Watching my sister Ros marry Shailan will live longer in the memory than anything I learned during the DPhil. Mum, Dad: thank you for the perennial support; I promise this is the last one.

Finally, and above all, my most heartfelt thanks to Apurva, for everything. It is almost as hard to believe that I started this journey not knowing you as it is to imagine finishing it without you quite literally by my side. I'm excited for all the journeys yet to come.

All intellectual work is necessarily indebted to the thinking and writing that has gone before. This thesis is no different, as the wide array of sources—academic and journalistic—that I have cited attests. I never fail to be impressed by the rigour, passion, and compassion that drives so much academic work. One inevitably spends large portions of a doctoral project alone with one's thoughts and ideas, but as long as one is surrounded by books, the process need not be lonely.

—Josh Cowls
April 2023

Table of Contents

Abstract....2

Acknowledgements....3

Index of Figures....8

Index of Tables....9

Chapter 1: Introduction....10

1.1 Background and rationale....12

1.2 Key concepts and questions....17

1.3 Epistemology, methodology, and limitations....20

1.4 Overview of argument and structure....27

Part I: Theory....31

Overview....31

Chapter 2: The "new governors"—platforms as governed and governing....33

Introduction....33

2.1 Internet governance: the shift to the private sector....33

2.2 Governing platforms....39

2.3 The discursive legitimation of platform self-governance....44

Summary....53

Chapter 3: Between the stack and a hard place—conceptualising and situating app store governance....55

Introduction....55

3.1 Cross-platform governance....56

3.2 App store governance....70

Summary....82

Chapter 4: The governance of hate speech in app stores....84

Introduction....84

4.1 Online hate speech: what is it, who does it harm, and how?....85

4.2 Platform self-governance of hate speech 95
4.3 Cross-platform governance of hate speech 101
4.4 App store governance of hate speech: framework or patchwork? 108
Summary 113
Part II: Case study analysis 115
Overview 115
Chapter 5: The "populist platform"—introducing Parler 117
Introduction 117
5.1 Parler: a chronology 117
5.2 Existing research on Parler: a systematised thematic review 129
5.3 Case selection 135
5.4 Methodological approach, data collection, and analysis 139
5.5 Ethical and social considerations 159
Chapter 6: Parler games—the rhetoric and reality of Parler's self-governance 161
Introduction 161
6.1 Technical walkthrough of Parler's functionality and features 162
6.2 The public performance of Parler's self-governance 169
6.3 Multimodal analysis of Parler's website 185
6.4 Parler's policies 190
Summary 197
Chapter 7: Parlous—the app store governance of Parler 199
Introduction 199
7.1 Parler in practice 199
7.2 App store governance of hate speech on Parler 206
7.3 Better late than (n)ever? Parler's removal from app stores 219
Summary 223

Part III: Implications, reflections, and conclusions 226

Chapter 8: Policy implications and impacts 226

Introduction 226

8.1 Synopsis 227

8.2 Who do you sue now? The platform power of app store operators 228

8.3 The impact of new legislation on app store governance 231

8.4 Reflections, limitations, and directions for future research 238

8.5 Closing remarks 242

References 243

Appendices 280

Appendix 1: Own work cited 280

Appendix 2: List of articles found in systematised thematic review 281

Appendix 3: Example of coding in mainstream media corpus 284

Appendix 4: Queries sent to APIs for hate speech classification 285

Index of Figures

Figure 1: Parler's first launch screen, May 2018 119
Figure 2: The home page of Parler's website in January 2021 129
Figure 3: PRISMA flow diagram depicting identification of records on Parler 131
Figure 4: Multi-level iterative coding schema for mainstream media analysis 147
Figure 5: Path diagram showing how the dataset was filtered and sampled from 156
Figure 6: Total number of posts and users sampled for each month 156
Figure 7: Parler's entry in the iOS App Store in January 2021 163
Figure 8: Parler's justification for asking for a new user's phone number 164
Figure 9: Parler's privacy policy displaced as a pop-up during registration 164
Figure 10: Suggested accounts to follow to "personalize your Parler experience" 165
Figure 11: Parler profile page for Representative Andy Biggs (R-AZ) as of June 30, 2020 168
Figure 12: Articles pertaining to different constitutional metaphors. 171
Figure 13: The basic structure of Parler's homepage between June 2019 and June 2020 187
Figure 14: Parler's homepage from June 2020 187
Figure 15: Parler's "values" as of June 2020 189
Figure 16: Parler's website as of October 2020 190
Figure 17: Parler's "Protected" and "Prohibited" content 192
Figure 18: Parler posts which contained at least one hateful or toxic post in the sample 201
Figure 19: Parler users who made at least one hateful or toxic post in the sample 202
Figure 20: The identity groups most frequently targeted on Parler. 205

Index of Tables

Table 1: Overview of the theoretical part of the thesis.....32
Table 2: Overview of research questions, data, and methods used in the case study..... 116
Table 3: Sources of data gathered for the environment of expected use analysis..... 141
Table 4: Summary of documents used in policy analysis. 150
Table 5: Industry and political stance of recommended Parler accounts to follow..... 166
Table 6: Proportions of hate and toxicity detected in the Parler post sample 201

Chapter 1: Introduction

On January 6th, 2021, the day that the United States Congress was due to formally certify Joe Biden's victory in the presidential election two months earlier, a mob of supporters of outgoing President Donald Trump stormed the Capitol complex. Stoked by an inflammatory speech given by Trump just outside the White House grounds, thousands marched to the Capitol to try to prevent the certification of the electoral count. Five people died during the violent attack, which forced Congress to halt the certification. The images and videos which rapidly spread around the internet served as an especially vivid manifestation of the threat of far-right extremism to twenty-first century American democracy and society (Bennett et al. 2021). More than 1,000 people have since been charged with crimes in relation to the insurrection (Hall 2023).

Starting two days later, the fringe social media platform Parler, which had been rapidly gaining popularity among conservatives and had even topped app store charts in the wake of Trump's electoral defeat, was hit by a series of suspensions by the infrastructural providers on whom it relied to stay online. On January 8th, Google announced that it was pulling Parler from its Play Store (Chen 2021). The same day, Apple demanded that the operators of Parler submit a revised plan to improve its moderation within 24 hours (Mac and Paczkowski 2021). Parler executives refused, and Apple, following Google, responded by removing the platform from its iOS App Store the following day (Peters and Lyons 2021). Parler suffered sanctions from other service providers in the following days, including its suspension by Amazon Web Services, which had the effect of taking the site temporarily offline altogether (Lyons 2021a).

The de facto shuttering of Parler by a small number of private infrastructural providers occurred at a moment of especially high political tension, and the contemporaneous justifications given by these companies reflected the immediacy with which they felt they had to act—and, perhaps just as importantly, their need to be *seen to be* acting. Taken together with similar recent actions by companies that provide core elements of the "tech stack" to social media platforms, the removal of Parler set a precedent with significant long-term consequences for internet governance.

This doctoral thesis is anchored in a case study of the app store governance of Parler in 2020, a period in which the platform grew rapidly despite concerns about the level of problematic material that it contained. The case study culminates with Parler's removal from Apple's App Store and Google's Play Store following the Capitol riot.[1] Through the case study, which employs app walkthrough and content analysis as its central methods, I explore questions such as: should Parler have been removed sooner, in view of app store governance guidelines and the hate speech that was said to be replete on the platform long before January 2021? What can the Parler case tell us about the responsibility and limitations of app store operators and other stack providers to uphold a safe and open internet, and how can they be held to account for doing so? Finally, what is the nature of app store governance, and how should this role be constrained or clarified through public policy?

In this introductory chapter, I elaborate on my motivations for undertaking the doctoral project, discuss my research questions, and explain the approach I take to answering them. In 1.1, I provide context for Parler as a platform and discuss my motivation for conducting a case study on it. The Parler case sits at an intellectually interesting point of intersection between the governance of platform users by platform operators, and the governance of platform operators by infrastructure providers. In 1.2, I introduce the various theoretical concepts that I will respectively engage with in Chapters 2 to 4, namely platform governance, app store governance, and hate speech, and explain how these translate into my overarching research question and three subsidiary research questions. In 1.3, I discuss my epistemological and methodological approach to approaching these questions and justify the selection of data that I use to answer them. Finally, in 1.4, I offer a detailed overview of the remaining chapters and outline how each contributes to my overarching research question.

[1] Parler was eventually reinstated to Apple's App Store in 2021 (Perez 2021) and Google's Play Store in 2022 (Grant 2022).

1.1 Background and rationale

Much has been said about what can be said online (Balkin 2018; Citron and Norton 2011; Douek 2022; Griffin 2022; Klonick 2017). A common area of scholarly focus is how—and how well—the operators of major social media platforms like Facebook, Twitter, TikTok, and YouTube manage the deluge of text, images and videos that are uploaded and shared every second on their platforms. This includes how platforms present content to users, as in the algorithmically powered "news feeds" that are now ubiquitous on social media (Binns et al. 2017; Milano, Taddeo, and Floridi 2020). It also includes how platforms detect content deemed objectionable and either "downrank" or remove it altogether, as part of a process commonly referred to as content moderation (Balkin 2017; Douek 2022; Gorwa, Binns, and Katzenbach 2020; Griffin 2022; Kettemann and Tiedeke 2020). In extremis, platforms may also suspend or ban specific users or groups outright in response to repeated violations (Rogers 2020). Taken together, these and other activities involved with managing a social media platform, and its design, users, and content, are often grouped under the heading "platform governance". However, since this term is also applied in other contexts (see Gorwa 2019), for clarity, I will hereafter refer to this process as *platform self-governance*.

One prominent focus of platform self-governance in scholarship and practice concerns forms of harmful content, including hate speech, that are created, shared, and amplified online. This is often done by individuals affiliated with amorphous movements like the "alt-right" (Daniels 2018; Ma 2021; Rieger et al. 2021). Faced with increasing public pressure and threats of tighter public policy, since 2016, major platform companies such as Meta have taken steps to tighten their rules against such forms of problematic content, and have shown more willingness to enforce them (Douek 2019; Kettemann and Schulz 2020). Platform companies have tended to take these steps only reluctantly, in response to outside pressure from the media, academics, and activists, as well as in the face of new regulation like Germany's NetzDG (Gorwa 2021; Heldt 2019; Schulz 2018). Meanwhile, there has been growing public pressure for firmer action to be taken by a

different[2] group of large technology companies with respect to hate speech and extremism. These companies, such as Amazon, Cloudflare, and GoDaddy, do not operate social media platforms themselves; instead, they provide various parts of the technical infrastructure that allows social media platforms to operate, often referred to as the "tech stack" (Cath and Ruddock 2022; Donovan 2019; Van Dijck, de Winkel, and Schäfer 2021). This infrastructure includes the cloud computing and reverse proxy systems that modern platform operators rely on to serve users data, the payment processing functionality that allows them to monetise, and the app stores through which they can make their platforms available to smartphones and other personal computing devices. Because of the infrastructural nature of their products (van Geuns and Cath-Speth 2020), these "stack providers" typically face pressure to intervene not against *specific users* or *communities*, but rather against entire *domains, websites, platforms* and *apps* that are considered dangerous, extreme, or otherwise harmful (Busch 2022; DeNardis and Bradshaw 2022; Van Dijck et al. 2021). I describe this form of governance as *cross-platform governance,* a concept that I explicate in Chapter 3.

The case of Parler represents an interesting point of intersection between these two forms of internet governance: *platform self-governance* on one hand and *cross-platform governance* on the other. Parler was[3] a social media platform with its own features, rules, and policies, and at least some aspiration to compete commercially with mainstream social media platforms. From its launch in 2018 through its rise to prominence in 2020, Parler was one of many would-be alternatives to dominant social media platforms like Facebook or Twitter. But from the outset, Parler's user growth was fuelled almost entirely by conservative Americans, many of whom were said to be disillusioned by the (to them) increasingly censorious governance decisions taken by Parler's larger rival platforms. This was no accident. From its launch, Parler proudly proclaimed that its mission

[2] It is worth noting at the outset that many of the same technology companies, and/or their subsidiaries, both operate social media platforms themselves and provide stack infrastructure to other social media platforms. This is the case with Alphabet, which operates the platform YouTube as well as the Google search engine and the Google Play Store.

[3] During the final stages of the preparation of this thesis, Parler was sold, with its new owners announcing plans to take it offline while it undergoes a "strategic assessment" (Weatherbed 2023).

was to safeguard "free speech", and it adopted a distinctively "heartland" iconography and right-populist rhetoric. Such discursive efforts were designed specifically to attract conservative American users that had become disillusioned with incumbent platforms. Therefore it is no surprise and no coincidence that the two largest spikes in Parler's popularity, each of which took it to the top of app store download charts, coincided with key political flashpoints: the defeat of President Donald Trump in early November 2020 (Hatmaker 2020); and the immediate aftermath of the Capitol riot (which resulted in Trump's permanent[4] suspension from Twitter), in January 2021 (Chen 2021). Shortly thereafter, Parler was removed from Apple's and Google's smartphone app store platforms, and had its hosting services suspended by Amazon, following allegations of the platform's complicity in the riot at the Capitol (Lyons 2021a).

Parler's rapid rise and even faster fall may be expected, a priori, to hold important implications for internet governance, in two key ways. The first relates to Parler's self-governance, i.e., how Parler as a company sought to govern its own users and content. Parler is one of the highest-profile and most popular cases of a social media platform designed to appeal specifically to right-wing users. With "the internet's town square" as its U.S.P., and the claimed safeguarding of "free speech" as its modus operandi, Parler sought to develop and legitimate what it claimed was a distinctively "freedom"-preserving approach to platform self-governance, in contrast to its allegedly more censorious rivals. The second key implication springs from how Parler was governed in turn by stack providers, including app store operators Apple and Google. Prior to its removal from app stores, Parler was one of the most downloaded apps in the United States, which inevitably invites questions about, on the one hand, whether its removal was justified given the impact on free expression—and on the other hand, whether it should have even been admitted into app stores in the first place, or removed from them sooner. With their removal of Parler, Apple and Google took what was perhaps their most prominent governance decision yet with

[4] Trump's ban was later reversed, but at the time of writing he has not returned to the platform.

respect to their app stores, which are the primary (on Google's Android) and sole (on Apple's iOS)[5] gateway for downloading and accessing applications on, taken together, the large majority of smartphone operating systems.[6] Therefore, to the extent that the Parler decision could serve as a "threshold" for app removals in future, gaining a clearer understanding of such a threshold and its implications is a worthy goal for research, and perhaps also for public policy.

Parler's retention in and eventual removal from app stores constitutes a bounded case with respect to the actors involved (Parler's leaders, its users, and Apple and Google), the timeframe (from its launch in mid-2018 until its removal in January 2021, with a focus on the year 2020), and the geographical and political context (the United States, at a time of considerable ideological polarisation, creeping extremism, and a heated presidential election). Absent in this discussion thus far, however, are normative and meta-theoretical considerations: for example, what are the societal values or ethical principles that *should* shape platforms' self-governance and, in turn, their governance by stack providers like app store operators; and how and by whom should these values and principles be determined and if necessary enforced? And when we talk about platform governance in its various guises, are we really speaking of "governance" in its conventionally understood sense? If Apple, Google and Parler are, in their own ways, "governors", then what is the nature of this governance, and how and by whom should it be held to account?

In order to more tightly bound the case study in view of such questions, I further narrow its conceptual scope in two ways. First, with respect to normative considerations, I focus specifically on the governance of *hate speech*. Hate speech is one of the most prominent, and trenchant, of the various forms of harmful content that afflict internet users, and it has particular relevance to my case study for two related reasons. Firstly because, as a consequence of its mission to be the home of "constitutional free speech", Parler's executives adopted an especially laissez-

[5] This may be set to change with Apple's plans to allow third-party app stores on its devices, in advance of forthcoming legislation. See Chapter 8.

[6] See https://gs.statcounter.com/os-market-share/mobile/worldwide.

faire attitude towards problematic content like hate speech. And secondly, because as I later discuss, the prevalence of hate speech on the Parler platform may have justified firmer governance action against Parler by stack providers, long before events spilled offline in January 2021.

Second, with respect to meta-theoretical considerations, in this study I adopt a sociotechnical approach to social media platforms that is rooted in a science and technology studies (STS) approach to internet governance (Epstein, Katzenbach, and Musiani 2016; Mager and Katzenbach 2021; Ziewitz and Pentzold 2014). This leads me to take a discursive approach to the data examined in the case study, including the policies, guidelines and explanations that constitute platform governance. This is not to suggest that platform governance only has "rhetorical" rather than very real and socially relevant impacts. But I argue that the complex nature of platform governance power is best revealed not by taking its textual manifestations at face value, but instead by situating them within the wider network of interactions between different actors that it involves and the array of motivations that it implicates.

In sum, this thesis project has theoretical, empirical, and political components. The thesis makes theoretical contributions at the point of intersection between, on the one hand, *platform self-governance* in the sense of social media platform companies governing their own users and content and, on the other, the *cross-platform governance* undertaken by infrastructure providers. Parler constitutes a relevant and important case study through which these theoretical contributions can be further elaborated and evidenced. And the focus on hate speech foregrounds normatively significant questions about the importance of efficacious platform governance and the possible role for public policy. The tripartite structure of the thesis reflects these three sets of contributions, with parts covering theory (Part I, Chapters 2 to 4), the empirical case study (Part II, Chapters 5 to 7), and normative considerations and reflections (Part III, Chapter 8), though there is of course some overlap. Each part has a subsidiary research question associated with it, which contributes to the answering of the overarching research question. In the remainder of this introductory

chapter, I lay further foundations for what follows, starting with an expanded discussion of key conceptual considerations, which leads to the identification of my research questions.

1.2 Key concepts and questions

At the highest level, this thesis is about *internet governance*: a large domain of research which incorporates perspectives from philosophy and ethics, law and governance, computer science, and the social sciences. More specifically, my topic is concerned with the subset of internet governance known as *platform governance*, which has been described as encompassing "the layers of governance relationships structuring interactions between key parties in today's platform society, including platform companies, users, advertisers, governments, and other political actors" (Gorwa 2019:854). More specifically still, my case study primarily involves two specific platform governance "configurations", or relationships: between (i) the social media platform Parler and (ii) its users; and between (i) app store operators and (ii) Parler. The first relationship is an example of can be called *platform self-governance*: a subset of platform governance which concerns the rules and procedures that platform operators adopt and implement with respect to their users' activity, behaviour, and content. Platform self-governance has been much studied, with much research focused on how the largest platform companies like Meta, Alphabet, and ByteDance undertake content moderation as a governance project on their respective platforms Facebook and Instagram, YouTube, and TikTok (Gillespie et al. 2020; Hosseinmardi et al. 2021; Kettemann and Schulz 2020; Zeng and Kaye 2022). Relatively less attention has been paid to what I call *cross-platform governance*, or the (often implicit) governance *of* platform companies by the third parties that provide various essential parts of these platforms' "tech stack" (Donovan 2019; Van Dijck et al. 2021). Cross-platform governance encompasses, among other things, *app store governance*: the governance role played by Apple and Google with respect to the apps, including social media platforms, that they enable users to install on their smartphones (Hestres 2013; Nieborg, Young, and Joseph 2020). Finally, circumscribing my scope still further, my focus is on the app store

governance of *hate speech* (Matamoros-Fernández and Farkas 2021), as one of the many forms of problematic and often harmful content that may warrant governance action by app store operators. I address each of these core concepts in depth in Chapters 2, 3, and 4.

To guide the study, I specify an overarching research question (ORQ), which reflects the aims and scopes of the project. This is as follows:

ORQ: How should app store operators responsibly govern hate speech on social media platforms?

The ORQ is designed to be open-ended enough to allow a comprehensive treatment of the topic, encompassing the theoretical, empirical, and normative considerations identified above. The use of the term "responsibility" is thus deliberately ambivalent. First and foremost, it is intended to capture the obligations and expectations that guide the app store governance of hate speech in principle, and how well these are applied in practice, as in the context of the case study on Parler. But simultaneously—if somewhat more implicitly—the ORQ is also designed to carve out space for reflection about the normative role that app store operators *ought to* play with respect to hate speech on social media, as well as how this responsibility could be clarified through policy.

Corresponding to the tripartite nature of the thesis project, the ORQ can be broken down into three subsidiary questions—one essentially theoretical, one essentially empirical, and one essentially normative—around which the thesis is structured. The first research question is theoretical and is designed to provide contextual grounding for the case study that follows, as well as prompting conceptual insights. It is primarily addressed in Part I (Chapters 2 to 4), and it asks:

RQ1: What responsibility do app store operators have to govern hate speech on social media platforms?

This question allows me to conceptually situate app store governance within the larger domain of internet governance, makes explicit my focus on hate speech, and creates a theoretical basis for the empirical case study that follows. The main contributions of this part to the thesis are (i) to theoretically clarify and situate key concepts and (ii) to provide theoretical foundations for the empirical case study. Specifically, I show that app store governance involves the delegation of responsibility for content moderation from app store operators to social media platforms. This

insight structures the case study analysis: I first examine Parler's platform self-governance (Chapter 6) before turning to Apple's and Google's app store governance (Chapter 7).

The second research question is anchored in the case study on the platform governance of Parler. It is designed to draw out empirical insights that can contribute beyond this specific case to the overarching research question more generally. It is primarily addressed in Part II (Chapters 5 to 7), and it asks:

RQ2: How well did Apple and Google meet their responsibility to govern hate speech on Parler?

This question is constructed to allow as direct as possible an interrogation of the available empirical data that I use in the case study. It is premised on the nature of the governance responsibility identified in Part I. The research question can therefore be broken down into two component parts, concerning (i) Parler's platform self-governance and (ii) Apple's and Google's cross-platform governance of Parler, which saw the platform eventually suspended from app stores.

The final research question is designed to prompt a synthesis of the prior theoretical and empirical insights, wider reflections, and consideration of possible public policy interventions to address the problems identified. It is addressed in Part III (Chapter 8), and it asks:

RQ3: What additional considerations should shape the app store governance of hate speech on social media?

Together, the three subsidiary research questions address different aspects of the overarching research question. RQ1 provides theoretical foundations, asking what responsibilities app store operators accord themselves with respect to governing of hate speech on social media platforms, and situating the case study on Parler within existing ethical, legal, and social requirements and expectations. RQ2 derives empirical evidence from the case of Parler, which, in turn, can be applied to the broader issue of the app store governance of hate speech. RQ3 builds on the prior theoretical and empirical insights to assess the extent to which forthcoming legislation will address the shortcomings of app store governance that I have identified. The three subsidiary research questions thus ask the overarching research question in three different ways: what responsibility do app store operators have to govern hate speech *in principle*? How responsibly do app store

operators govern hate speech *in practice*? And are there additional responsibilities that should be accorded to app store operators *in public policy*? In what follows, I turn from specifying these questions to explain how I will go about answering them, with a focus on the epistemological assumptions and methodological choices shaping the case study on Parler.

1.3 Epistemology, methodology, and limitations

In this section I discuss the approach to knowledge generation that I adopt in the thesis project, and the epistemological and ontological assumptions on which this approach rests. In 1.3.1 I situate my epistemology within science and technology studies broadly and within an infrastructural view of internet governance more narrowly, and tie this to the methodological approach I employ in the case study. In 1.3.2 I explain and justify my decision to use an empirical case study as my approach to the analysis. In 1.3.3 I offer some ethical reflections on the thesis as a whole. In each case, the discussion is necessarily only a preview, because it relies on definitions, arguments, and justifications that I provide in subsequent chapters, but it nonetheless provides useful introductory context to what follows.

1.3.1 Epistemological stance and limitations

Part I (Chapters 2 to 4) takes as its conceptual starting point the recent uptick in applications of science and technology studies (STS) to internet governance. By now, the theoretical underpinnings of "the sociotechnical", and the disciplinary debt that this terminology owes to STS, has to some extent "faded into the foreground" of contemporary internet governance scholarship (c.f. Epstein, Katzenbach, and Musiani 2016). This is evidenced by the common use of the term "sociotechnical" in recent work that addresses key concepts in this thesis, including platforms (Gorwa et al. 2020; Poell, Nieborg, and Van Dijck 2019) and content moderation (Gillespie 2018:197). However, it is still useful to (re)articulate the value of an STS approach to internet governance. Musiani (2020:85) characterises STS approaches to internet governance as centred on:

> "The agency of technology designers, policy makers, and users as they interact, in a distributed fashion, with technologies, rules, and regulations, leading to consequences with systemic effects

> that may, at times, be unintended. Social and political ordering is understood as a set of ongoing and contested processes … Thus, conceptually, *STS-informed Internet governance research* relies on understanding governance as a normative system of systems, and it acknowledges the agency, often discreet yet pervasive, of both human and nonhuman actors and infrastructures."

I take forward this characterisation of STS-informed internet governance research in what follows. Therefore, when I characterise a social media platform such as Parler as a "sociotechnical system", I mean to suggest that it *consists* at a fundamental level of both human participants and technical-infrastructural artefacts (Helmond 2015; Light, Burgess, and Duguay 2018), and moreover, that it is to some meaningful degree *constituted* by their intersection (Bucher and Helmond 2018; Waldherr, Geise, and Katzenbach 2019). "Technology has always both reflected and reorganised the social", as Katzenbach and Ulbricht surmise (2019). And as Poell and colleagues put it with respect to platforms specifically, "platforms transform cultural practices, and vice versa, evolving practices transform platforms as particular socio-technical constructs" (Poell et al. 2019:5). In short, the technical infrastructure of a platform is inextricably linked to—although not indistinguishable from—the ways in which it is used (Bucher 2012; Plantin et al. 2018).

These contributions have myriad implications in turn for how we are to understand the *governance* by, and of, these platforms. Many of the managerial and administrative activities undertaken by these platforms, such as detecting and removing content that violates community standards, are rightly seen as governance (Gillespie 2018; Gorwa 2019; Grimmelmann 2015). Following Taylor (2021), I will argue that the *private* nature of these governance activities—"private" in the sense both of "private sector" and in the sense of "unseen or secret"—invites questions about the democratic legitimacy of such "governance". And as I will show, this has in turn has prompted technology companies, especially social media platform companies who control and curate online public spheres, to seek to legitimate their governance role. In some cases, these efforts can be seen as legitimate in a normative sense, via legitimation relations as part of the

normative order of the internet (Kettemann 2020) but they often (also[7]) involve discursive attempts to cultivate legitimacy in the descriptive sense of the term. Internet governance can therefore be conceptualised in terms of what Ziewitz and Pentzold call "performing order in digitally networked environments" (2014). Musiani likewise foregrounds "controversies and their performative role" as a first-order consideration for internet governance, emphasising the significance of legitimation on the "processes by which norms are created, renegotiated, put to the test, and realigned and raise conflicts" (2020:93). Such controversies often arise over infrastructure (Musiani 2015), and specifically the internet's "control points" (DeNardis and Musiani 2014), of which, as I will argue in the case of platforms, app stores are one.

This performative conceptualisation of platform governance motivates the methodology that I adopt in the case study on Parler (Chapters 5 to 7). To summarise, the methodological approach I take in the case study involves an adapted version of the *walkthrough method* applied to the Parler app (Light et al. 2018), which is supplemented by an automated and manual content analysis of hate speech in a sample of posts made on Parler during 2020. The walkthrough method, which is grounded in STS, has become a popular way for researchers to perform "critical analysis of a given app" (Light et al. 2018). However, it was necessary to adapt this method to accommodate the fact that Parler had been taken offline at the time that I began to study it closely,[8] by adopting the techniques of web-archival research of social media (Brügger 2018). In order to assess the extent and nature of hate speech on Parler I used a combination of automated and manual classification (Rieger et al. 2021). I provide more detail about the specific data collected and methods employed in Chapter 5. Next, I turn to explain my choice of an empirical case study as an approach to knowledge generation.

[7] As in the case of Meta's Oversight Board, sometimes these efforts overlap in ways that are difficult to conceptually disentangle (Cowls et al. 2022).

[8] This was unfortunate but unavoidable insofar as the removal of Parler motivated my decision to study it.

1.3.2 The case study approach: utility and limitations

For the empirical component of the thesis project (Chapters 5 to 7) I adopt a case study approach, focusing on a social media platform, Parler, over the course of a year, 2020. Here I discuss the utility of taking a case study approach to this subject as well as noting common criticisms of the case study as an approach to knowledge generation. There remains some disagreement with respect to whether a case study is itself a "research method" (Yin 2018:15) or instead a research strategy which can in turn encompass multiple qualitative and/or quantitative methods (Kohlbacher 2006) and rely on multiple sources of evidence. Regardless, a key advantage of the case study approach to my thesis project is its epistemological flexibility and openness. In one well-known articulation, Yin (2018:15) describes a case study as "an empirical inquiry that investigates a contemporary phenomenon ("the case") within its real-life context, especially when the boundaries between phenomenon and context may not be clearly evident." This conceptual blurring of "phenomenon" and "context" is applicable to the case of Parler for several reasons. First, absent any context, "Parler" can refer to both the platform with a specific set of users, and to the company that provided the platform. In practice, these referents were often used interchangeably or ambiguously, particularly in discussions of Parler's platform governance.[9] This blurred the boundary between the phenomenon of "Parler" and its context. Second, Parler (in both senses just mentioned) was enmeshed within a wider web of relationships with other actors and networks, especially in relation to its governance. Here again, any attempt to separate the phenomenon from the context would have been futile and counterproductive. Third, as Yin describes, a case study typically "relies on multiple sources of evidence" (2018:15), as I do here.

Nonetheless, as with any approach to knowledge generation, the case study approach has its healthy share of detractors as well as advocates. Perhaps the most fundamental critique of the case study is that which takes aim at the usefulness of a single "case" to explain a complex

[9] This semantic conflation of "platform" as sociotechnical space and "platform" as the organisation that provides and manages the space arises in the case of other platforms too.

phenomenon, with some suggesting that this fatally limits its generalisability, and thus its utility as a form of knowledge generation. One traditionalist view of "one-shot case studies" lamented their "total absence of control" of variables, claiming them to "be of almost no scientific value" (Campbell and Stanley 2015:6). Yet as Flyvbjerg (2006:7) documents, Campbell would later entirely change his mind about the utility of case studies, describing "common sense naturalistic observation" as "the only route to knowledge". At least some of this disagreement over the utility of the case study seems to stem from the blurring of what is meant by "cases" on one hand and "observations" on the other. King, Keohane, and Verba's influential *Designing Social Inquiry* (2021) is often characterised as being critical of case study research, but in fact, it acknowledges that a single "case study" may consist of an "immense" number of observations which—from the authors' perspective—can suffice to allow for causal inference, and hence the testing of causative theories (King et al. 2021:52). These definitional discrepancies mask deeper disagreement over the epistemological value of different approaches, and in particular the attempt by King and colleagues to determine the usefulness of a study in terms of the number of "observations" (*N*) that it has collected.

While King et al. set out to make qualitative research more "rigorous" by borrowing tactics from quantitative methods, Mahoney and Goertz (2006) instead approach the quantitative and qualitative traditions as "alternative cultures", and argue that the typical goal of qualitative case study research is to explain particular outcomes. For this reason, qualitative researchers tend to deliberately select cases where the said outcome *has* occurred (though note that researchers sometimes *also* select a case in which an outcome has *not* occurred for the sake of comparison). By contrast, quantitative researchers consciously seek to *avoid* selecting for outcomes, by using random selection of cases in which an outcome of interest may or may not have occurred, to ascertain the causal mechanisms that explain the occurrences.

For Mahoney and Goertz, even a "within-case" analysis (i.e., where $N=1$) can still "provide substantial leverage for causal inference" (2006:239) in situations where the goal is to explain why

something happened in the way it did. And it is reasonable for cases to be selected because they are "substantively important" (2006:242–43)—again in contrast to purely quantitative approaches which do not treat any particular instance as ex ante important.[10] Flyvbjerg elaborates several ways in which a given case may be considered substantively important and therefore be worthy of dedicated study. This includes *extreme cases*, which can be "well-suited for getting a point across in an especially dramatic way"; *critical cases*, which may have "strategic importance in relation to [a] general problem"; and *paradigmatic cases*, which "highlight more general characteristics of societies in question" (Flyvbjerg 2006:14–16). These are prototypes, but they help to show the kinds of considerations that may justifiably shape a researcher's choice of case.

To summarise: its critiques notwithstanding, a case study is a valid and valuable research approach through which to explore and often to explain the occurrence of an emergent phenomenon within a particular social context and against the backdrop of existing theory, especially where there are multiple sources of evidence available (Stake 1995; Yin 2018). This is the overarching goal of my case study, which is to explore and explain the removal of Parler from app stores in January 2021, as an instance of the app store governance of hate speech. I provide further explanation and justification of my decision to study Parler as a case in Chapter 5.

1.3.3 Positionality and ethical limitations

Above I relayed Mahoney and Goertz's claim that the substantive importance of a case can reflexively motivate the decision to study it. Of course, defining "substantive importance" is an inescapably subjective judgment on the part of the researcher proposing to study it (and, ultimately, on the part of those peers who will judge its value). I have already begun to make the case for Parler's importance as a case worthy of objective study, i.e., an object of study. But before taking this justification further in what follows, it is important to briefly reflect on more subjective

[10] This of course assumes that qualitative researchers reflexively engage with the question of *why* they consider a particular case to be important, to which I shortly turn.

considerations, or why *I* specifically might consider the case to be of substantive importance—and to acknowledge some limitations of this conviction.

One factor concerns geography. There is no doubt that the Capitol riots were an event that sparked global attention, but they concerned the (mal)functioning of *American* democracy. Scholars have noted that the attention to questions of platform governance—in both media and research—skew towards cases in the global north (Arun 2021; Bloch-Wehba 2019), threatening to overlook other geographic contexts and undervalue the implications of platform governance decisions on other populations. This geographic skew can have deadly consequences, as in the case of genocide in Myanmar in which UN human rights experts found hate speech on Facebook to have played a role (Miles 2018). It in part reflects the dominance of American companies over platform governance, and, more broadly, the considerable attention paid to American politics, given the country's economic and geopolitical heft. I am not immune to such inequitable blind spots, and indeed, my interest in the Parler case stems in part from my earlier research in the context of American democracy and digital media, particularly on the trajectory of Donald Trump as a digitally facilitated right-populist (Cowls and Schroeder 2018). It also reflects the fact that I am anglophone and have lived, worked, and studied in the United States. Other subjective factors also undergird my perception of the Parler case as being of substantive importance. There is certainly some recency bias at play, in the sense that the Parler episode occurred at the "ideal" time during my doctoral study and emblemised many of the issues that I had already identified as being important in app store governance. In a way, then, it is not simply that researchers choose case studies, but that case studies also choose researchers—a claim reminiscent of Geertz's famous characterisation of "the field" as "assertive, demanding, even coercive" (Geertz 1996:119).

It is also important to reflect at this early juncture of my focus on hate speech as one of the many "problems" for internet governance. More than any other aspect of the study, hate speech implicates my positionality, specifically my status as a white, male, anglophone researcher. As I discuss in more depth later (see Chapter 4), most definitions of hate speech highlight its

embeddedness in social structures that facilitate the domination, exclusion, and marginalisation of members of specific identity groups, including but not limited to people of colour, queer communities, women, and immigrants. And identification with more than one of these identity groups creates the possibility for different and greater forms of intersectional prejudice and injustice to arise (Crenshaw 1990; Matsuda et al. 1993). As Daniels observed a decade ago, typically the "burden of noticing rac[ism] on the Internet has been left to 'minority participants'" (Daniels 2013:712). Countering online hate speech is certainly not a challenge that should not be left to only those subject to it—and in recent years, more scholars have begun to study hate speech. However, these efforts often lack reflexivity and an acknowledgement of positionality on the part of the researchers (Matamoros-Fernández and Farkas 2021). It is therefore important to acknowledge up-front that I do not belong to (m)any of the identity groups under the assault of hate speech in the case of Parler. This inevitably places some a priori constrains on my ability to fully articulate the effects of hate speech on those who belong to groups that are most often targeted by it.

1.4 Overview of argument and structure

In what follows, I provide an overview of each chapter and explain how each one contributes to answering the overarching and subsidiary research questions identified in 1.2. I restate my overarching research question here before introducing my thesis statement. The ORQ is:

> *ORQ: How should app store operators responsibly govern hate speech on social media platforms?*

In response to the ORQ, in the thesis I present and defend three constituent claims, which are as follows:

> (i) Apple and Google accord themselves a responsibility to govern hate speech on social media platforms that their app stores host. *However*,
>
> (ii) this responsibility is not always met efficaciously in practice. *Therefore*,
>
> (iii) this deficiency should be remedied through public policy.

Claim (i) is defended primarily in Part I (Chapters 2, 3 and 4); claim (ii) is defended primarily in Part II (Chapters 6 and 7); and claim (iii) is defended primarily in Part III (Chapter 8).

In the remainder of this section, I provide a chapter-by-chapter overview of the rest of the document.

In **Chapter 2,** I provide important background and context to the thesis by undertaking two important tasks: conceptualising app store governance and situating it within existing theoretical frameworks. I begin at the highest level of abstraction (LoA) that is relevant to the thesis, namely *internet governance*. I note that the term "governance" is often applied to entities other than governments and observe the increasing role that private technology companies play in internet governance. I then turn to *platforms* as simultaneously agents, targets, and sites of governance. I engage with contrasting definitions of what constitutes a platform, and discuss *platform self-governance*, or the governance roles that platform operators play with respect to their users. This in turn raises the question of what makes this form of governance legitimate, which I tie to my focus on the discursive legitimation of Parler's self-governance in the case study.

In **Chapter 3**, I introduce *cross-platform governance,* which refers to the governance that stack providers exert on platform companies. I situate *app store governance* as an instance of cross-platform governance. Through examination of their guidelines for app store governance, I conceptualise Apple and Google as the self-appointed stewards of their respective app stores. The two companies proscribe harmful material in social media platforms that their app stores host, but delegate practical responsibility for content moderation to platforms, adopting what I call a "meta-moderation" role. The nature of this delegation is unclear, and more generally, app store operators face little oversight or accountability for their governance role. I anticipate the implications of this insight for the remainder of the thesis.

In **Chapter 4**, I discuss the governance of hate speech on social media. I begin with a discussion of what hate speech is, and how and on whom it inflicts harm if left unchecked. I assess hate speech in platform self-governance and cross-platform governance contexts respectively, then focus specifically on the app store governance of hate speech. Drawing together conclusions from this and the previous chapters, I address RQ1. I argue that, at present, the responsibility in principle

for app store operators to govern hate speech is a patchwork of ambiguous, opaque, and often implicit obligations and assertions. In short, there is the lack of an overarching framework of responsibility for the app store governance of hate speech. I explain how this finding shapes the case study that follows in the remainder of the thesis.

In **Chapter 5**, I introduce the case study on Parler in 2020. First, I provide a chronological overview of Parler and a systematic literature review of research on it. I then discuss the advantages and drawbacks of taking a case study approach to research and explain why Parler is well-suited for a case study on this subject. I introduce the specific data and methods that I use in the case study, focusing on the utility of the walkthrough method, the analytical focus on discursive legitimation, and the mixed-methods content analysis of hate speech. Finally, I discuss the ethical, social, and logistical considerations that arose prior to and during the case study research.

In **Chapter 6**, I present the first set of empirical findings, which relate to Parler's performance of its platform self-governance. This involves analysis of four datasets as part of the walkthrough method. First, I report the findings of a technical walkthrough of Parler's infrastructural affordances. Second, I examine how Parler's self-governance was portrayed publicly, with a focus on a corpus of 248 news articles containing statements about Parler made during 2020 by the platform's co-founder and CEO John Matze. Using van Leeuwen's framework of discursive legitimation, I explore Matze's use of "constitutional metaphors" to legitimate Parler's "free speech"-centred approach to platform self-governance. Then, I conduct a multimodal discourse analysis of archival copies of Parler's website at various stages during 2020. I find that the centrality of "free speech" to Parler's portrayal of its platform self-governance is visually reinforced by the imagery suggesting Parler was a platform of, by, and for the American "heartlands", in implicit contrast to Parler's "coastal" Silicon Valley rivals. Finally, I analyse a corpus of Parler's publicly available governance documents that were in force at some stage during 2020, which consists of 18 documents of five different types, such as Parler's content guidelines and its terms of use. I again note the prevalence of "free speech" and related discursive tropes in

these documents. Drawing these four analyses together, I conclude that Parler sought to legitimate its approach to the self-governance of its platform by deploying "free speech" as a discursive totem.

In **Chapter 7**, I turn to the second tranche of empirical analysis, which relates to the implications of Parler's self-governance approach for the app store governance responsibilities of Apple and Google in turn. Building on my earlier conceptualisation of Apple and Google's responsibility towards governing hate speech in their app stores, I examine the performance of their "meta-moderation" role with respect to Parler during 2020. I first provide an overview of the extent of hate on Parler, drawing on a mixed-methods content analysis to classify the overall proportion of posts during 2020 that could be considered hateful, and examine longitudinal trends in the data. I supplement this with manual analysis of the sample, which provides additional insight into the nature and common targets of hate on Parler. Then, I synthesise these findings with analysis of Apple and Google's app store guidelines. I conclude that there would have been justification for Apple and Google, on the basis of their own rules, to remove Parler from app stores before January 2021. I then turn to the circumstances of Parler's removal from app stores shortly after the Capitol attack. I critically analyse the justifications that both companies provided for their decision. Given my findings about the extent and nature hate speech on Parler long before January 2021, I show how the Parler case surfaces limitations of app store governance.

In **Chapter 8**, I turn my attention to RQ3, which concerns additional considerations for app store governance. After a synopsis of the argument so far, I reflect further on the structural ambiguities of app store governance, arguing that avenues for legal redress for app removals are even less clear than those for content moderation. I turn to consider recent legislative steps that will affect app store governance, especially the EU's Digital Markets Act and Digital Services Act. I assess the extent to which these laws may address the problems associated with app store governance identified in the previous sections. I conclude with some closing reflections, acknowledgement of the limitations of the thesis, and some areas for future research.

Part I: Theory

Overview

The theoretical part of the thesis, in Chapters 2, 3, and 4, serves two overarching purposes. The first purpose is to conceptually situate the case study on the app store governance of Parler within the complex ecosystem of rules, norms, and expectations that collectively shape the governance of the internet. The second purpose is to answer my first research question:

RQ1: What responsibility do app store operators have to govern hate speech on social media platforms?

This necessitates the articulation and augmentation of existing internet governance theory. Each of the three chapters addresses one of the three core components of RQ1: social media platform governance (Chapter 2), app store governance (Chapter 3), and hate speech governance (Chapter 4).

In **Chapter 2** I discuss the governance of social media platforms by the private companies who operate them. First, I note the increasingly significant role in *internet governance* played by private companies. Second, I turn to the governance role that private platform operators play with respect to their users, which I refer to as *platform self-governance*. Third, I consider the legitimacy and legitimation of this form of governance and explore how platform operators seek to discursively legitimate their governance role through an array of communicative strategies.

In **Chapter 3** I discuss the governance of app stores, which is an example of what I call cross-platform governance. First, I define *cross-platform governance* as encompassing the governance of social media platforms by the companies that provide important infrastructure to them. Second, I conceptualise *app store governance* as an instance of cross-platform governance where an app store hosts a social media platform. I explain how the concept of *meta-moderation* intrinsic to app store governance shapes the structure of the case study on Parler.

In **Chapter 4,** I discuss the governance of hate speech by both social media operators (platform self-governance) and app store operators (cross-platform governance). First, I provide

an overview of hate speech, which covers its core characteristics, how it differs online from offline, and how and whom it most often harms. Second, I discuss platform self-governance in the context of hate speech. Third, I discuss cross-platform governance in the context of hate speech. Finally, I address RQ1 by considering what responsibility app store operators have to govern hate speech on social media platforms that they host.

Thus, by the end of this section part of the thesis, we arrive at the theoretical nub at which the empirical case study occurs. Part I is summarised in Table 1, which shows the key concept(s) that each chapter focuses on, how each chapter relates to a component of RQ1, and how the discussion therein is relevant to the case study.

Table 1: Overview of the theoretical part of the thesis

Chapter	**Key concepts**	**Component of RQ1 addressed**	**Significance for case study**
Chapter 2	Internet governance, platform self-governance, discursive legitimation	Platform self-governance	Situates social media platform as simultaneously governors and governed and explains the focus on the discursive legitimation of platform self-governance
Chapter 3	Cross-platform governance, app store governance, meta-moderation	App store governance	Identifies the meta-moderation model as central to app store governance and structures the case study as such
Chapter 4	The governance of online hate speech	Hate speech	Provides context for the empirical analysis of hate speech

Chapter 2: The "new governors"—platforms as governed and governing

Introduction

The first research question asks about the responsibility of app store operators to govern hate speech on social media platforms. The application of this responsibility in a specific case depends on how the social media platform in question governs its users. Therefore, to answer RQ1, it is first necessary to establish an understanding of the governance activities that social media platform operators themselves undertake. That is the central purpose of this chapter.

The governance undertaken by social media platform operators is an example of the broader shift of internet governance towards the private sector (2.1). Platform operators are simultaneously the agents (platform self-governance) and targets of governance (platform regulation). In some contexts, platform self-governance has become increasingly "state-like" and even "quasi-sovereign" (2.2). This raises the question of whether and in what circumstances platform self-governance activities are legitimate. Legitimacy is a complex concept that can be approached normatively or descriptively. I approach the concept descriptively, and show how, when it is in their interest to do so, platform companies seek to discursively legitimate their governance role, through a variety of communicative strategies (2.3).

I conclude by stating the contributions of the chapter to answering RQ1 and the implications of the chapter for the design of the case study. The governance responsibility of app store operators towards platforms must account for how the operators of these platforms self-govern. How platform self-governance is portrayed is just as important as how it is practiced. Therefore, the case study is designed to collect evidence of how Parler sought to discursively legitimate, as well as to undertake, its platform self-governance.

2.1 Internet governance: the shift to the private sector

The increased involvement of private companies in internet governance has occurred amid a wider reformulation of the meaning of "governance" itself. "Governance" is not a fixed or static term;

indeed, defining it can be "confusing", as it is "notoriously slippery" (Pierre and Peters 2020:1,7). However, this "conceptual vagueness" may be the "secret of its success" (Schneider 2004:25). At the most basic definitional level governance may be said to refer to "the development of governing styles in which [the] boundaries between and within public and private sectors have become blurred" (Stoker 1998:17). That the term has (in English) become unmoored in recent years from its etymological sibling "government" is in part a reflection of broader shifts in how decision-making power is allocated and asserted in an increasingly "globalised" world (Rosenau 2009), as the "institutional strength of the nation state has been challenged from several different sources" (Pierre 2000:1). This has resulted in a paradigm shift within policy studies, which has "undermined what had … been the basis of the model of political governance, i.e., the assumption that the state … is the *control centre of society*" (Mayntz 2003:29; italics added). Both shifts—empirical and intellectual—have led in the same direction: towards analysis of alternative forms of societal governance. This includes, for example, public-private partnerships and the governance role played supranational bodies like the European Union (Kohler-Koch and Rittberger 2006; Pierre 2000; Schmidt 2018). Thus, whereas *government* denotes a specific actor (usually, when unmodified by an adjective, a national government), *governance* is, for many scholars, actor-agnostic (Ansell and Gash 2008; Stoker 1998). Nonetheless, the purpose of governance remains rooted in the "provision of *publicly supported* goods and services", even if the actors providing these goods are not themselves public authorities (Lynn Jr et al. 2001:7; italics added). The semantic gap between *government* and *governance* makes it intelligible to speak of (and to study: Büthe and Mattli 2013; Graz and Nölke 2007; Vogel 2008) "private *governance*" as a form of governance, whereas "private *government*" as a form of government still feels uncomfortably close to a contradiction in terms. Likewise, the influential notion of "governance without government" (Rosenau, Czempiel, and Smith 1992) is understandable in a way that would not hold were the two terms entirely interchangeable.

Therefore, we can think of the semantic unmooring of *governance* from (nation-state) *government* as proceeding along two dimensions: sectorally ("horizontally") away from the state as

"the government" toward "a set of institutions and actors that are drawn from but also beyond government" (Stoker 1998:16), including private actors (Ansell and Gash 2008; Pierre 2000); and geographically ("vertically") away from states as the nuclei of atomistic political communities, toward institutions operating at both supranational and subnational levels (Rosenau et al. 1992). In the top-right quadrant of this hypothetical two-dimensional space, we can situate *transnational companies*, which play in increasingly significant rulemaking role (Strange 1996) in a phenomenon that has been dubbed *transnational new governance* (Abbott and Snidal 2009).

This more expansive and flexible notion of "governance" proves conceptually useful when we turn to consider *internet governance* specifically. The governance of the internet and related digital technologies involves an almost bewildering range of laws, rules, policies, standards, and mechanisms, and engages a heterogenous array of actors, making it "a difficult horse to catch" (Ziewitz and Pentzold 2014:1). Various concrete, albeit minimal, elements of internet governance are formally invested in specific institutions (Van Eeten and Mueller 2013) each with its own criteria for membership, means of decision-making, and organisational culture. This includes the Internet Corporation for Assigned Names and Numbers (ICANN), which assigns domain names and other unique identifiers (Klein 2002); the Internet Engineering Task Force (IETF), which handles the standardisation of the Internet's technical protocols (Cath 2021); and the Internet Governance Forum (IGF), which facilitates dialogue over key internet policy (Epstein 2013). However, as various scholars have noted, in practice much (if not most) internet governance occurs beyond these organisations (DeNardis 2012; Hofmann, Katzenbach, and Gollatz 2017; Kurbalija 2016; Van Eeten and Mueller 2013; Ziewitz and Pentzold 2014)—throwing into question whether internet governance writ large is really as "multistakeholder" as is often assumed (Morar 2018).

Therefore, as with *governance* per se, definitions of *internet governance* are not fixed, debates about it are not static, and neither definitions nor debates can be detached from the historical evolution of the internet as a technically, politically, and culturally complex system (Dutton and

Peltu 2007; Hofmann et al. 2017; Ziewitz and Pentzold 2014). Many early visions of the internet saw it as a space beyond the scope of state authority: consider John Perry Barlow's oft-cited declaration, which taunted "governments of the industrial world, you weary giants of flesh and steel" that "you have no sovereignty where we gather" in cyberspace (Barlow 2016), or Nicholas Negroponte's pointed claim that "the Internet cannot be regulated [… because] the nation-state is not relevant" (quoted in Drezner 2004:481). Less lyrical analyses from the time likewise highlighted some of the key features of the internet, particularly its apparent lack of territorial boundaries, that would seem to undermine the ability of nation-state governments to govern (Goldsmith 1998; Johnson and Post 1996; Lessig 1999). However, as internet use grew in the following decade, nation-states were observed to have "struck back", reasserting their control through legal, economic and technical means (Christou and Simpson 2009; Deibert 2008; DeNardis 2012; Drezner 2004; Goldsmith and Wu 2006; Morozov 2011).

Whereas this first wave of internet governance scholarship crested on the rocks of nation-state power, a second wave of scholarship on governance in the digital age is less concerned with power struggles between the nation-state on one hand and a "self-sovereign" internet, or global community of internet users, on the other. It is more concerned with the governance role played by large, increasingly powerful private companies, most of them headquartered in the US and China (Moore and Tambini 2018; Rahman and Thelen 2019). The growing geographical and sectoral breadth of a handful of these companies has positioned them firmly at the apex of global communication, content, and commerce (Fuchs 2021; Rahman 2017; Srnicek 2017); this has been achieved via an array of corporate strategies, including but not limited to the data-driven profiling of users and monetising of their attention (Fuchs 2012; Zuboff 2019), the devaluation and casualisation of labour (Wood et al. 2019), aggressive acquisition strategies (Perez 2020), and creative tax accounting (Action Aid 2020). The economic heft that such strategies have yielded has subsequently been funnelled into political influence, with the help of large lobbying efforts at national and supranational levels (Nicolás 2021; Swindells and Corbineau 2021), the inculcation of

tacit allegiance of these companies' users (Culpepper and Thelen 2020) and their ability to negotiate (or withdraw from) favourable investment deals with municipal governments (Gupta 2019; Mann et al. 2020).

A recent, especially vivid example of the paramountcy of private companies in the digital sphere is the development of apps used to trace the physical proximity of smartphone owners to one another in order to track the spread of Covid-19. The de facto control exerted by Google and Apple, as operators of the two dominant smartphone software platforms, over national governments' development of apps designed to run on these platforms represents what Tamar Sharon has called "sphere transgression", whereby *domain* expertise (e.g., in public health) is "crowded out" by *technological* expertise (Sharon 2020). Though the decentralised, privacy-protecting model insisted upon by Google and Apple was superficially laudable, the capacity and readiness of these private companies to reject the app features requested by democratically elected governments exposed "a power play between sovereign states and the corporations, in which sovereign states had little say". This reflects a new "governance surplus" held by big tech (Sharon 2020).

The (accurate) claim that *sovereign* states had little *control over* the functions of a *public* health system during a deadly pandemic may seem like a contradiction in terms—or at the very least, it would seem to destabilise the understanding of sovereignty as something that only the governments of nation-states can hold. What makes a governing actor *sovereign* is beyond present scope (c.f. Floridi 2020), but it is increasingly clear that large technology companies occupy what Pohle and Thiel call a "quasi-sovereign" role (2020:7), reigning supreme over the design and deployment of digital technologies in at least some societally significant circumstances. Combined with increasing globalisation, which has resulted in a situation in which, as Hurrell and MacDonald (2012:561) state:

> "States are no longer the sole bearers of the kind of decision-making power that have traditionally been subject to legitimating political control. Instead, non-state actors such as multinational corporations (MNCs), NGOs, and hybrid institutional actors constituted by both state and non-

> state actors, now wield many of these forms of decision-making power … alongside a plurality of states".

This does not, however, necessarily mean that multinational corporations such as technology companies are *legitimate* governors (Buchanan and Keohane 2006; L. Taylor 2021), nor that these private forms of internet governance power are practices that should be unduly reified—considerations that I take forward in 2.3.

In view of the distinction between *government* and *governance*, it is notable that the largest technology companies have begun to rhetorically embrace the nomenclature of the former as well as the latter. Meta CEO Mark Zuckerberg, for instance, has claimed that "in a lot of ways Facebook is more like a government than a traditional company" (quoted in Foer 2017; see 2.3). Meta created a quasi-independent Oversight Board in 2020, which many have dubbed "Facebook's Supreme Court" (c.f. Cowls et al. 2021; Klonick 2021). Moreover, research suggests that large technology companies, including platform operators, are increasingly benefiting from the "tacit allegiance [of] consumers" which can provide "a formidable source of opposition to [state] regulations that threaten" these companies (Culpepper and Thelen 2020:288). And most importantly for this thesis, the powerful role played by the operators of social media platforms in mediating online communication has led to them being referred to as "the new governors" of online speech (Klonick 2017).

In sum, "internet governance" today evidently extends far beyond the multistakeholder fora that are empowered to make decisions about specific aspects of the internet's core functions, to encompass not only the attempts of states to control aspects of internet technology (Deibert et al. 2008), nor the coordinating role played by intergovernmental fora such as the IGF (Epstein 2013; Hofmann et al. 2017), but also the roles increasingly played by large technology companies. This mirrors the semantic journey of "governance" more generally, which has come to encompass much more than the activities of nation-state governments alone. In the years since 1996, when Barlow pitched the battle for control over the internet as between weary governments of flesh and

steel and the internet itself as a "self-sovereign" space, the growing sectoral scope and sociopolitical heft of technology companies has situated them as among the most powerful actors involved in internet governance. Nowhere is this clearer than in the governance of the platforms that several of these companies have created.

2.2 Governing platforms

In this section I introduce platform governance, a subset of internet governance that has been described as encompassing "the layers of governance relationships structuring interactions between key parties in today's platform society, including platform companies, users, advertisers, governments, and other political actors" (Gorwa 2019:855). I consider several popular definitions of "platforms" in the literature. While my present focus is on *social media* platforms, I find that more expansive definitions advance theoretical claims that will also be important to the thesis.

"Platforms", argue van Dijck, Poell and de Waal (2018:2), "have penetrated the heart of society—affecting institutions, economic transactions, and social and cultural practice—hence forcing governments and states to adjust their legal and democratic structures". But what exactly is a platform? Reminiscent of the previous discussion of "governance", the definition of "platform" is contested, contingent and contextual. It has a computational dimension which centres "programmability" (Bogost and Montfort 2009; Helmond 2015). This is encapsulated by Marc Andreesen's definition of a platform as "a system that can be reprogrammed and therefore customized by outside developers [aka] users," and his mantra that "if you can program it, then it's a platform. If you can't, then it's not" (cited in Helmond 2015). This definition seems too restrictive, however, in view of how broadly the term is applied in common parlance today. Few users of social media platforms would consider themselves "developers" or think that sharing a status update or video constitutes an act of "reprogramming". This is not an accident: "platform" was helped to escape from its original definitional straitjacket thanks in part to the strategic

discursive deployment by social media companies, who spied a semantic opportunity to deflect responsibility for content posted on these sites (Gillespie 2010).

To reflect the shift in meaning of "platform" away from merely programmability, consider three more recent, prominent definitions. Tarleton Gillespie (2018:18; italics added) defines platforms as: "online sites and services that host, organize, and circulate users' *shared content or social interactions for them*, without having produced … that content, built on an infrastructure, beneath that circulation of information, for processing data for customer service, advertising, and profit."

Meanwhile, José van Dijck and colleagues (2018:4; italics added), define an online platform as "a programmable digital architecture designed to organize interactions between users—*not just end users* but also corporate entities and public bodies".

Finally, Nick Srnicek describes platforms generally as "digital infrastructures that enable two or more groups to interact", by providing "the basic infrastructure to mediate between different groups" (2017:43–44).

There are risks here of focusing on just a handful of definitions, no matter how prominent. For one thing, each emerges out of the global north and may thus overlook other considerations, values and priorities implicated by platforms elsewhere (Arun 2021; Bloch-Wehba 2019). Moreover, beyond these definitions there exists a larger body of literature which sits at adjacent research fronts such as "platformisation" and "platform studies" (Bogost and Montfort 2009; Helmond 2015; Plantin et al. 2018) and critiques thereof (Apperley and Parikka 2018). But for present purposes, it is nonetheless illuminating to consider the scope and implications of these three prominent definitions of platforms—the distinctions between which suggest that a shared, all-encompassing definition may still be some way off.[11] For his part, Gillespie focuses on what were previously (and are sometimes still) known as *social networks* like Facebook, Twitter, and YouTube. He regards "users' shared content or social interactions" (2018:18) as the core currency

[11] This is not to suggest that arriving at a shared definition is necessarily a desideratum (Ziewitz and Pentzold 2014).

of platforms. By contrast, van Dijck et al. include in their scope not only social media platforms but also what they call "sectoral platforms" that operate "in market sectors such as transport and hospitality [and] public sectors such as education and health" (2018:4). Thus, while van Dijck and colleagues use hospitality site Airbnb as an exemplar of their characterisation of the "platform society", for Gillespie, by contrast, Airbnb and similar "exchange platforms … do *not* neatly fit [my] definition of a platform" though they do face similar "challenges of content moderation in *platformlike* ways" (Gillespie 2018:18; italics added). Srnicek's definition is at first glance the broadest but is narrowed by his identification of the centrality of "network effects" as an "essential characteristic" of digital platforms, and points out that through their intermediary position at the centre of a network, platforms maintain "control and governance over the rules of the game" (Srnicek 2017:47).

These divergent conceptions of what makes a platform a platform in turn affect what the respective authors see as the key sociopolitical considerations that platforms provoke. For van Dijck et al., platforms are not merely "technological tools that allow [users] to do things online [like] chatting, sharing, commenting", and so on; they also "shape the way we live and how society is organized" (2018:9) in sectors and industries as diverse as health, education, and transport. By contrast, Gillespie's conceptual aperture, which sees platforms chiefly as online venues for user-to-user communications—i.e., *social media* platforms—foregrounds the narrower but no less weighty consideration of content moderation. To wit, the final essential element of Gillespie's definition of platforms holds that "platforms[12] *do, and must,* moderate the content and activity of users, using some logistics of detection, review, and enforcement" (2018:21; italics added). Content

[12] There is some discrepancy in the literature about whether "platforms" should refer to the service provided to users by technology companies, or in a de facto sense to the technology companies themselves. And this has become more confusing as single-platform companies, like YouTube and Instagram, have been acquired by companies that operate multiple platforms, like Google (parent company now Alphabet) and Facebook (parent company now Meta). As far as possible in this thesis I follow Gorwa (2019) in using *platforms* to refer to the specific sites and services that are provided by *platform operators*, except where referring to or quoting from existing literature.

moderation "is not an ancillary aspect of what platforms do. It is essential, constitutional, definitional … they are not platforms without it" (2018:21).

Does content moderation as it is characterised here amount to "governance"? There are strong reasons to conclude that it can indeed be characterised as such. First, content moderation can be seen as just as socially important as traditional regulation. As Gillespie (2017:12) argues:

> "[Platform operators'] ceaseless and systematic interventions cut much deeper than the law requires. Both in terms of their impact on public discourse, and for the lived experience of its users, the rules these platforms impose themselves probably matter more than the legal restrictions under which they function."

Moreover, content moderation consists of governance mechanisms (Grimmelmann 2015). As Srnicek observes, platform operators set the "rules of the game" by "having designed [the] core architecture that governs the interaction possibilities" between users (2017:48). However, this inevitably begets questions in turn about *how*, exactly, platform operators should govern in principle and how they do govern in practice. These and related debates have given rise to *platform governance*, a research front at the intersection of law, social science and computer science that is dedicated to asking and answering such questions. In a foundational 2019 article which addresses the question "what is platform governance?", Gorwa characterises the concept as consisting of three core claims: (i) that "platforms govern", (ii) that "platforms are governed", and (iii) that "platform companies are companies" (2019:857).[13] Platform governance as a focus of enquiry, encompassing Gorwa's senses of governance *by* platforms and governance *of* platforms, now attracts a growing community of interested scholars and has its own network and conference.[14]

However, despite this clear typology and growing scholarly interest, as an emerging concept, platform governance is still open to ambiguities with respect to its definition and scope. In light of the ambiguities of "governance" discussed in 2.1, the use of the term (as opposed to,

[13] The distinction between "platforms govern" and "platforms are governed" recalls earlier distinctions, drawn by for example DeNardis and Hackl (2015) and Gillespie (2017), between governance *by* platforms and the governance *of* platforms.

[14] https://www.platgov.net

say, "regulation" or "control") to capture the actors in play and issues at stake in the context of platforms should not be adopted uncritically (Ziewitz and Pentzold 2014). And indeed, as Gorwa's claim that "platform companies are companies" reminds us, conceptual danger arises from taking too far the characterisation of platform operators as governors. The conceptual lines between *governance* and *government* can easily blur, and this in turn may complicate efforts to understand, and critically engage with, "platform governance". In the case of social media platforms, this blurring is, in part, the result of what Kettemann and Schulz (2020:9) succinctly characterise as "a normative phenomenon that is unique to communication practices in the digital age: private actors set the rules for privately owned communicative spaces". This is not, of course, to suggest that social media platform operators like Meta and ByteDance exist in a legal or moral vacuum: they are still subject to external guidelines like the "Ruggie Principles" (see Chapter 4) and have been subject to considerable and growing scrutiny from regulators, politicians, the media, academics, and the public (see, e.g., Paul 2019). But as Kettemann and Schulz state, platform operators are nonetheless left "with substantial freedom to normatively structure their platform through terms of service within the limits of the … national jurisdictions in which they are active" (2020:9).

It is beyond present scope to detail the considerable array of problems, challenges, and scandals that have arisen in recent years with respect to how social media platform companies operate their services, and the impact of their decisions on democracy, human rights, online safety, and much else. But to summarise the argument thus far: on one hand, the scale and scope of platforms in modern digital life is such that, in a purely descriptive sense, platform operators do indeed undertake governance activities with respect to their platforms as communicative spaces. This form of governance—which I hereafter refer to as *platform self-governance*—has significant societal impact. But on the other hand, platform operators are not governments—and seeing them as such may ultimately diminish the ability to challenge these platforms' "governing" role. This creates a conceptual gap. If platform operators are not governments, then what justifies and

stabilises their powerful role? A fruitful way to excavate this tension is through the lens of legitimacy.

2.3 The discursive legitimation of platform self-governance

In this section I examine the legitimacy and legitimation of platform self-governance. I begin by identifying two contrasting paradigms of legitimacy: a "hard" normative/political-philosophical paradigm that tends to assume that political legitimacy is solely or primarily the preserve of the state; and a "soft" descriptive/social-scientific paradigm, which is agnostic regarding who may bear legitimacy, and focuses more on how and why (rather than by whom) it is cultivated (2.3.1). I show how these two paradigms lead to contrasting approaches regarding how to conceptualise the legitimacy of platform self-governance. I then focus on efforts by platform companies to discursively legitimate their self-governance, which I root in the "soft" sociological paradigm for legitimacy (2.3.2).

2.3.1 Legitimacy and legitimation: two paradigms

Legitimacy is a complex and contested concept; "many researchers employ the term, but few define it" (Suchman 1995:573). At a high level, scholars even disagree about whether the concept should be understood *descriptively*—regarding whether, how, and why individuals understand an authority to be legitimate, and/or to act legitimately—or *normatively*, regarding the justification(s) for political authority that those who are subject to the authority ought to accept and thus comply with (Buchanan and Keohane 2006; Peter 2017; Tallberg, Bäckstrand, and Scholte 2018:6; Westergren 2016:46–49). Analogously, a distinction has been drawn across disciplines: between *political-philosophical* approaches, which involve the iterative development of substantive standards by which the relative legitimacy of an actor can be judged from "the outside"; and *social-scientific* approaches, which seek to understand how and why an actor is regarded as legitimate from within

(Beetham 2013).[15] This high-level distinction between normative/political-philosophical (or "hard") and descriptive/social-scientific (or "soft") approaches to understanding legitimacy frames in different ways the question of *who* is, or can hope to be, *legitimate*. For many political philosophers, legitimacy is bound up with the nation state—the nation state is the "locus of political legitimacy", as Nagel (2005:113) puts it. However, within this normative paradigm, some attempts have more recently been made to expand legitimacy beyond the state (Black 2008; Bodansky 1999; Buchanan and Keohane 2006; Hassoun 2012; Hurrell and Macdonald 2012).

By contrast, descriptive approaches see legitimacy as something asserted by an actor, rather than outwardly accorded to them. The high priest of this approach is commonly taken to be Max Weber (Black 2008; Peter 2017), whose *Politics as a Vocation* set out his view of the state as the entity that successfully claims a "monopoly of the *legitimate* use of physical force within a given territory" (Weber 1994; italics added). Weber thus characterises legitimacy as something that can be strategically cultivated and obtained by actors including, but certainly not limited to, public authorities. If we thus turn, as Weber recommends, from seeing legitimacy as something that is awarded extrinsically and a posteriori (and with a whiff of ceremony) like a Nobel Prize, to something that must be outwardly fought for in order to be won, like a boxing belt, we can more readily think of it in terms of being "open to political manipulation" (Tallberg et al. 2018:6).[16] Thus, organisations can "instrumentally manipulate and deploy evocative symbols in order to garner societal support". Legitimacy is an "operational resource … that organizations extract—often

[15] This bipartite distinction between normative and descriptive approaches to conceptualising political legitimacy is the most common in the literature (Black 2008; Peter 2017). A different approach is taken by Franck (1990), who likewise characterises legitimacy as having "many mansions", but who outlines a tripartite schema, orthogonal to the one I outline, which consists of procedural (Weberian), procedural-substantive (Habermasian), and outcome-based (neo-Marxian) approaches.

[16] Another type of terminological molehill that arises in discussions of legitimacy is whether we should be talking narrowly of "*political* legitimacy" or of "legitimacy" more generally. This is complicated further by the substantive nature of the debate between the normative and descriptive paradigms: normative contributions may be naturally more predisposed to refer to *political* legitimacy because the terrain they cover is more definitively *political*, i.e., concerning the state (e.g., Peter 2017). I use "legitimacy" unmodified here, because that is how the bulk of the literature refers to it, but evidently, we are very much in the realm of the political.

competitively—from their cultural environments and that they employ in pursuit of their goals" (Suchman 1995:572–76).

These paradigms for understanding legitimacy per se lead in different directions when evaluating the legitimacy of platform self-governance. Some internet law and communications scholars have sought to develop and apply objective standards to assess the legitimacy of platform companies, for instance in terms of how well they safeguard rights (Suzor, Van Geelen, and Myers West 2018), or procedural values (Suzor 2018), and with respect to their throughput legitimacy (Haggart and Keller 2021). A promising approach is provided by Linnet Taylor, who argues that how these companies are regulated "should be [re]shaped according to how state-like and potentially domination-related their actions are in relation to the public" (L. Taylor 2021:906). Taylor thus argues that we need to "thicken" the concept of legitimacy as it is applied to technology companies who intervene so thoroughly in society and identifies public consent as a potential basis for this, while also noting the problematic absence of genuine (as opposed to merely "tick-box") public consent for the role played by technology companies in modern life. Taylor thus proposes a combination of deliberative processes, specialist watchdogs, and civil society input.[17]

In the specific case of social media platform self-governance, it is possible to foresee direct user involvement in how governance decisions are made, beyond thin "tick-box" consent (and beyond "mere compliance" with public regulation: Floridi 2021a). Over the past decade there have been both conceptual and empirical advances in understanding how this might work. Helberger and colleagues (2018) frame platform self-governance as in essence a question of responsibility, and ascribe a share of responsibility to users themselves, in their capacity to upload and share content. This has echoes of Hofmann et al.'s view of governance as a matter of "reflexive coordination" (Hofmann et al. 2017:1414). However, the concept remains aspirational. As

[17] Taylor's focus on consent and accountability mechanisms, in which both the state and civil society play a crucial role, betrays a (merited) scepticism for governance initiatives that emerge from the private sector, such as "voluntary" codes of conduct.

Helberger et al. acknowledge, it is not necessarily the case that the responsibility of users is balanced or co-equal to that of platforms in practice. It has also been suggested that, even in the absence of direct democracy on platforms, it may be possible to divine the opinions of users about how platforms ought to govern (Myers West 2018; Zarsky 2014). Users are often involved in micro-level enforcement of governance guidelines; indeed, some platforms such as Reddit rely on this as a form of labour (Gillett and Suzor 2022; Matias 2019).

Given its longstanding status as the largest platform, how Facebook involves users in its governance has attracted particular attention. The company's first experiment with platform democracy saw it allow users ostensibly to vote on key policy decisions for the site. This experiment lasted for three years, but ended in ignominy in 2012 when it was voted out of existence, after a too-small minority of the platform's user base (the threshold was defined by Facebook) participated (Leetaru 2019; Robertson 2018).[18] More recently, the company has involved both users and experts in its platform self-governance on what it calls a "proxy" basis. Kettemann and Schulz (2020) provide hermeneutic and ethnographic insights from "first-of-a-kind access to the internal processes of policy development" at Facebook, which documents how outside stakeholders were actually consulted in the process of developing the platform's Community Standards. This was in service of Facebook's stated purpose to base its policies "on feedback from community representatives and a broad spectrum of the people who use our service" through a "proxy-oriented approach" (2020:15-16). Also in 2020, Facebook announced the creation of an "Oversight Board" (OB),[19] a body tasked with providing retrospective judgments about Facebook's content moderation decisions, which Facebook (now Meta) pledged to enforce (Douek 2019; Klonick 2021). Because of its expert membership, appellate function, and claimed independence, the OB has commonly been referred to as "Facebook's Supreme Court";

[18] In absolute terms, the "turnout" of 665,654 people was rather impressive, albeit this was only 0.3% of Facebook's 200 million users at the time.

[19] https://oversightboard.com

I and colleagues found over 300 references to the OB as a "Supreme Court" in mainstream media sources since the idea was first floated by Facebook executives (Cowls et al. 2022).

The phenomenon by which the OB has come to be commonly referred to as "Facebook's Supreme Court" shows how platform self-governance can be legitimated. In the above study, we argue that the evocation of a "supreme court", as a "constitutional metaphor", accords the OB—and by extension Meta—unearned legitimacy in a descriptive sense. Platform operators like Facebook may thus seek to cultivate legitimacy by appropriating the symbols and meanings associated with traditional governance, through employing "constitutional metaphors", such as that of a Supreme Court. However, these companies' embrace of the tropes of government, beyond their inevitable but narrower governance role, should be viewed with cynicism (see also Celeste 2019). Constitutional metaphors raise the spectre of, to paraphrase Foucault, the point at which a particular discourse emerges and comes to be seen as true. If platform operators come to be seen as legitimate decision-makers in their own right—equivalent to, on par with, or in some circumstances even superior to states—the possibility of robust public control over them may be reduced. This, in turn, may allow platform operators to sidestep, downplay, or outright ignore larger structural questions (Bietti 2020; Wagner 2018), such as the need for antitrust measures, or their disregard for users' data privacy, by funnelling attention instead towards important but not (for them) existential content moderation questions (Gillespie 2023).

Constitutional metaphors for platform self-governance like "Facebook's Supreme Court" can therefore be seen as analogical insinuations as much as they are actual implementations of participatory platform governance. The deployment of constitutional metaphors which evoke, but do not necessarily entail, the democratic involvement of users, may prove strategically beneficial, helping Meta to legitimate its self-governance in the eyes of users and others in a descriptive sense, potentially forestalling state-based regulation (Culpepper and Thelen 2020; Wagner 2018). As I will show, attempts to cultivate legitimacy were also pivotal in the case of Parler. It is for this reason that, at the fork in the road between the descriptive and normative approaches to conceptualising

the legitimacy of platform companies (and *pace* the important normative work detailed above), I share Black's (2008:144) view that "it is empirical questions that we need to ask" about legitimacy "before we can approach the normative". Specifically, I adopt a discursive approach to the legitimation of platform self-governance, which I explicate in what follows.

2.3.2 The discursive legitimation of platform self-governance

Social media companies employ various discursive strategies not only to promote their platforms to prospective users, but also to legitimate their governance of the platforms (Haupt 2021; Hoffmann, Proferes, and Zimmer 2018a; Mager and Katzenbach 2021; Natale, Bory, and Balbi 2019; Ziewitz and Pentzold 2014). Such discursive framing extends beyond the slick keynotes and glitzy adverts through which platform companies explain what their platforms are for (i.e., their affordances), to also frame how they can and cannot be used (i.e., their rules). On its face, the proposal to treat not merely the rhetoric, but also the rules, of social media *as discourse* may seem radical. After all, platforms' rulemaking (or lack thereof) has real effects on people's rights, such as free speech, and their expectations, such as safety (Balkin 2018; Duguay, Burgess, and Suzor 2020; Gillespie 2018). Platform self-governance is more than empty gestures. But a discourse-based approach does not deny these real-world impacts of platforms' rules; by conceptualising platform governance as an assemblage of rhetorical constructions, one can gain an insightful view of platform power.

Some work in this area pre-dates the modern platform company. For instance, Timothy Kuhn's "communicative theory of the firm" sees companies as "not merely economic actors or bundles of resources, but as complex social systems that incorporate a variety of interests" (Kuhn 2008:1231). The communicative activities that these companies undertake are "constitutive", in that it is these "very process[es] by which organisations are called into being" (2008:1233). Modern technology companies—and especially social media platform operators in a governance role—are some of the most well-practiced exponents of such discursive construction. Corporate communications by social media companies—such as in the quarterly earning calls that public

companies are obliged to stage, or the legislative hearings at which tech CEOs are increasingly being compelled to testify—garner widespread attention. Moreover, in the case of social media companies, CEOs can make statements directly to users via their own platform. However, the most important discursive "venue" for shaping how technology is understood remains the mainstream media. As Hoffman et al. (2018:201) observe, technology companies, like any corporate actors, may not "hold a monopoly on how technologies are understood, but they are in a position to leverage the power and reach of mass media to promote particular views of a technology", including their own. As we will see, the mainstream media was especially crucial to the discursive construction of Parler in 2020.

Whatever their venue, the corporate communications of technology companies may be thought of as outwardly poetical and implicitly political. CEOs frequently employ grandiose, visionary language to characterise the affordances and social value of their platforms, but this is often designed to serve narrower corporate interests, including to obtain greater legitimacy for their platform governance role (Hoffmann et al. 2018a; Rider and Murakami Wood 2019). When it involves the leaders of companies, it contributes to the construction of what has been called the "business celebrity", which consists of "the orchestrated co-production, cross-promotion, and circulation of images, narratives, and personal appearances of such figures via a wide range of media platforms" (Guthey, Clark, and Jackson 2009:36). The presentation—and confected "celebration"—of business celebrities such as tech figureheads marks "an effort to legitimate concentrated business power by putting a human face on otherwise faceless corporations" (Streeter 2015:3109). This phenomenon, variously referred to as "entrepreneurial storytelling", "cultural entrepreneurship", "corporate narrative construction" and "narrating corporate reputation" (Garud, Schildt, and Lant 2014; Lounsbury and Glynn 2001; Vaara and Tienari 2011; Vendelø 1998), is, again, not novel to the tech world (Beckert 2016; Guthey et al. 2009). But the rapid global growth and intangible, flexible and ever-evolving product offerings of social media companies seem especially well-suited to carefully crafted discursive strategies, as they allow

companies to steer audiences towards seeing these amorphous offerings in specific ways (Cowls and Ma 2023).

With respect to how the operators of social media companies, specifically, seek to legitimate their platform self-governance, much scholarly attention has been paid to Facebook. Several such studies draw on the Zuckerberg Files,[20] a dedicated online archive consisting of "all public utterances" of Facebook's co-founder from 2004 to 2021 (Haupt 2021). Hoffman and colleagues (2018) draw on the Zuckerberg Files to argue that Zuckerberg's utterances are "part of Facebook's strategies for stabilizing the meaning and potential uses of the platform" (2018:200). Other scholars have focused more narrowly on single documents as artefactual emblems of Zuckerberg's broader vision. Rider and Murakami Wood (2019) examine a single 6000-word open letter, "Building Global Community", published by Zuckerberg in 2017, shortly before a series of scandals would rock the company. They argue that Zuckerberg's framing of Facebook as a historically novel global community is deliberately designed to distance his company from its competitors, elide Facebook's status as first and foremost a profit-making enterprise, and shrug off demands for accountability.

Some discursive strategies look backwards, to conjure images or metaphors from the past to explain or justify digital phenomena of the present (Katzenbach and Larsson 2017). For instance, technology companies are often subject to "genesis myths" or "creation myths" (Poulton 2005), which ascribe special reverence to the founders of start-ups that go on to become dominant companies (Natale et al. 2019). A prominent example is Apple's creation myth, which as Belk and Tumbat (2005) document, foregrounds the role of Steve Jobs and to a lesser degree Steve Wozniak as singular innovators. Myths such as these are not necessarily *fictional*—Jobs and Wozniak are formally recorded as Apple's founders[21]—but they may serve to drastically simplify the complex process by which companies emerge. Genesis myths such as these are underpinned by what

[20] https://zuckerbergfiles.org

[21] Though so too is the far lesser-known Ronald Wayne.

Marwick (2013:50) more generally calls Silicon Valley's "almost mythological trust in entrepreneurialism", which, she argues, serves to perpetuate the male dominance at the upper echelons of the tech industry. Genesis myths are especially significant when a company's founder continues to manage the day-to-day operations of the company as CEO (or in the case of Jobs, leaves, then returns to further celebration: Streeter 2015) and to control the direction of their company through dual class share structures that vest them with disproportionate voting power.

Another discursive strategy involves the construction of sociotechnical imaginaries, defined by Jasanoff and Kim (2009:120) as "collectively imagined forms of social life and social order reflected in the design and fulfilment of nation-specific scientific and/or technological projects." Yet as Mager and Katzenbach (2021:225) note, sociotechnical imaginaries today are "multiple, contested and commodified". It is now more often private companies, rather than states, that are the "primary agents of powerful imaginaries", yet these imaginaries are also the subject of contestation (Lehtiniemi and Ruckenstein 2018). In one recent study investigating "Mark Zuckerberg's discursive construction of a better world", Haupt (2021) conducted a discourse analysis of Zuckerberg's public statements, and identified "global community" and "global connectivity" as the predominant themes deployed. Haupt argues that these themes rely on existing values already inherent in Silicon Valley and that they are "explicitly *corporate*" (2021:239; italics in original). By presenting a "teleological view of human history" (2021:251), the purpose of these future imaginaries, Haupt suggests, is in effect to narratively "narrow down the *contingency of the future* in a way that is in line with the company's business objectives" (2021:239; italics added), rendering Facebook a "prophetic corporation" (2021:237).

Scholars further argue that not merely a company's public communications, such as in the media, but also its governance documents such as terms of use and community standards, should be studied as discourse. For Gillespie, "articulating the rules is the clearest opportunity for the platforms to justify their moderation efforts as legitimate (2018:45)." He goes on to characterise social media governance documents as "*discursive performances*, *first and foremost* … that make legible

the anxieties and assumptions of the platform providers" (2018:48; italics added). Likewise, van Dijck and colleagues note the need to see understand platforms' discourses as "performative" (2018:25).

Finally, Ziewitz and Pentzold (2014) conclude their article tellingly titled "In search of internet governance" by asserting the fundamentally performative nature of internet governance—not only as it is undertaken, e.g., by social media companies, but also as it is researched, e.g., by academics. I postpone reflexive considerations about how I view my own role as a researcher to the introduction to the case study (Chapter 5). But looking at the research front of internet governance tout court, Ziewitz and Pentzold argue that "the 'reality' of governance is not given but [instead] achieved in our readings of discussions, seminar presentations, articles or policy briefs". Therefore, internet government studies do not "just point to some otherworldly existence but are situated constructions of our own making" (2014:12).

As this survey makes clear, platform companies adopt an array of communicative strategies to situate, stabilise, and legitimate their platform self-governance role. In the case study I adopt an equivalently expansive approach to analyse the discursive legitimation of Parler's platform self-governance.

Summary

To assess the governance responsibility of app store operators, it is necessary first to understand the governance role that social media platform operators themselves play with respect to their own platforms. Platform operators play a powerful governance role, but the legitimacy of this role is contested. In considering the legitimacy and legitimation of platform self-governance, I have highlighted the fundamentally performative nature of platform self-governance as a form of discursive construction.

This discussion shapes the design of my case study, in which I collect evidence of how Parler sought to legitimate its platform self-governance through an array of discursive strategies.

Compared with mainstream platform companies like Meta, Parler's approach to legitimation was distinct as a result of its smaller size and fringe appeal. But before embarking on the case study, it is necessary to assemble the two other constitutive components of RQ1, starting with app store governance.

Chapter 3: Between the stack and a hard place—conceptualising and situating app store governance

Introduction

Before assessing what governance responsibility app store operators have with respect to hate speech specifically, it is necessary to consider at a more abstract level the nature of their responsibility. This is the task I undertake in this chapter. I examine what has been described as the "expanding content moderation universe" (Busch 2022:32); that is, the emergence of a governance responsibility for the companies that provide the "stack" infrastructure, including app stores, on which social media platforms rely.

Platform governance has tended to focus on two key relationships: between platform operators governing users (platform self-governance); and between platform operators as governed by public authorities (platform regulation). A third form of platform governance, which I call cross-platform governance, arises in cases in which one platform operator governs another (3.1.1). Compared with platform self-governance, cross-platform governance involves different actors, mechanisms, and sites of governance, including app stores (3.1.2), and introduces sociopolitical considerations that are distinct from other forms of platform governance (3.1.3).

App stores control the distribution of software packages, or apps, on digital devices. Apple and Google have long held a strong duopoly over the consumer smartphone market. Though the two companies' approaches to app distribution are superficially different, both exert substantial control over the installation and use of apps on their respective platforms (3.2.1). Three characteristics of app store governance are most relevant to the case study. First, Apple and Google outwardly seek to justify and legitimate their control over app store governance by discursively portraying themselves as stewards of their app stores. Second, the two companies delegate the responsibility for content moderation of the social media platforms in their app stores to the operators of these platforms (i.e., platform self-governance) in the first instance. However, they retain the authority to sanction these platform operators if these moderation efforts are deemed

inadequate. Several ambiguities attach to the application of this "meta-moderation" concept in practice. Third, corporate self-interest dissuades Apple and Google from acting against apps that fall foul of their rules in all but the most extreme cases; excessive sanction against apps invites greater public scrutiny of their platform governance power (3.2.2). I conclude by stating the implications of these insights to the design and scope of the case study.

3.1 Cross-platform governance

3.1.1 Governing the tech stack and the origins of cross-platform governance

The dual view of platform operators *as governing* and *as governed* (DeNardis and Hackl 2015; Gillespie 2015; Gorwa 2019) reflects two visions of what may be meant by the term *platform governance*, with platform operators pitted as either the agent or the subject of governance depending on the context. However, this dichotomy is complicated by cases in which platform operators govern, and are governed, by one another.

As discussed, the focus of platform governance scholarship on the "platforms govern" or *platform self-governance* side has been oriented towards platform operators' governance of their own users' content and communication (Citron 2014; Douek 2022; Gillespie et al. 2020; Grimmelmann 2015; Roberts 2019). This is understandable because in practice, the greater share of governance decisions taken by platform companies are indeed taken with respect to users—whether, as we saw in Chapter 2, on the micro-level of individual content moderation decisions, or with respect to macro-level community standards and terms of use that apply to these users (Kettemann and Schulz 2020; Klonick 2017). And it is sensible from a normative standpoint because of the very high stakes involved in governing individuals' behaviour on platforms, which may allow dangerous content to go unchecked if standards are insufficiently enforced—or may impinge on fundamental rights such as free speech if they are excessively restrictive (Balkin 2017; Kaye 2018). The social and ethical stakes involved with platform operators governing users, in terms of the impact of this

form of governance on the rights, expectations and wellbeing of the individuals and communities that use these platforms, are therefore high.

Meanwhile, on the "platforms are governed" or *platform regulation* side, scholars have so far tended to focus on platforms as the subject of governance by public authorities, such as nation-states and the European Union (e.g., Gorwa 2021; Schulz 2018). This may reflect a latent assumption, rooted perhaps in the conceptual genealogy covered in Chapter 2, that *governance* is interchangeable with (nation-state) *government*. But, as discussed there, from the 1980s onwards the term "governance" began to slacken to include a wider assortment of agents: companies as in "private governance" (Büthe and Mattli 2013), intergovernmental organisations as in "governance without government" (Rosenau et al. 1992), transnational companies as in "transnational new governance" (Abbott and Snidal 2009), and so forth. In principle this definitional shift would seem to offer conceptual space for discussion of private companies governing one another, but (with some exceptions, to be discussed) this does not yet seem to be a major focus of the platform governance literature, where greater attention has been paid to platforms as governed by public authorities and coerced by law (Balkin 2014; Heldt 2019; Kosseff 2019; Schulz 2018).

A third configuration of platform governance, which concerns platform operators governing not end-users, but one another—what I hereafter call *cross-platform governance*—is an increasingly significant aspect of internet governance. It ought to be stated at the outset that many of the most prominent examples of this form of platform governance are the result of dominant companies seeking to gain competitive advantage over rivals, in contexts where the latter are reliant on the former, for corporate reasons. This includes well-publicised spats between Apple and media distribution platforms in the Apple ecosystem, like Spotify (White 2021), Netflix (Welch 2018) and Epic Games (Epic Games 2021), regarding financial and technical issues like in-app payment tariffs and operating system integration. To this we may add controversial practices like companies bundling and pre-installing their own apps on devices (Hardwick 2022), making their apps the default option on operating systems (Gurman 2020), and other anticompetitive steps, such as

Google favouring its own comparison-shopping service in its search results, which resulted in a £2bn European Commission fine (Wakefield 2020). Recalling Gorwa's (2019) observation that "platform companies are companies", it is important not to overlook how these commercial incentives and corporate interests affect the governance activities taken by these large platform companies, particularly when the company being governed is a potential rival, as for example Spotify is for Apple. And the evidence I marshal in my case study will offer reason to think that Apple and Google indeed acted at least partially out of such corporate self-interest, albeit of a different sort, when they decided to remove Parler from their app stores. Equally, however, narrow corporate self-interest is not the *only* motivation behind platform governance, either in platform self-governance (Gillespie 2017) or, as we will see, in cross-platform governance. As Gillespie (2019) states with respect to platform self-governance, platforms "do, and must, moderate"—but *when*, *where*, *how* and indeed *why* they moderate are complex questions that should not be reduced simply to corporate self-interest or mere legal compliance (Gorwa 2021).

In brief (see 3.1.2 for a formal definition), cross-platform governance consists of cases in which the proliferation of offensive material or harmful behaviour on a social media platform falls into the purview of the operator of a different, "upstream" platform, such as an app store. As is the case with platform self-governance, at issue is not typically the *existence* of problematic phenomena on a smaller platform—seemingly every platform is afflicted, to at least some extent, by some form of objectionable content—but the *persistence* of it, resulting from a failure of a platform's own governance, owing to the inability or unwillingness on the part of the social media platform to respond to it.

The platform governance literature, including the various foundational definitions of "platforms" introduced in Chapter 2, provides some pointers for what cross-platform governance might be thought to consist of. As discussed there, in contrast to Gillespie's focus on social media platforms specifically, van Dijck et al. and Srnicek offer broader definitions which encompass platforms in other domains like hospitality (e.g., Airbnb) and transportation (e.g., Uber). In

substantive terms, these other domains are outside the scope of this research, which is concerned more narrowly with hate speech on social media platforms. Yet in a structural sense this literature provides useful pointers regarding how governance can occur across, and between, platforms. van Dijck and colleagues provide a conceptual basis for the governance of smaller platforms by dominant platforms by distinguishing between *infrastructural* and *sectoral* platforms, paraphrasing Orwell in observing that "some platforms are more equal than others" (2018:12). Meanwhile, Srnicek's depiction of platform ecosystems—sustained by network effects and dominated by a single central platform, which creates the architecture of rules that other participants (including both users *and* other platforms) must abide by—is also pertinent to many instances of cross-platform governance, including app stores. Therefore, while encompassing a wider array of platforms beyond social media, these two definitions also provide important conceptual grounding for what follows.

To conceptualise cross-platform governance it is necessary to also draw on literature concerning what is often called the "tech stack". In lay terms, a tech stack can be defined as "the combination of technologies a company uses to build and run an application or project."[22] In the case of a social media platform, then, the tech stack encompasses all the technologies and tools that the operator needs to keep its platform online. If it wishes to make its platform available as a smartphone app on one or both major platforms, Android and iOS, then that company's tech stack necessarily includes the Google Play Store and/or the iOS App Store. Social media platforms have been drawn into myriad controversies concerning their governance of the content on, and the users of, their own platforms. But increasingly, the technology companies that provide infrastructural services to these platforms—such as app stores as well as cloud computing, online payment processing, domain registration, and other functions elsewhere on the tech stack

[22] https://heap.io/topics/what-is-a-tech-stack

(Donovan 2019)—are also being drawn into[23] de facto governance roles when deciding whether to permit or remove offensive sites and platforms on their services (Cath and Ruddock 2022; DeNardis and Bradshaw 2022; Van Dijck et al. 2021).

The "infrastructural turn" in internet governance is nothing new (DeNardis 2012; DeNardis and Musiani 2014; Musiani et al. 2016; Plantin et al. 2018; Plantin and Punathambekar 2019), but controversies regarding the control of the technical infrastructure on which social media platforms rely to function, has taken on greater prominence of late. What DeNardis in 2012 dubbed the "hidden levels of internet control" have, in the decade since, come further into view, owing both to DeNardis' and others' scholarship, and to a series of high-profile flashpoints that relate to privately controlled infrastructure. These have included:

- cloud company Cloudflare's decision to terminate its infrastructure services to the fringe alt-right message board site 8chan, after the posting on the site of manifestos and video footage from three shootings in the United States and New Zealand between March and August 2019 (Benedek and Kettemann 2020; Kelly 2019);
- a similar decision by content delivery network company Diamwall to withdraw services from trolling and doxing site KiwiFarms (Taylor 2022);
- social network OnlyFans' initial decision—later reversed—to ban adult content on its platform after pressure from banking services (Bronstein 2021);
- PayPal's withdrawal of service from 34 hate groups after the deadly 2017 Charlottesville riot (Tusikov 2019b); and
- Google and GoDaddy's decision to terminate domain registration services for the neo-Nazi site The Daily Stormer, also in the wake of Charlottesville (Romano 2017).

These and other high-profile incidents have helped to bring "the turn to infrastructure in internet governance" (DeNardis 2012) into closer contact with the ambit of *platform* governance, not only regarding legal obligations such as intellectual property rights and state requests for censorship (Kettemann 2020:222), but also regarding content moderation. This point of intersection between infrastructural internet governance and content moderation has variously been characterised as

[23] Whether or not this amounts to the conscious co-optation of existing infrastructure for the advancement of political and economic power—in short, whether said organisations are *pulled* into such governance decisions or *push* their way into them—is subject to debate (Cath 2021; DeNardis and Musiani 2014).

"the expanding content moderation universe" (Busch 2022), "widening the lens on content moderation" (Ruddock and Sherman 2021), "navigating the tech stack" (Donovan 2019), "moderation at an infrastructural level" (Gillespie et al. 2020), and "content moderation through internet infrastructure" (Cath and Ruddock 2022).

In their work on alt-right platforms as "parallel ports", Donovan and colleagues trace the creation of "forked" versions of social media platforms to serve the needs of far-right communities that have been "deplatformed" from mainstream social networks. As they conclude (Donovan, Lewis, and Friedberg 2019:62; italics added):

> "No single user or platform can act in isolation given the architecture of the internet. That is to say, while [social media] platforms may be organized as parallel ports, which can function independently of one another, they must be plugged into other internet services such as service providers, domain registrars, and cloud services. As a result, the terms of service for companies that are *deep in the stack* may become the *ultimate arbiters* of what content gets to stay online."

This observation is crucial. Notwithstanding the scrutiny of what can and cannot be said and done on a *particular* social media platform—as determined and enforced by that platform's operator—the rules adopted by commercial stack providers affect multiple social media platforms (even if the *enforcement* of these decisions is often ad hoc and inconsistent in practice). At the same time, technical advances have lowered the necessary level of skill, knowledge and resource that is required to establish a functional social media platform. As such, the ability for particular communities—including unpleasant ones—to establish *their own* platforms, supported by a growing alt-tech infrastructure (Zuckerman and Rajendra-Nicolucci 2021), has increased. In my terminology, this subjects them to *cross-platform governance* (as a platform), instead of *platform self-governance* (as a community on an existing social media platform).

Compared with the content moderation that platforms perform with respect to their own users, cross-platform governance involves a different set of "governors" and a different set of governance mechanisms—distinctions on which several scholars have remarked. In recent work on global platform governance from the perspective of infrastructure, Laura DeNardis and

Samantha Bradshaw (2022; italics added) observe, in the context of moderating disinformation, that:

> "Unlike social media platforms that host highly visible human-facing content, internet intermediaries are *more hidden and much less accountable*, and their moderation practices often *involve greater overreach*, such as blocking access to an entire website, deleting an app or shutting down all mobile and internet traffic in an entire region."

Likewise, as Cath (quoted in Walter 2021) has observed:

> "Sometimes these infrastructure providers make active political decisions about what content should be available online and sometimes we collectively can get behind the politics of their decisions. But the ease with which they can actually preside over online content can nonetheless be somewhat worrying. Especially because these companies tend not to be subject to the same kind of scrutiny that social media platforms are."

These observations show how the stakes of cross-platform governance are in some ways even higher than those of platform self-governance—potentially involving greater overreach, with less oversight.

In a similar distinction, van Dijck and colleagues (2021) distinguish between[24] the technique of *deplatforming*, employed by social media platforms against their own users (e.g., the removal of Donald Trump from Twitter and Facebook), and what they call *deplatformization*, an "implied governance strategy" (2021:3) which they define as "an effort to push back encroaching extreme (right-wing) platforms to the fringes of the ecosystem by denying them access to the infrastructural services needed to function online" (2021:2). In what follows, drawing on this existing literature, I define cross-platform governance.

3.1.2 Defining cross-platform governance

Stated formally, I consider cases to be cross-platform governance when two minimal conditions are met. The first condition is that the *agent* of governance and the *subject* of governance must both be platform providers, the latter of which is a social media platform operator. Thus, a cloud hosting platform withdrawing its infrastructural services from a social media platform would fall under this

[24] C.f. Zuckerman and Rajendra-Nicolucci (2021), who state that "it's not just people who get deplatformed: so do platforms."

definition. However, other governance configurations would not—of which two are worth noting here, for the sake of clarity. Out of scope, first, are cases in which the *subject* of governance is an individual user or community, no matter how prominent (up to and including "influencers", or former presidents,[25] say); this is the realm of platform self-governance. Also out of scope are cases in which the *agent* of governance is not a technology provider. This is the case with, say, a national government deciding to ban an app, such as India's ban of TikTok (Brandom and Christopher 2023). This is the realm of platform regulation.

The second condition is that the *site* of governance must be an integral component of the tech stack—such as an app store, a cloud service, a domain name registrar, or an internet service provider—that a social media platform relies upon to function. In any given case, this component of the tech stack, or "gateway", need not be the only way for a particular organisation to reach a particular audience or segment of internet users; otherwise, cross-platform governance would only apply very narrowly to cases where a single provider has absolute, monopolistic control over some part of the tech stack. That said, we should anticipate that the most socially significant and intellectually interesting cases of cross-platform governance will involve cases wherein the gateway in question provides a common or convenient means to reach an audience; this may be quasi-monopolistic or duopolistic in nature. This dynamic is clearest in the case of smartphone app stores, which are dominated by the Apple/Google "dual monopoly", to which I turn in 3.3. But, as an alternative example, there are a sufficiently small number of robust and reliable cloud hosting and payment platforms that an agreement by all major operators of these platforms would have the de facto effect of taking (and keeping) a particular platform offline. As van Dijck and colleagues put it, although one:

> "cannot merely suspend a platform from all sites … mainstream actors (big tech platforms) can consistently and systematically undermine a platform's material basis for connectivity, hence blocking its entry to data flows that function like oxygen" (2021:6-7).

[25] However, the allusion to Trump here extends only to his use of (and suspension from) mainstream social media platforms like Facebook and Twitter. By contrast, in his capacity as founder of his own platform, Truth Social, he became subject not to platform self-governance but to cross-platform governance (Cowls 2022; see Chapter 8)

Even outside of cases where there is across-the-board consensus about the unacceptability of a particular platform among all feasible providers of a crucial stack component, considerable switching costs attach to platform companies that are forced to switch to less discerning (and often less reliable) service providers at short notice (Chang et al. 2019; Tusikov 2021).

This definition and conceptualisation of cross-platform governance is not wholly novel.[26] In particular, there are similarities between my conceptualisation and van Dijck and colleagues' (2021) concept of "deplatformization". Though the two terms could perhaps be used interchangeably in many cases, I use the terminology of cross-platform governance, as defined above, for two main reasons. First, deplatformization as van Dijck and colleagues conceptualise it is just one act, or at best an inchoate strategy, of governance by major tech companies. A focus on when these companies do "deplatformize" platforms risks overlooking cases in which they *do not* act—or act in other, more nuanced ways, such as Apple's attempted "negotiation" with Parler (see Chapter 5). The more inclusive conception of cross-platform governance allows us to examine instances in which actors take punitive action either too late or not at all, because of competing priorities. Second, van Dijck et al. state that deplatformization is limited to pushing back "radical right-wing platforms". Although my case study concerns a platform, Parler, that fits this description, my conception of cross-platform governance is designed to also encompass cases in which left-wing or other movements are the ones that are pushed back to the fringes. With the ascendancy in Silicon Valley of avowedly right-wing figures such as Peter Thiel (Helmore 2022),

[26] There has long been interest in legal and economic issues adjacent to cross-platform governance. For example, the first great regulatory action against a digital giant, Microsoft, at the turn of the century, was sparked by the issue of its bundling a web browser and media player in its Windows operating system platform, and the effects that this had on competition in the nascent web browser market. More recently, the European Commission has applied the label of "gatekeeper", and with it, new obligations to certain large technology companies, in its economically minded Digital Markets Act (see Chapter 8). And there is already a considerable body of academic legal research that bears upon antitrust concerns in the context of multi-service platforms (Bauer and Prado 2020; Fox 2019; Hylton 2019), but this is mostly an issue of corporate (anti)competitive practices, rather than sociopolitical questions concerning, for example, online safety. Yet a purely legal or economic approach is insufficient to capture the issues related to cross-platform governance. Therefore, much as scholars have already argued for in the case of platform self-governance (Gillespie 2018; Gorwa 2021), one should anticipate a similar need for cross-platform governance to include an array of perspectives and considerations beyond economics and the law.

not to mention the effects of Elon Musk's takeover of Twitter (Shephard 2023) the potential for the pendulum to swing in the other direction should not be overlooked. This is already the case in other jurisdictions, such as Russia, where Apple and Google were forced by the Kremlin to remove the opposition-led Smart Voting app (Cowls and Morley 2022).

With the general meaning and scope of cross-platform governance clarified, one can be more precise about what shape it takes in practice. The core architecture of cross-platform governance (CPG) consists of (i) agents of governance, (ii) subjects of governance, and (iii) sites, gateways, and "chokepoints" (Tusikov 2019a) at which stack providers exert control. I have already established that both the agents and the subjects of CPG are platform providers (in practice these are most often private companies). Sites of CPG include (but are not limited to) app stores, search engines, payment processing services, domain registrars, and cloud servers. *App stores* represent an especially salient instance of CPG, but before turning to this in 3.2, it is useful to outline other key instances. *Search engine services*—most prominently Google—also figure prominently in CPG. In 2015, Google briefly suspended 8chan, a hotbed of objectionable content and communications, from its search index, meaning that a search for anything on the 8ch.net domain retrieved no results (Machkovech 2015). Companies providing *payment processing* and *domain registrar* platforms have also been pulled into governance decisions: PayPal withdrew service to Alex Jones' conspiracist InfoWars site, claiming that the site's "discriminatory intolerance … run counter to our core value of inclusion" (Kan 2018), while domain registrar GoDaddy took down gun forum AR15.com following the Capitol riot, citing the website's violation of GoDaddy's policy on inciting violence (GoDaddy 2021). Finally, *cloud computing* platforms, including Amazon Web Services, have also been pulled into CPG decisions, including in the Parler case (Lyons 2021a); smaller cloud service provider Cloudflare has been the subject of repeated controversies for continuing to provide web infrastructure to violent far right sites like (again) 8chan and The Daily Stormer (Roose 2019).

On a different dimension, one can categorise cases of CPG in terms of the governance mechanisms or strategies associated with it. van Dijck et al. (2021:7) identify three: blocking access to networked distribution; demonetisation; and disabling infrastructural services. Different agents may be able to draw on different mechanisms based on which gateway or chokepoint they control.

Therefore, each example of CPG introduces different considerations, depending on—besides the actors involved—which part of the tech stack or site at which the decision took effect, which mechanism was used (e.g., blocking or demonetisation), and the nature of the policy violation which led to the sanction. (e.g., untrammelled hate speech: see Chapter 4).

Advancing the notion that stack providers govern platform operators is an essential prerequisite for my case study, in which I investigate the withdrawal of a platform, Parler, by the operators of the two major smartphone app stores. When we try to make sense of the Parler case, platform governance in the *platform regulation* sense—which generally assumes that the agent of governance is a government, and that the form of governance is regulation—is clearly ill-fitting: the "governors" in question are private stack providers, not governments. The applicability of platform governance in the *platform self-governance* sense, meanwhile, is in this case limited to the important but by itself incomplete question of how Parler governed its own users. The concept I have developed in this section, *cross-platform governance*, provides important theoretical grounding for my case study. CPG involves elements of both platforms as governors and platforms as governed. Yet as the case study will help to show, CPG is conceptually distinct from both, raising novel sociopolitical considerations and, ultimately, posing new normative challenges.

3.1.3 The sociopolitical stakes of cross-platform governance

Debates over internet governance do not occur in a sociopolitical vacuum (Gillespie et al. 2020; Mager and Katzenbach 2021; Ziewitz and Pentzold 2014). In the case of platform self-governance, the providers of major social media platforms provide services that they claim benefit the public, create public "value", or advance social "good". Yet they are not neutral, and nor are they publicly owned or managed institutions; they are, generally, for-profit private corporations. The friction

between platform operators' allegedly publicly beneficent and privately enriching functions has, in recent years, seemed to come close to being unsustainable. A series of scandals related to data protection and privacy (Cadwalladr and Graham-Harrison 2018), opinion manipulation (Howard 2020), and myriad online harms (Nash 2019) have undermined platform operators' claims to serve the public interest, just as their skyrocketing profits have under*lined* the lucrative nature of their enterprise.

The same is true for cross-platform governance: legal norms, geopolitical factors, and of course economic interests play a role. Consider, for example, the shift in PayPal's stance regarding its service provision in Russia. In 2015—just a year after Russia's invasion of Crimea—the company blocked an account that funded the posthumous distribution of a report by the recently assassinated Kremlin critic Boris Nemtsov, citing that it refused to process "donations for political parties or [opposition] political causes in Russia" (Dolgov 2015). Yet in 2022, following Russia's invasion of further Ukrainian territory, PayPal shut down its entire Russian service in protest (Dave 2022). Similarly, as we have seen, external "real-world" events such as the riots at Charlottesville or the Capitol, or the Christchurch Mosque shooting, have served as a spur for stack providers to respond, which often involves retrospectively enforcing existing policies. This reactive, often reputationally minded approach recurs throughout CPG as a whole and is especially resonant in my case study (see also the discussion of HKmap.live in 3.2.2).

Because CPG involves one technology company exerting a governance role over a smaller one in cases where the latter is reliant on the former, it is tempting to focus on the respective size and scope of platform companies as the key dimension of comparison between them. Take for example the European Commission's recent Digital Services Act (DSA) and Digital Markets Act (DMA). The former introduces a series of "cumulative obligations" based on the size of a platform, and the latter creates a binary "gatekeeper" designation for platforms meeting a certain threshold (European Commission 2023c; see Chapter 8). However, recent history furnishes us with examples of distinctions that have less to do with the size or scope of platforms than with more qualitative

or contextual factors; such episodes caution against viewing platform operators as monolithic or indistinguishable in terms of the ethical stances they adopt. Consider, for instance, Meta's long-simmering battle against Apple, which, to be sure, results primarily from competitive rivalry, but which is also grounded in substantive disagreements over how the privacy of users' data should be weighed against the value of free access to services (Hern 2021). And the rise to popularity of non-Silicon Valley platforms, especially TikTok—a spin-off for the west of Chinese video-sharing app Douyin—introduces the possibility of cultural differences also shaping how platforms govern their users, as exemplified by Mark Zuckerberg's (again, self-interested) warning in 2020 about Chinese-operated TikTok not sharing Meta's commitment to "freedom of expression" (Wells, Horwitz, and Viswanatha 2020). Consider here too the different approaches that platform operators have adopted to deal with the demands of Chinese regulators: Facebook and Google's search engine are banned in the country, and Microsoft has limited LinkedIn's networking features (Weise and Mozur 2021), whilst Apple has acceded to Beijing's demands to store user data in Chinese data centres (Nicas, Zhong, and Wakabayashi 2021). As these examples show, beyond their size and reach, there is much else that separates platform companies in terms of how they both govern and are governed—even though all the companies thus far mentioned would, in view of their size, be subject to the same, highest level of "cumulative obligations" in Europe's proposed DSA, and most would be considered "gatekeepers" in the context of the DMA.

CPG also raises new concerns with respect to substantive political rights. Take the broad debate between, on the one hand, the need to safeguard free expression on platforms, and on the other hand the need to keep people safe online, by for example addressing the harms that arise from hate speech. Such debates are hardly novel to the internet (Matsuda 1993; Waldron 2012) but the stakes may be higher in the case of online platforms, given the anonymity and spontaneity with which (hateful) speech can be freely expressed (Brown 2018; see Chapter 4). In some cases of CPG, however, it is not clear that free expression is the only right we should be concerned with—nor are the traditional mechanisms of dealing with hate speech the same for cross-platform

governance as for conventional on-platform governance. Consider a small fringe platform populated by ideologically like-minded individuals. Were an app store operator to arbitrarily ban this platform from its app store, it is not clear that freedom of expression is necessarily the only right that would be most egregiously violated; indeed, the users of the fringe platform would presumably remain free to use other platforms to express their opinions elsewhere online. It may be the right to free assembly and association of those app users—their ability to assemble in a particular place online and communicate within their own rules of association—that would have been undermined just as much, if not more. And on the other hand, were an app store operator to fail to take responsibility to sanction an egregiously hate-filled app that was being regularly used to stoke trouble and cause harm *outside* the app (offline or elsewhere online), such as coordinating a campaign of online abuse or offline violence, the (unmet) responsibility of the operator differs from a case in which it had (not) banned a user or group from its own platform. Again, stack providers are never neutral, and thus while "some of their decisions will be justified … some will not, [which] leads us into unknown territory" given the possibility of arbitrary censorship that this form of governance may entail (Benedek and Kettemann 2020:146). As the hypothetical example above anticipates, it is into this rather unknown territory that we will soon tread—but it in its present prototypical form this example suffices to demonstrate that CPG invites fresh consideration of existing rights and responsibilities online, as well as potentially surfacing distinct ones.

In sum: platform operators are not simply the agents or the subjects of governance, but instead are participants in a complex ecosystem of rules, norms and practices (Poell et al. 2019), with their own governance decisions undoubtedly guided by corporate self-interest, but also shaped by differences in culture and ideology. As a configuration of platform governance, CPG surfaces distinct sociopolitical considerations and reframes existing debates. This is evident in app store governance as an instance of cross-platform governance.

3.2 App store governance

App stores are a prominent and perhaps even paradigmatic instance of cross-platform governance. App stores, themselves, are platforms (Flew and Martin 2022:6; Gillespie 2018:18; Poell et al. 2019) which control the hosting, marketing, and distribution of "apps" including social media platforms, on consumer devices like smartphones and tablets. The fact that both (i) app stores themselves, and (ii) many of the apps that they host and distribute are both platforms satisfies the first of the two criteria for CPG in the definition I laid out above. Meanwhile, the duopolistic nature of the smartphone market, and the control that Apple and Google exercise over their own smartphone platforms, makes app stores an important if not unavoidable gateway for platform operators who wish to release their platforms as smartphone apps—thus satisfying the second criterion.

In this section I first provide some historical background to smartphone app stores, the longstanding duopoly that Google and Apple hold over them, and the responsibility they have historically taken for governing them (3.2.1). I then surface three key characteristics of app store governance that are most relevant to the case study on Parler.

3.2.1 Apple and Google as app store duopolists

One of the most important affordances of modern personal computing devices is their ability to run software developed by parties other than the manufacturers of the hardware or developers of the core operating system (OS) software. Historically, this affordance has not been without controversy. For instance, two of the most significant legal cases of the early phase of personal computing turned on whether Microsoft was justified in bundling its Internet Explorer web browser and its Windows Media Player with its dominant Windows operating system (Economides and Lianos 2010). In both cases, Microsoft's bundling was alleged to have advantaged its in-house software, for example by giving Internet Explorer a foothold over rivals such as Netscape in the so-called "browser wars" of the 1990s (Windrum 2004). Likewise, the more recent rise of smartphones that can download and execute third-party software packages, or *apps*, has prompted debate over the appropriate level of control that smartphone manufacturers and OS developers

ought to hold, and the responsibilities they ought to bear, concerning what users can do with their devices (Beres 2021). However, in contrast to the allegedly monopolistic desktop era, the smartphone era involves a *duopoly*, held by two US-based "big tech" companies, Apple and Google, whose software powers 99.3% of smartphones worldwide.[27]

App stores are platforms whose operators, such as Apple or Google, host—and, via their review processes, govern—other platforms in turn, such as Snapchat or Instagram. App stores are thus the main interface between device owners and third-party smartphone app developers, which manifests both economically as a two-sided market, and politically as a governance framework structuring interactions between these groups (Poell et al. 2019). On the governance side, app store operators have the capacity—and have on occasion shown the willingness—to reject or remove social media platforms from app stores due to the perceived inadequacy (or absence) of those platforms' own content moderation practices.

Though borne of quite different philosophies about the smartphone experience, both Apple's iOS and Google's Android platforms have an official app store from where most apps (on Android) and virtually all (on iOS) are downloaded. Because it does not permit users to download and install apps from elsewhere, Apple has what it claims is a strenuous process for reviewing every app submitted to the App Store; Google also reviews apps submitted to its Play Store, while still permitting the "side-loading" of apps to its devices outside its official app store. Both companies reserve the right to suspend and remove apps that breach respective guidelines from their app stores, and both have shown a willingness to do so, including for popular social media platforms and similar apps (Hestres 2013). For example, Apple removed blogging platform Tumblr in 2018 after an audit found child sexual abuse imagery on the service (Porter 2018); Google removed Gab, a social network popular among the far right, in 2017 for violating its (Google's) hate speech policy (Price 2017); and Apple removed a platform designed for vaccine

[27] See https://gs.statcounter.com/os-market-share/mobile/worldwide.

sceptics in 2021 (Hill 2021). And as far back as 2012, both companies removed a streaming app released by the Lebanese political and militant group Hezbollah (Cooper 2012). These cases constitute an indirect form of content moderation that we can call "meta-moderation", wherein app stores set standards for content moderation and assess whether a subject platform meets these standards (see 3.2.2).

As Nieborg and colleagues note, "despite the ubiquity of app stores"—and despite the aforementioned high-profile controversies—scholarship on them has been slow to emerge" (Nieborg, Young, and Joseph 2020:2; c.f. Hestres 2013). Some exceptions are provided by scholars of platform studies, such as Dieter and colleagues, who characterise app stores as "key gatekeepers" for the accessing of apps, and further note that each app store has its own "affordances, built-in logics, and mechanisms" (Dieter et al. 2019:2–3). Poell and colleagues also use app stores as an exemplar of platformisation (2019). Although Dieter et al. point out, rightly, that Apple's and Google's app stores are just two of many online app repositories—alongside app stores for desktop devices and game stores, for example—in what follows, I focus solely on the iOS App Store and Google Play Store. This is for three reasons: the sheer volume of smartphone apps (a reflection of the ubiquity of smartphones) compared to apps developed for other devices such as tablets or smartwatches (with fewer users); the duopoly that Apple's iOS and Google's Android together hold in the smartphone software market; and the dominant position that the app stores occupy within their respective smartphone operating systems. A short history of each smartphone app store is provided.

Apple's first iPhone device, which was announced in January 2007 and released later that year, launched without an app store. The company's initial focus was on its pre-installed apps, with third-party developers limited to developing web applications to be accessed through the in-built Safari web browser (Apple 2008). The ability for users to install and run third-party apps arrived a year later with Apple's launch of its App Store for the iPhone's iOS software (henceforth "App

Store")[28]. Since then, App Store apps have steadily gained greater affordances within iOS, including integration with the notification system and access to a user's location data, camera, photos, etc. Apps are lucrative for Apple: the App Store allows developers to submit apps for iPhone users to download, either for free or for a fee, and Apple takes a 30% cut from App Store revenue raised.[29] The company has received criticism and, in some jurisdictions, has faced legal action (including in the Netherlands: Khalid 2022; Sterling 2022) regarding the excessive revenue cut it takes from developers. It has also been alleged that Apple systematically disadvantages rival software developers through onerous restrictions (Robertson 2021), such as preventing developers even from informing users that they can secure a cheaper subscription rate (i.e., minus Apple's fees) if they sign up for their service elsewhere (e.g., through the open web). Apple's iOS is also designed to prevent users from "sideloading" apps—that is, installing apps to their devices in any way other than through the official App Store (e.g., direct downloads from the web, from another device, or through a third-party app store), marking a restrictive approach to device management.[30] Facing renewed pressure to open up its app architecture, Apple published a "threat analysis of sideloading", in which the company claims that enabling the practice "would cripple the [iPhone's] privacy and security protections", suggesting that users would be left more exposed to harmful apps (Apple 2021:2). Apple claims that its longstanding resistance to allowing sideloading owes to its allegedly safer and more secure App Store—though such claims must be viewed sceptically, in the context of the significant revenue Apple makes from the cut it takes of developers' app income. Apple claims that it "review[s] every app before it becomes available on the App Store to ensure

[28] Versions of Apple's App Store app are also pre-loaded on its tablet, desktop, watch and TV device software, but only the apps that each device is capable of installing are presented to users. Therefore, for simplicity, I treat the iOS App Store as distinct. However, it is important to note that the removal of an app from one app store has the effect of removing it from all of Apple's platforms.

[29] In 2017 Apple updated its subscription model: thereafter, it would receive only a 15% cut for subscriptions lasting more than a year. In 2020 it lowered its cut to 15% for all revenue raised by app developers making less than $1m USD per year.

[30] This is not to suggest that installing third-party apps or app stores on an iPhone is impossible—only very cumbersome. "Jail-breaking" an iPhone gives device owners considerable free rein over what software they can install, but according to Apple violates the device warranty. Other approaches, such as altstore (altstore.io) also exist, but involve several limitations and difficulties.

it is free of malware and accurately represented to users, and [that it] swiftly remove[s] apps from the App Store if they are found to be harmful" (Apple 2021:4). It has also taken a strong stance against apps containing content such as pornography and the promotion of drug use. Indeed, former CEO Steve Jobs claimed that Apple had a "moral responsibility to keep porn off the iPhone" (Siegler 2010). In this way, Apple has made itself, in Gillespie's words, "the marketplace, the clearinghouse, and the editor's desk, all in one place" (Gillespie 2018:81). However, independent analysts suggest that, despite these supposedly careful controls, the App Store is still littered with malicious apps, such as those that scam unwitting users out of hundreds of dollars (Hollister 2021; Lin 2020). Despite modest concessions of late, such as reducing its revenue cut, generally the company has retained its tight control over how users install third-party software on their devices. Because the official App Store remains[31] the only straightforward way for users to install apps, this raises the sociopolitical stakes for Apple's responsible management of its App Store (Gillespie 2018). If an app is rejected or removed from the App Store, it cannot be used on iPhones, potentially limiting users' access to specific forms of information or communication, depending on what the app offers. At the same time, the admission and retention of an app in the App Store implies Apple is satisfied that the function and content meet its standards for safety and social acceptability, raising questions when "approved" apps inflict harm on users or societies.

The Android mobile operating system has a roughly 70% global smartphone market share to Apple's 29%. It is borne of a superficially different philosophy. Android's core code is based on open-source software, and it is developed and maintained by the Open Handset Alliance (OHA), a consortium of over 84 software and hardware manufacturers, which announced Android in 2007 as the "first truly open and comprehensive platform for mobile devices" (Open Handset Alliance 2007). From the start, however, Android has been dominated by Google. The Android name is trademarked by Google, and most Android smartphones come pre-installed with Google's

[31] This is set to finally change with Apple allowing third-party app stores: see Chapter 8.

proprietary apps (Edelman and Geradin 2016), such as its browser Google Chrome and, most significantly, Google's app store, known as the Play Store. Though Android users are, unlike iPhone users, permitted to sideload apps from outside the Play Store, the process is inconvenient, and the ability to do so is disabled by default. Moreover, members of the OHA are forbidden from forking Android to create alternative versions, as Amazon (a non-OHA member) has done with its FireOS. This approach resulted in the $5.05bn fine handed down by the European Commission to Google in 2018 (Brodkin 2018), on the basis of Google's requirements that manufacturers pre-install Chrome as a condition for licensing the Play Store (and the millions of apps it offers), as well as the inducements it offered manufacturers to pre-install Google apps and its prevention of other manufacturers from pre-installing apps on devices running other Android forks (European Commission 2018). Even since the Commission's record fine, Google has continued to run afoul of authorities. Later in 2018, the third-party app store operator Aptoide won a case against Google in a Portuguese court which ordered that Google to stop removing Aptoide's app store from Android devices, which Google had been doing on alleged "safety" grounds (Reuters 2018). A year later, Aptoide claimed that Google was still systematically disadvantaging its rival app store, by flagging it as "unsafe" and forcing users to change multiple system settings just to allow the app store to function (Lomas 2019). Much like Apple, Google has historically taken a 30% revenue cut on digital purchases made through the Play Store (albeit offering similar concessions in recent years),[32] whereas Aptoide only takes a 15% cut.

Despite the superficially divergent approaches to device management taken by the world's two dominant smartphone operating systems with respect to sideloading, we can see significant commonalities between them. Apple and Google both exert considerable control over their respective architectures, albeit in ways that are commercially and technically different. Apple takes a "hard" approach to prevent sideloading, by making it technically impossible for users to do

[32] In 2021 Google followed Apple in halving its revenue cut to 15% on the first $1m of Play Store purchases on a per-developer basis.

without foregoing the warranty of their smartphone devices. Google takes a superficially "soft" approach, using default settings and warnings to dissuade users from doing so; however, it sometimes also acts unilaterally to remove third-party app stores altogether, as in the case of Aptoide. This has landed each company considerable criticism and, in the case of Google, sizeable financial penalties; Google's willingness to incur these penalties further demonstrates the importance that the company places on its app store as central to the Android experience.

Taken together, the ubiquity of smartphones, Google and Apple's shared dominance over smartphone software, and the centrality of app stores that these companies enforce on the smartphone experience suggest that how these companies *govern* their app stores is a subject of considerable legal and political importance (Hestres 2013; Hofmann et al. 2022; Nieborg et al. 2020; Poell et al. 2019). This form of internet governance encompasses the overarching policies and guidelines that determine what sorts of apps are allowed, and the processes put in place to ensure that these standards are enforced. Unsurprisingly yet rather belatedly, the dominance of Apple and Google within their respective smartphone software architectures has begun to attract greater attention from legislators, including those in Europe and the United States, which I discuss in Chapter 8. A remaining question is how to conceptualise app store governance.

3.2.2 Conceptualising app store governance: stewardship, delegation, self-interest

App store governance seems to pose a conceptual challenge. App store operators like Apple and Google clearly have the capacity to refuse or remove apps that do not conform to their respective guidelines or terms of service. But when, how and why do they intervene to enforce these rules in practice? In this section I characterise the governance of smartphone app stores as consisting of the interplay of three concepts: stewardship, with respect to how the companies portray their governance role publicly; delegation, with respect to how they undertake their governance role in the case of social media platforms; and self-interest, with respect to the underlying motivations of their governance role.

First, an assessment of how Apple and Google portray their role in managing their app stores finds that they rhetorically embrace a *stewardship* model for app store governance (ASG). Both companies justify their restrictive approaches on the basis that the feasibility and sustainability of their smartphone systems, and the safety of their users, rely on well-managed app stores. Take Apple first. At the time of writing,[33] the strapline fronting the App Store section of Apple's website (aimed at users, not developers) promises "the apps you love. From a place you can trust".[34] The App Store, Apple claims, "has proved to be a safe and trusted place to discover and download apps", thanks, we are told, to Apple "ensuring that the apps we offer are held to the highest standards for privacy, security and content." By touting its privacy, security, and content standards, the implication is that downloading apps via the App Store provides a far safer experience for users than downloading software from elsewhere (which, of course, users cannot really do). The first line of Apple's Introduction to its App Store Review Guidelines restates this notion to third-party app developers, stating that "the guiding principle of the App Store is simple—we want to provide a safe experience for users to get apps and a great opportunity for all developers to be successful" (Apple 2022). This appeal to "safety" also recurs throughout Apple's App Store documentation; most of the rules relevant to objectionable content are under the section of the Guidelines titled "Safety" (see Chapter 4).

With its claims that its purpose is to ensure "safety" and "trust", Apple frames its role with respect to the governance of its app store as a stewardship responsibility. In general terms, stewardship can be thought of as consisting of "the careful and responsible management of something entrusted to one's care" (Merriam Webster 2023). It has been applied to as diverse contexts as the management of cultural artefacts and health care, but is perhaps most often used with respect to the natural environment (see e.g., Berry 2006). In different ways, both Apple itself, and its detractors, often use environmental metaphors to discursively portray its role viz-a-viz its

[33] Apple's rhetorical presentation of its services is wont to change.
[34] https://www.apple.com/uk/app-store/; accessed 24th March 2023

App Store. Apple refers to the App Store as a "trusted ecosystem" (Apple 2021)—while the term "walled garden" is commonly used by outsiders to characterise, and criticise, the control that Apple exerts over its hardware and software (e.g. Bohn 2021). Whether to legitimate or to criticise Apple's role, the invocation of such environmental metaphors creates an implicit delineation between "in here" and "out there" with respect to Apple's package of hardware and software. Rhetorically, Apple can thus be seen as claiming a stewardship responsibility for what is and is not allowed inside its walled garden, legitimating its responsibility.

Google's approach to managing its smartphone software is, as we have seen, at least superficially less restrictive, insofar as device owners may install software packages downloaded from outside the Play Store as apps. But this does not stop Google from using similar—in fact, almost identical—terminology to characterise its ASG role. The strapline on the homepage of Google Play's Developer Policy Center is "providing a safe and trusted experience for everyone" (Google 2023a), while Google has likewise used terms like "ecosystem" to refer to its smartphone software, and has positioned its Developer Guidelines as "rules of the road to keep the store secure" (White 2022). Thus, Google shares with Apple a reliance on safety and trust as rhetorical touchstones, and likewise uses the environmental term "ecosystem" to justify its stewardship role, in addition to using the metaphor of built infrastructure in its promulgation of the "rules of the road".

To be sure, "safety" and "trust" are more than just discursive calling cards for Apple and Google, just as "connection" has salient meaning for Facebook. Both companies do, in fact, dedicate significant resource to protect the users of their devices, under the rhetorical guise of stewardship. But, just as with Facebook (Rider and Murakami Wood 2019), it would be naive to accept stewardship at face value as the sole modus operandi in Apple and Google's approach to app store governance. How the two companies handle the presence of *social media platforms* as apps in their app stores is especially instructive. It is one thing for an app store operator to review a relatively "static" app such as, for the sake of example, an app which lets users play a simple

noughts-and-crosses game. This sort of app can be reviewed when the app is submitted to each app store, and again each time an update to the app (e.g., with the addition of functionality for a checkers game). By contrast, social media platforms provide a much more complicated challenge for ASG, insofar as the "content" that appears on the app is constantly being generated by its users. It would obviously be infeasible for app store governors to review every piece of content submitted to a social media platform, which is a vexing enough challenge for social media platform operators themselves. Instead Apple and Google adopt a delegated model, which we can call *meta-moderation*, to handle this task.

In this approach, the companies have specific rules for the moderation of what they call "user-generated content" in apps that feature it. In its App Store Review Guidelines, Apple notes that "apps with user-generated content present particular challenges", and insists that such apps include mechanisms for filtering objectionable material, for reporting offensive content, and for blocking abusive users from the service" (2022:S1.2). Each of these stipulations obliges platform companies to add features or mechanisms to their own platforms, in effect pushing the burden of content moderation (though the term is never explicitly used) onto platform companies themselves, or, in cases where platforms already have content moderation systems in place, demanding that these be maintained. By implication, it is only if platforms are unwilling or unable to discharge these responsibilities that Apple will (in theory) intervene. Google's rules around user-generated content are similar, but, if anything, more demanding. Its policy states that platforms that must "implement robust, effective, and ongoing user-generated content moderation", and specifies functions and mechanisms to those demanded by Apple. It goes further than Apple by explicitly stating that platforms' rules that "define objectionable content and behaviors … in a way that complies with [the] Google Play Developer Program Policies".[35]

[35] Both companies' policies are critically examined in greater depth in Chapter 7.

This meta-moderation model is a crucial aspect of app store governance as it relates to objectionable content on social media. But before turning to the app store governance of hate speech specifically, it would be remiss not to consider a final important factor, namely *self-interest*. Both companies are keen to highlight the economic benefits that can accrue to app developers from well-designed apps. The strap line on Apple's App Store landing page for app developers is "Together we turn apps into opportunities", advertising the App Store as "help[ing] you … to grow your business", and claiming that "Apple has paid out over $320B to developers selling digital goods and services since 2008."[36] What goes unsaid here is the cut that Apple and Google, as middlemen, takes from each transaction between developer and user. Operating an app store is a lucrative endeavour for both companies. In the fourth quarter of 2022, Apple broke its revenue record for its "Services" sector, which includes App Store tariffs. At $20.8bn, its revenue from Services alone accounts for more than one-sixth of its overall revenue (Apple 2023)—an increasingly important source of revenue for the company given the material costs of its hardware devices. Google parent company Alphabet's revenue is still dominated by its search advertising business, but its "Google other" sector, which includes Google Play revenues as well as its home devices and other subscription services, still accounted for around 10% of its overall revenue in 2022 (Alphabet 2023).

This scale of profit-making explains why the companies have so doggedly defended their app store business models. Apple was, at one stage in 2022, willing to accept a fine of €5m for ten weeks in a row, levelled by Dutch antitrust regulators, for refusing to allow third-party payment options for dating apps (Reuters 2022). Apple appeared to relent by allowing developers of dating apps in the Netherlands to charge users using third-party payment systems. But it came with a catch: Apple still demanded a 27% cut of the transaction. This was only a small discount on its standard 30% cut, an amount which Apple said "excludes value related to payment processing and

[36] https://www.apple.com/app-store/developing-for-the-app-store/

related activities" (Dillet 2022). Apple thus claims that, less the cost of payment processing, the intrinsic value of its app store management efforts amounts to 27% of transactions. Of course, Apple has not been obliged to test whether the market agrees with this price-point; users and developers have had no choice but to accept this cut if they wish to buy and sell iPhone apps. In any case, one might expect that a 27% cut would buy well-defended walls for Apple's garden.

Corporate self-interest motivates app store governance in other ways, beyond the direct profits that spring from facilitating transactions between app store developers and users. For different reasons, Apple and Google both benefit from (at least) maintaining or (ideally) growing their share of the smartphone market. For Apple it is especially apparent why: its iPhone line brings in extraordinary sums, by itself accounting for more revenue than all of Apple's other products and services combined (Apple 2023). Google also manufactures and sells its own Android smartphones through its Pixel line. But it benefits more substantially from the continued ubiquity of its search engine advertisement business on Android phones. Fining Google €4.34bn in 2018, the European Commission found that Google had "use[d] Android as a vehicle to cement the dominance of its search engine" on smartphones (European Commission 2018).[37]

With the maintenance of this dominance in mind, Poell and colleagues note that, as app store operators, Apple and Google "constantly tinker with their governing instruments to keep end-users and complementors tied to the platform" (Poell et al. 2019:8). Noting Google's tendency to tinker with how it ranks apps, they argue that "for developers and end-users, the Play Store is [thus] a typical black box, as app[']s rankings are based on opaque and largely invisible algorithms". But another form of "black box" are the "often opaque moderation practices" that both companies undertake (2019:8).[38]

[37] Meanwhile, with respect to the share of the smartphone market that it does not more directly control, Google pays Apple an estimated $15-20bn per year to remain the default search engine on iPhones (Moreno 2021).
[38] This "dual opacity" of recommendation and moderation systems for app stores, as a form of cross-platform governance, is analogous to the interplay of opaque systems for recommendation and moderation that individual platforms deploy in platform self-governance.

Corporate self-interest extends beyond the safeguarding of revenues that flow directly from app store transactions. Apple's removal of Hong Kong mapping service HKmap.live from its app store is a case in point. The app, which gathers reports on police patrols in Hong Kong, was removed from the App Store in October 2019 amid mass protests in Hong Kong against a bill to amend extradition procedures. In a statement explaining the removal, Apple claimed that "the app has been used to target and ambush police [and] threaten public safety"; this was vehemently refuted by the app's developer (Mihalcik 2019). Apple was roundly criticised for the move, with critics noting the apparent double standards concerning the inclusion of app Waze, which also shows the location of police, in app stores. A contemporaneous report observed that "no multinational company arguably has as much at stake in China as Apple, which assembles nearly all its products there and counts the country as its No. 3 market after the United States and Europe" (Nicas 2019). As this episode demonstrates, political and commercial considerations evidently play a role in app store governance, even when decisions can be ostensibly justified in terms of guidelines. This suggests, in short, that the rules of app store governance are just as performative as those of social media platforms.

The three characteristics of app store governance highlighted here—the characterisation of app store operators' role as stewardship, the delegation of content moderation responsibilities to platforms, and corporate self-interest—cannot be easily disentangled in practice. But each one affects how the two companies address specific forms of problematic content, including hate speech.

Summary

In this chapter I have introduced cross-platform governance, which encompasses the governance of social media platform operators by the platform companies that provide essential infrastructure to them. App store governance is one of several instances of cross-platform governance. Three key characteristics of app store governance that are relevant to the case study are the stewardship

role that Apple and Google discursively construct for themselves, the meta-moderation approach that they adopt for social media content moderation, and the corporate self-interest that motivates their continued dominance of smartphone software.

The identification of the meta-moderation model structures my dual focus in the case study, wherein I examine the intersection of two modes of platform governance: Parler's platform self-governance and Apple's and Google's cross-platform governance of Parler. I assess what the retention of Parler through 2020 and its eventual removal in January 2021 reveal about the circumstances in which, and reasons for which, Apple and Google intervene in app store governance.

This chapter has outlined the governance responsibility that app store operators adopt with respect to the social media platforms that they host. This leaves the elaboration of the third central concept in RQ1, namely hate speech.

Chapter 4: The governance of hate speech in app stores

Introduction

In this chapter I consider the challenge that hate speech poses for internet governance generally and app store governance specifically. Together with the theoretical development of the previous two chapters, this allows me to answer the first of my three subsidiary research questions:

RQ1: What responsibility do app store operators have to govern hate speech on social media platforms?

Hate speech involves the intentional expression of a negative sentiment against a group identity; the internet allows for the anonymous, instantaneous, and widespread expression of hate speech (4.1.1). Hate speech can inflict harm directly by causing fear and trauma; by inspiring physical violence; and by silencing the voices of those targeted (4.1.2). In the US context, the amorphous alt-right movement has used internet technologies to monetise and amplify harmful hate speech (4.1.3).

The rise of online hate speech implicates the operators of social media platforms in speech governance. Laws proscribing hate speech differ across jurisdictions, but transnational platform companies have signed up to voluntary agreements and standards-setting initiatives to define and deal with hate speech. The longstanding shield against liability for illegal content has recently been challenged by the German NetzDG law. Platform operators also generally prohibit hate speech in their extra-legal terms of use and community standards (4.2).

Stack providers are becoming increasingly implicated in hate speech governance responsibilities, as individuals and communities "deplatformed" from mainstream platforms seek refuge in more distant corners of the internet. Each form of harm that hate speech can inflict creates responsibilities in principle for stack providers who indirectly sustain its expression; these must be set against other claims and interests (4.3).

Through their app store governance guidelines, Apple and Google have accorded themselves a responsibility to govern hate speech by (i) prohibiting it outright in apps that their

stores host and (ii) obliging apps with user-generated content to undertake content moderation. However, the extent of this commitment is substantively and procedurally ambiguous, and cannot be easily disentangled from the two companies' self-interest. In answer to RQ1, I characterise the governance responsibility of app stores towards hate speech as a patchwork of rules and rhetoric that is ambiguous in terms of its definition, implementation and motivation (4.4).

In concluding Part I of the thesis, I turn from answering RQ1 to explaining how this answer motivates and shapes the empirical case study which follows in Chapters 5 to 7.

4.1 Online hate speech: what is it, who does it harm, and how?

4.1.1 What is (online) hate speech?

Preventing the spread of hate speech on the internet is a major challenge for public authorities, nongovernmental organisations, and technology companies alike (Citron 2014; Citron and Norton 2011; Daniels 2013; Matamoros-Fernández and Farkas 2021). In 2019, in a blog post entitled "The wildfire of hate speech", the United Nations Secretary-General decried "a menacing wave of intolerance" online and appealed for "an all-out effort to end hate speech globally" (Guterres 2019), marking the launch of the UN's Strategy and Plan of Action on Hate Speech (United Nations 2019). Meanwhile, at a domestic level, civil society organisations such as ADL in the US (2022) and the Centre for Countering Digital Hate in the UK (2021), have campaigned for platforms and public authorities alike to take greater action to tackle the spread of hate speech online and its harmful effects. The extent of online hate speech only worsened during the Covid-19 pandemic, with charity Ditch the Label finding a 28% increase in ethnicity-based and racist hate speech between 2019 and 2021, amidst a pronounced rise in anti-Asian hate speech in particular (2021). A nationally representative survey found that 65% of Americans who belong to marginalised groups received hate-based harassment in 2022 (ADL 2022).

Any attempt to proactively detect and remove online hate speech first requires a clear understanding of what it is, and the circumstances in which it is prohibited by law (Cowls and

Vidgen Under Review). The extent to which national law prohibits hate speech varies by jurisdiction, even among ordinarily like-minded liberal democracies, reflecting differences in history and culture (Howard 2019). Typically, hate speech laws only offer protection against overt and/or inflammatory forms of hate speech. This may include expressions that explicitly incite hatred or violence against a group, as in the UK (The Law Commission 2014); expressions advocating genocide, inciting hatred that breaches the peace, or wilfully promoting hatred, as in Canada (Government of Canada 2023:§318-319); or (inter alia) expressions that maliciously malign segments of the population, as in Germany (Federal Ministry of Justice 2021:§130). Some states criminalise certain expressions for historical reasons, such as the ban on Holocaust denial in Austria (*Verbotsgesetz* 1945; revised 1992). By contrast, in its approach to hate speech, the United States has taken "the road least travelled" by adopting a form of "free speech exceptionalism" (Sumner 2013:377) which protects hate speech in almost all cases (Howard 2019). The protection from government restrictions on free expression broadly is enshrined in the First Amendment to the US Constitution; and in *Brandenburg v Ohio* (1969), a ruling that still has effect (Parekh 2006), the US Supreme Court ruled that the First Amendment protects even speech that would inspire criminal violence.

Though these cross-national differences are significant (Gagliardone et al. 2015), intergovernmental bodies have sought to create more consensus around what constitutes hate speech, in order to coordinate more effective responses to it at the international level. In a 2019 report, the United Nations acknowledged that "there is no international legal definition of hate speech", but provided its own working definition (United Nations 2019:2) as:

> "any kind of communication in speech, writing or behaviour, that attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of who they are, in other words, based on their religion, ethnicity, nationality, race, colour, descent, gender or other identity factor."

Similar clarificatory efforts have been made under the auspices of the Council of Europe (Weber 2009), while the European Commission in 2021 announced plans to include hate speech and hate

crime as "EU crimes" (European Commission 2021) without providing a precise definition.

Meanwhile, there is likewise no clear definition of hate speech in academic scholarship (Siegel 2020); hate speech is "anything but simplistic" (Kirtz and Talat 2023:468). Descriptions of it often underscore the subtle and complex ways in which hate can be expressed through speech (Parekh 2006). Simpson (2013:701–2) describes hate speech as "identity-prejudicial abuse and harassment, certain uses of slurs and epithets, some extremist political and religious speech … and certain displays of 'hate symbols'". Simpson suggests that "we classify such activities as hate speech if, and insofar as, they convey the idea that belonging to a particular social group warrants someone's being held in or treated with contempt." Nockleby offers a similar definition, defining hate speech as "any communication that disparages a person or a group on the basis of some characteristic" (2000:1277).

Several points of agreement can be found across these legal and academic definitions of hate speech, despite the lack of an overarching consensus definition (Sellars 2016). First and most straightforwardly, hate speech involves the expression of something negative—such as contempt, disparagement, derogation, demonisation, insult, animosity, hostility, or bias (Delgado 1982; Parekh 2006). Second, hate is directed against the identity of a group (Marwick and Miller 2014; Parekh 2006; Rieger et al. 2021). Identities can be understood as the social groups and affiliations that individuals belong to and are associated with. Identities are often complex, shifting and contingent, and individuals typically hold multiple, intersecting identifications at once (Roccas and Brewer 2002:89), creating the potential for unique forms of intersectional prejudice and injustice to arise (Hancock 2007). For instance, a Black woman may suffer forms of abuse which are irreducible to her identity as either a woman or a Black person alone (Crenshaw 1990; Matsuda et al. 1993). Third, the expression of hate is an intentional act, insofar as the speaker intends to express something negative, and intends for it to harm: for example, to "ambush, terrorize, humiliate, [or] degrade" (Matsuda et al. 1993:1). This is an important consideration because space must be made for democratic debate and free expression in civic discourse, or what Imhoff and

Recker describe as "legitimate critique" of certain groups (2012).[39] In practice, hate speech is highly context-dependent (Delgado 1982; Parekh 2006) and the line between hate and not-hate is often unclear; furthermore, establishing a definitional boundary which is too broad may mean that the term loses its analytical power (Fortuna and Nunes 2018; Salminen et al. 2018).

The apparent difficulty of establishing a clear, universally applicable definition of hate speech in the abstract has significant implications for how it is detected and dealt with when it is expressed and spread on the internet (Gagliardone et al. 2015; Siegel 2020; Waseem 2016). In addition to disagreements over what constitutes hate speech per se, further debate considers the ways in which *online* hate speech differs from its offline equivalent. In answering this question, Brown (2018) identifies several distinctive features of online hate speech, including (i) the ease with which purveyors of hate can access audiences (because it is straightforward and 'low cost' to post content online); (ii) the potential size of the audiences they reach; (iii) their anonymity (i.e., they do not need to reveal any personal information); and (iv) the instantaneousness which with hateful content can be communicated (see also Rieger et al. 2021). The ease with which certain content can reach large audiences has been examined in other areas of internet studies, as in "context collapse", wherein content is shared in settings far beyond the original intended audience (Marwick and boyd 2011). And, to Brown's second point, because content can go "viral", spreading far beyond the original intended audience, hate that is aimed at "just" one person may be seen by many (Hasinoff and Schneider 2022). This can be unintentional, such as when hate expressed in a small conversation between like-minded people is shared more widely. Or it can be intentional. Compared with the logistical barriers of offline content creation, it is easy for individuals to create and share hateful content online, often requiring just a few clicks (Brown 2018). Bots may also be used to scale up the spread of this content (Ziems et al. 2020), creating further challenges for content moderation (Gorwa et al. 2020). Thus, even though a large

[39] This should not be taken to suggest that unintentional expressions cannot also be harmful in some circumstances.

proportion of hateful content may not reach a sizeable audience online, this may be counterbalanced by the far higher rate at which it can be produced, as well as how long it can stay online (Gagliardone et al. 2015).

Anonymity (and pseudonymity, when individuals adopt alternative or fake identities: Farkas, Schou, and Neumayer 2018) is another important factor in online communication, as it creates distance between the speaker and the audience (Brown 2020). Citron and Norton contend that the anonymous and pseudonymous nature of social media discourse "can just as easily accelerate destructive behaviour as it can fuel public discourse" (Citron and Norton 2011:1477), as a result of a so-called 'disinhibition effect' (Wachs and Wright 2018). Brown suggests that disinhibition occurs because the internet enables individuals to engage in more spontaneous and unconsidered communications, with little time spent reflecting on how their behaviour might impact victims (Brown 2018). Experimental research suggests that disinhibition in online environments may arise chiefly from a lack of eye-contact between users (Lapidot-Lefler and Barak 2012).

In addition, many forms of online communication have little or no offline precedent, either in terms of the forum (e.g., pseudonymous online servers, such as Discord) or the medium (e.g., gifs, emojis (Kirk et al. 2022), and face filters: Nast 2017). This presents new challenges for conceptualising, analysing, and identifying online hate. For instance, memes are a multi-modal form of communication which combine an image with text. Many are innocuous, while others contain and convey prejudiced ideas in certain communicative contexts or when combined with certain words (Kiela et al. 2020; Lamerichs et al. 2018; Winter 2019). Finally, the architecture of social media platforms inescapably affects how and by whom hate speech is both encountered and countered (Gillespie 2018; Matamoros-Fernández and Farkas 2021; McMillan Cottom 2020; Roberts 2019).

Drawing these definitions and distinctions together, *online hate speech* can be defined as a communication on the internet which intentionally expresses negative sentiment against a form of

identity. Online hate speech can take the form of, among other things, derogatory, demonising, and dehumanising statements, threats, identity-based insults, pejorative terms, and slurs. With the key components of online hate speech identified, we can turn to consider the ways in which it is harmful.

4.1.2 How and on whom does online hate speech inflict harm?

Whether or not harm is a necessary component of hate speech remains subject to a degree of academic disagreement (Marwick and Miller 2014; Parekh 2006; Sellars 2016) which I do not seek to resolve here. What is clearer though is that hate speech, including online hate speech, *can* be harmful in at least some cases, and that harm can be inflicted in various ways. Hasinoff and Schneider distinguish between two related ways in which harm can arise from online hate speech: "individually targeted harassment" on one hand and "harmful content … targeted at a group or at no one in particular" on the other (Hasinoff and Schneider 2022). Similarly, Hope Not Hate's 2020 report The Good Web notes the potential for harms from hate speech to both "occur on, or [be] facilitated by, the online world" (Hope Not Hate 2020:6). The UK Law Commission's 2020 scoping report on hate crime similarly identifies three core motivations for countering hate speech: the prevention of "emotional and psychological harms of hate speech", the "protection of groups from violence", while also highlighting the importance of the "prevention of the social exclusion and marginalisation of vulnerable groups in society" (The Law Commission 2020:69).

This suggests several possible taxonomies for conceptualising the harms of online hate speech: between harm arising directly online, and harm arising indirectly offline; or between verbally (Matsuda et al. 1993) and violently (Benesch 2012) assaultive speech, for example. The taxonomy I use takes inspiration from Woods (2020), who identifies "three broad categories of harm" that can arise from dangerous misinformation online: *direct harm* to those exposed to or targeted by it; *indirect harm* arising from a change of behaviour triggered by exposure to it; and *harms to society* at large. This tripartite framework can be adapted to identifying the harms of hate speech.

First and foremost, online hate can inflict instantaneous emotional harm on victims (Delgado 1982)—both on individuals who are directly and deliberately targeted, and on anyone else who views the content incidentally (or as part of their job: Roberts 2019). This can cause an immediate, unpleasant and deeply affective feeling of fear and anxiety: a metaphorical "gut punch" (Matsuda 1993:23). Being targeted by online hate can be highly affective for victims, given the fear that further negative consequences may arise, the evocation of a previous negative experience such as a physical attack (which may have been experienced by the individual or by people they know), and/or being reminded of one's negative or subaltern status in the context of broader historical inequity and oppression (Delgado 1982). Sadly, this is nothing new: studies suggest that experiencing racism is associated with substantially worse physical and mental health outcomes (Anderson 2013), and that lesbian, gay, bisexual and trans people can suffer from "social stress", which can even result in suicide (Meyer 2003). But growing social media use creates further opportunity for racism, homo-, bi-, and transphobia to be expressed—while at the same time providing supportive venues for identity expression (Miller 2017).

Second, being targeted by hate speech online can also lead individuals to suffer longer-term forms of harm—both mental and physical. People who are victimised can suffer ongoing emotional and mental health challenges, such as continued experiences of anxiety and trauma, with negative long-term effects (Delgado 1982; Wypych and Bilewicz 2022). Research highlights, for example, the stark impact of Islamophobic hate on British Muslims. A 2017 review from the anti-Islamophobia charity Tell MAMA reported that "fear of victimisation can limit the activity of British Muslims who may avoid using public transport, leaving home, or even leaving the neighbourhoods in which they feel safe" (Tell MAMA 2017:11), findings echoed by a report by the Muslim charity MEND (Ingham-Barrow 2018). This impacts not only individuals who have personally been targeted by online hate but also other members of their communities. Anyone who bears the identity of a group which suffers widely from discrimination and prejudice knows that they could be attacked at any time. Other work likewise shows the effects of hate speech on

the mental health, life choices and wellbeing of those targeted (Dickter 2012; Reichelmann et al. 2021; Tell MAMA 2017; Twomey 2017). Crucially, online hate can instigate physical violence and abuse offline (Awan and Zempi 2016; Müller and Schwarz 2021; Williams et al. 2020). Susan Benesch has examined how speech becomes dangerous, showing that certain discourses have encouraged mass acts of violence, such as the Rwandan Genocide and the Holocaust (Benesch 2012). In terms of social psychology, Bilewicz and Soral find that "exposure to derogatory language about immigrants and minority groups leads to political radicalization and deteriorates intergroup relations" (Bilewicz and Soral 2020:3; see also Soral, Bilewicz, and Winiewski 2018). This highlights the very real risk of physical violence resulting from online hate speech.

Third, online hate speech creates harm at the societal level. As a result of experiencing online hate speech, members of targeted groups often do not feel safe joining public spaces or contributing to forums of civic discourse. This is costly to them, but also robs society of their participation. As one respondent to a survey by EVAW and Glitch! about experiences of online abuse during COVID-19 wrote, "if [social media] wasn't [also] used to connect to some good voices and work, I would leave" (End Violence Against Women Coalition and Glitch UK 2020:28). Similarly, female politicians report that they have been targeted by online abuse, much of which is hateful in nature, leading them to consider leaving frontline politics (Gorrell et al. 2018; Margetts, Vidgen, and Hale 2019). In this sense, the sad fact that victims of hate speech may be "daily silenced [and] intimidated by racist assailants who employ words and symbols to suppress and subordinate" (Matsuda et al. 1993:7) does not only harm those targeted—it impoverishes democratic debate and representative government. This poses some interesting implications for debates over free speech and open debate. It may be, for instance, that the absence from debate of groups victimised by online hate, owing to intolerable harassment, is more harmful to the overall health of democratic debate than would be the removal from debates of the purveyors of online hate speech—often decried as "censorship". Free speech, it could therefore be argued, must be protected for groups who are so frequently targeted by online hate that they withdraw from public

debates, at least as robustly as it is for individuals who repeatedly engage in it. Similarly, the European Court of Human Rights has found that Article 17 of the European Convention on Human Rights, which prohibits the abuse of rights, can be used as a basis to exclude certain forms of hate speech from freedom of expression protection (Benedek and Kettemann 2020:149).

The likelihood of and extent to which each of these three forms of harm arise from any given expression of hate speech online is highly dependent on context (Kirtz and Talat 2023). Thus far, the discussion has remained at a high level with respect to the different forms of identity that hate speech may target, which includes gender identity, sexuality, and religion (Sellars 2016). In practice, however, much scholarship on online hate speech concerns an especially prevalent form of hate, namely racism, targeted at specific identities. In the US context, racist hate speech has become associated with the so-called "alt-right".

4.1.3 Hate speech and the alt-right

In *Excitable Speech: A Politics of the Performative*, Judith Butler argues that the impact of hate is a product not only of what was said, but also of the relation between (i) the identity of the speaker and (ii) that of the victim (Butler 1997). Butler argues shows that the force of hate is a result of the broader social and linguistic structures in which it is embedded (Livingstone and Das 2013). The importance of context is reflected in the use of slurs and pejoratives, which are often used to express deep animosity against targeted groups; as Matsuda et al. put it in their work on "words that wound" (see also Delgado 1982), hate speech is an "injury to a group. To privatise it ignore[s] the greatest part of the injury" (1993:8). Butler likewise argues that although they are "just words", the harmful power of hate speech comes from the broader social and historical structures in which these words are embedded.[40] She contends that "the racial slur is always cited *from elsewhere*, and in

[40] It is important to also acknowledge that hateful terms are often reclaimed by those who have either personally been targeted by hate or who belong to communities which have been targeted (Cepollaro and Zeman 2020). Such language has a fundamentally different meaning when it is reclaimed and used by the targeted groups: it is no longer hateful as the reappropriation "undermin[es] the signal strength of the slurring term" (Bolinger 2017:18). This of course has implications for accurate hate speech detection (Waseem 2016).

the speaking of it, one chimes in with a *chorus of racists*". This, in turn, creates an "imagined relation to an historically transmitted community of racists" (Butler 1997:81; italics added). Therefore, although the extent to which harm is experienced will depend to a degree on the target's circumstances and life experiences, the potential for harm relies on reference to a shared social and historical context of discrimination and prejudice.

How has the internet transformed the expression of racist hate speech? The connections between networked online communication and racism have been the subject of considerable scholarship over the past two decades (Daniels 2009, 2013; Donovan 2019; Munn 2019; Phillips 2015; Whine 2003). This research has shown how white supremacist groups have adopted new technologies "to amplify, monetize, and mask their ideologies" (Ma 2021; see also Ganesh 2020; Massanari 2017). Far-right actors are motivated not only by their political agendas but also by the online "attention economy", which rewards "engagement" and thus incentivises the creation of sensational, provocative, and hyperactive content (Daniels 2018; Harsin 2015; Kirtz and Talat 2023; Persily 2017; Vaidhyanathan 2018). The affordances of the internet and the incentives of social media have allowed this "community of racists" to be much more than (to paraphrase Benedict Anderson: 2006) just imagined, as users assemble to organise, and "troll" other internet users in racist and sexist ways (Ortiz 2020). A recent systematic literature review of racism on social media concludes with a plea for future research to interrogate "how race is baked into social media technologies' design and governance rather than just focusing on racist expression in these spaces" (Matamoros-Fernández and Farkas 2021:218). This conclusion builds on important prior work which highlights how the ill-considered design of sociotechnical systems like social media platforms as well as search engines (Noble 2018) can promote and support "toxic technocultures" (Massanari 2017). The "baking in" of race and racism to social media design and governance motivates my study of Parler's attempts to discursively legitimate its "free speech" approach. This follows research into such "dark platforms" as 8kun and Gab, which can be conceptualised as "digital platforms that are less regulated and moderated, hence can be used for hosting content

and content creators that may not be tolerated by their more mainstream counterparts" (Zeng and Schäfer 2021:1321)

In recent years, terms like "alt-right" and "alt-lite" have emerged as distinguishing markers for various online far-right "communities", especially white supremacists, in the US context. Hawley (2018) dates the first reference to "alt-right" to the title of a talk by far-right ideologue Paul Gottfried in 2008, which was followed by the launch of the website AlternativeRight.com by fellow far-right figurehead Richard Spencer in 2010. As Hartzell (2018:19) observes, Spencer's website focused on "arguing for the importance of embracing pro-white racial consciousness." Building on the legacy of white supremacist organising through the 20th century, the "alt-right" embraces a "victim ideology" that frames advancements won by Black people, and other people of colour, as discriminatory attacks on the rights of white Americans (Berbrier 2000; Ma 2021), even raising the false but inflammatory prospect of "white genocide" (Ganesh 2018). While many self-identified members of the "alt-right" are explicitly white supremacist, the term has also been used in journalistic and scholarly accounts to describe a broader swathe of reactionary actors online. And some of the less extreme personalities in this fractious coalition have in turn been deemed "alt-lite": a pejorative term coined by "alt-right" figures to refer to those who broadly align with their ideology but do not openly advocate for white nationalism (Ma, 2021).

The emergence and evolution of the "alt-right" is just one of the myriad manifestations of racist hate speech online, many of which—notwithstanding the Anglospheric skew of scholarship (Matamoros-Fernández and Farkas 2021)—occur in contexts outside the US (e.g., Fink 2018; Miškolci, Kováčová, and Rigová 2020). Wherever it arises, networked racism today implicates platform governance.

4.2 Platform self-governance of hate speech

The shift to the internet of speech in general and hate speech in particular inevitably implicates technology companies, especially social media platform companies, in speech governance

(Alkiviadou 2019; Balkin 2004; Citron and Norton 2011; Klonick 2017). As Suzor and Gillett recently put it, "platforms play an important but fraught role in setting and enforcing the boundaries of acceptable speech" (Suzor and Gillett 2022:272). Established requirements for private operators to counter hate speech on their platforms have traditionally taken the form of private standards-setting (see below) accompanied by voluntary agreements, rather than legal requirement. This reflects the broad liability exemption carved out in Section 230 of the US Communications Decency Act 1996 that exclude them from treatment as "publishers" of their users' content, alongside equivalent protections in other jurisdictions (Schulz 2018); it may also reflect the sheer difficulty in legislating for the governance of complex problematic material like hate speech online (Ganesh 2018; Nash 2019). In one prominent example of a voluntary agreement to counter hate speech, in 2016, the operators of major social media platforms signed up to a voluntary Code of Content on Countering Illegal Hate Speech Online with the European Commission. The Code requires removal of hate speech as defined by the European Union—raising questions about both the appropriateness and the applicability of a European definition of hate speech to worldwide platforms (Aswad 2016). Separately, the Parliamentary Assembly of the Council of Europe has also called for companies to adopt self-regulatory codes of conduct in support of freedom of expression online (Benedek and Kettemann 2020:85). Platform operators, as companies, also have a duty to uphold the UN Guiding Principles on Business Human Rights (informally the "Ruggie Principles"). The Principles state that "business enterprises should respect human rights", which entails that "they should avoid infringing on the human rights of others and should address adverse human rights impacts with which they are involved" (United Nations 2019:13). In the context of platforms this responsibility involves ensuring that terms of service and other user-facing policies respect human rights (Aswad 2016). The principles are only "cautious and aspirational" (Benesch 2020:93), but they serve the purpose of tying the conduct of business to the UN's longstanding human rights instruments, not least the International Covenant on Civil and Political Rights, which provides inter alia for bans on speech that advocates "national,

racial or religious hatred that constitutes incitement to violence, discrimination or hostility" (United Nations 1966:§20). Finally, Google, Meta and Microsoft (but not, Amazon, Apple or Twitter) are members of the Global Network Initiative, a voluntary cross-sectoral group which aims to prevent online censorship by governments by safeguarding free expression, and to protect privacy online.

To these voluntary international initiatives and platform self-governance rules has more recently been added, in some jurisdictions, proposals for de novo legislation to *compel* platforms to deal with online hate speech. Germany's Network Enforcement Act (NetzDG), which came into effect in 2018, has been described as "probably the first regulation in the world to directly proscribe how large user-generated content platforms moderated harmful content" (Gorwa 2021:2). The relevant political context to the law included rising hate speech against minority groups online following the arrival of a large number of Syrian refugees to Germany in 2015-16 (Heldt 2019). The NetzDG requires large social media platforms (those with over 2 million users) to remove evidently unlawful content, including hate speech, from platforms within 24 hours and all other illegal content within seven days following the receipt of a complaint (Schulz 2018).[41] The definition of hate speech for the purposes of the act is deferred to existing German penal code.

Initial interpretations of the NetzDG, prior to its implementation, noted its potential incompatibility with human rights, especially freedom of expression (Kaye 2017). Critics noted that the short turnaround time that platform operators are given to remove manifestly illegal content risks incentivising platforms to adopt a "just in case" approach by taking down flagged content when in doubt (Schulz 2018). The law also contains a reporting requirement, but an ex post analysis found that the "transparency" reports on content take-downs that platforms were required to provide had "low informative value", thus making it harder to work out the extent to which expressive rights were being impinged (Heldt 2019). Another analysis of the early period of

[41] The NetzDG does not include app stores in its remit.

the NetzDG's enforcement likewise criticised the inconsistency of reporting, and found that hate speech was the most common basis for removed content (Tworek and Leerssen 2019). Characterising the NetzDG as "an incremental, narrow approach to a complex sociotechnical problem", a more recent policy analysis calls for a more holistic regulatory approach to online hate speech, more geared towards public interest considerations (Griffin 2022). This would involve, for instance, giving platform operators the ex-ante responsibility to "discourage" objectionable content instead of just the ex-post requirement to delete it.

Though the NetzDG remains the most prominent domestic regulatory framework for tackling online hate speech, other jurisdictions have adopted laws with similar provisions, despite significant controversy. France's "Loi Avia", which had similarities to the NetzDG but focused solely on online hate speech, was passed in 2020, but was struck down by France's Constitutional Council shortly thereafter for violations of proportionality. Earlier versions of the UK's Online Safety Bill—which has had a rather Homeric journey[42] to adoption (Nash 2019)—originally provided for restrictions on "lawful but harmful" online speech which have since been dropped (Vallance and McCallum 2022). Finally, the European Union's recently implemented Digital Services Act also accords obligations for online speech governance; I reserve discussion of the Act until Chapter 8.

Above and beyond provisions regarding manifestly illegal content, platform operators themselves, especially those that operate mainstream platforms, also prohibit "lawful but harmful" hate speech in the context of platform self-governance. For example, Meta's current[43] policy regarding hate speech on Facebook and Instagram declares simply that "we don't allow hate speech on Facebook" (Meta 2023). It goes on to define hate speech as:

[42] As yet incomplete.

[43] It is worth noting that these policies are subject to constant revision, and Meta provides a changelog for its policy on hate speech stretching back to their initial introduction in 2018.

> "A direct attack against people—rather than concepts or institutions—on the basis of what we call protected characteristics: race, ethnicity, national origin, disability, religious affiliation, caste, sexual orientation, sex, gender identity and serious disease."

Meta in turn defines "attacks" in the context of hate speech as:

> "Violent or dehumanising speech, harmful stereotypes, statements of inferiority, expressions of contempt, disgust or dismissal, cursing and calls for exclusion or segregation. We also prohibit the use of harmful stereotypes, which we define as dehumanising comparisons that have historically been used to attack, intimidate or exclude specific groups, and that are often linked with offline violence".

Google likewise states that "hate speech is not allowed" on its YouTube video platform, and pledges to "remove content promoting violence or hatred against individuals or groups based on" any of a series of identity attributes including age, ethnicity, and sex/gender (Google 2023b). By contrast, the operators of smaller platforms, especially those catering to a political or ideological niche of users, often adopt starkly different approaches to policing hate speech. Indeed, they sometimes reject altogether any responsibility for doing so, even making this part of their "USP" (see Part II and Trujillo et al. 2020; Zannettou et al. 2018; Zeng and Schäfer 2021).

However, it is important to keep in mind that mainstream operators like Meta and Google, who have rules in principle against hate speech, face competing incentives and technical challenges when implementing these rules. For one thing, the common business model for social media platforms involves keeping users "engaged"—that is, keeping them on the platform, where they can be advertised to. But the use of algorithmic recommendation systems that are optimised to keep users engaged may have the inadvertent effect of "recommending" increasingly hateful or radicalising content (Tufekci 2018) (albeit empirical studies refute such claims: Hosseinmardi et al. 2021; Ledwich and Zaitsev 2019) among a host of other technical and ethical challenges associated with content recommendation (Milano, Taddeo, and Floridi 2019). Moreover, the use of algorithms for content moderation, including hate speech detection, also poses weighty technical and political challenges (Gorwa et al. 2020; Waseem and Hovy 2016), despite ongoing improvements in what is an active area of research (Röttger et al. 2021; Vidgen et al. 2021). In practice, effective content moderation also relies on an enormous, under-resourced and often

mentally endangered workforce. A 2022 *Time* investigation, titled "Inside Facebook's Africa Sweatshop", documented how Kenyan employees of a company to whom Meta outsources its content moderation are paid as little as $1.50 per hour, and are frequently victims of trauma, intimidation, and suppression of attempts to unionise (Perrigo 2022).[44]

These two sociotechnical architectures—determining what users can and do see, and what they can't and don't—are, therefore, often at cross-purposes. But crucially, this is a set of problems of the platforms' own making. In a recent short article, Tarleton Gillespie (2023) urges reflection on "the fact that content moderation exists". Gillespie considers what the very need for content moderation to tackle problematic content like hate speech implies, and who it implicates—including non-profit groups, content creators, fact-checkers, and regulatory agencies. Gillespie warns that taking too much of the fact of content moderation as a given is "an enormous mistake" (2023:2). Platform companies are, after all, founded on the "engagement" of their users, and, as Zuboff's famous work on "surveillance capitalism" documents, maximising user "engagement" relies on the collection of ever more granular data about users' preferences and interests, in order to enable more targeted, and more thus lucrative, advertising (Zuboff 2019).

To Gillespie's list of the diverse actors who are implicated in content moderation, such as fact-checkers and regulators, we need also add stack providers. Both in the case of mainstream platforms trying but failing to adequately moderate hateful content, and in the case of fringe platforms wilfully refusing to do so (and all those in between), it is increasingly falling on stack providers to determine what is acceptable, in terms of their own rules (DeNardis and Bradshaw 2022; van Geuns and Cath-Speth 2020). Preventing the spread of hate speech is thus within the ambit of not only platform self-governance but also, in extremis, a matter of cross-platform governance. However, the nature of these respective obligations varies. When it comes to enforcing policies against hate speech, the conceptual complexities of cross-platform governance

[44] The company in question, Sama, has since stopped providing its content moderations services to Meta.

discussed in Chapter 3 introduce considerations that are distinct from those associated with platform self-governance, such as the use of algorithmic approaches to detect hate speech. Companies situated higher up the tech stack have been drawn into (often reluctantly, as in the case of Cloudflare: Lee 2017), discussions, and on occasion decisions, about the problematic content that prevails on the platforms to whom they provide their services. Typically, these decisions are made ad hoc and post hoc, arising only after shocking "real-world" events like the Charlottesville or Christchurch attacks, often amid public backlash. In these cases, stack providers can be seen as "governors of last resort"—acting in cases where platform operators themselves have failed to effectively deal with online hate speech (or have brazenly refused to even try). But what would a more procedurally robust and normatively grounded approach to the cross-platform governance of hate speech look like? That is the question I tackle in the next section.

4.3 Cross-platform governance of hate speech

In what follows I use the three forms of harm inflicted by online hate speech that I identified in 4.1.2—instantaneous emotional harm, offline violence, and societal exclusion—as the basis for initial consideration of the governance responsibility that stack providers in a cross-platform governance capacity may have, with respect to hate speech on the platforms to whom they provide essential infrastructure. Consideration of the three forms of harm in turn yields three possible obligations. This points to a clearer framework for the responsible cross-platform governance of hate speech, in concert with other claims and interests.

4.3.1 Preventing instantaneous emotional harm

The most direct form of harm inflicted by online hate speech is the "metaphorical gut punch" to those inadvertently exposed to or deliberately targeted by it. As we saw in 4.1.2, this can cause immediate emotional effects like fear and hurt and can also contribute to longer-term psychological problems such as anxiety and trauma. This form of harm from hate speech tends to arise on large, "general-purpose" platforms frequented both by a minority of users (and bots) creating and

spreading objectionable content like hate speech, and ordinary users who may be exposed to and harmed by it, resulting from what Marwick and boyd have referred to generally as "context collapse" (Marwick and Boyd 2011).

To decide which content to display to users and how to rank and order it, different platforms use different data signals: some voluntary, such as a user's choice to follow another user or join a group, and others inferred, such as showing "suggested posts" based on personal preferences deduced from click-data or internet speed (Lada, Wang, and Yan 2021). Some platforms (e.g., Twitter) allow users to further modify the "feed" of content to which they are exposed by blocking or muting users or specific topics or letting them turn off algorithmic ordering altogether. Most platforms also offer search functionality, allowing users to actively search for a particular topic or form of content. And as well as designing the architecture that determines the content displayed to users, large social media platforms also have rules and processes in place to decide what users cannot see or say. To enforce their rules prohibiting hate speech and detect violations, platform operators tend to take a multi-pronged approach, increasingly using both automated methods (Gorwa et al. 2020) and humans, including contracted content moderators (Jereza 2022; Perrigo 2022; Roberts 2019) and volunteers (Matias 2019; Seering et al. 2019), each of which have their own functional and ethical drawbacks.

Given that most large social media platform operators already attempt, however imperfectly, to deal with the spread of hate speech on their platforms, is there a normative duty for stack providers to prevent the instantaneous harm inflicted by hate speech, and if so, in what circumstances does it arise? Some stack providers, notably including app store operators Apple and Google, appear to tacitly accept such a duty, by stipulating in their own terms of service that platforms should have robust content moderation systems—effectively affording themselves the status of "meta-moderators" (see Chapter 3). However, to operationalise such rules and hold platforms accountable in practice, stack providers require some yardstick by which a platform can be judged as (i) sufficiently replete with hate speech, and/or (ii) sufficiently incapable of or

unwilling to tackle it, to warrant sanction. Deciding this threshold may prove difficult to establish in practice. By way of illustration, consider the case of Mumsnet, a British internet forum established in 2000 for expectant and current parents, which can also be accessed via dedicated smartphone and tablet apps. Mumsnet is a large, popular platform, claiming 1.3 billion page views from 119 million unique users in 2018, yet as its name suggests, it is populated mostly by mothers and is characterised by a focus on parenting issues. Notably, despite what may seem on the surface a rather innocuous and benevolent community of interest, since at least 2018 Mumsnet has been the subject of controversy over the prevalence of transphobic hate speech on the platform. It was forced to tighten its content moderation rules regarding the discussion of trans issues after campaigners pursued a demonetisation strategy against the platform, contacting companies who advertised with the site (Ntim 2018).[45] However, although one advertiser pulled an advertising partnership with Mumsnet as a result (Lothian-McLean 2019), there is no evidence that Mumsnet faced large-scale sanctions from stack providers, such as (the threat of) being removed from app stores, perhaps because of how it responded to the controversy.

The prevalence of transphobic hate speech on Mumsnet highlights the current lack of clarity regarding the threshold at which a platform can be considered "sufficiently" hateful, and/or the operator of a platform considered "sufficiently" unwilling or unable to counter the spread of hate speech, to warrant intervention by stack providers. Whether or not stack providers should have intervened in this specific example is beyond present scope, but it is worth noting that they technically could have: Mumsnet has a smartphone app (albeit this is not the only, and probably not the most common, way for members to use the platform) and as with any large platform it relies on various infrastructural and hosting services to stay online. In principle, then, a responsibility for stack providers to prevent emotionally harmful hate speech on platforms they serve exists; this takes the form, for example, of stipulating that subject platforms have sufficiently

[45] The level of transphobic hate speech on Mumsnet led one commentator to make the striking claim that "Mumsnet is to British transphobia [what] 4Chan is to American fascism" (Miller 2018).

effective content moderation systems in place to prevent the spread of hate speech (alongside other objectionable content). It is less clear, however, what yardstick is used to decide what is "sufficiently effective", and even less clear how willing stack providers are, in practice, to sanction violating platforms in cases where direct harm may arise from hate speech, given the ramifications for freedom of expression that such restrictions would involve.

4.3.2 Preventing aggravated offline violence

The second form of harm that can result from online hate speech concerns the risk that online hate spills into offline hate crime, including physical violence. This poses different governance considerations from those discussed in 4.3.1. Whereas the previous concern was to prevent the exposure of vulnerable users to hate speech on platforms that are frequented by both the perpetrators and targets of hate speech, preventing hateful online discourse from motivating aggravated offline attacks more often concerns homogenous online communities (Gallacher and Bright 2021), which are increasingly developing and using own platforms.

A prominent, tragic example of offline spill-over from online case is the Christchurch mosque shootings in March 2019, which left 51 Muslim worshippers dead. The atrocity had a strong online component; the shooter, who identified as part of the "alt-right", shared his "manifesto" on both mainstream and far-right social media platforms before the attack, and live-streamed the attack on Facebook Live. The attack and its aftermath highlighted the starkly different approaches taken by mainstream social media sites and far-right platforms respectively. Facebook claimed that the original video was viewed 4,000 times before it was removed and that it thereafter blocked 1.5 million attempted re-uploads of the video, while Reddit permanently banned certain subreddits, including r/WatchPeopleDie, for glorifying the violence.[46] By contrast, both the manifesto and the footage were freely shared—and widely lauded—on the far-right platform 8chan, and the perpetrators of several subsequent attacks would later take to 8chan to

[46] This is not to suggest that Facebook's or Reddit's moderation was necessarily adequate.

claim inspiration from the Christchurch mosque attack for their own deadly shootings. The Christchurch shooting and subsequent far-right atrocities in 2019, in Poway, California in April and El Paso, Texas in August, led several stack providers to take action against 8chan. This included Cloudflare, which prior to 2019 had stubbornly resisted pressure to act, but finally, and reluctantly (Roose 2019), withdrew its anti-DDoS services to 8chan in the wake of the El Paso shooting (Kelly 2019). At the same time, domain name registrar Tucows pulled its DNS support for 8chan, forcing 8chan to move to Epik's BitMitigate.[47] And telecommunications companies in Australia also blocked access to the domains of platforms like 4chan and 8chan that were hosting footage of the massacre. As Vodafone, one of the companies that collectively initiated a domain block, indicated, "this was an extreme case which we think requires an extraordinary response", acknowledging "discussions at an industry level" to coordinate the block (AAP 2019). The massacre also led to intervention from public authorities, including a bill hurriedly passed by the Parliament of Australia to require "internet service providers" and companies that "provide a content [or] hosting service" to "ensure the expeditious removal" of "abhorrent violent material" (Parliament of Australia 2019)

Far-right atrocities such as these emblemise, with tragic starkness, the danger that online hate speech poses in causing offline violence (Williams et al. 2020), and the longer-term risk of a vicious self-sufficient cycle of hateful discourse and either copycat or retributive violence emerging (Williams and Burnap 2016). The stakes of cross-platform governance in this category of harm are clear, and the existence of a consequent responsibility for stack providers is evident. However, the ambit of cross-platform governance in the context of this form of harmful hate speech is different from that discussed in 4.3.1. Whereas the responsibility to prevent instantaneous emotional harm that is inflicted by hate speech involves focusing on platforms frequented by *both* the perpetrators *and* common targets of hate speech, preventing offline violence more often involves

[47] This, in turn, prompted server company Voxility to stop leasing its servers to BitMitigate, which forced the latter to relent, in an interesting example of "nested" or multilevel cross-platform governance.

"conspiratorial" platforms like 8chan, whose U.S.P is precisely their ideological homophily (Zeng and Schäfer 2021). As a result, such platforms are typically far less willing to negotiate, let alone collaborate, with stack providers to improve their content moderation systems; platforms like 8chan can scarcely be described as having a content moderation system at all (Rieger et al. 2021). Therefore, in comparison to preventing instantaneous emotional harm to the victims of hate speech, the responsibility to prevent the aggravation of offline violence is distinct, often involving different platforms altogether. However, as will be seen, in the case study on the app store governance of Parler, these two distinctive obligations blurred.

4.3.3 Preventing exclusion

The third and final form of harm of online hate speech identified in 4.1 concerns the longer-term harm to society that results from the groups that are most frequently targeted by hate speech withdrawing from public spaces and fora. The effect of intimidating and silencing those repeatedly targeted by hate speech robs society of important perspectives. Perhaps because it is gradual and cumulative, effective responses to this form of harm are the most difficult to identify in existing examples of cross-platform governance. But this is not to say that a normative responsibility to tackle this form of harm does not exist. Stack providers acknowledge this duty themselves, albeit inadvertently. For example, the CEO of Cloudflare, Matthew Prince, described himself in 2017 as "almost a free-speech absolutist" (Lee 2017) when defending what he hoped would be a one-off decision to cut off neo-Nazi site the Daily Stormer. But as I observed in 4.1.2, it is equally possible to argue that a commitment to free speech for all creates an obligation for *greater*, not lesser, restrictions on certain forms of speech, e.g., speech that systematically excludes or alienates users. Preventing this systematic exclusion could enrich the health of public discourse and serve the aims that defenders of the classic libertarian flavour of free speech often claim to support. However, to a greater degree than the previous two forms of harm, it is an open question as to what responsibility stack providers may have in this respect.

4.3.4 Towards a basis for the cross-platform governance of hate speech

To summarise, the distinct ways in which online hate speech inflicts harm each creates possible obligations in turn for cross-platform governance. Instantaneous emotional harm is most likely to arise on mainstream platforms frequented by both the perpetrators and targets of hate speech. By contrast, offline violence is commonly (though not exclusively) the result of the radicalisation and incitement that occurs between like-minded individuals on fringe, extremist platforms. And the impoverishment of perspectives from those most often targeted by hate speech serves to complicate the common justifications for permissive approaches to hate speech. Of course, these forms of harm, and the concomitant responsibilities that accord to stack providers, tend to blur in practice. Targets of hate speech may unwittingly "stumble upon" fringe platforms that perpetuate instantaneously harmful hate speech or be exposed to such content through sharing across other platforms or websites. Radicalisation occurs not solely on fringe platforms, but also in the darker corners of the largest and most mainstream networks like Facebook (Hutchinson et al. 2021; Scrivens and Amarasingam 2020). And the withdrawal from the public sphere of those most likely to be the targets of hate speech is hard to systematically demonstrate and, therefore, to prevent. My case study on Parler is situated amid this conceptual fuzziness. Throughout its initial growth, Parler executives sought to portray the platform as an upstart rival to mainstream platforms, championing its "unbiased" approach to platform self-governance and its cross-ideological membership to lure new users. But this inclusive rhetoric obscured its reality as a platform beset by objectionable content, including hate speech, and the homogenous far-right skew of its membership.

To the extent that hate speech can, at least in the "right" circumstances, inflict any of an array of potential harms, a responsibility in principle to counter it—not only by the operators of social media platforms on which it prevails, but also, if necessary, by stack providers on whose services these platforms rely—seems at least plausible. But this obligation is both set against other,

contrasting, normative impulses such as the safeguarding of free speech, and hampered by more pragmatic considerations. This is, in short, the realm of the political (Gillespie 2010).

While moral justifications for countering hate speech may be evident, the political and legal basis for doing so is much less concrete. This is in part because of competing claims and responsibilities, not least safeguarding free expression. But it also reflects the fundamentally ambiguous status of both platform operators, and stack providers, as *governors*, which I discussed in Chapter 2. In the final section I wrestle with this ambiguity in the specific case of app store operators.

4.4 App store governance of hate speech: framework or patchwork?

In this section I draw on the theoretical development of this and the previous two chapters to answer my first subsidiary research question, which concerns the responsibility that Apple and Google as app store operators have to govern hate speech on social media platforms that they host. I look briefly at the relevant rules and processes that Apple and Google have in place for reviewing apps submitted by developers to their app stores, and identifying some challenges that may arise from this governance responsibility, tying this to the limitations of ASG more generally as identified in Chapter 3. I conduct a more in-depth examination of these rules in the context of Parler in Chapter 7.

As we have seen, Apple and Google provide guidelines for developers seeking to submit apps to their respective app stores. Although the philosophies of Apple's and Google's smartphone ecosystems are at least superficially different, in practice, both rely on app stores that are highly-populated yet well-managed—and both companies cite the costs of maintaining a safe and trustworthy app store when seeking to justify the sizeable cut that they take from apps and subscriptions sold therein.

Both sets of guidelines have policies relating to hate speech, though these are often ambiguous. Apple's guidelines are vague (Hestres 2013; Hofmann et al. 2022)—in places to the point of satire. In the introduction to its guidelines, Apple notes as a "point to keep in mind" that:

> "We strongly support all points of view being represented on the App Store, as long as the apps are respectful to users with differing opinions and the quality of the app experience is great. We will reject apps for any content or behavior that we believe is over the line. What line, you ask? Well, as a Supreme Court Justice once said, "I'll know it when I see it". And we think that you will also know it when you cross it" (Apple 2022).

The tongue-in-cheek reference to US Supreme Court Justice Potter Stewart's definition of pornography ("I know it when I see it") is ironic, and not merely given Apple's strict ban on pornographic material in the App Store (Siegler 2010). For what is supposed to be a usable "guideline"—and also, presumably, an enforceable standard—it is almost self-parodically opaque and unaccountable (only Apple can decide what is "over the line": Hestres 2013; Poell et al. 2019).

In addition to Apple's guidelines concerning user-generated content (see Chapter 3), the company also identifies a set of "objectionable content" that it prohibits. This is summarised as "content that is offensive, insensitive, upsetting, intended to disgust, in exceptionally poor taste, or just plain creepy". This may include "defamatory, discriminatory, or mean-spirited content, including references or commentary about religion, race, sexual orientation, gender, national/ethnic origin, or other targeted groups, particularly if the app is likely to humiliate, intimidate, or harm a targeted individual or group."

The situation is similar for Google's Play Store. "Hate speech" is included as a form of "Inappropriate Content" that Google prohibits in Play Store apps, which it defines as "apps that promote violence, or incite hatred against individuals or groups" based on a series of identifying characteristics; and "violence", i.e., "apps that depict or facilitate gratuitous violence or other dangerous activities" (Google 2023c). As we have seen, Google's guidelines on user-generated content stipulate that platforms must define and prohibit objectionable content, establish ongoing moderation of user-generated content, and include a reporting mechanism for violations (Google 2023c).

Taken together, the demands contained in these guidelines are significant. They constitute an avowal by Apple and Google of the power not only to determine the suitability not merely of the *pre-loaded* content that apps contain upon installation (the usual material on which app store reviews are conducted), but also dictate how *user-generated* content is to be moderated in, and by, apps that contain it. In other words, through these guidelines Apple and Google accord themselves the role of meta-moderators, overseeing the enforcement of the requirements that they impose on platforms' own content moderation, and reserving the ability to intervene when they fall short—a power they have sometimes exercised, as in the cases of Tumblr (Porter 2018) and Gab (Price 2017).

However, it is unclear how the precise circumstances in which Apple and Google exercise this meta-moderation role are to be objectively evaluated. At a high level, neither company releases reports or datasets detailing the rejections or removals of violating apps, nor which rules have been violated in each case (in contrast to, say, Meta's quarterly reports on standards enforcement). It is also uncertain how Apple and Google reconcile their rules on objectionable content such as hate speech with their rules on user-generated content and how effectively apps should manage it. More specifically, it is unclear at what point a platform would be adjudged to have become sufficiently replete with hate speech (or another form of harmful content) that Apple and Google take it upon themselves to sanction it. Consider Facebook, a platform which houses millions of pieces of offensive content, including hate speech.[48] There are two principal reasons that may explain why Facebook (the app) has not been sanctioned by app store operators on the grounds of hate speech.[49] First, because its parent company Meta attempts (with efficacy that varies, especially

[48] Meta "took action" on 11 million pieces of hate speech content on Facebook and 4.7 million pieces on Instagram in Q4 2022. These counts obviously exclude content that was not detected by or reported to its moderation team. https://transparency.fb.com/data/community-standards-enforcement/hate-speech/facebook/

[49] This speculation that follows should not be taken as an assumption either that Meta does a satisfactory job moderating its platforms, nor that other considerations do not also affect Apple's and Google's retention of Facebook and Instagram on app stores. Corporate concerns on the part of Apple and Google undoubtedly also play a role; Apple, for example, may be reluctant to alienate its customers or be wary of antitrust or other lawsuits were it to remove Facebook from its app store.

across jurisdictions) to remove objectionable content and sanction offenders. Second, because the proportion of the platform that contains objectionable material and illicit behaviour is judged to be sufficiently small. Were Facebook to consist mostly, or to a sufficiently large extent, of hate speech, and/or were efforts to prevent and remove it deemed insufficient, one might expect that it would be removed from app stores. Similar dynamics may help explain why, drawing on the earlier example in 4.2, Mumsnet was not sanctioned. Though transphobia appears to have been common—or even, as some claimed, rife—on that platform, it may not have been *predominant*. In any case, Mumsnet introduced new rules regarding the discussion of transgender issues after issues were brought to light.[50]

This discussion brings to the fore an important, as yet unasked question, which animates the case study on Parler. What is the threshold, in terms of the proportion of hate speech or some other metric, at which a hate-*filled* platform—i.e., a platform that is frequented by hate-spewing individuals among others—comes to be treated by app store operators as a hate*ful* app tout court? Although the specific point at which app store operators refuse or remove an app that indulges hate speech is not outwardly clear, the notion of such a threshold suggests that in principle, the proportion of a platform's content that is objectionable is objectively relevant to how it is treated.

However, this principle runs up against a further consideration. Recall that above I identified the most immediate form of harm inflicted by hate speech, on those either inadvertently exposed to or directly targeted by hate speech. As discussed, the harm of any given instance of hate speech is bound up with the group identities of those involved, specifically those of the speaker and receiver—such that in Butler's evocative phrase, in the speaking of a racial slur "one chimes in with a chorus of racists", creating an "imagined relation to an historically transmitted community of racists" (1997:81). Therefore, with respect to the harm *instantaneously* inflicted by hate speech, what matters is not only *how many* users are encountering hate speech, but also more

[50] The caveat in the previous footnote likewise applies to Mumsnet.

precisely *who* is encountering it. This complicates the notion that there is a simple (even just in theory) threshold at which a platform may be sufficiently laden with hate speech to warrant intervention from app store governors.

This should not for a moment be taken to suggest that there is "nothing wrong" with platforms containing only like-minded users sharing hateful views only among themselves. As discussed in 4.2, these homogenous platforms also have great potential to inflict harm—albeit often in different ways, such as by instigating violent offline attacks, or by serving as an organising base for the spread of harmful hate speech on more mainstream platforms. However, there is little distinction made between these conceptually distinct forms of harm in app store governance guidelines. Therefore, just as it is not obvious what "threshold" of hatefulness a platform must reach in order to be sanctioned on the grounds of preventing *online* harm from hate speech, it is also unclear how concerns over *offline* violence instigated by hate speech weigh into app store governance decisions.

These ambiguities are pivotal to the story of Parler, and as the trajectory of this discussion implies, it will soon be time to turn to the case study. But first, it is time to advance an answer to RQ1, which concerns the responsibility that app store operators have, in principle, to govern hate speech on social media platforms. Apple and Google have accorded themselves responsibility to govern hate speech in social media apps, via the rules prohibiting hate speech that both companies adopt in their app store guidelines. As Gillespie puts it, Apple and Google "share a sensibility that, as providers of the devices and the operating system, they are to some degree responsible for the apps that run atop them" (Gillespie 2018:81). And this responsibility is underscored more generally by the rhetorical emphasis that both companies put on the role they play keeping app stores safe and trustworthy. However, the nature of this responsibility is ambiguous, in three key ways.

The first ambiguity is substantive. It is unclear how different elements of Apple's and Google's governance guidelines that concern hate speech are connected. In particular, the

threshold at which an app is adjudged to have unacceptable levels of hate speech and/or insufficient self-governance, warranting direct action by app store operators, is unclear.

The second ambiguity, which springs from the first, is procedural. Without publicly available reports on enforcement, there is no way to trace the application of app store governance guidelines to specific cases in a systematic way. Nor, consequently, is there the possibility for shared reflection on the implications of app removals for online safety, free speech, and other relevant considerations. Instead, via their stewardship model of governance, Apple and Google ask civil society to "trust us" when it comes to balancing competing values and interests in their app store governance.

The third ambiguity relates to Apple and Google's self-interest. It is difficult to disentangle rhetorical commitments to trust and safety from the strong economic incentives that each company has to maintain their dominance over their smartphone software, which relies heavily on their app stores.

Drawing these points together, in answer to RQ1, I argue that, while Apple and Google have indeed accorded themselves a governance responsibility with respect to hate speech on social media platforms that they host, this responsibility is founded on a *patchwork* of rules and rhetoric (Duguay et al. 2020) that is ambiguous in terms of its definition, implementation and motivation. This falls far short of an overarching governance framework that, if well-designed, could foster coherence, clarity, and participation in an aspect of internet governance that has often hidden, yet high, social stakes. The theoretical investigation I have undertaken in these chapters has identified areas in which app store governance could be improved, which I consider in Chapter 8, after providing evidence arising from the empirical case study on Parler.

Summary

In this chapter, I have discussed online hate speech and the responsibilities of platform companies, including app store operators, to address it. After providing a definition of online hate speech, I

explicated key ways in which it inflicts harm. The governance responsibility of social media platform operators with respect to hate speech rests on agreements, operators' own rules, and laws in some jurisdictions. By contrast, the basis on which a responsibility for app store operators, as stack providers, to govern hate speech is situated in a patchwork of rules and rhetoric. App store operators are not signatories to the European Commission's Code of Conduct on countering hate speech online, and laws such as NetzDG do not apply to app store governance. Thus, we are left with the operators' own rules and rhetoric, which are ambiguous with respect to the circumstances in which app stores intervene and the applicability of rules to different platforms.

The answer to RQ1 motivates and shapes the empirical case study to which I now turn. In the absence of a clear framework of responsibility for the app store governance of hate speech, the case of Parler provides evidence for, and thus helps to explicate, the theoretical ambiguities that Part I has identified.

Part II: Case study analysis

Overview

The case study is designed to answer my second research question, which is:

RQ2: How well did Apple and Google meet their governance responsibility towards hate speech on Parler?

In Part I of the thesis, I showed that Apple and Google have adopted a "meta-moderation" model for governing of hate speech on social media, delegating this responsibility to platform operators in the first instance but retaining the ability to remove violating platforms as they see fit. This shapes the case study analysis, which first covers Parler's self-governance, and second covers Apple's and Google's app store governance of Parler.

I begin the case study with a comprehensive introduction to Parler in Chapter 5. This includes a chronology of the platform from its launch in 2018 to its removal from app stores in January 2021. I also report the findings of a systematised thematic review of existing studies that examine Parler. I detail the data collection and analysis to be undertaken in the two subsequent chapters and identify ethical and logistical considerations that arose from the case study.

I then turn to the analysis. Answering RQ2 first necessitates understanding how Parler performed its platform governance with respect to hate speech. This can be phrased as a question:

RQ2a: How did Parler perform its platform self-governance of hate speech?

This question is primarily descriptive, and I address it in Chapter 6. Using the walkthrough method for the critical analysis of apps, I gather and analyse an array of data which describes Parler's approach to the governance of hate speech on the platform. Besides the walkthrough itself, the other data that I draw on in this chapter is (i) 186 news articles covering Parler and the statements of its CEO John Matze, (ii) 12 archived pages from Parler's website, and (iii) 18 iterations of archived policy documents from Parler's website.

With this question answered, a second component question can be considered, which concerns the role of Apple and Google:

RQ2b: Should Apple and Google have removed Parler from app stores prior to January 2021 on the basis of their guidelines on hate speech?

This second question is primarily interpretive, and I address it in Chapter 7. I consider Apple and Google's app store governance responsibilities with respect to the failure of Parler's governance of hate speech. The data I primarily draw in this chapter is (i) a random stratified sample of 19,237 Parler posts during 2020 and (ii) Apple's and Google's app store governance guidelines that were in force during 2020. By combining these two forms of data, I make a best effort to apply Apple's and Google's guidelines to the findings regarding hate speech on Parler and its moderation of it. I find that Apple and Google would have been warranted in removing Parler sooner than January 2021 on the basis of hate speech. This finding surfaces questions for app store governance in turn, which I consider in Chapter 8.

An overview of the questions, data, and methods used in the case study is presented in Table 2.

Table 2: Overview of research questions, data, and methods used in the case study

Chapter	**Research question**	**Data**	**Methods**
Chapter 5	n/a	Corpus of 43 existing studies on Parler	Systematised thematic review
Chapter 6	RQ2a: How did Parler perform its platform self-governance of hate speech?	Technical walkthrough; corpora of 186 news articles; 12 archived webpages; 18 policy documents	Walkthrough method: technical walkthrough and critical discourse analysis
Chapter 7	RQ2b: Should Apple and Google have removed Parler from app stores prior to January 2021 on the basis of their guidelines on hate speech?	Random stratified sample of 19,237 Parler posts made in 2020; app store governance guidelines	Mixed-method content analysis; policy analysis

Chapter 5: The "populist platform"—introducing Parler

Introduction

The purpose of this chapter is twofold: to provide a short history and literature review of Parler; and to describe and justify the design of the case study that I take forward in the following two chapters. A basic chronology of Parler is presented in 5.1, including its rapid rise and fall, which provides important context and background to the later analysis. I also discuss core elements of Parler's functionality. 5.2 reports the results of a systematised review of existing scholarship on Parler. This identifies 47 studies that I group and discuss in terms of their focus. 5.3 explains the suitability of Parler as a case study. 5.4 explains how the study is conceptualised and designed in order to address RQ2, drawing on the theoretical developments of the earlier chapters. RQ2 is broken down into two component questions which are tied to the analyses undertaken. 5.5 offers further reflections on the broader ethical, social, and logistical considerations that emerged from conducting the case study.

5.1 Parler: a chronology

In this section I present a chronological overview of Parler from its founding in mid-2018 to its shuttering in January 2021. Sections 1.1-1.4 respectively provide an account of Parler's key periods: its founding and emergence; its rise to prominence in early 2020; its apogee in late 2020; and its removal after the Capitol riot in January 2021.

5.1.1 2018 and 2019: Founding and early growth

Parler[51] was founded in 2018 by John Matze and Jared Thomson, computer science graduates from the University of Denver.[52] The role of a third founder, Rebekah Mercer, was to be revealed only in late 2020 (Horwitz and Hagey 2020). In a Parler post in November 2020 following reports of

[51] It seems that Parler was originally intended to be pronounced "par-lay" as in the French term "to speak", but had switched to "par-luh" by mid-2019 (Schreckinger 2019).
[52] https://web.archive.org/web/20180709190428/https://home.parler.com/team/

her early involvement, Mercer acknowledged that she had, with Matze, "started Parler to provide a neutral platform for free speech, as our founders intended, and also to create a social media environment that would protect data privacy" (Benveniste and Yurieff 2020). Mercer's invocations of political neutrality and respect for privacy protection must be viewed sceptically given her political profile. Hailed as "the First Lady of the alt-right" (Swenson 2019), Mercer, along with her father, the billionaire hedge fund manager Robert Mercer, are highly influential figures in US conservative politics. The pair donated $25m in the 2016 presidential election, first to right-wing primary candidate Ted Cruz, and then to support Donald Trump's campaign against Hillary Clinton in the general election (Schouten 2018). As well as campaign contributions, the Mercer family has given money to other radical conservative causes, such as giving $11m to far-right news site Breitbart (Kutner 2016) as well as to various political advocacy groups (Schouten 2018). Ironically, the Mercers were investors in Cambridge Analytica, the data mining and political analytics company responsible for the eponymous data privacy scandal that exposed Facebook's lax privacy standards for user data, eventually resulting in a $5bn fine for the company, as well as the collapse of Cambridge Analytica itself (Wong 2019). This is important context not only for Mercer's retrospective claims about her motivation for co-founding Parler, but also for John Matze's ongoing attempts during 2020 to portray Parler as a "neutral" platform that handled user data more securely than its rivals like Facebook (see Chapter 6).

The first archived versions of Parler's website date to June 2018. A user navigating to www.parler.com at that time would have been met with a holding page (see Figure 1), which, despite its parsimony, contained an embarrassing error with the misspelling "free speech" as "free speach": ironic given the company's claimed concern for free expression. Despite the misspelling, this emphasis on free speech, and the claimed influence of the US Constitution on its platform self-governance, would figure prominently in the portrayal of Parler. Also notable at this early juncture is the red background colour, similar to the traditional colour of the US Republican Party, which would remain Parler's primary colour throughout.



Figure 1: Parler's first launch screen, May 2018. Source: The Internet Archive, https://web.archive.org/web/20180705162002/http://parler.com/

Parler appears to have "soft-launched" the following month. The Internet Archive reveals that on or before[53] July 9th, 2018, www.parler.com began redirecting to home.parler.com, a new homepage that was initially geared as much towards publishers as to end-users. Its strapline pledged to "increase reader engagement, grow audience exposure, and drive site traffic", and offered prospective publishers a demo.[54] At this early stage, the website already included links to download the Parler app on Google's Play Store and Apple's App Store. Elsewhere on the website around this time, Parler is described as providing "Social News and Content Engagement for media outlets, social users, writers, emergency services and politicians to convey news and share content in real time"—clearly indicating Parler's early desire to emulate Twitter—as well as "enterprise tools to enhance online blogs, media and websites with direct social integration."[55] The site's "Team" page at this time (the first time this information was available)[56] listed Matze as "Co-Founder, CEO", Thompson as "CTO & Co-Founder", Robertson Williams as "VP of Business

[53] See 5.4 for a discussion on the utility and limits of web archive research with respect to dating archival materials.
[54] https://web.archive.org/web/20180709190113/https://home.parler.com/
[55] https://web.archive.org/web/20180709190428/https://home.parler.com/team/
[56] Mercer's involvement is not acknowledged on the website until 2020.

Development", and six other employees in software development positions.[57] It appears that all nine employees at the time were young men, most of them white (and judging by their surnames, two of the senior developers were likely related to John Matze). Many of them were reported to have attended the University of Denver along with Matze and Thomson (DiStefano, Spence, and Mac 2019), who had been roommates in college (Mac and Gray 2021). This level of demographic homogeneity is far from a rare occurrence for a tech start-up, but it is again significant to what Parler claimed was its "neutral" approach to content moderation, including with respect to racist hate speech (see Chapter 6).

Notwithstanding its initial spelling error, the company's rhetorical commitment to the concept of "constitutionally free speech" is clear across early iterations of Parler's website. The first available version of the website's blog page, titled "Parler Newspeak",[58] contained two posts, both apparently published on June 26, 2018, and both written by Williams. The more pertinent[59] of these posts is titled "Permitted Speech", which lays out the nascent platform company's commitment to "take the concept of constitutionally free speech seriously" (Williams 2018a). The blog post anticipates the "constitutional metaphor" of free speech protection that Matze would publicly deploy throughout the platform's rise to prominence in 2020 (see Chapter 6). As "an American based company", Williams claimed that it was Parler's "responsibility to uphold true American values and promote the American tradition of free discourse." This US-centricity casts a notable contrast to mainstream technology companies' typically more globally-minded mission statements (Rider and Murakami Wood 2019).[60] Users would not be banned "off of their viewpoints" since it is "our believe [sic] that through free dialogue people of differing opinions we

[57] https://web.archive.org/web/20180709190428/https://home.parler.com/team/

[58] Whether the Orwellian allusion was deliberate is unclear.

[59] The other, titled "The Thing About Passwords", explains why "we will not be using passwords because they are too much of a burden to implement and they only add insecurity to our platform" (Williams 2018b). Parler would later adopt passwords, but—while the platform's Parler's lax security practices are beyond the scope of this project—this early post unwittingly hints at some of the problems that were to follow.

[60] See e.g., Google's mission to "organise the world's information and make it universally accessible and useful" (https://www.google.com/intl/en_uk/search/howsearchworks/our-approach/) and Meta's mission to "give people the power to build community and bring the world closer together" (https://about.facebook.com/company-info/)

may attempt to understand positions on both sides of any argument." However, users would also be able to "self filter and create content based platforms surrounding their desired level of discourse" and "there will be tools offered to users to allow them to protect themselves from internet "trolls" [and] people they disagree with." Thus, Parler claimed to adopt a laissez-faire approach to what may be posted on the platform, while empowering users to set their own limits on what they would see. Free speech was not the only value signalled in the early iterations of Parler's website. The Team page also claims that co-founders Matze and Thomson "decided to create [Parler as] an alternative … after being exhausted with a lack of transparency in big tech, ideological censorship and privacy abuse". Each of these concepts would continue to be drawn upon through 2020.

In this early blog post, Williams cited "obscenity", and specifically the Federal Communication Commission's definition of "obscene, indecent and profane broadcasts" (US Federal Communications Commission 2019), as Parler's own dividing line for what content was to be allowed on the platform. This was a curious choice insofar as the FCC does not regulate online content (US Federal Communications Commission 2022). Williams justified this appeal to regulatory authority on the basis that Parler did not "want user content to be so obscene that it drives away other users, advertisers, investors, partners, and overall ruin[s] the reputation and community of Parler". It instead allowed "Parler to remain a neutral party whose responsibility is to uphold the law" (Williams 2018a). In practice, however, this gave Parler considerable leeway to define obscenity in ways that suited its corporate objectives. Even the FCC's *own guidelines* on obscenity recall (as do Apple's App Store guidelines) Justice Potter Stewart's[61] axiom that "I know it when I see it". Therefore, despite Parler's claims to "neutrality" and its commitments to free speech—which were to be repeated ad nauseam throughout its rise—its nascent content policy

[61] Unfortunately, the FCC's guidelines erroneously rechristen Justice Potter Stewart "Justice *Stewart Potter*". The overarching conclusion I have drawn from so often encountering Stewart's decades-old aphorism in the course of researching this thesis is that it ought perhaps to be quietly retired.

gave it considerable scope for interpretation of what constituted "obscenity".

When retrospectively examining content such as this, the benefits of hindsight must be acknowledged. Parler is surely not the first, and nor will it be the last, start-up to have cobbled together its approach to content moderation at pace—and even established platform companies, such as Meta, have taken years to fully reckon with the sociopolitical implications of the development and enforcement of their rules, or lack thereof (Vaidhyanathan 2018). However, for all its *functional* incoherence as actionable policy, what is striking in retrospect about Parler's approach to content moderation is that the *rhetorical* touchstones that would be consistently drawn upon in the years to come had been laid at this very early stage. In contrast to many of its rivals, for Parler content moderation (or lack thereof) was not an occasional source of irritation or embarrassment—it was a cause celebre, and the platform's unique selling proposition.

Parler's first eighteen months, from the middle of 2018 until early 2020, were marked by occasional spikes of publicity driven by right-wing figures and conservative media outlets. These usually coincided with perceived "censorship" of content or bans of users by other social media platforms, causing users to (or at least claim to) "migrate" to Parler. A December 2018 article by the conservative Daily Caller website, the first such article I was able to identify, provides a template for many future such stories. It describes Parler as "the new app that some are calling the answer to the ongoing battle between conservatives and Twitter … promising to support free expression online". This suggests that Parler's claimed commitment to free speech had already been established in conservative media discourse by this early stage (Ryan 2018). The article identifies several figures who had joined Parler and spoke favourably of it. One early such figurehead was Candace Owens, a right-wing media commentator who had had her Twitter account suspended in August 2018 before it was reinstated by the company after an apology. In a December 2018 tweet she called Parler "the new Twitter," claimed to have been "the first conservative to join," and said Parler's emergence felt "like a long-overdue social media rebellion" (Ryan 2018). Matze would later claim that this tweet alone drove 40,000 users to the platform

(Schreckinger 2019). Parler had at this time already become the seventh most popular app in the News category of the iOS App Store in the US, and the article also quoted Matze describing Parler as an "alternative platform," lauding those "bold [enough] to switch," claiming that "big tech is not too big to topple," and accusing rival platforms of "ideologically targeting groups" (Ryan 2018). Casting Parler's appeal in terms of its opposition to larger platforms, in order to draw both high-profile and ordinary users to his platform, remained a calling card for Matze throughout Parler's rise (see Chapter 6).

Further media coverage of Parler's growth throughout 2019 followed a similar pattern. A move by Twitter or Facebook that was perceived as "censorship" would drive fresh attention, and new users, to Parler. Most new Parler users were conservative Americans claiming to be driven by the behaviour of excessively "liberal" Silicon Valley platforms (Culliford and Paul 2019); a Politico article in June 2019 reported that the head of President Donald Trump's re-election bid was "checking out" Parler as a potential platform for the campaign (Schreckinger 2019). 2019 also saw the apparent migration to Twitter of an altogether different group of users: supporters of Saudi Arabian Crown Prince Mohammed bin Salman. Analysis by Reuters and the Citizen Lab verified that 200,000 Saudi accounts joined Parler in June 2019, allegedly as a reaction to clampdowns by Twitter on an "electronic army" backing the Saudi government in the wake of the murder of journalist Jamal Khashoggi (Loveluck and al-Alwani 2018). Matze zealously welcomed the arrival of the Saudi users, whom Reuters suggested were "unusual bedfellows" for the primarily conservative American user base, and had received "mixed reactions" from existing users, with some "expressing anti-Muslim sentiments" (Culliford and Paul 2019).[62] The official Parler account on Parler lauded the "surge of new users [which] shows how many people of the world share our Free Speech values."[63] Yet even by the end of 2019, Parler had fewer than one million users

[62] The Reuters article covering the story suggests that the "new neighbours" were met with mixed reactions from users, with some echoing the site's welcome while others expressing anti-Muslim sentiments. By 2020, anti-Muslim rhetoric was increasingly commonplace on Parler (see Chapter 7).

[63] Thus associating "free speech" with the state-sanctioned assassination of a journalist.

(Newhouse 2020) and remained barely a blip on the radar of the social media landscape—something that was set to change with the tumultuous events of 2020.

5.1.2 January to June 2020: Rise and rise

Examination of an archived version of Parler's website from early 2020 suggests that at this time the company continued to tout the core tenets that it had first espoused publicly in 2018. Its homepage as of April 2020 described Parler as a "user driven platform", with "social news" and "community commenting", offering users the ability to "read news, speak free [sic], [and] control your experience".[64] Its Company page characterised Parler as "a non-biased, free speech driven entity," pledging that "we are not regulators. We are not governors. We are a community" and that "Parler accepts your right to express your thoughts, opinions and ideals online."[65] It also retained the original text regarding Parler's "founding story" from 2018 in response to "a lack of transparency in big tech, ideological suppression and privacy abuse."

Parler's spokespeople, especially John Matze, were to frequently restate these tenets during 2020, amid a series of seismic events in American society. The first Covid lockdowns in March, the murder of George Floyd by a Minneapolis police officer in May, and the run-up to the presidential election in November each served as social media flashpoints, with mainstream platform operators increasingly forced to act on the spread of problematic material including health misinformation, racist hate speech, and the communiques of an ever more erratic president. Parler's positioning of itself as a platform that provided an "alternative" to mainstream social media stood it in good stead to gain greater attention and more users, as a result of what were perceived, by some, as the increasingly restrictive, "censorious" policy decisions taken by Parler's larger rivals (see e.g., Schultz 2020). The most promising avenue for Parler was rival platforms'

[64] https://web.archive.org/web/20200412220730/https://home.parler.com/
[65] https://web.archive.org/web/20200417170346/https://home.parler.com/about/

"censorship" of President Trump. Over the course of 2020, Twitter and Facebook[66] gradually imposed greater restrictions on Trump's communications. On 26th May, Twitter for the first time appended a warning label to a tweet by Trump falsely claiming that mail-in ballots for the forthcoming election would be fraudulent, based on its recently updated civic integrity policy (Ghaffary 2020). This led Trump's campaign manager, Parscale, to complain that "we always knew that Silicon Valley would pull out all the stops to obstruct and interfere with President Trump getting his message through to voters." On 28th May, amid ongoing protests after the murder of George Floyd days earlier, Trump darkly tweeted that "when the looting starts, the shooting starts," a nod to an inflammatory racist proclamation by a Miami police chief in 1967. In response, Twitter hid this tweet behind what it called a "public interest notice" for its violation of the platform's policy on the "glorification of violence" (Twitter Comms 2020), requiring users to click through the notice to see the tweet. Twitter applied further warnings to various other of Trump's tweets throughout the summer according to its policies on manipulated media, abusive behaviour, and misleading claims (BBC 2020).

Each of these moves were met with what could be called the "performative migration" of conservative commentators and elected officials to Parler, among them California Representative Devin Nunes in February (Levy 2020) and Texas Senator Ted Cruz (Lima 2020). These moves drove ordinary users to Parler in greater numbers; Parler doubled its number of all-time downloads in a single four-week stretch between mid-June and mid-July (Brustein 2020). Yet as a contemporaneous Bloomberg article observed, while Parler was "a good measuring stick for the amount of anger conservatives online have for mainstream tech companies," few of these figures committed to using Parler *exclusively*, with many of them continuing to post much more on Twitter than on Parler even after they had claimed to have switched (Brustein 2020). Meanwhile, just as

[66] For historical consistency, I use "Facebook" here to refer to the name of the parent company that is now known as Meta. As such, some of the decisions described here as having been taken "by Facebook" also had effect on its subsidiary Instagram.

Saudi Arabia had provided an unlikely boost to Parler's user base in 2019, at least part of the wave of sign-ups in mid-2020 came not from the US but from Brazil. One of Parler's new users in July 2020 was Brazilian president Jair Bolsonaro (Fordham 2020), who, like Trump, had endured restrictions on several of his tweets containing misinformation related to the pandemic (Lyons 2020). The influx of users in this period meant that Parler was estimated to have 1.5 million users by the end of June 2020 (Levy 2020).

5.1.3 July 2020 to January 2021: The beginning of the end

Parler's popularity increased as the presidential election drew near, but so did the challenges and contradictions involved with its operation and governance. Matze was forced to issue several policy updates to users over the course of 2020, including, in one especially memorable post in July, asking users not to share pornography, post photos of faecal matter, or to "threaten to kill anyone in the comment section" (Smith 2020). Less scatological yet more serious were reports that Russian election disinformation and hate speech were running rife on the platform. For example, in early October the *Washington Post* reported that a "mysterious European named Leo"—as it transpired, an asset of a Russian disinformation campaign—had made false claims about mail-in voting on Parler. While mainstream platforms removed "Leo's" account, Parler refused to act, claiming it had not been contacted directly by US law enforcement (Timberg 2020). Later that month the same outlet described accounts related to the shadowy conspiracy theorist "QAnon" movement as "thriving" on Parler, again in contrast to the restrictive steps taken by mainstream platforms (Timberg and Stanley-Becker 2020).

The peak of Parler's popularity and attention occurred in the aftermath of the election on November 6th, when it had become clear that Trump had been voted out of office, a verdict Trump flatly refused to believe and fought against. As the site almost doubled its user base from 4.5 million to 8 million users over the course of a week after the election (Isaac and Browning 2020), the app rose to the top of the iOS App Store and Google Play Store's app charts; an analysis found that the site added 7,029 users per minute during the election itself (Stanford Internet Observatory

2021). As results showed a close but clear Biden victory, Parler was one of several fringe platforms to play host to the amorphous "Stop the Steal" movement designed to delegitimise the results, providing an outlet for the false claims that larger platforms were clamping down on (Dwoskin and Lerman 2020; Kelly and Schiffer 2020).

The storming of the US Capitol complex occurred on January 6, 2021, the day that Congress gathered to formally certify the results of the presidential election. Trump supporters, including a loose assortment of insurrectionists and far-right anti-government groups such as the "Oath Keepers" and the "Proud Boys", gathered to hear the outgoing president speak at a park adjacent to the White House (Blake 2022). Trump inflamed the crowd by declaring that "if you don't fight like hell, you're not going to have a country anymore," urging them to "walk down to the Capitol," and promising that he would join them (Naylor 2021).[67] The mob that descended on Congress, estimated to be in the thousands, had breached the Capitol complex by the early afternoon, rampaged through the building, and, in surreal scenes, reached the floor of the Senate chamber (Bennett et al. 2021). For his role in stoking the attack, and condoning it after the fact, Trump was indefinitely banned from Twitter and Facebook on January 8th. Following this move, Parler gained a further wave of sign-ups, which the company claimed took it above 15 million total users (Hagey and Horwitz 2021), and on the same day it again rose to the top of the iOS App Store chart (Shieber 2021). But this was to be the briefest of apogees.

5.1.4 January 2021: Parler's shuttering

That Parler was one of the social media platforms used to organise the storming of the Capitol on January 6th is not in dispute. A *Washington Post* story filed the day before the riot listed Parler alongside other popular platforms like Telegram and thedonald.win as places where "talk of guns and potential violence [in Washington DC the next day] is rife" (Timberg and Harwell 2021). The article picked up specific Parler posts, including one which stated, "To all the Patriots descending

[67] He did not.

on Washington DC on #jan6 … come armed …" and another which warned "the American people on the ground in DC today and all over this great nation, [to] be prepared for anything." Contemporaneous news reports like this are especially useful insofar as they demonstrate that reports of high levels far-right radicalism and extremism on Parler were present *prior to* the Capitol riot itself. Retrospectively, however, in a March 2021 letter from Parler's lawyers to the US House Committee on Oversight and Reform, the company contended it played a "positive role … in the days and weeks leading up to January 6th."[68] The letter includes the claim that ahead of the riot, "Parler referred violent content from its platform to the FBI for investigation over 50 times." This, the company claimed, exculpated Parler from the role of "far-right instigator and rogue company that Big Tech has portrayed [it] to be". By way of example, Parler claimed that on January 2nd it had reported to the FBI a post on Parler stating that "I trust the American people will take back the USA with force and many are ready to die to take back #USA."

The removal of Parler from app stores on January 8th is covered in Chapter 6. Though there was no immediate *technical* effect of the removal of Parler from app stores on the company's public-facing website, it immediately preceded, and likely helped to precipitate, Amazon's withdrawal of cloud services from Parler a day later. On January 9th, Amazon announced that it would be withdrawing its Amazon Web Services (AWS) cloud hosting from Parler, effective the following day, citing the calls for violence that were prevalent across the platform. Amazon also cast doubt on Parler's moderation capacity, stating that "it's clear that Parler does not have an effective process to comply with the AWS terms of service" (Paczkowski and Mac 2021). As a result, Parler's website went offline on January 10th. It would be a week until Parler returned to the internet in any form—with a simple static webpage at www.parler.com containing a message purportedly from John Matze regarding "technical difficulties" (see Figure 2).

[68] https://www.justsecurity.org/wp-content/uploads/2021/06/Jan6-Clearinghouse-Parler-Letter-to-Chairwoman-Maloney-March-25-2021.pdf

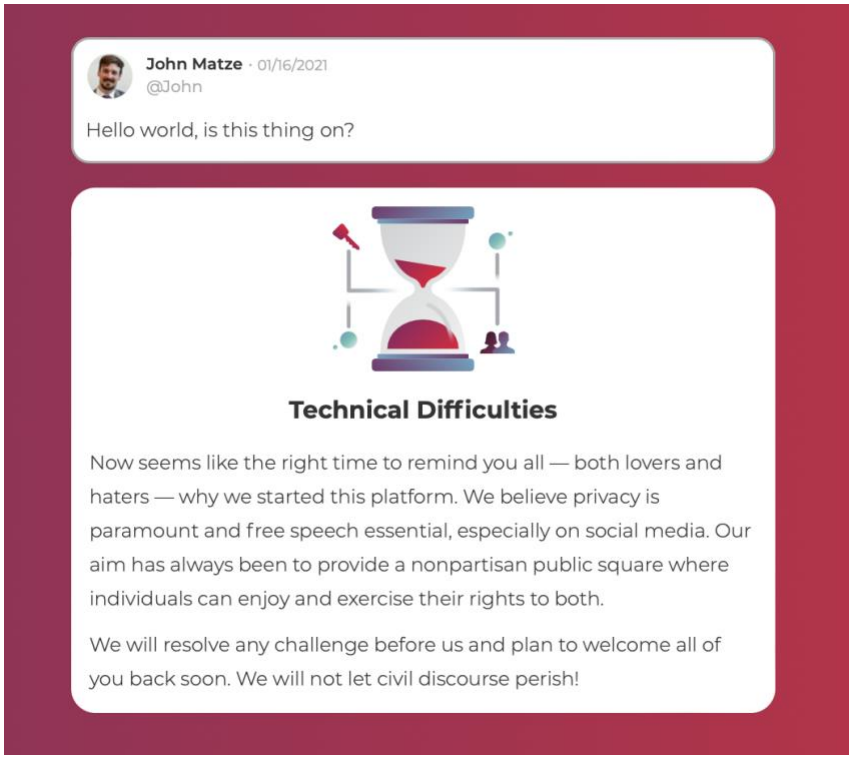


Figure 2: The home page of Parler's website after the restoration of its service in January 2021. Source: The Internet Archive, https://web.archive.org/web/20210117083806/https://parler.com/

Using the same colour palette and drawing once more on the rhetorical touchstone of free speech (albeit spelled correctly this time), this static landing page is visually and symbolically reminiscent of Parler's first webpage, which appeared in June 2018. But as this section has documented, the thirty months between the appearance of these two online artefacts saw the concept of untrammelled "free speech" on social media being stretched to its limit on, and by, Parler.

5.2 Existing research on Parler: a systematised thematic review

In order to gain an overview of existing scholarship on Parler, much of which has emerged since its removal from app stores in January 2021, I undertook a systematised review of recent articles discussing Parler (Grant and Booth 2009). This consisted of two phases. In the first phase,[69] I searched two large academic databases, Scopus and Web of Science, using structured queries

[69] The original searches were conducted on 6 June 2022. They were conducted again and updated where necessary based on a search on 3 April 2023.

designed to find articles mentioning "Parler" in their title, abstract or keywords.[70] I supplemented this with a broader search of Google Scholar,[71] which does not allow specific searches only of titles, abstracts or keywords, but instead returns results based on "relevance". All searches were conducted on papers published from 2018 onwards. Because "parler" is a French word, all searches were filtered to include only English-language articles.

The searches of Scopus and Web of Science yielded 84 and 39 matching results respectively. I manually filtered these results for relevance, removing several false positives, most of which[72] had been included due to their use of the word "parler" as a French word even in English-language works.[73] This left nineteen results from Scopus and nine from Web of Science; all the latter results were contained in the former, indicating internal consistency across the two database queries. I then reviewed the first 100 results from the Google Scholar search for relevance, identifying 24 additional academic works with relevance to Parler. 22 of these 24 true positives were contained within the first 50 results, suggesting diminishing returns from further searching beyond 100 results. In attempting to access the articles, I found that one was inaccessible (Boyd 2020) and that one was written in Italian despite its English-language title (Puddu 2021). I

[70] For Scopus, the query was: "TITLE-ABS-KEY (parler) AND PUBYEAR > 2017 AND (LIMIT-TO (LANGUAGE , "English"))". Query URL: https://www.scopus.com/results/results.uri?cc=10&sort=plf-f&src=s&st1=Parler&nlo=&nlr=&nls=&sid=b20adf82eff444959c62949abbc36685&sot=b&sdt=cl&cluster=scolang%2c%22English%22%2ct&sl=40&s=TITLE-ABS-KEY%28Parler%29+AND+PUBYEAR+%3e+2017&ss=plf-f&ps=r-f&editSaveSearch=&origin=resultslist&zone=resultslist
For Web of Science, the query was: "Parler (Title) or Parler (Abstract) or Parler (Keyword Plus) and English (Language)" with Timespan: 2018-01-01 to 2022-06-01 (Publication Date). Query URL: https://www.webofscience.com/wos/woscc/summary/4dd1a9c6-556e-4d8b-894f-2f6165611840-39c837fd/relevance/1

[71] The Google Scholar search was for "Parler" in works between 2018 – 2022, sorted by relevance, and filtering for English pages. Query URL: https://scholar.google.com/scholar?lr=lang_en&q=Parler&hl=en&as_sdt=0,5&as_ylo=2018&as_yhi=2022

[72] I also removed from consideration an article for which I was an author (Cowls, Morley, and Floridi 2023) and an article written by my supervisor (Floridi 2021b).

[73] For example, one such result was titled "Why we need to talk about preventing depression [De la nécessité de parler de la prévention de la dépression]".

supplemented this original search with a follow-up,[74] which yielded a further eight relevant studies. The identification of studies is shown in Figure 3 below based on the PRISMA flow diagram developed by Page et al. (2021).

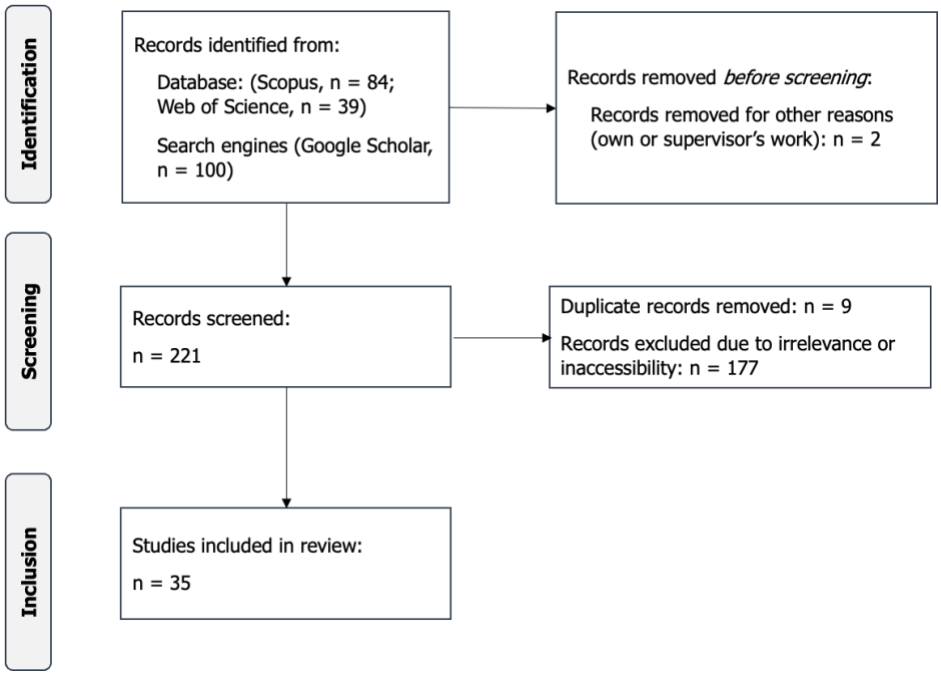


Figure 3: PRISMA flow diagram depicting identification of records on Parler (first phase).

The combined database of 43 non-duplicate research works (see Appendix 2 for the full list) provided a substantial but manageable set of resources with which to identify key insights and debates relating to Parler. The 43 research works can be grouped in terms of their focus. A foundational paper which many subsequent papers cite, and draw upon, is the massive Parler dataset presented by Aliapolulios and colleagues (2021). Their publicly available dataset of 183 million Parler posts and data for 13.25 million users offers various insights into the platform, including the domination of conservative topics and voices, and gives an overview of various features of Parler including flags, badges, and the use of hashtags.[75]

[74] The original search was conducted on 6 June 2022 and the follow-up was conducted on 3 April 2023. In contrast to the original search, in the follow-up search I reviewed all results meeting the criteria that were published since the original search was conducted. This amounted to 58 further results, of which I deemed an additional eleven as relevant, three of which also arose in the Scopus/Web of Science searches.

[75] I discuss my own use of this dataset in 5.4.6.

Other articles use this dataset or self-collected data to offer other descriptive insights about Parler, often using other platforms like Twitter and Gab for the sake of comparison. Unsurprisingly, many such studies focus on the general prevalence of hateful, violent, and conspiracist content on Parler. A thesis project by Ward (2021) found that around 2% of parleys and comments contained a hate word, while Israeli and Tsur (2022) find that "hatemongers" made up 16% of Parler's active user base, and were more prolific and central to the network than non-hateful users. Esser (2021) finds that the most popular topics of conversation on Parler were centred on Trumpism, conspiracy theories and far-right tropes, and suggests that radical(ised) users deployed strategies to create an "us versus them" dynamic to legitimate violence. Meanwhile, Sipka et al. (2021) compare the spread of QAnon-related content on Parler, Gab, and Twitter, finding that although Parler had on average a higher level toxicity of posts containing the #Qanon hashtag, this content was more related to support for Trump than QAnon conspiracism, compared with Gab (Park et al. 2022). Other papers focus more narrowly on high-profile Parler users. In their computational content analysis of the websites shared on Parler by a selected group of political users, Catipon and Savas (2021) find that websites shared on Parler were more extreme-right than on Twitter. Otala et al. (2021), meanwhile, analyse the use of Twitter and Parler by US politicians with accounts on both platforms, finding that these politicians had not fully "migrated" from Twitter but continued to use both platforms—indeed, Twitter more so.

Other studies in the sample focus directly on the use of Parler in preparation for, and during, the Capitol riot. Munn (2021) argues that Parler served as "preparatory media" for the riots, exploring how an online platform such as Parler can frame events and identify and set a shared agenda for violent offline action, making Parler fundamentally "more than a mob". Hitkul et al. (2021) find that a significant proportion of Parler content at the time cast doubt on the validity of the 2020 election, the main trigger for the riot. Norgaard and Walbert (2023) use a variety of sentiment analysis to measure the increase of group polarisation that arose during the riot, detecting some dissent within the Parler "echo chamber" on and shortly after January 6th.

Several papers looking at Parler's role in the riot deploy novel methods to drive insights. Jakubik and colleagues (2022) claim to identify the emotions that were most prevalent among Parler users on January 6th 2021. Ng et al. (2021) use text similarity metrics to identify coordinated narratives among clusters of users in the immediate run-up to and aftermath of the riot, and later work by Ng and colleagues (2022) explores the dissemination of disinformation on Parler during the riot, finding that sources of information consumption differed between Parler and Twitter users, despite similar narratives spreading. Tachaiya et al. (2021) use NLP techniques to classify the sentiments and stance of users towards political figures, finding that users turned against Vice President Mike Pence on and after January 6th, following his role in certifying the election results.

Several recent studies explore the effect of the mass deplatforming of individuals from mainstream social media platforms that followed the January 6th riots. In a wide-ranging study, Buntain and colleagues (2023) detect further evidence of what I above call the "performative migration" of departing users announcing their intent to join other platforms on existing mainstream platforms. Agarwal et al. (2022) meanwhile use user mobile data to find that after its shuttering, a section of Parler users migrated to Telegram, a trend driven mostly by highly engaged and partisan users. This is echoed by work by Horta Ribeiro et al. (2023) which finds that the removal of Parler caused users to increase use of other fringe social media sites such as Gab and Rumble.

A separate set of research articles look at the role of Parler in the spread of disinformation. The above-mentioned paper by Pieroni et al. (2021) fuses UX design and data science to evaluate key features of the Parler platform, from its colour palette to its presentation of content, and explain how factors like anonymity, the prioritisation of images, and the weight given to "influencers" made Parler especially conducive to the spread of disinformation. They remark on Parler's "rustic appearance", and its colour scheme dominated by reds and greys, noting the use of red to evoke strong emotions, and speculated that the persistence of glitches and lack of polish on Parler may have been designed to bolster its "outsider" status in comparison to its rivals. They also

highlight the prominence of images on Parler's profile pages and the relatively muted contrast on text which "pulls attention away from the source, allowing the user to focus more intently on the content of the parley, regardless of source reliability" (2021:2-3). This, they suggest, increased the perception of anonymity on Parler and made the platform more conducive to the spread of false information.

Greenhalgh et al. (2021) use Parler as a case in support of greater efforts by educators to combat what they call "miseducational" informal learning online, while Dori-Hacohen and colleagues (2021) likewise use the case of Parler as argument for the deployment of novel content moderation methods to better detect online disinformation. In a similar vein, Fedoruk et al. (2021) use Parler as a basis for developing what they call a "plebeian algorithm": a social media algorithm which optimises for the reduction of misinformation. Other studies focus on disinformation specifically related to Covid-19. In an article in *Vaccines,* Baines and colleagues (2021) draw on a random sample of Parler posts using popular vaccine-related hashtags to conduct a descriptive analysis of the themes (such as vaccine refusal and alleged side-effects) and terms used in discussions of vaccines (c.f. Dasgupta 2021). Yang and colleagues (2021) find that compared with Twitter, debate about Covid was less balanced on Parler, with particular negativity among Parler users towards lockdowns and mask mandates. By contrast, a study by Noor and colleagues found a lower level of toxicity regarding Covid-19 discourse on Parler as compared to Reddit (Noor et al. 2023).

Some of the articles found are primarily methodological, using elements of Parler as the basis for developing tools or strategies for researching phenomena on mainstream or fringe social media platforms. Ravazzi and colleagues (2022) present an approach for detecting bot accounts on fringe platforms before the malicious content they spread reaches mainstream platforms. In work on "Alt-tech social forensics", Johnson and colleagues (2022) study how Parler and other fringe platforms handled user data, while Torene et al. (2022) use Twitter and Parler data to test the automatic creation of semantic hierarchies in user-created hashtags.

Finally, some of the most insightful findings are offered by reports and policy papers. In work published by Stanford University's Internet Observatory, Thiel, DiResta and Grossman (2021) present a helpful overview of key dynamics on Parler, noting how Parler was moderated (it had 802 moderators for 13 million users as of January 2021), and highlighting the distinct communities of Saudi Arabian and Brazilian users mentioned in 5.1. In a report for a Brazilian research foundation, Ruediger et al. (2021) focus on Parler's Brazilian user base and its relations with other users, identifying a "transnational network engaged in the reproduction of discourses aiming to corrode the credibility of democratic institutions". Meanwhile, a ProPublica hub titled "What Parler Saw" provides a curated set of videos posted on Parler during—and from within—the Capitol riots (Groeger et al. 2021). The sample is rounded off by a series of rapid-reaction policy briefings on different aspects of Parler's chronology by research consultancy Oxford Analytica (Oxford Analytica 2021c, 2021b, 2021a).

5.3 Case selection

In this section I first explain the suitability of Parler as a case study. The decision to conduct a case study on Parler reflects three main epistemic considerations. First, using George and Bennett's (2005) terminology, the case that I study can be thought of as an *instance* (the removal of Parler from app stores) of a *class* (app store governance). George and Bennett consider a *class* to be any "phenomenon of scientific interest, such as revolutions, types of governmental regimes," and so on (2005:17). As I argued in Part I, app store governance is a particular "regime" of internet governance. In defining a case study as a "well-defined *aspect of* a historical episode that the investigator selects for analysis" (2005:18; italics added), George and Bennett draw on the example of the Cuban missile crisis, which they describe as a "historical instance of many different classes of events" including coercive diplomacy and crisis management (2005:18). While holding somewhat less existential stakes, my case study can be considered ontologically similar: Parler's

removal from app stores occurred in the immediate aftermath of the Capitol riots and was just one aspect of this broader, multifaceted historical episode.

Second, the case study on Parler is strictly speaking a "within-case" analysis insofar as I did not directly compare the case with others. However, there are two caveats to this classification. Firstly, I situate the removal of Parler within the broader ecosystem of rules and norms that constitute cross-platform governance, and where necessary draw on other proximate examples of gatekeeper governance—such as the removal of the opposition app Smart Voting from Russian app stores—to draw out pertinent insights. Secondly, the case study was structured in such a way as to enable a "before and after" analysis of Parler's removal. The data I analyse in Chapters 6 and 7 cover the year *before* Parler's removal, which allow me to ask not just "why then?" (i.e., January 2021, immediately after the Capitol riots), but also "why not sooner?". This nuances the "within-case" classification of the study, and can be considered a chronological equivalent of Mahoney and Goertz's (2012) observation that researchers sometimes deliberately select a case in which an outcome of interest has *not* occurred.

Third, the Parler case meets one of Flyvbjerg's (2006) several possible purposes for the reflexive selection of a particular case. Superficially, the Parler removal could be thought of what Flyvbjerg considers an "extreme" case. This has resonance in the sense that Parler was removed in the immediate aftermath of a shocking historical event motivated by extremist ideology. But the removal of Parler, while certainly occurring after a striking event, was not "extreme" in the history of app store governance or cross-platform governance more broadly, which (sadly) furnishes several other examples of platform shutdowns in response to shocking, extremist events, such as Cloudflare's termination of its service for 8chan following the El Paso shooting in 2019 (see Part I). Instead, my selection of Parler more precisely meets the definition of a "critical" case, which Flyvbjerg defines as "having strategic importance in relation to the general problem" (2006:14), in the sense that cases are strategically selected in order to permit logical "if … then" deductions—along the lines of "if this is (not) valid for this case, then it applies to all (no) cases" (2006:34). The

ability this affords researchers to make such logical deductions of this sort enhances the generalisability of the case study method.

Critical cases can be of either "most likely" or "least likely" forms. To paraphrase Flyvbjerg, "most likely" cases can be thought of as having the form "if some action did *not* occur in *these* circumstances (in which it is most likely to have occurred), then it can be also expected *not* to occur in circumstances with a similar or less likelihood." Applying this logic to my study, a priori it seems there were grounds on which Parler could have been proactively removed sometime before January 2021—evidenced by, among other things, the speed with which it was removed in the immediate aftermath of the Capitol riots—but it was not. If the case study supports this a priori assumption, then it is reasonable to assume that all other platforms with similar and lower levels of problematic content were also not removed from app stores. This is a formal way of stating a more intuitive point about the hoped-for contribution of this case study to app store governance research. As I discussed in Part I, there is considerable ambiguity with respect to how and why app store operators like Apple and Google decide to remove apps from app stores on the basis of hate speech. In this context, the Parler case can be seen as providing a precedent for the retention, rather than the removal, of problematic platforms in app stores. This exercise is of course naive, in view of the other motivations and incentives—such as corporate reputation—that also influence app store governance in practice (see Chapter 3). Yet this naivety is a useful starting strategy, insofar as it helps to document potentially insufficient app store governance standards and/or a failure by app store operators to enforce them—in short, to enable the critical outside scholar to ask: "why was action not taken sooner?".

There is an additional, unanticipated way in which the Parler case ended up fitting Flyvbjerg's definition of a "critical" case, this time in the inverse "least likely" form. Although the classifications of cases as "extreme", "critical", "paradigmatic", and so on are designed to guide the strategic selection of cases, Flyvbjerg notes that in practice, the classification of cases can change *during* the case study in unforeseen ways. Flyvbjerg relays the "frustrating" experience of

initially selecting a particular case study as an example of a critical case, only to discover that it was in fact an extreme case after further facts became available during the study (2006:15). A similar experience occurred during my case study. The retention of Parler in Apple and Google's app stores until after the explosive events of January 2021 seemed to have provided me with the ideal "critical" case study, in the "most likely" sense discussed. Yet the events of the following months would conspire to complicate this classification. In May 2021, a new, allegedly better moderated version of Parler was re-admitted onto Apple's, but not Google's, app store. Google then followed suit more than year later, in September 2022. Apple's reverse move was an unexpected (by me) development, given that the political fall-out from the Capitol riots and Trump's attempt to violently retain power were still very vivid four months on. On a conceptual level it meant that the case study ought to be partially reclassified. A priori, Parler's reinstatement in May 2021 appeared as a "least likely" case, in the sense that it seemed unlikely that a platform that had been removed just months earlier for its alleged role in fomenting violent insurrection had (in Apple's case) been so swiftly restored. As a corollary to its removal, the unforeseen and outwardly "unlikely" reinstatement of Parler in turn provided an implicit precedent for the reinstatement of similar platforms that were equally as, or less, problematic.

It is important not to overstate the generalisability of this case study, nor indeed to overstate the benefits of generalisability as a goal for social scientific research. As Flyvbjerg notes, "formal generalization is only one of many ways by which people gain and accumulate knowledge" (2006:10). And indeed, there are many factors which make the Parler case distinctive and indeed unique—which, in turn, may make it difficult to generalise to other "more likely" or "less likely" cases. Nonetheless, as a starting strategy, I found Flyvbjerg's framework of case-classification useful to reflect on and anticipate the ways in which I felt that the Parler case might hold lessons for app store governance more generally—depending, of course, on what I was to find.

5.4 Methodological approach, data collection, and analysis

In this section I introduce the methodological approach taken to the case study and describe the data that I gathered and analysed in service of this approach.

In Chapter 6 I use the "walkthrough method" to analyse Parler's basic affordances, vision, and governance in order to answer RQ2a, which asks how Parler performed its platform self-governance. In 5.4.1 I introduce the walkthrough method, and in 5.4.2-4 I discuss the different forms of analyses that I undertook to the data gathered as part of the walkthrough method.

In Chapter 7 I report the findings of a content analysis of hate speech on Parler that combines automated and manual classification, in order to answer RQ2b, which asks whether Parler ought to have been removed from app stores sooner given app store governance guidelines around hate speech. I discuss this process in 5.4.6.

5.4.1 The walkthrough method

The overarching approach I took to collecting and analysing data is inspired by an approach to the study of smartphone apps known as the walkthrough method (Light et al. 2018). This method is designed to allow researchers to "perform a critical analysis of a given app" (2018:881). The method is grounded in STS and is founded on several assumptions which are salient to the theoretical discussion in Part I and relevant to the case of Parler. First, the method is designed to interrogate the ways in which "technologies service the cultural aspirations of their creators, who often accrue power by oppressing particular groups." This is accomplished by assessing how "technologies are deployed with symbolism and rituals to reinforce these aims" (2018:887). Second, the method is situated within existing research on infrastructural systems (2018:887) and extends these analyses to encompass how "qualities specific to apps, such as … distribution through app stores, also feature in cultural struggles among a technology's economic, political and social players" (2018:888). Third, the method does not take how an app is governed as a given, but instead includes its governance in the analytical framework (2018:890-891).

The walkthrough method consists of two main steps, which can be undertaken consecutively or concurrently: (i) establishing the app's "environment of expected use", which "points researchers towards pivotal aspects of an app's context for analysis in conjunction with its technological architecture" (2018:889); and (ii) conducting a "technical walkthrough", which "involves the researcher engaging with the app interface, working through screens, tapping buttons and exploring menus" (2018:891). I used this bipartite structure as the basis for my data collection. In what follows, I take forward the division between, first, the Parler smartphone app as a source of data (analysed via a technical walkthrough), and second, all other materials found beyond the app itself, including Parler's public-facing website (analysed to establish the app's environment of expected use). However, at this juncture, it is important to clarify this distinction, and to acknowledge one sense in which it is ambiguous. As I will discuss, Parler's website (www.parler.com) served an important public-facing purpose for the company, promoting the platform's affordances and encouraging users to join. The most common way for users to sign up and actively use the Parler platform was via Parler's smartphone app, and—because the scope of this thesis is the governance of smartphone app stores—it is the smartphone app that forms the basis of the technical walkthrough component of my analysis (see below). However, it was also possible for users to access and use the Parler platform through a web browser on a desktop computer, via a "Log in" option on Parler's website. Thus, in the analysis that follows, when I discuss "Parler's website" I am referring to the part of it that was publicly accessible and relevant to establishing the app's environment of expected use, and thus excluding the desktop version of the platform which users had to sign into in order to access.

In the case of establishing Parler's "environment of expected use", I gathered as much material as possible to understand, as Light and colleagues put it, "how an app's designers, developers, publishers and owners expect users to receive and integrate it into their technology usage practices" (2018:889). They identify three components that constitute this phase of the analysis. The first is an app provider's *vision*, which "involves its purpose, target user base and

scenarios of use, which are often communicated through the app provider's organisational materials" (2018:889). These materials may include app store entries, webpages, company blogs, press releases, and public statements. This helps to "establish an app's discursive and symbolic representation (e.g., logos, colour schemes, images), which is carried through to its technical interface" (2018:889). The second component is an app's *operating model*, which "involves its business strategy and revenue sources, which indicate underlying political and economic interests". Relevant materials may include technology industry sources and public market information. The third and final aspect of an app's environment of expected use is its *governance*, which involves "how the app provider seeks to manage and regulate user activity to sustain their operating model and fulfil their vision" (2018:890). Relevant materials include an app's rules, guidelines, and governance procedures.

I used this guidance to order my data collection. I drew materials from three sources: mainstream media coverage of Parler; archived versions of Parler's public-facing website; and archived copies of Parler's policy documents. These sources and the specific datasets they yielded are summarised in Table 3. Each of the datasets was applicable to more than one component of the environment of expected use, as indicated in the table. All three datasets touched on Parler's vision and its governance. With respect to Parler's operating model, mainstream media coverage included discussion by Matze of the platform's "influencer network", while some of Parler's policies touched on the role of its influencers as part of its operating model.

Table 3: Sources of data gathered for the environment of expected use analysis

Data source	**Data set**	**Vision**	**Operating model**	**Governance**
Mainstream media coverage of Parler	186 news articles	X	X	X
Archive of Parler's website	12 archived webpages	X		X
Archive of Parler's policy documents	18 policy documents	X	X	X

The second step of the walkthrough method is the technical walkthrough, in which "the researcher assumes a user's position while applying an analytical eye to the process of acquiring the app, registering, accessing features and functionalities and discontinuing use" (2018:891). This aspect of the analysis was harder to accomplish for the simple but problematic reason that Parler was removed from app stores during the early stages of the empirical research (recall that it was Parler's removal that primarily motivated my initial interest in it as a case). Though I was able to gather and analyse data from a technical walkthrough during this short time-window, I was unable to do so as systematically or comprehensively as I would have liked. Therefore, where necessary, I have supplemented aspects of the technical walkthrough with other found sources, including screenshots from elsewhere on the web.

The article introducing the walkthrough method provides pointers for how researchers should conceptually approach the data that they gather. However, it leaves decisions open to the researcher about how exactly to analyse the various sorts of found material. This is a useful allowance insofar as the data gathered can be—as it was in my case—rather heterogenous. Therefore, I assessed each form of data differently, as I discuss in more detail in each of the following three subsections. However, what unites them is my use of discourse analysis in each case. The past decade has seen greater interest in discourse-analytical approaches to internet governance, with studies focusing primarily on multistakeholder fora for internet governance such as the IGF and associated discursive artefacts such as mailing lists (Epstein 2013; Epstein et al. 2016; Hofmann 2020; Padovani, Musiani, and Pavan 2010; Ten Oever, Milan, and Beraldo 2020). Recent work has also considered the discursive efforts of private platform governors, such as Facebook's Mark Zuckerberg (Haupt 2021; Hoffmann et al. 2018a; Rider and Murakami Wood 2019) as part of growing interest in the narratives and imaginaries that technology companies, among other actors, deploy (Mager and Katzenbach 2021; Natale et al. 2019). Because the focus of this discursive work is legitimation, the specific form of discourse analysis that I undertake is inspired by Theo van Leeuwen's framework for discursive legitimation. For van Leeuwen,

discursive legitimation involves "the legitimation of the practices of specific institutional orders" (Van Leeuwen 2007:92); in this case, the specific "institutional order" in question is the governance of Parler as a platform. The legitimation components of a particular text[76] are those that concern the question "why": why something should be done, and why in some particular way, within the institutional order in question (Van Leeuwen 2007:93). In the following sections I explain how I operationalised this framework of discursive legitimation for each form of data that I analysed.

5.4.2 "The Matze files": Analysis of media coverage of Parler's CEO

The first set of data that I gathered as part of establishing the environment of Parler's expected use concerns the public statements made by its co-founder and CEO, John Matze, during 2020, in which he discussed the platform's self-governance. Since its founding, Parler's "free speech" credentials have been at the heart of its strategy to legitimate its self-governance and differentiate itself in the competitive social media landscape, and these featured prominently in coverage of Parler, even in mainstream outlets (e.g., the BBC). To gauge the extent to which, and the mechanism by which, "free speech" and other tropes were employed to promote and legitimate Parler in 2020, I chose to focus on the media appearances Matze as Parler's serving CEO. At the early stage of Parler's growth, Matze's primary purpose in 2020 was to win over new users to the platform, and the media provided a crucial avenue for Parler to reach these users, particularly via conservative outlets. Matze was very much the public face of Parler during this period, giving scores of TV interviews and quotes to press articles. This is to be expected, as the founder-CEOs of tech start-ups tend to enjoy a high degree of reverence with respect to the companies that they found, and empirical examinations of what they say publicly can lead to new insights (Haupt 2021; Hoffmann et al. 2018a; Rider and Murakami Wood 2019). Moreover, 2020 saw an unprecedented degree of mainstream media focus on issues raised by the governance of social media platforms governance. In particular, the steady increase in restrictions that Facebook and Twitter imposed

[76] Note that "text" here is meant flexibly, as van Leeuwen's framework is multimodal (see 5.4.3).

on the communications of President Donald Trump sparked opposition, both from Trump himself and from his supporters. Having honed the distinctiveness of its "free speech" appeal since 2018, Parler stood rhetorically well-placed to benefit in 2020, and Matze was more than willing to draw the contrast between his platform and its larger and allegedly more censorious rivals.

With the goal of gathering a comprehensive dataset of mainstream media coverage in 2020, I turned to two existing large databases of news media coverage: Media Cloud and Google News. The Media Cloud project is the result of a collaboration of several Boston-area universities and describes itself as an "open-source platform for media analysis."[77] Using its Explorer tool, I queried the Media Cloud "United States" news media database for stories whose text mentioned both "Parler" and "John Matze", between Jan 1 2020, and Jan 1 2021. This returned 198 results from a variety of news sources, the majority of which were in English. To ensure that as many relevant stories as possible were captured in my data set, I performed the identical query on Google News (US) across the same period, resulting in 105 stories. Happily, search queries made on both Media Cloud and Google News return stable URLs for each story, so I was able to easily compare the URLs to detect duplicates, and in doing so I identified 21 stories that appeared in both datasets, as well as 17 stories that were duplicated internally within one of the datasets. After removing these duplicates, this resulted in a combined dataset of 248 news articles with unique URLs that matched the search criteria. For these 248 articles I created a "clean" summary spreadsheet including the publication, title, date, and URL for the story, as well as a unique file identifier.

The next step was to create a complete database including the full text of the news articles, since the results returned from Media Cloud and Google News did not contain the body of text in each the story.[78] To build the full database I used the qualitative data analysis software NVivo,

[77] https://mediacloud.org

[78] For clarity it should be noted that, although the full text was not returned in the search results, Google News returns matching results from the full text of stories, not just headlines.

creating a new file for each of the 248 stories corresponding to the ID applied in the summary spreadsheet. I manually entered the URL for each story and copied across the title and all relevant text (i.e., excluding comments and other webpage artefacts) to populate each file. Nvivo automatically strips out non-text content, such as photos and videos, when copying from webpages, but certain text content unrelated to the story itself (such as "Read more" interstitials with links to other content) found its way into the files. Because I would be analysing each story personally and individually and would not be performing any automated labelling on or analysis of the corpus as a whole, this did not pose a problem; I simply ignored these irrelevant strings during the coding process.

The primary purpose of this process of accessing each page and copying across content was simply to create a full database of raw text files, but it had two further effects. First, it allowed me to exclude several more stories from the total of 248 for reasons which only became clear when reading the stories themselves (as opposed to the metadata). These included syndicated stories, where text was copied verbatim (usually from an agency such as the Associated Press) and republished elsewhere under a different URL (N=9); broken URLs (N=2); paywalled articles that I was unable to access with existing institutional subscriptions (N=16); articles which, despite the search parameters, fell outside of the time period under analysis (N=4) or did not contain the searched-for text strings (N=3); and non-English language articles (N=26). This resulted in a smaller dataset of 186 accessible, time- and content-relevant, English-language stories, for which I had copied the full-text into the NVivo for analysis. Second, accessing each story and copying across the relevant text was the first time I was exposed to the underlying content that I would later be analysing in earnest. Therefore, while I did not perform any coding, it was at this juncture that, inescapably, the iterative process of interpretation, sense-making and pattern-detection of qualitative analysis began. Notably, it was also during this preliminary examination of the dataset that I decided to use constitutional metaphors as the analytical frame for this analysis. Building on existing work on the media portrayal of Meta's Oversight Board as a "supreme court for

Facebook" (Cowls et al. 2022) as an example of the legitimation of platform self-governance, I similarly found that Parler's CEO Matze leant heavily on a variety of constitutional metaphors in his discursive portrayal of Parler.

With this preliminary coding schema and analytical frame established I began to code the data. This involved carefully reading through each story, and applying none, one, or multiple codes to each passage.[79] Prior to coding I developed a preliminary coding schema, whose contents was influenced by the scope and focus of the study, my reading of wider material about Parler, Matze, and the events of 2020, and the aforementioned exposure to the material when extracting it. This schema was iterative (see below). The initial scheme had a hierarchical structure, with three top-level "nodes" or code groups. First, I followed Hoffman et al. (2018) in distinguishing whether a particular passage was a ***Direct quote,*** either by Matze or from Parler as a corporate entity or one of its other public figures.[80] My second top-level code group was ***External entities***: other organisations who were frequently mentioned by Matze with respect to Parler. The initial codes in this group were for *App stores & gatekeepers*, *Rival networks*, and *Donald Trump and/or Republicans*; the *Mainstream media* was added during the process of coding. The third top-level code group reflected the aim of the study to understand Matze's invocation of constitutional metaphors to characterise the affordances and governance of Parler. Based on my earlier research and preliminary data exploration I had already noted several symbolic touchstones that Matze and Parler tended to employ, including descriptions of Parler as a "*town square*" where *free expression* was protected; claims of *truth and transparency* in terms of how content was moderated and served to users; and Parler's supposedly superior protection of users' *privacy and personal data*. Through the coding process I also observed other distinct values and symbols arise in the text, including appeals to *legal proceduralism*,

[79] Most codes were applied across an entire sentence and sometimes a whole paragraph, to ensure maximal context when reviewing all coded material; occasionally only a single clause was highlighted where the context was sufficiently clear and self-contained.

[80] Although many passages that were not quotes were also coded as a useful indicator of how Parler's communications were framed by the news media, my analysis is based solely on Matze's own direct quotes except where stated otherwise.

as well as references to more traditional *right-wing values*; these were all appended as additional codes. Figure 4 below shows the overarching coding schema and indicates which codes were amended (^) and appended (^^) as necessary during the process. With the preliminary coding scheme in place, I proceeded to code passages in each of the articles in the corpus. An example coded passage is in Appendix 3.

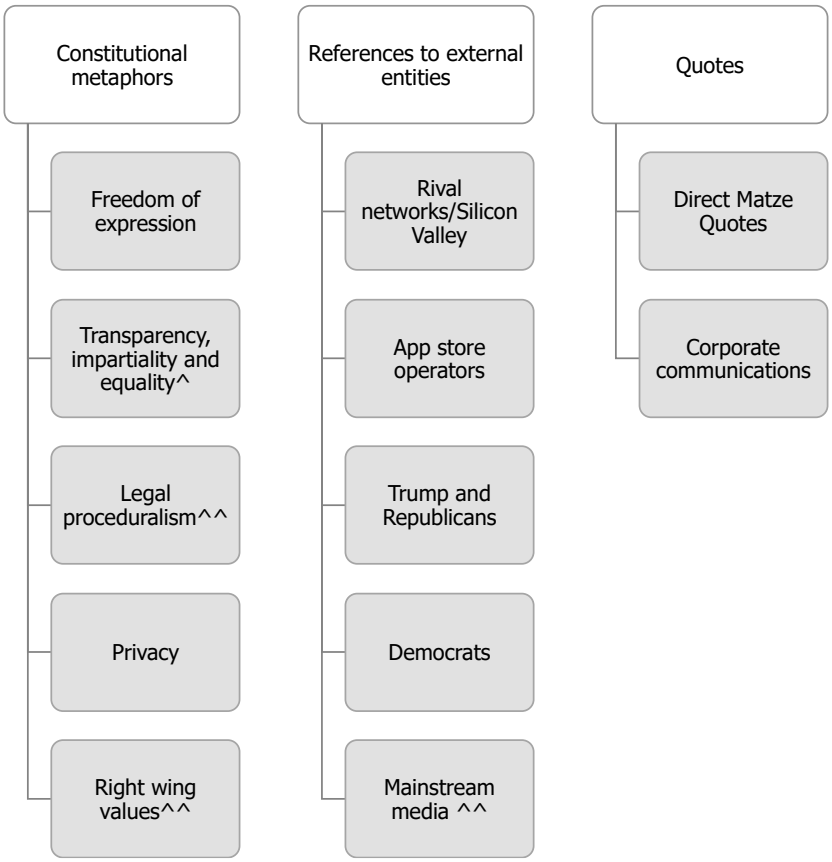


Figure 4: Multi-level iterative coding schema for mainstream media analysis

5.4.3: Multimodal discourse analysis of Parler's website

The discourse analysis extended to a multimodal examination of archived copies of Parler's website. At this juncture I briefly discuss the use of the archived web in research, which also applies to the policy analysis described in 5.4.4. I then discuss the analysis I performed.

As the internet has become an increasingly important object of historical and social study, web archives have served as a valuable resource for scholars. Many bespoke web archives exist,[81]

[81] For example, the UK Web Archive maintained by the British Library: https://www.bl.uk/collection-guides/uk-web-archive.

but by far the largest web archive is the Internet Archive, whose Wayback Machine allows for navigation of its vast set of archived websites. To build a corpus for my visual analysis of Parler's website, I used the Wayback Machine to navigate versions of the website that were preserved during 2020, and ultimately accessed different versions of eight archived webpages from across Parler's website, starting with archived versions of its homepage, www.parler.com. Web archives are a valuable resource, but come with various logistical challenges and epistemic limitations (Ben-David and Amram 2018). The Internet Archive's "crawler" navigates across the open web, using the structure of embedded links (rather like a search engine index) to access and download as much website content as possible (Schafer, Musiani, and Borelli 2016). This process of data collection is not therefore random and may overrepresent better-connected domains and sites in the network; consequently, the resulting archive is not a "random sample" in a statistical sense, which limits the empirical inferences that can be drawn from it. The crawling procedure also introduces other quirks. It is a convention that blogs, news websites, and social media sites include the date that a particular piece of content or post was published or shared online. By contrast, artefacts in the Internet Archive are arranged in terms of the date that they were accessed and archived. Thus, for a given archived webpage, the retrospective researcher cannot be assured of the date at which it was published online, only that it (and not some earlier or later version) was live at the specified URL at some point on the given date (and, for URLs that were crawled multiple times on the same day, within a particular time window). This also means that there may be iterations of webpages that were never crawled and archived and thus are lost to the virtual sands of time. Despite these limitations, web archives including the Internet Archive have been used in creative ways to shed light on the recent past. The internet itself is, understandably, a common object of web-archival research—but increasingly, scholars are also using web archives as a dataset with which to analyse broader historical and social phenomena out with "the internet" per se (Brügger et al. 2017).

Webpages (whether live or archived) are "seen" before they are "read". Websites are built from text, using various coding languages, and most pages contain at least some text designed for

an audience to read. At root, however, webpages are visual artefacts, and for all but the most barebones webpage, it is not fruitful (nor enjoyable) to try to separate textual elements from their visual context. As such, I extended van Leeuwen's discursive legitimation framework, which is explicitly multimodal (2007:107), to analysis of the archive of Parler's website. For each page in the corpus, I attended to how different text and images were used to convey meaning to Parler's intended audience.

5.4.4: Policy analysis of Parler's governance guidelines

Finally, the discourse analysis was extended to the third dataset: a collection of Parler's policy documents, each of which was in place for at least some part of 2020. The corpus of policy documents that I assembled is a rich source of data, encompassing rules, and the justification of these rules, in relation to Parler's self-governance. In order to obtain all archived material related to Parler's policies and their enforcement, I used the Internet Archive's "wildcard" function, appending an asterisk at the end of "legal.parler.com/documents", which my preliminary research had suggested was where Parler's key documents were deposited within the structure of its website.[82] This yielded 26 unique URLs,[83] many of which were duplicates.[84] After deduplication, I identified five types of available governance documents: "Community Guidelines"; an "Elaboration" on these guidelines; various versions of Parler's "Privacy Policy"; its "User Agreement" (which later became known as "Terms of Service"); and a factsheet about Parler's "Community Jury" procedure. Table 4 summarises these governance documents and includes a count of the number of versions that were possibly in force at some point during 2020.[85]

---

[82] I based this starting assumption on the fact that the three policy documents that were made available in the footer of Parler's homepage— "User Agreement", "Privacy Policy", and "Community Guidelines"—were all contained in the legal.parler.com/documents subdomain. See e.g., https://web.archive.org/web/20200902003443/https://parler.com/

[83] Accessible at https://web.archive.org/web/*/https://legal.parler.com/documents/*

[84] This is a result of how the Internet Archive crawls the web. For example, one of the "unique" URLs was a copy of Parler's privacy policy, http://legal.parler.com/documents/privacypolicy.pdf, followed by a query string (i.e., a question mark symbol followed by a long string of characters.

[85] Erring on the side of caution—because we do not know when archived websites were published, only when they were archived—in relevant cases I included in the sample the final iteration of each document archived in 2019.

Table 4: Summary of documents used in policy analysis.

Document name(s)	Description	Count (unique)
"Community guidelines"	Guidelines on permitted and prohibited content	6
"Elaboration on guidelines"	Further detail on permitted and prohibited content	1
"Privacy policy"	Procedures for handling and processing of users' personal data	4
"User agreement"/ "Terms of service"	Conditions users must accept for using Parler	6
"Parler community jury"	Explanation of the platforms' "Jury" review procedure	1

As discussed in 5.4.3, in web archival research, it is only possible to tell when each webpage was archived, not when it was published, and the possibility exists that certain iterations of webpages at the same URL were never archived. This has particular relevance for the rules contained in Parler's policy documents, which were often wont to change considerably from one iteration to the next. Happily, most of the documents I found are dated, which tells us when they became effective. But it is not possible to say with certainty whether there are iterations of each document missing from the sample, and thus not possible to add an effective "end date" to these documents.

After carefully examining each of the documents, I decided that two types, the privacy policy and the user agreement/terms of service, were not within the scope of my research. Though interesting,[86] the two types of documents did not touch on issues related to the moderation of hate speech. Therefore, in Chapter 6 I explore each iteration of the remaining three sets of documents that were germane to Parler's platform self-governance: its "Community guidelines" (which had six iterations), an "Elaboration" on these guidelines, and an explanation of Parler's "Community Jury" procedure. These were the documents included in the analysis.

[86] Parler's privacy policy during the period makes interesting retrospective reading in light of the platform's problems with data security and privacy.

5.4.5 A (time-constrained) technical walkthrough of Parler's app

As discussed in 5.4.1, the other major component of the walkthrough method is (as the name suggests) the walkthrough itself. In the short time window between the Capitol riot and Parler's removal from app stores and removal from the wider web, I was able to undertake an abbreviated technical walkthrough of the app, which I supplemented with screenshots of Parler's smartphone app found elsewhere on the web.

Light et al. (2018) centre the technical walkthrough on what they call *mediator characteristics*. These include an app's *user interface arrangement*, which concerns the placement of buttons and menus to guide users; its *functions and features*, such as pop-ups and compulsory fields; its *textual content and tone*, which concern menu options and categories and how these may exact "discursive power to shape [the app's] use"; and its *symbolic representation*, which involves examination of the app's look and feel, as well as its "likely connotations and cultural associations with respect to the imagined user and ideal scenarios of use" (2018:891-2). This schema beckons the researcher to examine how an app's mediator characteristics relate to, and embedded in, "cultural texts and understandings that exist outside the app" (2018:891). It proved highly suited to the case study on Parler, as it allowed me simultaneously to assess Parler as a smartphone app in the context of app store governance, and to tie this assessment to relevant political and cultural factors, such as the implicit positioning of Parler as a conservative alternative to mainstream social media platforms.

Light et al. also identify three main stages of a user's "lifecycle" with an app: registration and entry; everyday use; and app suspension, closure, and leaving. Owing to Parler's shuttering, my analysis of each of these stages had varying stages of completeness, which I discuss in turn below with reference to the mediator characteristics identified above. As far as possible in each of these stages, I followed the suggestion to "assume a user's position while applying an analytical eye", paying special attention to the "app's materiality", and drawing on "textual and semiotic analysis [to recognise] indicators of embedded cultural discourses" (2019:891). The findings of the analysis are reserved for Chapter 6.

I was able to *register for and enter* the app successfully. Although it is possible to register for Parler via a desktop web browser, I decided to register directly inside the app to tie my experience as closely as possible to the focus on app store governance. This involved, first, downloading the app from Apple's iOS App Store, where I took note of the visual and written presentation of Parler to would-be users.[87] I then registered for the app, using a pseudonym as a handle, and agreeing to, and recording, the platform's governance documents (which were presented as full-screen pop-ups within the app). Then, I recorded the steps involved with "entering" the app, which consisted of an option to "personalize your Parler experience" by following popular users (this important feature is covered in Chapter 6). I was then presented with my first Parler post, which was autogenerated for all users.

I was also able to undertake some *everyday use* of the app while it was available. Light et al. note that "not all functions may be easily accessible to a researcher who does not wish to or is not able to interact with other users, pay money or execute certain activities," but nonetheless that "walking through even the app's basic functionality provides a sense of what activities it enables, limits and guides users towards" (2018:893). I took the latter route, observing key mediator characteristics without interacting with other users or transacting with the app. This nonetheless allowed me to engage with each of the app's basic functions. I undertook this systematically by examining the screens associated with each of the options on the app's two main menus: the bottom menu, which included icons for accessing content feeds, notifications, and search; and the sidebar menu, which included various settings. I was also able to view the screen where "parleys" are created, and, because of the auto-generated first post, I could see the "engagement" screen offering stats like a view count for this post. Besides the unanticipated shuttering of Parler which

[87] It is interesting to consider in which part of the walkthrough method—establishing its environment of expected use, or the technical walkthrough—an app's listing in an app store should be situated. Light and colleagues discuss app store entries as a basis of an app's vision as part of the environment of expected use, but also more implicitly note that "initial contact with an app may … be part of a larger infrastructure", which could include an app store. This choice is not mutually exclusive; as they also note, "an app's expected use is often strongly communicated during registration". I thus incorporate analysis of Parler's app store entry as I encountered it prior to downloading and registering for the app as part of the technical walkthrough.

limited how many times I was able to walk through the app, constraints on this phase of the technical walkthrough also arose from my decision not to engage with the app's users, i.e., not following anyone or posting any content. This meant that I did not experience a fully featured content feed, or to be notified about interactions with other users. However, this limitation had a useful upside. It meant that (with the sole exception of the "personalize your Parler experience" pop up), I experienced the app without exposure to the app's content, which as noted was highly skewed towards right-wing and often fringe viewpoints (see Chapter 7). I found this to be a useful way to isolate the app's mediator characteristics, especially its tone and symbolic representation, and how these related to Parler's *imagined* and implicitly *ideal* users and scenarios of use, absent the decidedly messy reality of the app's *actual* use and users, which are discussed separately (see 5.4.6 and Chapter 7).

Finally, I was not able to record any data regarding *suspension, closure and leaving* of Parler. In the walkthrough method, these terms refer only to actions taken by users themselves to exit an app, but because of Parler's sudden and unexpected shuttering this was not possible in my case.[88]

5.4.6 Content analysis of Parler posts

In Chapter 7 of the thesis, I turn to analyse the level of hate speech on Parler during 2020. To do so I used a time-bound subset of an existing dataset of 183 million posts and comments published on Parler between 2018 and 2021 (Aliapoulios, Bevensee, Blackburn, De Cristofaro, et al. 2021). The dataset was collected and shared online by Aliapoulios and colleagues shortly after Parler was taken down in January 2021. A revised version of this data paper was presented at a leading web and social media conference science later that year (Aliapoulios, Bevensee, Blackburn, Bradlyn, et al. 2021). Since then, the dataset has been used by other researchers to answer various research questions about Parler (Catipon and Savas 2021; Israeli and Tsur 2022; Noor et al. 2023; Ward 2021). This dataset consists of "classic" social media data, including date and timestamps, the

[88] The app would only return to full functionality after the end of the period of study, and was thus out of present scope.

content of each post, embedded links contained in the post, and so on. Its creators describe the dataset as built from crawls of the "undocumented but open" Parler API (building on prior knowledge provided by the pseudonymous donk_enby: 2020) and released according to the FAIR principles of scientific data sharing.

As in all studies drawing on "found" data, the circumstances in which the data was originally collected and collated—and any limitations that result—are imposed on the secondary researcher. My design choices with respect to this component of the case study analysis were constrained as such. Most importantly, the researchers make clear that the dataset is neither inclusive of, nor statistically representative of, the entire universe of Parler posts made in the period (Aliapoulios, Bevensee, Blackburn, Bradlyn, et al. 2021). One cannot therefore make definitive statements about what Parler "was like" based solely on this dataset.

Moreover, the sheer size of the dataset imposed technical challenges. It is important to restate that the dataset is very large: totalling more than a terabyte when unzipped, and with its 183 million posts and comments across 167 files, it is by far the largest available dataset consisting of Parler posts made prior to the platform's temporary shutdown in January 2021. Therefore, after downloading the data from its public repository, there were several steps required to prepare it for analysis. First, I familiarised myself with the nature of the data, exploring the several variables (i.e., columns) that it included for each Parler post, and deciding which were most relevant for my purposes. Second, I decided how exactly to analyse it in view of RQ2a. With the caveat about statistical representativeness noted, I decided that as far as possible I wanted to "make sense" of what Parler was like during the period. I employ the phrase "make sense" here deliberately, to signal the qualitative inflection of the research, even when confronted with such a large dataset as this. In answering the question regarding Parler's platform self-governance of hate speech on Parler, I sought not just a simple statistical measure of "how hateful" Parler was and how this was dealt with, but also a more qualitatively informed sense of the nature of hate speech on Parler.

Therefore, this component of the analysis was best suited to a mixed-methods approach, discussed below.

Ahead of the analysis, I created a filtered subset of the dataset, in three steps. First, I restricted the sample to contain only *posts*, excluding *comments* (i.e., responses to posts) which entail more contextual understanding that would not necessarily be captured in the automated classification (a decision also taken in a similar study by Zeng and Schäfer 2021). This reduced the overall size of the dataset from 183,063,302 to 98,516,118 million entries. Second, I sampled only posts made between 1 Jan 2020 and 8 Jan 2021. This reduced the size further, to 91,317,161 million posts. Third, I removed blank posts, i.e., those in which the user had added no text.[89] This resulted in a "universe" of 21,764,396 (23.8%) entries that met each of these three criteria.

I then created a random sample of this filtered dataset of posts. For every month from January 2020 to January 2021 (inclusive), I created a random sample of 0.1% of the posts made in that month. This yielded a sample of 21,798 posts. After removing non-English language posts and API requests that returned an error (see below), I was left with a final filtered sample of 19,237 posts, made by 13,936 unique users. This was the basis for all analysis using this dataset. My choice of 0.1% as a sampling proportion was designed to strike a balance between ensuring a reasonably representative sample[90] and allowing me to read each sampled post. The multistage filtering and sampling process is summarised in Figure 5, and the number of posts analysed for each month is shown in Figure 6 below. This shows a sharp rise in the number of posts across the period, which is to be expected, as the amount of activity on Parler grew substantially over the course of 2020 (see 5.4.1).

[89] Most "blank" posts were "echoes"—Parler's equivalent of retweets.

[90] For clarity, I mean here representativeness in terms of the dataset itself, bearing in mind the limits on the representativeness of the dataset as a whole discussed above.

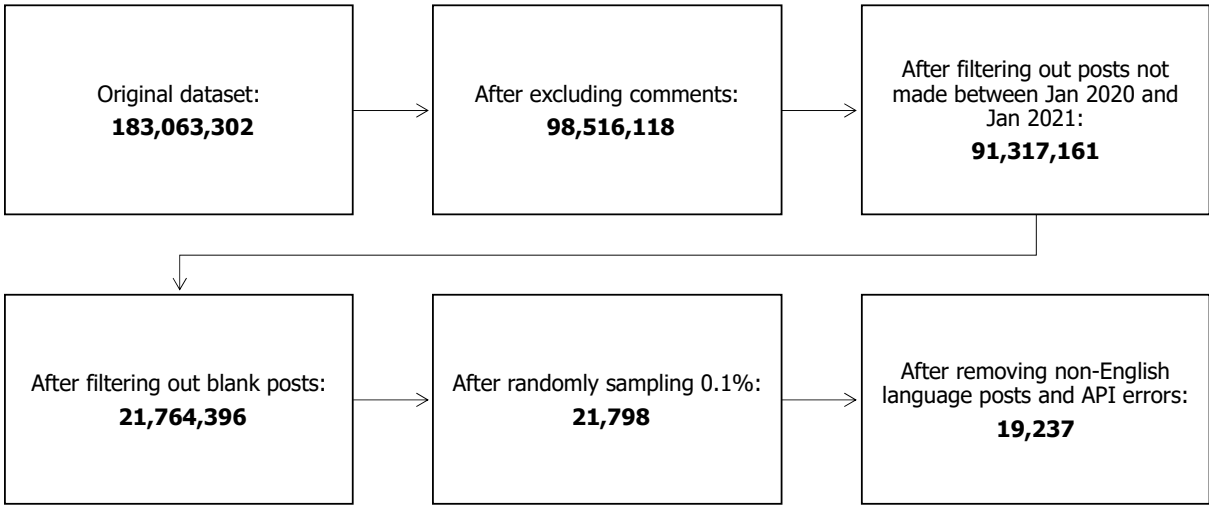


Figure 5: Path diagram showing how the dataset was filtered and sampled from.

Month	Sampled Posts	Sampled Users
Jan-20	219	168
Feb-20	186	153
Mar-20	187	155
Apr-20	175	148
May-20	368	317
Jun-20	1,361	1,255
Jul-20	2,211	1,970
Aug-20	1,615	1,363
Sep-20	1,544	1,308
Oct-20	1,749	1,477
Nov-20	5,737	5,265
Dec-20	3,135	2,762
Jan-21	750	725
Total	**19,237**	**13,936**

Figure 6: Total number of posts and users sampled for each month.

I decided to employ a mixed-methods approach which employed both automated and manual classification of content. In doing so I followed the approach taken in similar work to assess hate speech in fringe communities (Rieger et al. 2021), with the key difference that the manual component of the analysis was qualitative in this study (see below). Given both the complexity of hate speech discussed in Chapter 4, and the respective limits of both automated systems of hate speech classification and the biasing effects of my own lived experience on my perception of what "counts" as hate speech, this form of triangulation was employed to provide a more holistic view of the nature of online hate (Poole, Giraud, and de Quincey 2021). This reflects

Denzin's classic definition of between-method triangulation as a "complex process of playing each method off against the other so as to maximise validity" (quoted in Flick 2017:181).

For the automated classification I used two different application programming interfaces (APIs) platforms to triangulate further. These were the Rewire API and Google's Perspective API. Rewire was[91] a start-up founded by computer science and social science researchers which offered a state-of-the-art, machine-learning-based API to detect toxic content.[92] The API provided four classifications; I used "hate", which is described as "abuse targeted at a protected group or at its members for being a part of that group. Protected groups are based on characteristics such as gender, race or religion."[93] Perspective API is an established service which likewise uses machine learning techniques to classify various forms of toxic content in text corpora. I used its flagship toxicity attribute, which is described as encompassing "a rude, disrespectful, or unreasonable comment that is likely to make someone leave a discussion".[94] I chose this attribute, which is broader than Rewire's "hate" metric, in order to reduce the risk of false negatives in Rewire's narrower definition.[95] In both cases, I used a threshold score of 0.7, which is in the range recommended by Perspective API for social science researchers.[96] This means that for any given post to be classified as hate it needed to have a greater than 70% likelihood that it would be perceived as containing hate/toxicity.

The combined wall-clock time for sending and receiving the 21,798 queries to and from each API was approximately 60 hours (spread over multiple sessions), which gives some sense of the computational intensity that the classification of the sample involved. Querying the APIs sometimes resulted in errors, resulting from server time-outs and the APIs being unable to read

[91] Rewire has since been acquired by ActiveFence.
[92] Further information about the performance of the Rewire API is available at https://web.archive.org/web/20221005175805/https://rewire.online/introducing-the-rewire-api/.
[93] https://web.archive.org/web/20221005173618/https://rewire.online/api-documentation/
[94] https://developers.perspectiveapi.com/s/about-the-api-attributes-and-languages?language=en_US
[95] Although Perspective also offers a narrower "identity attack" attribute akin to what Rewire considers "hate", based on its description I was concerned that it would only include direct attacks, which may not be a useful metric given Parler's homogenous user base.
[96] https://developers.perspectiveapi.com/s/about-the-api-score?language=en_US

certain characters in Parler posts. The overall error rate was 8.3%, which meant that the total number of posts for which a score was successfully returned by the APIs was 19,237. The code that was used to query the APIs is included in Appendix 4 (see supplementary material).

I then manually analysed each post to seek to understand the extent and nature of hate on Parler. Compared with the automated binary classification, in this qualitative analysis I did not classify posts as being either "hateful" or "not hateful". Instead, I sought to understand the nature of the hate that was expressed in the sample of posts. This posed a trade-off in the design of this component of the study. I could have used a filtered dataset of only posts considered hateful in terms of the automated classification, and thus risk missing out on false negatives (i.e., posts that were hateful despite being classified to the contrary). Or I could have studied the entire dataset (N=19,992) at the cost of being unable to analyse each post in quite so much depth. I chose the latter strategy, because the purpose of undertaking mixed-methods research was to triangulate between automated and manual assessment of hate, which would have been undermined by using the automated classifications as a basis for the manual analysis. Therefore, this component of the analysis took the form of a *directed* qualitative content analysis (Hsieh and Shannon 2005). Instead of starting with a completely open approach to the corpus, I began with a structured coding scheme for the targets of hate, based on six common identity groups targeted by online hatred that were captured in topic modelling of a Gab corpus by Gallacher and Bright (2021) and iteratively expanded this to encompass three further identity groups targeted on Parler when I encountered it—namely transphobia, anti-Latinx hatred, and anti-Asian hatred.

When navigating the corpus, I considered not only the question of who the target of the hate was, but also what the implicit purposes for expressing hate in each case may have been. This was especially important in view of the findings of Benikova, Wojatzki, and Zesch (2018), who suggest that automated classifiers struggle to detect implicit hate speech. I analysed the entire sample chronologically, and although I did not use the automated classifications to filter or sort the data, the scores for each post were visible to me while I was performing the manual analysis.

Manually analysing the corpus surfaced the considerations already noted about my positionality as a researcher and the stakes of this topic, and form, of research. Such reflections extend beyond the present methodological discussion to encompass ethical and social considerations of the case study to which I now turn.

5.5 Ethical and social considerations

In addition to the broader ethical reflections associated with the thesis project as a whole (see Chapter 1), the empirical research that I undertake in the case study raises more specific research ethics considerations. This study received ethical clearance from the University of Oxford's institutional review board,[97] which permitted me to undertake the data collection and analysis described in the previous subsection. The various components of the walkthrough approach involved entirely publicly available data, contained either in published news articles and broadcasts and from publicly available archives of Parler's website and app. The datasets that I use for the automated and manual analyses of hate speech, described in 5.4.6, are more ethically complex. In contrast to the focus of the earlier analyses on elite figures like Parler CEO John Matze, this dataset concerns posts made by individual, "ordinary" users of Parler. This data is potentially identifiable insofar as it also contains the user handle (@username) of the user making each post. As with social media platforms like Twitter or Reddit, some users voluntarily use their own real name as their handle (e.g., @JoshCowls), though this was not a requirement for using the platform. The posts in the dataset are all public, because posts marked private, and direct messages, were either not available or not archived (Aliapoulios, Bevensee, Blackburn, Bradlyn, et al. 2021). Nonetheless, since Parler was a small, niche network it may not be reasonable to assume that users making public posts had the expectation they were, or would be, genuinely publicly available. As I make clear above, the purpose of the study is not to reidentify specific users. As an extra precaution, I do not include the text of any Parler posts in my presentation of the analysis. This decision also

[97] Ref: SSH_OII_CIA_22_111

reflects the caution advised by Matamoros-Fernández and Farkas (2021:214) not to unduly amplify or "overshare" hate speech content in research.

A related consideration that arose given that the focus on hate speech concerned the emotional and psychological effects on myself as a researcher. This has been one of the major motivations for automating hate speech detection, in order to remove humans from the "loop" of constant exposure to hateful content (Roberts 2019). Nonetheless, it is important to recognise the emotional toll that both the ex-ante training of automated systems and the ex-post monitoring of their efficacy continues to have on human labellers—many of whom are poorly paid and tend to receive little to no psychiatric support. By comparison, my relative level of risk is considerably reduced. As stated above, as part of this component of the analysis, I personally assessed posts that were assigned as hateful by the automated classifier. However, in absolute terms this sample was small, and I took several steps to protect my own psychological wellbeing for this task. This included not following out-links to content, to ensure that I was only exposed to text and not visual content (the database itself is purely textual), and only analysing small subsets of posts in any one session. It is also important to acknowledge my positionality as a white male researcher in conducting this research, especially in view of the differential effects of hate speech on different identity groups that I discussed in Chapter 3. Due to Parler's demographic and ideological skew, it is unlikely that I would have fallen into (m)any of the identity groups likely to be under verbal assault in the dataset. This, in turn, surfaces one of the tensions associated with combatting online hate that I identified in Chapter 4. There is a difficult balance to be struck between, on the one hand, protecting those identity groups that are most targeted and harmed by online hate speech, and, on the other, ensuring that members of these same groups are not excluded from the design of equitable and responsive systems for managing hate speech. As such, I intend for the findings of this analysis to contribute to far broader, inclusive, and representative efforts to rid the internet of harmful hate speech.

Chapter 6: Parler games—the rhetoric and reality of Parler's self-governance

Introduction

In this chapter I embark on the first component of my case study analysis, in which I assess how Parler performed its platform self-governance. This involves a technical walkthrough of Parler, and discourse analyses of media coverage about Parler, Parler's public presentation on its website, and its archived policy documents. The focus of this analysis is how Parler sought to discursively legitimate its platform self-governance in these public-facing settings. The analysis spans the year 2020, during which Parler rose from a fringe platform to top app store download charts. The performative nature of Parler's platform self-governance, expressed in rhetorical and visual communication choices and its policy documents, can be contrasted with the prevalence of hate speech on the platform and how it was moderated.

The findings of the technical walkthrough of Parler's functionality and features are reported first. Parler resembled a "Twitter clone" to some extent, albeit with important differences in some of its mediator characteristics, including its recommendations of who to follow, colour scheme, and voting system (6.1). Next, the discourse analysis of mainstream media coverage of Parler is presented. The company's CEO John Matze invoked an array of constitutional metaphors to lure more users to the platform and to legitimate its platform self-governance (6.2). This is supplemented by analysis of archived copies of Parler's public-facing website during the same period. Parler's website foregrounded visual tropes relating to debate and disagreement as well as "heartland" iconography, apparently pitching the platform at "patriotic" conservative users (6.3). Finally, the dataset of governance guidelines demonstrates the platform's struggles inchoate approach to hate speech (6.4).

Overall, the analysis finds that Parler's approach to platform self-governance was discursively legitimated through the strategic deployment of constitutional metaphors and "heartland" imagery. These techniques sought to situate Parler's minimalist platform self-

governance in the traditions of American freedom and democracy, contrasting the platform with the coastal elites of Silicon Valley. This confected view which casts "the people" against "the elites" is a cornerstone of the populist radical right (Mudde 2016:4). Key mediator characteristics of Parler's smartphone app reinforced the platform's right-wing skew. Thus, I argue, Parler sought to portray itself as a "populist platform", both to appeal to disaffected right-wing users of other platforms, and to legitimate its inchoate approach to platform self-governance.

6.1 Technical walkthrough of Parler's functionality and features

Like any social media start-up, Parler's functionality and features evolved over the course of its growth. Here, I draw on contemporaneous news reports and screengrabs of Parler's onboarding process taken shortly before its removal from app stores, to characterise the affordances of the platform in terms of the technical walkthrough method (see Light, Burgess, and Duguay 2018 and Chapter 5). Unless otherwise noted, these descriptions relate to Parler's functions and features as of early 2021, shortly prior to its removal from app stores.

Parler's entry in the iOS App store is the first time that many of its mediator characteristics are presented to the user, albeit these take the form of static screenshots from the app itself (see Figure 7). In terms of textual content and tone, the strapline "Unbiased Social Media" reinforces the app's claimed advantages over allegedly censorious rivals (see 6.2), while the imperative to "Read News, Speak Free" reflects what Light et al (2018:892) describe as the "discursive power" of embedded text to "shape use".[98] In terms of symbolic representation, this is also the first time (outside of app store search results) that the user is presented with Parler's default red colour palette, as depicted in its icon, as well as in the background to the featured screenshots. Light et al. argue that visual components such as colour can reveal an app's "likely connotations and cultural associations"; the dark red colour dominant here and recurred throughout the app reveals

[98] Ironically, the "use" that the app's discourse seeks to shape here is itself a particular kind of discourse in the form of "free speech".

and reinforces Parler's association with the traditional colour of the US Republican Party and thus conservative politics.

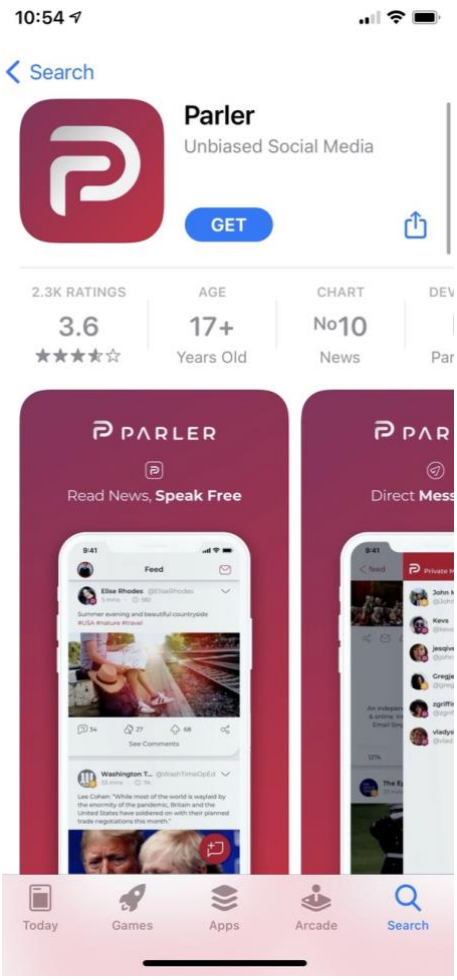


Figure 7: Parler's entry in the iOS App Store in January 2021. Source: author's screenshot.

Once the app had been downloaded and installed, joining Parler through the app's onboarding process was free. However, this required providing a valid email address and phone number to sign up (in contrast to Twitter, which requires only one or the other), through compulsory fields. In an optional pop-up window, the app reassured new users that it "NEVER shares your information", which was required solely for authentication and security (see Figure 8). In the next registration step, the user was required to agree with Parler's privacy policy, its terms of service, and its community guidelines. I postpone substantive discussion of these documents

until 6.4, but it is worth noting that it was possible to view them in full, as a pop-up dialogue, within the app prior to agreeing to them (see Figure 9).[99]

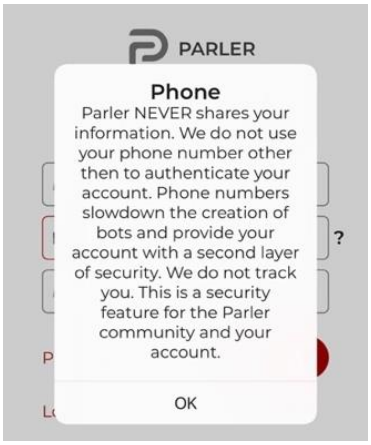


Figure 8: Pop-up dialogue showing Parler's justification for asking for a new user's phone number. Source: author's screenshot (image cropped).

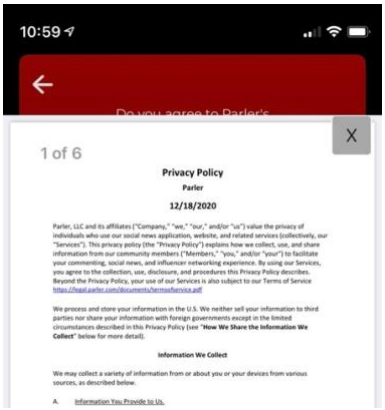


Figure 9: Parler's privacy policy displaced as a pop-up during registration. Source: author's screenshot (image cropped).

Once sign-up was complete, a further option was to "verify" one's account and get a "Verified" badge (see below), which could be accomplished by uploading a photo and a scan of government-issued identification.

After signing up for Parler, the onboarding process invited a new user to "Personalize your Parler Experience" by suggesting other accounts for new users to follow. The button at the top-right of the screen which allowed users to "skip this step" was relatively small, revealing the intention to guide users to select users to follow from the start of their use of the app. This is

[99] Several other quirks associated with the registration process, such as storing potential passwords in memory and showing them as tooltips, led one security-focused developer to contemporaneously describe the registration process on Parler's smartphone app as "extremely suspicious", concluding that "Parler isn't safe" (Taylor 2020).

significant, because as Figure 10 and Table 5 below illustrate, the suggestions that were presented to me in January 2021 consisted almost exclusively of right-wing figures, such as Fox News presenters Sean Hannity, Maria Bartiromo, and Tucker Carlson, and Republican elected officials Devin Nunes and Kristi Noem.

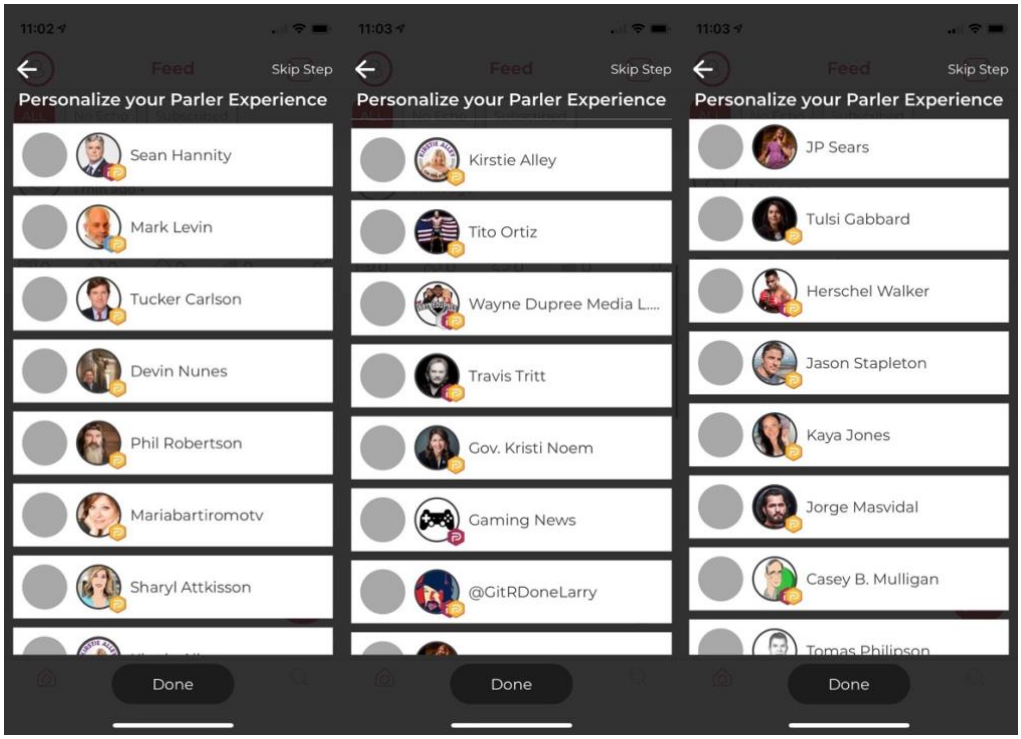


Figure 10: Suggested accounts to follow to "personalize your Parler experience". Source: author's screenshots.

The clear right-wing skew of the recommended accounts makes this "personalisation" feature a notable mediator characteristic. It is in part a reflection of Parler's ideological skew—but also serves as evidence that the platform sought to maintain this dynamic by essentially giving free promotion to specific right-wing accounts. This exemplifies what Light et al. (2018:293) describe as a "preferred user practice" that is often presented in the registration process, which often "go[es] hand-in-hand with subtle suggestions about the ideal user". It is apparent from this analysis that the ideal user is one expected to be sympathetic or at least receptive to right-wing causes and personalities.

Table 5: Industry and political stance of recommended accounts to follow when joining Parler

Name	**Industry**	**Political stance**
Sean Hannity	Broadcaster	Right-wing
Mark Levin	Broadcaster	Right-wing
Tucker Carlson	Broadcaster	Right-wing
Devin Nunes	Elected official	Republican
Phil Robertson	Reality TV star	Right-wing
Maria Bartiromo	Broadcaster	Right-wing
Sharyll Attkisson	Journalist	Anti-vaccine
Kirsty Alley	Actor	Trump supporter
Tito Ortiz	Martial artist and elected official	Republican
Wayne Dupree	Online broadcaster	Right-wing
Travis Tritt	Musician	Trump supporter
Kristi Noem	Elected official	Republican
Gaming News	n/a	n/a
Larry the Cable Guy	Comedian	Libertarian
JP Sears	Comedian	Right-wing
Tulsi Gabbard	Elected official	Democrat
Herschel Walker	Sportsman	Trump supporter
Jason Stapleton	Influencer	unknown
Kaya Jones	Musician	Right-wing
Jorge Masvidal	Martial artist	Trump supporter
Casey B Mulligan	Economist	Trump appointee
Tomas Philipson	Economist	Trump appointee

Once the user had installed the app and registered with the platform, they had five main options. Four icons—depicting a house, a compass, a bell, and a magnifying glass—were arrayed on the bottom menu bar, and their meaning (respectively a home feed, a discover feed, a notifications tab, and a search function) would have been immediately apparent to the users of other social media apps, especially Twitter. The fifth was represented as a speech bubble with a plus sign, which was positioned in the bottom right of the screen—precisely the same position as

the "quill" button to create a tweet is situated on the Twitter app. Therefore, at the level of symbolic representation, Parler's icons thus drew on existing cultural representations that (at least for those already versed in social media platforms) were already bound up in the meaning (in an online context) of symbols such as a magnifying glass. Moreover, at the level of function, the way these features as mediator characteristics were arrayed and presented on the app's primary menu also closely mimicked existing social media platforms. Thus, although Parler's CEO would take every opportunity to criticise rivals like Twitter during 2020, in terms of the functions and features that guided the everyday use of the platform, Parler's app interface brings to mind the sense that imitation is the best form of flattery. Such similarities were contemporaneously noted by several outlets, which characterised Parler as most closely resembling Twitter (Brown 2020; Courty 2020; Sullivan 2020). This emulation is unsurprising, given that Twitter has historically been seen as the social media platform most akin to a "town square" (Haggart et al. 2022; Pariser 2022)—the moniker that Parler sought to adopt—and has traditionally hosted a high proportion of influential and engaged "elite" users from politics, academia and the media (Blank 2017; Dagoula 2019; Dubois and Gaffney 2014). As discussed, it was from Twitter that most high-profile conservative users claimed to be "fleeing" to Parler, owing to that platform's increasing constraints on certain content and users over the course of 2020.

Similarly, the functions and features contained on Parler's profile screens and as enabled by the speech bubble icon would have been familiar to users of other social networks and, again, these most closely resembled Twitter. A user's profile screen showed their screen name and handle, a profile photo and cover photo, the user's join date, and a character-limited biography, as well as an index of their "parleys" and follower information. Meanwhile, Parleys could contain up to 1,000 characters (an increase on Twitter's character limit of 280), and could contain hashtags with the # symbol, tags of other users with the @ symbol, emojis, gifs, and uploaded photos. They could be "echoed" or shared by another user (equivalent to retweeting), commented on, voted on, and shared outside the site. Parleys could only be upvoted, akin to tweets, but *comments* on parleys could

be both upvoted and downvoted, mimicking Reddit's upvote/downvote functionality and reflecting the fact that comments on Reddit (but not posts) can be downvoted below zero. Given the widespread functional similarities between Twitter and Parler it is notable that the latter departs from the former in this respect, allowing the expression of disagreement or disapproval with respect to comments. There was also support for direct messaging between users, as well as the ability for users to mute, block, and follow others. Parler offered limited support for searching, as searches were limited only to specific #hashtags and @users, rather than the content of parleys; comments were also not searchable.

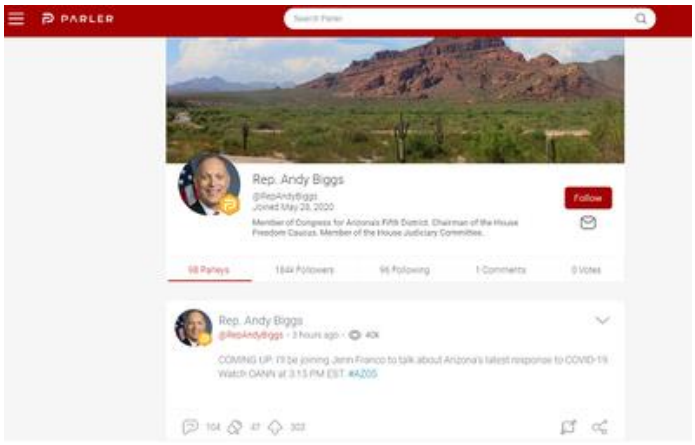


Figure 11: Parler profile page for Representative Andy Biggs (R-AZ) as of June 30, 2020. Source: https://en.wikipedia.org/wiki/Parler#/media/File:Andy_Biggs'_Parler_feed.png

As Figure 11 above also shows, Parler assigned an array of "badges" to certain users. The badge system again presumably took inspiration from Twitter's "blue tick" system, which between 2009 and 2023 was used to verify that users had met certain criteria for authenticity, notability, and activity. However, the volume and diversity of Parler's badges represented an advance on Twitter; Aliapoulios et al. (2021) identified ten badges that users could be assigned,[100] including a red "Verified" badge, a "Gold" badge meant to indicate users at risk of impersonation (most closely akin to Twitter's blue tick), as well as more novel badges, such as those for parody accounts and early adopters. The widespread use of badges created a sophisticated hierarchy of users, with the

[100] By far the most common badge to have been assigned by the start of 2021 was the "Private" badge, which marked over 922,000 users as having a private account.

gold badge in particular used to indicate esteem—as revealed by the fact that almost all of the accounts recommended to follow had gold badges (see Figure 10).

This technical walkthrough of the app has suggested that in many ways Parler was indeed a "Twitter clone just for conservatives", as its new owners bluntly described it after its acquisition (Binder 2023). In particular, many of the app's functions and features would have been very familiar to Twitter users, betraying the clear technical influences of its larger rival. However, the app also consisted of distinctive mediator characteristics that reveal how the company imagined, and encouraged, the platform to be used by conservatives specifically. Its red colour palette marked a change from the soft blues of rivals Twitter and Facebook and immediately associated it with conservative American politics. This was reinforced textually in its app store entry by its commitment to "unbiased news" and its exhortation to users to "speak free"—tropes that would also recur in the company's public communications (see 6.2). The platform's ideological skew was further reflected, and reified, by the "personalisation" feature that greeted all new users, which this analysis reveals was (at least in my encounter with the app) very clearly right-leaning. And the app's close functional similarity to Twitter, underscored by the familiarity of its menu interface and icons, suggests that the platform was intended for established social media users for whom the underlying meaning of such symbols and utility of features were already well-established, i.e., for the "extremely online".

6.2 The public performance of Parler's self-governance

As we saw in Chapter 5, Parler experienced a dramatic rise to prominence over the course of 2020. This occurred against the backdrop of other platform operators "deplatforming" right-wing American figures, and steadily increasing their restrictions on the social media accounts of President Donald Trump. This context is important for understanding the purpose of John Matze's media appearances during the year. Both the volume of his media appearances and the venues in which he most often appeared suggest an attempt to capitalise on conservative

disgruntlement with mainstream platforms to pitch his platform as an alternative. And Parler's claimed commitment to "constitutional free speech"—which, as we have seen, was established as the site's U.S.P. at its launch in 2018—provided Matze with a powerful metaphorical basis on which to tout the platform's distinctive approach to self-governance (Cowls and Ma 2023).

This context sets Matze's appearances apart from discourse analyses of public statements by the leaders of larger and more established social media platforms. Prior research focused on Facebook's Mark Zuckerberg reveals his intention to justify and defend Facebook's existing dominance, and to project it into the future as a "prophetic corporation" (Haupt 2021; see also Hoffmann, Proferes, and Zimmer 2018; Rider and Murakami Wood 2019). By contrast, even at its zenith, Parler's base of 15 million users remained significantly smaller than those of not only the largest platforms like Facebook, but also the somewhat smaller rivals with whom Parler was more directly competing, above all Twitter (Hagey and Horwitz 2021). As such, it is reasonable to presume that Matze's core motivation throughout 2020 was to attract as many new users to the platform as possible, in part through the promotion and legitimation of Parler's supposedly more freedom-focused approach to platform governance.

In this section I analyse the key legitimating themes that Matze emphasised in his mainstream media appearances discussing Parler, focusing specifically on his deployment of constitutional metaphors. Figure 12 summarises the metaphors that I found Matze to have evoked across the corpus. The number given in the y-axis indicates th*e number of articles in the dataset* that *contain at least one direct quote from Matze* that were coded as *relating to the respective legitimating theme*.

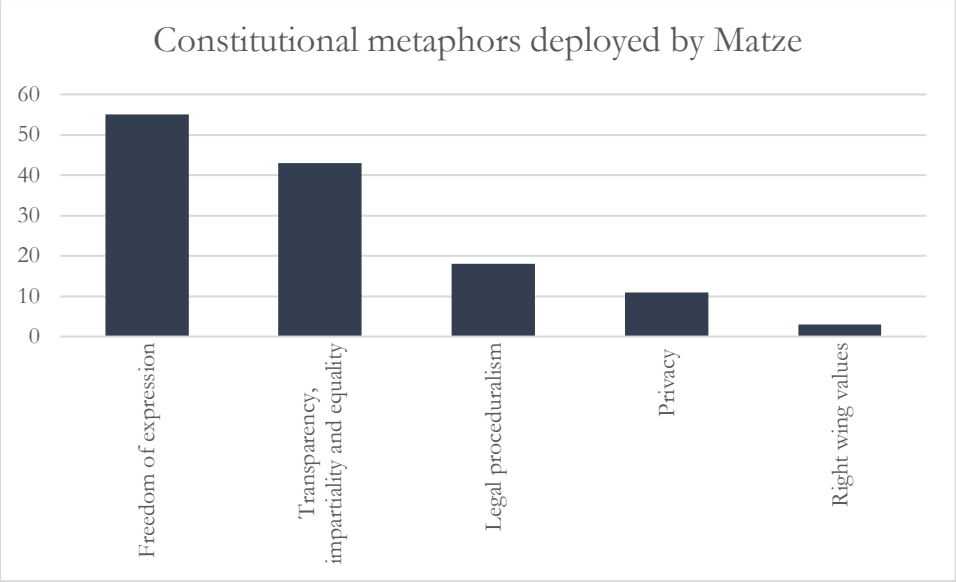


Figure 12: Number of articles in the dataset containing at least one direct quote from John Matze pertaining to different constitutional metaphors.

Because of the broad range of news outlets in the sample and the varied reception—from the credulous to the critical—that Matze received in each instance, this summary tells us about the themes that Matze sought to convey, rather than necessarily how they were received and subsequently framed in media coverage. To explore *how* these constitutional metaphors were used to legitimate Parler, I used van Leeuwen's (2007) framework for the discursive construction of legitimation in public communication Van Leeuwen (2007:91) identifies four categories of legitimation:

1) "authorization, or legitimation by reference to the authority of tradition, custom and law, and of persons in whom institutional authority is vested;
2) moral evaluation, or legitimation by reference to discourses of value;
3) rationalization, or legitimation by reference to the goals and uses of institutionalized social action, and to the social knowledges that endow them with cognitive validity; and
4) mythopoesis, or legitimation conveyed through narratives whose outcomes reward legitimate actions and punish non-legitimate actions."

In my analysis, I found that the deployment of each constitutional metaphor tended to incorporate several of these forms of legitimation, a possibility that van Leeuwen admits (2007:92).

Van Leeuwen characterises legitimation as "an answer to the spoken or unspoken 'why' question—'why should we do this'?" (Van Leeuwen 2007:94). The authoritising form of legitimation is founded on the response "because I say so", or because someone or something

else, in whom authority is invested, says so (2007:94). This can include impersonal authority, i.e., "because the laws (the rules, the policies, the guidelines, etc) say so" (2007:96). In the case of Matze's constitutional metaphors, the laws in question were most often the US Constitution, but also on occasion other government rules such as the FCC guidelines.

Moral evaluation as a form of legitimation functions differently. Moral discourses are often only hinted at through related terms like "health" and "hygiene", but this is not always the case: van Leeuwen gives the counterexample of President George W. Bush's explicit proclamation of America's enemies as an "axis of evil". Moral evaluation can thus take the form of analogies; as van Leeuwen notes, the use of "comparisons in discourse almost always [has] a legitimatory or de-legitimatory quality" (2007:99). In such cases, the implicit answer to the question "why should we do this" is "because it is not like another activity which is associated with negative values". In this vein, Matze frequently drew morally negative analogies to the platform self-governance of larger rivals—at various points characterizing them as "technoauthoritarians", "technofascists", and "socialist dictatorships"—in order to legitimate Parler's supposedly more patriotic approach to platform self-governance by way of contrast.

Rationalisation formed the smallest component of Matze's legitimating efforts. As van Leeuwen notes, "moralization and rationalization keep each other at arm's length" (2007:101) and thus, given Matze's heavy dependence on moralisation, we should expect to find less in the way of rationalisation. On occasions where Matze did resort to rationalisation, it most often took the form of predictive legitimation, which is based not on authority but on expertise. In this vein, Matze occasionally predicted the arrival of floods of new users to Parler because of the platform's superior self-governance.

Finally, legitimation can be achieved through storytelling, or more precisely, what van Leeuwen calls mythopoesis. This was an essential part of Matze's utilisation of constitutional metaphors in his discursive legitimation of Parler's platform self-governance. This is unsurprising since the US Constitution, on which Matze's metaphors often drew, arose from a specific, well-

known historical context, the American War of Independence. This provides a ready-made "cautionary tale", which Matze frequently invoked, casting larger rivals as tyrants in the mould of the colonial British authorities, with Parler positioned as patriot revolutionaries. As van Leeuwen states, the protagonists of cautionary tales "engage in deviant activities that lead to unhappy endings", and this storytelling structure undergirded many of Matze's rhetoric.

In what remains of this analysis I explore how Matze deployed specific constitutional metaphors, in each case drawing on several of the forms of legitimation just discussed. As Figure 12 shows, the constitutional metaphor that Matze most often deployed was *freedom of expression*. In such cases, Matze drew on the proud American tradition of free speech contained in the First Amendment to the US Constitution, as well as making moral claims about the utility of "healthy debate" in the "town square" to counter harmful content. To a somewhat lesser degree, Matze also laid claim to *transparency and impartiality*, portraying Parler as an impartial arbiter of content, often appropriating language from the preamble to the US Declaration of Independence, which holds that "all men are created equal". A somewhat less common cluster of themes concerned *legal proceduralism*, which appealed most often to the concept of jury trial enshrined in the Sixth Amendment to the United States Constitution. Matze also, on occasion, cited *privacy of personal data* as something that his platform allegedly safeguarded. This drew (albeit much less explicitly) on constitutional provisions for personal privacy that are implied by the Fourth Amendment's assertion of "the right of people to be secure in their persons, houses, papers and effects" and its prohibition on "unreasonable searches and seizures", and have been extended analogously to the online world.[101] Only very occasionally did Matze make appeals to constitutional provisions with exclusive appeal to right-wing audiences, such as the "right to bear arms" as contained in the Second Amendment to the US Constitution. Throughout the corpus, Matze contrasted uplifting depictions of life on Parler with the perceived failings, weaknesses, and inequities of larger rivals

[101] For an overview of cases involving Fourth Amendment privacy provisions in online contexts, see e,g, https://epic.org/issues/privacy-laws/fourth-amendment/

like Facebook and Twitter, often using historical analogies to portray Parler as a revolutionary, patriotic alternative to the supposedly "tyrannical" governance of mainstream social media platforms.

In what follows, I explore Matze's deployment of each of these constitutional metaphors in media appearances, assessing how each was used to promote Parler, to win new users, to criticise rival platforms, and ultimately to discursively legitimate its approach to platform self-governance.

6.2.1 Freedom of expression

Freedom of expression was the constitutional metaphor that Matze most frequently deployed across the dataset. In a typical example of how Parler's "free speech" credentials were framed by conservative media, in March 2020 article, the American "alt-right" news site Breitbart covered a speech given by Matze at the high-profile Conservative Political Action Conference. It quotes Matze as claiming that "we created Parler with free speech … in mind", drawing a contrast with rivals who "don't respect that". Both Matze's speech and the article itself actively encouraged prospective users to join the platform. Matze is quoted as urging the audience to "go to your app store, type in Parler, it's the red app, you can download and join, you can follow me there @john", and the article ending with a link to Parler's website.

Given Breitbart's status as a key node in the online right-wing US ecosystem, its uncritically favourable coverage of Matze, and especially his totemic "free speech" claim, is unsurprising. Other, more mainstream outlets approached Matze's free speech claims somewhat more critically. Matze was quizzed on the meaning and implications of Parler's "pro-free speech" stance in an interview with Forbes published in June. The article begins with a vignette: Matze apologizes for eating during the interview— "I hope you don't mind … I haven't eaten all day"—then explains why: he is "sitting here like, banning trolls." The "trolls" Matze is banning are what he calls "teenage leftists" who had descended on Parler in recent days. The article takes at face value Matze's claim that these users are troublemakers, who are, it suggests, "swamping them with

messages that make it unpleasant for the app's conservative users to post and interact with each other".

The apparent contradiction here—a platform described as being "dedicated … to the promulgation of free speech", whose CEO personally bans users he deems "trolls" for "swamping" the network's "conservative users with messages"—is not fully explicated in the article. However, elsewhere in the interview, Matze is pushed further, by being "asked if there ever might be an instance when the N-word would be appropriate" on Parler, and has this answer:

> "It depends on the context. If they just said that word alone, I don't think we would touch it." He thinks a couple minutes longer, then restates his opinion. 'If somebody came on there and said the N-word to somebody, and they got very upset as a result of that, then it would get taken down."

Here, in initially trying to remain consistent to Parler's free speech absolutism, Matze invoked a hypothetical scenario in which the N-word was simply stated on its own, devoid of context, intention, or malice. After further reflection, however, he seemed to reckon with the reality that language is deployed within specific social contexts and between people occupying different social positions, and rather often with harmful effects. As such, Matze concedes that whether and how someone is affected by a hateful slur *could* affect how a message would be treated by the site. This episode is, however, a rare exception to Matze's more general tendency to propound the virtue of untrammelled free speech on the platform without being forced to reckon with its implications.

Many of the articles in the sample that referenced Parler's "free speech" credentials also featured Matze attacking Parler's larger rivals. "Censors" and "censorship" are words that routinely recurred when Matze assailed the moderation activities of rival platforms like Facebook and Twitter. A sympathetic Fox News interview between Laura Ingraham and Matze in October 2020 hailed Parler as a contrast to what the Fox News on-screen chyron described as "social media's war on free speech". The pull quote used for the published version of the interview reinforces this contrast, quoting Matze as claiming that "instead of censoring people more, we're advocating for people to give information". The interview includes Matze's allegation that "Twitter is getting involved and is actively editorializing this topic … to manipulate what the public thinks".

In reality, the examples of "shameful censorship" that Matze assailed were mild, ad hoc attempts by platform operators to dampen the spread of harmful misinformation relating to topics like the Covid-19 pandemic and the presidential election.

However, while Matze emphasised that it was "conservatives" who "have been repeatedly censored and disenfranchised on Twitter and Facebook", he nonetheless maintained that the platform was intended for users of all political stripes, frequently using the metaphor of Parler as a "town square". In an interview in June 2020, Fox News host Neil Cavuto pushed Matze on who the platform was for, asking whether "you are more of a conservative site, or are liberals welcome, everybody's welcome, how would you describe it?". Matze responded, "Well we're a town square, that's how I view us. So, everyone's welcome." However, as Matze went on, "we're seeing a lot of popularity with conservatives. They seem to be the ones most affected by Twitter's censorship or Facebook's censorship." This image of a town square is one of the most prominent mythopoeic narratives that Matze evoked to portray the culture he claimed that Parler represented.[102] However, the latent tension between Parler as a refuge for "disenfranchised" conservatives on one hand and Parler as an all-inclusive "town square" on the other goes unresolved.

In several interviews, Matze referenced the First Amendment to the US Constitution, which constrains the government from infringing citizens' speech rights, as having shaped the rules that govern Parler as a private platform. In a Fox News interview in November 2020, above the chyron "the conservative alternative to big tech", host Tucker Carlson begins by offering Matze "congratulations [on] the massive surge of popularity in your company". Carlson asks Matze how he would respond to facing "massive pressure" to "censor" voices on the platform. Matze responds:

[102] According to Matze, Parler was, variously, a "neutral town square", a "community town square", and a "free and open town square"; the strapline on one version of Parler's homepage likewise described the platform as "the world's town square".

> "it's not against the law to have [unpopular] opinions. It's not against the law to express yourself, you know. And if you like one political candidate or another, or you believe or don't believe in climate change, or whatever it might be, you shouldn't be taken offline because of it."

Carlson is fawning in his response. "Exactly right … exactly", he has already started to reply before Matze finishes speaking. He goes on to tell (a beaming) Matze: "Thank you. You just restated the traditional *American* understanding of free speech which the left defended for decades—apparently they didn't mean it—but it's left to you to continue that noble defence". Placing emphasis on the word "American", Carlson drew on the authoritisation of the American political tradition to endorse Parler's and Matze's "distinctive" approach to platform governance.

This grounding in American "freedom" might have held superficial appeal for prospective users aggrieved by opaque, arbitrary content moderation and community standards enforcement elsewhere online. However, as we have seen, in practice Matze was forced to announce ever more caveats to this supposedly straightforward and transparent policy over the course of 2020. But even when he was defending fresh restrictions on speech, *rhetorically* Matze held firm to free speech as a cause célèbre, justifying every policy change as furthering his site's mission to safeguard rather than squash free expression. For example, when justifying the platform's new restrictions on spam in June 2020, Matze describes spam as "malicious" and as "trying to take away [users'] voice". A *Washington Post* article the following month noted that while Parler "ha[d] become the poster child for free expression online", Matze had recently announced some rather peculiar rules on the platform, among them: "when you disagree with someone, posting pictures of your faecal matter in the comment section will not be tolerated", and "you cannot threaten to kill someone in the comment section." Yet in response to a question about the new rules, Matze again resorted to the mythopoeic vision of a town square and the constitutional metaphor of free speech, quoted in the story as saying that the purpose of these rules was to help "create a proper town square without people ruining it by violating it with speech not protected by the First Amendment or FCC guidelines". "Ruining it" is obviously a subjective view dressed up as a moral evaluation, but the fact that Matze was able to iteratively adjust his metaphorical portrayal of Parler as a free and open

town square for free speech, even to defend speech restrictions, reflects the discursive flexibility of the legitimating strategies that he employed.

6.2.2 Transparency, impartiality, and equality

Another constitutional metaphor that Matze regularly deployed was transparency, impartiality, and equality. Unlike freedom of expression, these concepts do not have concrete analogues in the US Constitution, but Matze nonetheless evoked the mythopoeic history of American independence when discussing them. Matze made much of Parler's feed, which displayed content from everyone a user followed in a simple reverse-chronological timeline—mimicking features that Twitter and Facebook had long since abandoned in favour of algorithmically-powered timelines prioritizing "most relevant" content.[103] Matze drew on this contrast as part of a broader critique of these rivals, who he suggested were "manipulating people" with their feeds. An appearance on Fox & Friends in November 2020 began with a softball question, with the host claiming that "the president himself has been censored 52 times [on other platforms] since election day—how do you do things differently over there on Parler?" As a scrolling walkthrough of Parler's homepage was visually overlaid behind him, Matze restated Parler's claimed commitment to free speech, but went further:

> "Parler is of course about free speech, making sure people have a voice again … but we're so much more than just free speech too, because people also want freedom from the data manipulators, they want change in social media in general. … People want a social media like Parler that works the way they thought, where it truly is by the people, it's private, you choose the content you wanna see, and you get the content that you've chosen to see. And so there's so much more than just free speech."

With his reference to "freedom from data manipulators", Matze framed one of Parler's design choices—its reverse chronological feed—as a matter of fundamental liberty.

When Parler briefly overtook Twitter on the top chart for News apps on Apple's App Store, Matze posted a valedictory message on Parler that was picked up by conservative site RedState. It read:

[103] Twitter displays its algorithmic "For you" feed by default, but its reverse-chronological feed is still (at the time of writing) accessible.

> "This is our victory as a community for free speech over the tech tyrants. This is a symbol that we want a Town Square for free discussion. We the people do not want to be told what to think, we do not want to be manipulated, and we want our data to be private. We reject technofascism and those who think they are the sole arbiters of truth. We reject their biased editorial panels, we reject their "fact checkers" and we reject censorship."

Drawing on the cornerstone of American liberty, here Matze hails "we the people"—aping the US Declaration of Independence—for rejecting "technofascism", an umbrella term which includes "biased editorial panels" as well as "fact checkers", a reference to the content moderation and anti-disinformation efforts of other platforms. In contrast, as Matze explained in an email interview with Fox News's "Media Angle", Parler was a haven for free thought and information, leaving fact-checking "to the responsibility of the community". Users should not "be worried about receiving inaccurate information on Parler", Matze went on, because "our users are wise enough to filter and judge information themselves". This places the responsibility of deciphering fact from fiction squarely on individual users, thus conveniently letting Parler off the hook for its role in facilitating the spread of harmful disinformation.

Fortunately for Matze, the events of 2020 provided ample scope for critique of mainstream social media platforms. The release in September of the documentary *The Social Dilemma* offered insight into how the design of social media platforms and algorithmic systems affect social and political life. In an interview with Fox News shortly afterwards, Matze seized on this more widespread critique of his rivals to take a retrospective victory lap:

> "As I was watching the movie, I kept thinking we encountered this decision [at Parler] and could have done that but we decided not to … We've really done a good job avoiding a lot of these problems."

Restrictions on right-wing American users, up to and including President Trump, by Parler's rivals also allowed Matze to go on the front foot, assailing the larger companies for what it claimed was excessive moderation, while sidestepping the issues that plagued Parler's own "laissez-faire" approach. In an October 2020 blog post on Parler's website that was picked up by outlets including Newsweek, Matze claimed that, by restricting the spread of a *New York Post* article about Hunter Biden's laptop, "Facebook and Twitter have proven that they are agenda-driven publishers who

want to control information and trample on our right to read what we want and decide for ourselves what to think". By casting Facebook and Twitter as "publishers," rather than "platforms," Matze aligned these companies with establishment media outlets, which have long faced claims by US conservatives that they are ideologically biased towards liberal elites (Hemmer 2016; Lane 2019). The Newsweek piece quotes Matze's blog post at length, focusing largely on Facebook's and Twitter's justifications for their restrictions rather than on Parler's own content moderation failings.

With these discursive efforts, Matze tapped into both longstanding perceptions among American conservatives of tech platforms' bias against them, and opacity regarding how content on social media platforms is ranked, filtered, and presented to users by algorithmic systems. As a report on the alleged bias against conservative social media users observes, "there are no empirical studies that definitively assess the claim of anti-conservative bias", but nonetheless, what evidence is available suggests that US conservative voices enjoyed considerable attention on mainstream social media platforms like Facebook during 2020 (Barrett and Sims 2021:14). Whatever their merit or lack thereof, claims of anti-conservative bias were rife throughout the period, and Matze frequently drew upon them in making the case for Parler. Meanwhile, when it came to the opacity of the algorithmic systems employed by sites like Facebook, YouTube, and Twitter to rank and filter content, Matze had chanced upon an issue that had already elicited valid grievances from many sides, not just the American right.

Matze also frequently painted Parler as a cross-ideological or even post-partisan arena for political discussion—even though in practice, its membership leaned heavily right-wing. Tacitly acknowledging this reality, Matze claimed in a Fox News interview in June 2020 that "you're going to see a lot of people on the other side of the aisle coming over very soon … in fact, we're seeing them in waves now". Again, self-aggrandisement regarding Parler's cross-ideological membership was never far removed from criticism of Parler's larger rivals and predictions of their coming demise. In an interview with CNBC, Matze is portrayed as non-partisan: "Matze doesn't want

[Parler] to be just an echo chamber for conservative voices. Personally, he says he doesn't like either political party and he wants to see more healthy debate." The article goes on:

> "Matze expects Parler to become a more attractive site for a more diverse audience over time because he sees Twitter continuing down a path of alienating right-wing voices, and 'no one is going to want to stay on Twitter if the conservatives are gone.'"

But for all his attempts to present Parler as an "impartial" platform, in reality, the vast majority of new Parler users were disgruntled US conservative émigrés from larger platforms (Brustein 2020), whose predominance on Parler created not some post-partisan paradise but instead an "echo chamber" for the American right—or perhaps a platform-wide "echo mansion". This led Parler to resemble what Titley (2020:108) has characterised as a "discourse laboratory": a space where participants critique media coverage and generate counter-narratives to oppose "the presumed hegemony of progressive, 'liberal' positions in public culture." Matze may have trumpeted the impartiality and transparency of Parler's platform self-governance and the equal treatment of its users in constitutional terms—but in practice, these "features" appealed almost exclusively to US conservative users who subscribed to a belief in Silicon Valley's bias against right-wing voices.[104]

Sometimes the constitutional allusions were more barefaced. "A Declaration of Internet Independence", published on Parler's website in June 2020,[105] and picked up by several articles in the corpus, consciously mimics its 1776 inspiration, beginning:

> "When in the Course of technological development, it becomes necessary for free people to reject the Terms of Service which have connected them with a platform, and to take back their Constitutional and Human rights, our great American traditions call on us to declare the causes which impel the separation."

The "Declaration", it becomes clear, is targeted at Twitter, which "once pledged to be a Public Square" but has "long ceased to be … They are now merely a publisher. And a bad, biased

[104] This is backed up by contemporaneous polling, such as one poll in August 2020 which found that 90% of Republican voters "say it is likely that social media sites censor political viewpoints", versus 59% of Democrats (Vogels, Perrin, and Anderson 2020).

[105] An archived version is available on the Internet Archive: https://web.archive.org/web/20200715181055/https://news.parler.com/email-letters/declaration-of-internet-independence.

publisher at that." Twitter has "become a Tech Tyrant", which "threaten[s] all of us with digital extermination should we dare to challenge them, chilling debate and forcing the People to self-censor". This heady rhetoric—whose analogical connection to 1776 is rather weakened by Trumpian phrases like "a bad, biased publisher at that"—bears similarities to the victim ideology and apocalyptic imagery of the "alt-right" and other far-right groups (Berbrier 2000). Thus, to complete the mythopoeic analogy with American independence, the "tech tyrants" like Facebook and Twitter assume the role of the British government, whose oppressive behaviour necessitates the popular uprising that the Declaration of Independence sparked.

6.2.3 Legal proceduralism

In addition to making much of the claim that Parler safeguarded *substantive* rights such as free speech and free thought, Matze also suggested that Parler users enjoyed *procedural* rights, such as due process, equal treatment, and jury trial, that were not enjoyed by the denizens of other platforms. Such appeals were again often couched in terms of American constitutionalism. We have already seen the authoritising role that the First Amendment to the US Constitution played in Matze's portrayal of Parler as a free speech haven, but Matze also alluded to the governance institutions charged with defending the Constitution, especially the US Supreme Court, as in his interview with Fox News' Laura Ingraham:

> Ingraham: So no censoring—if someone says something inciting violence, you will leave it up?
> Matze: Well, see, that's actually against the law. So anything that we get involved with has to have some sort of Constitutional violation, some sort of Supreme Court premise, something that says, hey, that's really not all right. There has to be a premise for it, though. We really don't want to get into the business of determining what is and is not allowed to be discussed.

Here, Matze outlines a view of moderation that claims to be heavily influenced by American constitutional governance. Likewise, in an email to the *Washington Post*, Matze cited "speech not protected by the First Amendment or FCC guidelines" as a basis for determining the sort of content that Parler would sanction, echoing Parler's pledge to use FCC guidelines in a blog post shortly after the site's launch (see Williams 2018a and Chapter 5). The *Post* article noted that in jurisprudential terms this was an odd basis for standards-setting, since the First Amendment

generally only prohibits censorship by the *government* not by private companies, and since FCC guidelines only concern content broadcast on public airwaves, not online platforms. Yet when such claims are viewed in symbolic rather than concrete terms, Matze's appeals to the Constitution, Supreme Court, and even the FCC as "precedents" for Parler's own rulemaking can be seen as fitting with his broader attempt to paint Parler as a bastion of support for the foundational democratic values of the American republic.

Moreover, these symbolic allusions even translated into concrete procedure, in the form of Parler's embryonic "Community Jury" system for judging the appropriateness of content (see 6.4.3), which in the summer of 2020 was said to consist of a roster of 200 unpaid volunteer Parler users. In the November 2020 interview with Tucker Carlson, the host frames Parler as being "created as an alternative to Twitter", before asking Matze "what is it you offer that [rivals] don't?" Matze replies:

> "What we're seeing is a massive explosion in growth because people trust that Parler is gonna do the right thing. So as opposed to these other companies, where moderation seems to be the norm, on Parler we have a Community Jury. This is where the people decide what is allowed and what is not allowed. You're judged by your peers just like our government allows for people. You're innocent before [*sic*] proven guilty, unlike these other platforms, that are colluding to, I guess, find things to find you guilty for."

Similarly, as the *Washington Post* noted in a December 2020 article flagging Parler's problem with moderating pornography on the platform, the rather grandiose guidelines on the Parler Jury's official page assert that:

> "No user shall be stripped of his parleys or comments, nor shall he be suspended, banned, or deprived of his standing in any other way, except by the conscientious judgment of his equals".[106]

By invoking the presumption of innocence and the common law tradition of jury trial contained in the Sixth Amendment to the Constitution, Matze sought to associate Parler's processes with longstanding hallmarks of American legal proceduralism.

[106] While the First Amendment is neutral with respect to gender, Parler's Jury guidelines substitute the masculine "he" and "his," revealing the imagined publics invoked by Parler's policies.

6.2.4 Privacy

A less frequent yet still notable value that Matze claimed Parler safeguarded was the privacy of users' data. The idea that Parler, in contrast with its rivals, treated the privacy of users' data as sacrosanct regularly appeared in the company's corporate communications, including its founding myth that co-founders Matze and Jared Thomson set up Parler because they were "exhausted with a lack of transparency in big tech, ideological suppression and privacy abuse", which was present on the platform's website from its launch until Matze's departure as CEO in early 2022. Matze reinforced this notion in media appearances, contrasting Parler's approach with that of rivals even more forcefully than with the other values he discussed. Although it was overall deployed less often overall than free speech claims, when Matze did discuss data privacy he framed it as being just as important to Parler, stating at one point that "we're so much more than just free speech … because people also want freedom from the data manipulators" and claiming that users were fed up with "manipulative algorithms built on data mining". At one point Matze even drew on the notion of sovereignty as it concerned data, suggesting Parler's ethos was "free speech, privacy, and data sovereignty. The idea that you own your own data and your own experience. And once again, events concerning larger rivals conspired to allow Matze to portray Parler as a smaller but morally superior alternative. After the release in September of *The Social Dilemma*, which raised issues of extractive data practices by social media companies, Matze seized on this to suggest that the documentary vindicated Parler's approach to "not just free speech [but also] privacy and ownership of your data". Even the filing of a lawsuit against Google in June regarding allegedly illegal data collection through its Chrome browser provided fodder for Matze, who in a statement suggested that Google "violated the trust and privacy of millions of users" and "trespassed on their user's privacy".

Again, however, the haven of privacy and data sovereignty that Matze seemed to imagine fell far short in practice. More critical reports at the time noted that for all Matze's protestations about privacy abuses elsewhere, Parler's own practices were invasive even by industry standards,

with a driver's license scan and photo required to verify a user's identity, and even tax identification information required to join the app's "influencer network". And this is not to mention the disastrous data breach that occurred in January 2021, putting exposed reams of user data including users' GPS coordinates (Cameron and Mehrotra 2021).

Overall, Matze's embrace of constitutional metaphors, which characterised Parler's platform self-governance as distinctly American, suggested the cultivation of a democratic culture that was almost Tocquevillian—set against the faceless, fact-checking, coastal (or colonial) elites. As we will see, the reality of Parler was very different. But even taken on its own terms, Matze's frequent invocation of a supposedly "simpler time"—the founding of the American Republic—as the basis for his legitimation of Parler's self-governance elides the gender- and race-based exclusions that are inextricably linked to this historical touchstone (Fraser 1990). In 1776, the ambit of free speech—and even physical freedom—extended only to a select stratum of American "citizens", and it would be nearly another century until the widespread institution of slavery was abolished in the United States. Matze's attempts to discursively legitimate Parler's self-governance with such appeals must be seen in this historical context.

6.3 Multimodal analysis of Parler's website

The framework for discursive legitimation in public communication presented by van Leeuwen is multimodal and can thus be adapted to visual contexts: "stories, for instance, can be told visually [and] moral evaluations can be connoted visually or represented by visual symbols" (van Leeuwen 2007:107-108).[107] This is reflected in the attention to visual elements of the corpus just discussed, such as videos of Matze's TV appearances embedded in articles in the sample. I conclude this discourse analysis by examining a separate corpus which contains visual as well as linguistic

[107] The reverse is also true—discourse analyses are a valid form of visual methodology (Rose 2022).

elements: archived copies of Parler's publicly facing website during 2020, which was another venue for Parler's performative legitimation of its platform self-governance.[108]

In June 2019 Parler's homepage received a revamp, similar iterations of which were to remain in place until Jun 2020 (see Figure 13); I therefore begin the analysis here. In contrast to its initial focus on attracting publishers to the platform (see Chapter 5), by June 2019 the site is squarely aimed at users. Parler is described as a "user driven platform", which allows users to "read news, speak free" and to "control your experience". The "read news, speak free" formulation recalls the revolutionary mantra "live free or die". Despite the red colour palette, the homepage contains no explicit references to the platform having an ideological skew, much as Matze would continue to portray the platform as open to users of all ideologies throughout 2020, even as conservative users were joining the site en masse.

However, the ideological skew and aggressive tenor of the platform itself is also hinted at. A prominent "Heated Discussion" label overlays a cascade of Parler posts with various degrees of visibility. The sentiment expressed in each post is superficially neutral, reflecting the platform's stated commitment to impartiality, but the focus is clearly skewed towards Republican politics and a combative communicative style. The foregrounded post focuses on Donald Trump and states that "I know exactly where my vote goes"; the second hints at the fierceness of debate on Parler with a reference to "getting popcorn"; and the third is less visible but shows a link to conservative news website Breitbart. The availability of Parler as an app on both iOS and Android devices is also featured prominently, with direct download buttons which take users to the respective app stores, showing the centrality of the app experience to the platform as a whole.

[108] The artefacts contained in the third corpus—Parler's policy documents—were also on Parler's website during 2020, and were also, as I will argue, performative attempts to legitimate Parler's self-governance. However, as these are solely textual policy documents, I treat them separately in the next section.

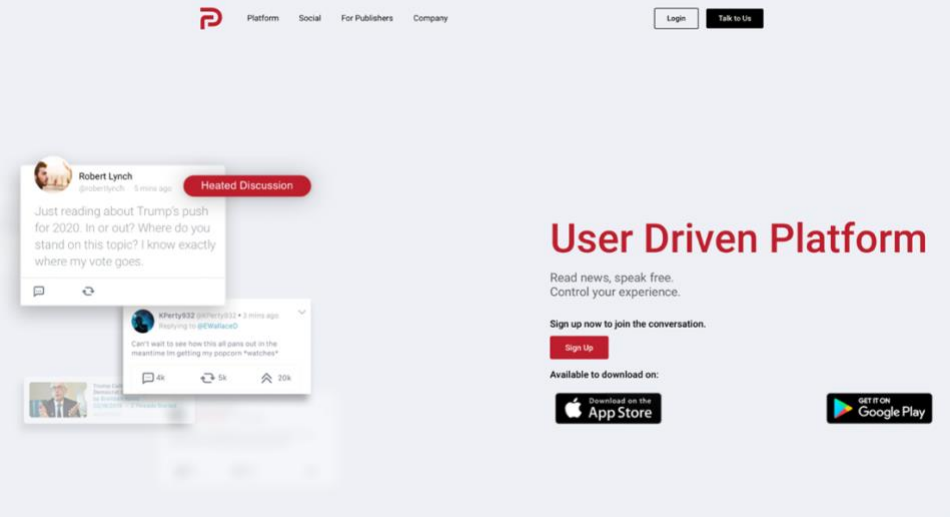


Figure 13: The basic structure of Parler's homepage between June 2019 and June 2020. Source: The Internet Archive, https://web.archive.org/web/20200619180729/https://www.parler.com/auth/access

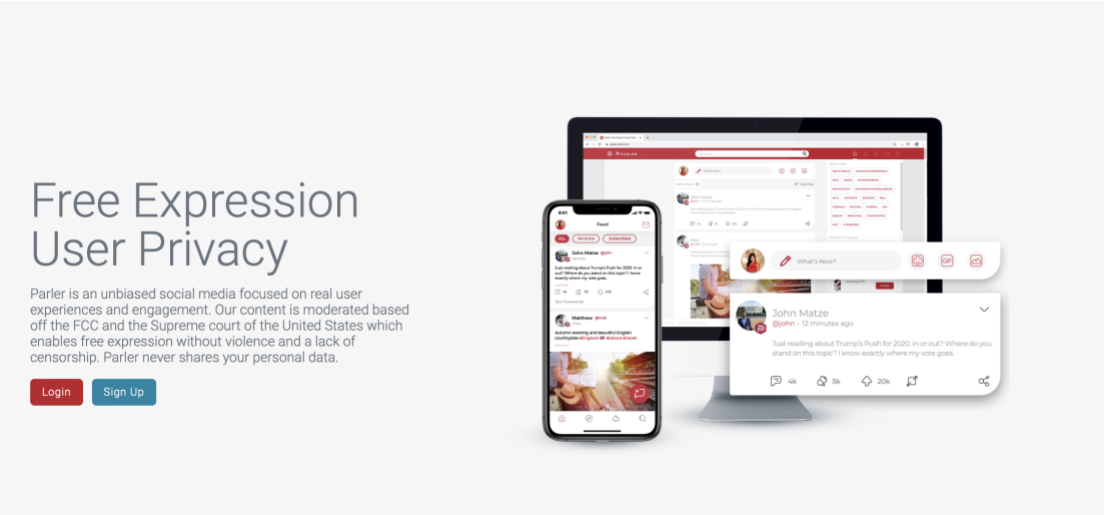


Figure 14: Parler's homepage from June 2020. Source: The Internet Archive, https://web.archive.org/web/20200619180729/https://www.parler.com/auth/access

Parler's home page was modified in the middle of 2020 (see Figure 14). This version has a similar design, with screengrabs and some text, but there are notable differences. First, the strapline now reads "Free Expression[,] User Privacy", focusing on the supposedly superior affordances that Matze was contemporaneously touting in media appearances. The text below also reflects Matze's public rhetoric; Parler is "an unbiased social media", and its content moderation is "based off the FCC and Supreme court of the United States which enables free expression without violence and a lack of censorship". Constitutional precedents are thus apparent here too. Meanwhile, the chosen images now emphasise that Parler can be accessed both on desktops and on smartphones (an

iPhone in this case). Notably, the foremost Parler post presented in Figure 13 above appears here too, but is now attributed to CEO Matze. This underlines Matze's centrality to the company, yet the authenticity of the post itself is obviously undermined by its reattribution.

For most of 2020 Parler's website remained basic. Other than the homepage there were three further pages, "Social", "For Publishers", and "Company". The "Social" page at this time[109] provided somewhat more information about the platform's affordances and expected uses, with Parler described as "a non-biased free speech social media focused on protecting user's rights and providing transparency in technology"—highlighting many of the same features and values Matze had touted publicly. The "For Publishers" page was, as the title suggests, aimed at the news brands whom Parler sought to still bring to the site; notably, the mocked-up screengrabs used the Epoch Times, a far-right, conspiracist website, as the example of a publisher on Parler. Finally, the "Company" page[110] contained a statement of the company's four "values", accompanied by rather obviously stock images (see Figure 15 below). Notably under the "Build Communities" pane, it pledges: "We are not regulators. We are not governors. We are a community", alongside a photo of a circle of people happily arm in arm. This marks an attempt to recast the role of not just Parler the platform, but also Parler the company, as a community, with echoes Mark Zuckerberg's discursive construction of Facebook as a "global community" (Haupt 2021; Rider and Murakami Wood 2019). It echoes the hands-off approach to governance which Matze frequently celebrated. In the next pane, under "Free Speech", Parler is described as "a non-biased free speech driven entity", with a line of people apparently chanting or protesting; one speaks into a megaphone. In contrast to the depiction of Parler as just one community, the value of "Free Speech" contains the potential for dissent.

[109] https://web.archive.org/web/20200626204912/https://home.parler.com/social/
[110] https://web.archive.org/web/20200605170304/https://home.parler.com/about/

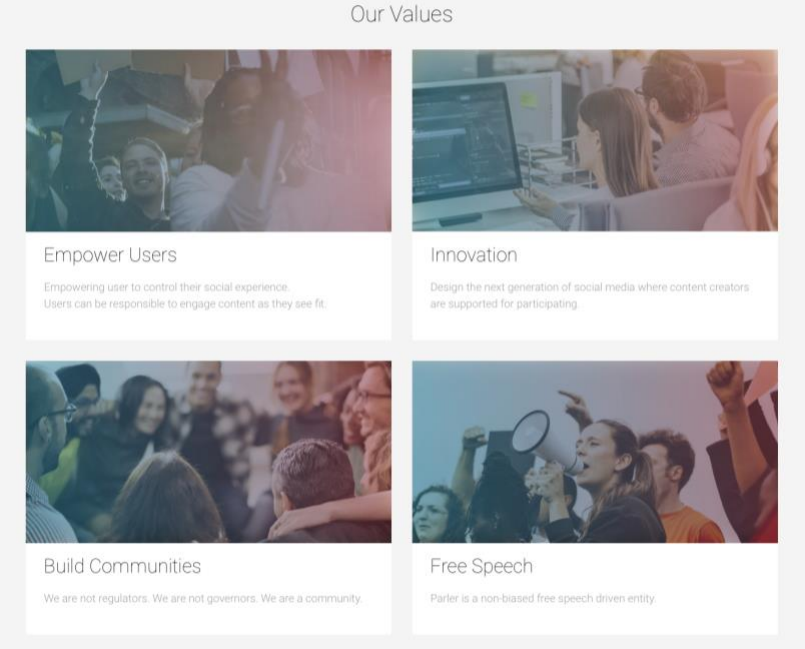


Figure 15: Parler's "values" as of June 2020. Source: The Internet Archive, https://web.archive.org/web/20200605170304/https://home.parler.com/about/

Parler's home page was once again revamped in late 2020 and the visual iconography changed considerably (see Figure 16). A new strapline declares Parler to be "the world's town square", employing the same rhetorical touchstone that Matze used in scores of media appearances, Meanwhile, the description nods to the "deplatforming" that had recently occurred on rival social media platforms. Most striking in visual terms is the choice of background image: a young white woman, wearing dungarees over a red checked shirt, and sturdy checked boots, gazing into the distance. Behind her, a soft-focus background resembles a rural setting with what looks like a green hedgerow. The storytelling at work here is not terribly subtle. Parler is middle America: the country's heartland, where the real hard work gets done and food gets grown, geographically and culturally far from the west coast where Silicon Valley elites reside and rule.

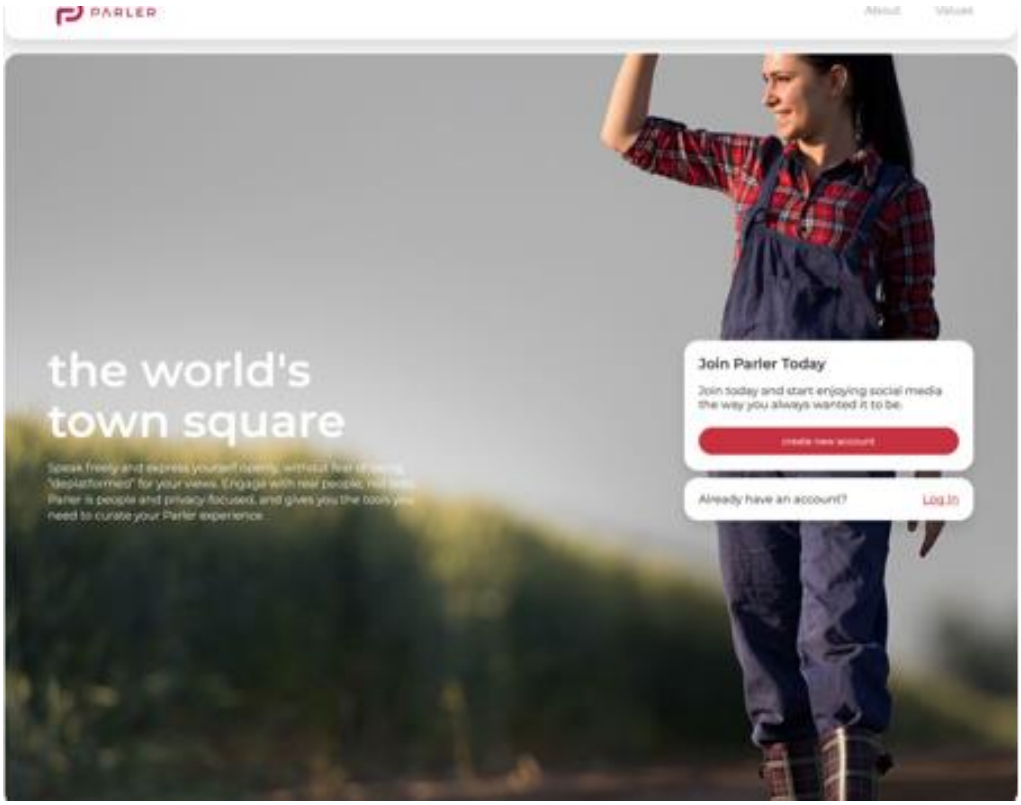


Figure 16: Parler's website as of October 2020. Source: The Internet Archive, https://web.archive.org/web/20201026205033/https://www.parler.com/auth/access

When drawing overarching conclusions from these messages and images, one must be cognisant of the context, and to the multiple purposes to which Matze's media appearances and Parler's public web presence were being put. During 2020, Parler was targeting multiple audiences—ordinary users, elected representatives, and publishers—at once. Just as significant was the gulf between the ideological diversity of users Matze purported to want, and the users Parler actually had, revealed and reinforced by the "personalisation" component of the technical walkthrough. In the controlled communicative settings of largely sympathetic media appearances and the company's public presence, these contradictions could be accommodated. But this was a harder balance to strike in the platform's policy documents.

6.4 Parler's policies

In this second component of the analysis, I turn from how Parler's self-governance was presented during 2020 to explore evidence closer to the "shop floor" of Parler as a platform in practice in the same period. I analyse three types of policy document that were in effect in the period:

"Community Guidelines" (for which I collected six versions); an "Elaboration" on these guidelines; and a description of Parler's "Community Jury".

6.4.1 Community guidelines

Various versions of "community guidelines" appear to have been in force on Parler throughout 2020, but their evolution through the period reflects the difficulty Parler faced in articulating a coherent and consistent set of rules for its users to follow, at a time of considerable growth and outside scrutiny.

The first version archived in the year 2020 (on 17 April) is undated. As a result, it is possible that the last version of the guidelines that were archived in 2019 (dated as 26 June 2019; archived on 29 June 2019) were in force at the start of 2020; I therefore begin the analysis here. In this June 2019 version, the guidelines are described as being "intended to provide insight on what kinds of interactions on our platform may warrant negative attention and action" by Parler. Notably, the constitutional inspiration for Parler's platform governance that was cited widely by Matze in mainstream media appearances is apparent in its Community Guidelines too. Parler, it is stated, "stand[s] behind the definition of free speech as laid out in the United States Constitution". As a result, Parler considers itself merely "a neutral party whose responsibility is to uphold the law as defined by" the Constitution, US Supreme Court and FCC guidelines on obscenity. I have already discussed the curious application of constitutional and regulatory guidelines to Parler as a private platform in 6.2. But this document goes further by defining, "for the purposes of clarity", examples of content that is and is not permitted in Parler. An included chart lists a series of "cases" concerning "Protected Content" and "Prohibited Content" (see Figure 17).

Case	Protected Content	Prohibited Content
1	Reasonable Discourse	Spam
2	Political Discourse	Terrorism/Terrorist Organizations
3	Political Views	Unsolicited Advertisements
4	Satire	Defamation, including libel and slander
5	Dark humor	Fighting words
6	Offensive speech	Blackmail
7	Hate Speech (As long as it not considered obscene by the FCC.)	Threats against a group or others
8	Disagreements	Solicitation to commit crime
9	Direct debate with users	Impersonation of individual or company
10	Sarcastic suggestions	Pornography
11	Talking positively or negatively about Gender, Orientation, Religion, and Race	Nudity (No exceptions for art, breastfeeding, etc.)
12	Fake News (Even if it is misleading)	Obscenity (Defined by the FCC)
13	Real News	Plagiarism
14	Positive or Negative speech about the president or any elected official of the United States	Bribery
15	Talking positively or negatively about the Parler platform	References to killing groups of people or individuals
16	Your right to be anonymous to other users of the platform	Any other speech that is federally illegal in the United States of America

Note: We previously included racial slurs in prohibited content. We have removed it all together from the chart. We will take racial slurs on a case by case basis until we can properly provide automated tools for the community to opt in/out of racially motivated comments. We are not in the business to define what racism/racial slurs are so we would like to leave you automated tools for the community to do so.

Figure 17: Parler's "Protected" and "Prohibited" content, as of its Community Guidelines, archived 29th June 2019.

This long, scattershot list of examples complicates the notion of Parler as the self-proclaimed home of free speech. Included in "prohibited content" are forms of content that were permitted on (some) rival social networks, such as pornography and nudity (permitted on Twitter), with "no exceptions" for non-sexual content such as images of breastfeeding. Yet "hate speech" per se is *explicitly permitted*, and so too is "talking … negatively" about characteristics such as gender and race. Even given the prohibitions against "fighting words", "threats", and "references to killing groups of people or individuals", the explicit permissibility of hate speech evidences a very specific form of free speech maximalism and, importantly, runs in direct contrast to app store governance

guidelines on this point. Also significant is the caveat beneath the chart, which addresses "racial slurs", stating that Parler is "not in the business to define … racism [or] racial slurs". This rather conveniently exonerates Parler from ruling on the use of, for example, the N-word, reflecting the equivocations Matze offered on the subject in media appearances.

An updated "Community Guidelines" document was archived by the Internet Archive from the same URL on 17 April 2020. In this iteration, the list of "permitted content" has been removed, and the document instead focuses exclusively on content that is either prohibited or discouraged (the distinction is not always clear). This version echoes the earlier reliance on American constitutional and regulatory doctrine, stating that the provisions "are defined and updated by the FCC and Supreme Court case outcomes", and this time, "related" Supreme Court cases are included to legitimate several of the outlawed forms of content. In contrast to the previous set of guidelines, this document contains scant reference to hate speech explicitly. Additionally, the caveat on the use of "racial slurs" has vanished. Thus, in contrast to the previous version, hate speech is no longer explicitly permitted. The document discourages use of "fighting words", which are defined as "incitements to violence that produce a clear and present danger". Examples of "fighting words" are said to include "any direct and very personal insult … which would lead to violence if you were to talk in that fashion in person", as well as "a threat or advocation of violation [sic] against an individual or group". This suggests that words targeting specific groups are not allowed if, and perhaps only if, they would lead to violence.

A later version of the Community Guidelines, dated 21 August 2020, changed Parler's policies on hate speech once again. The document restates that Parler's goal is to "provide all users with a welcoming, nonpartisan Public Square," and declares that Parler is "view-point neutral", echoing Matze's contemporaneous public comments. This version of the document acknowledges, for the first time, that "the First Amendment does not apply to private companies such as Parler" but nonetheless that "our mission is to create a social platform in the *spirit* of the First Amendment" (italics in original). In this version, the range of prohibited content is much reduced

to just two "principles". The first principle concerns content that "the law requires us to exclude", among the "obvious examples" of which is "content posted by or on behalf of terrorist organizations, child pornography, and copyright violations." The company also suggests that it would sometimes remove content even when not required by law, "especially when these interfere with our mission of providing a welcoming, nonpartisan Public Square"; such content is said to include "criminal solicitation, fraud and nuisance." The second principle involves rules against "posting spam and using bots." But beyond these minimalist principles, the more significant addition at this juncture was the addition of a link to a separate document, entitled "Elaboration on Guidelines", which is said to offer "a detailed discussion of the types of actions encompassed by these two principles".[111]

6.4.2 Elaboration on guidelines

In the version that I was able to obtain from the Internet Archive (archived 1 November 2020), the undated "Elaboration on guidelines" document discusses the prohibition of ten forms of problematic content on Parler. The document addresses some of the more notorious content challenges Matze had acknowledged publicly, including bans on excessively "gratuitous" depictions of "death", "serious physical violence", and "bodily fluids or organs of any kind (blood, feces, etc)". These tweaks to Parler's rules are roughly contemporaneous with Matze's public discussion of them, suggesting alignment between Parler's rules and Matze's public attempts to discursively legitimate them. The Elaboration document also reveals more substantial updates to the company's professed approach to hate speech. In contrast to the April-archived document in which they are discouraged, here, "fighting words" are explicitly *permitted* and deemed "not a violation of our Guidelines". However, a subset of "fighting words" are disallowed, namely those that contain a "serious expression of an intent to commit an act of unlawful violence to a particular

[111] Thereafter, only very small changes were made to the community guidelines in the remainder of 2020. The "no spam" principle which underlined the platform's commitment to its "influencers", and references to "users" were changed to "community members", likely reflecting Matze's disavowal of the term after the release of *The Social Dilemma* (see 6.2).

individual or group of individuals, with either the *intent* or *reckless disregard* as to whether the communication will "place the victim in fear of bodily harm or death" (italics in original). I quote this rule at length because it mimics a somewhat obscure legal text: a dissent by Supreme Court Justice Clarence Thomas, against the Court's decision to deny a writ of certiorari to (that is, a refusal to hear the case of) *Kansas v Boettger*, a case concerning the right of a state government from criminalising threats of violence where the person making the threat acts in reckless disregard to causing fear in the subject of the threat. What is notable for present purposes is less the specifics of this legal case, but more the fact that Parler directly draws on an unheard Supreme Court case as the basis for its own rulemaking.

In addition to Justice Thomas (who, since the death of Justice Antonin Scalia, is considered the most consistently conservative member of the US Supreme Court: Gonzalez and Alberti 2022), the document assembles a broader array of libertarian voices in its citations, including libertarian legal scholar and blogger Eugene Volokh (2018), a blog post by the conservative Cato Institute (Samples 2018), and a book on hate and free speech by the former president of the ACLU, Nadine Strossen (2018). What explains the sudden addition of legal citations? This appears to be the result of the hiring by Parler in August of someone Matze described as "a chief policy officer who's a real lawyer", which meant that it was no longer Matze "writing the community guidelines" (Blake 2020). We can therefore infer that the "Elaboration" document was the brainchild of Parler's Chief Policy Officer, Amy Peikoff. The hiring of Peikoff—a lawyer and Ayn Rand devotee—as Chief Policy Officer evidently represented an attempt to tighten up (and smarten up) Parler's rules and policies. This had an immediate impact with the platform's rather more considered guidelines, a reflection of Peikoff's scholarly publication history on topics such as the right to privacy (Peikoff 2008). Regardless of the legal merits of Parler's (or Peikoff's) arguments, this "Elaboration" document represents an effort to further ground Parler's rules in American "constitutional" precedent, marking an advance on Matze's uneven and often inchoate discursive commitments, through references to (specifically right-wing and libertarian) legal scholars and judges. The

appearance of the Elaboration document in the latter half of 2020 thus suggests a more focused and informed effort by Parler to clarify its rules and procedures.

6.4.3 Parler's Community Jury

Finally, my crawl of Parler's website archive yielded a copy of a short document entitled "Parler Community Jury". The document describes the functioning of Parler's "jury" process: "a user reports a piece of content", which is then "sent to the Jury Portal for review by members of the Community Jury". In their review, the "jurors … determine whether the Community Guidelines have been violated, and also whether they believe the content is illegal" and should therefore be removed from the platform. This proceeds through a "quorum system", in which up to five jurors are shown the content in question and decide whether to cast a "guilty" or "not guilty" judgment; four out of five "guilty" verdicts are required. If the post or comment in question is considered "guilty" of a violation, then the user who posted it receives "points" by way of sanction. Points lapse 90 days after they are assigned, but "users accumulating 20 or more points at any time will be banned". Finally, the document states that the Community Jury "is composed of volunteer, verified Parleyers, who participate in regular training sessions to address any questions they have about applying the Community Guidelines".

This document echoes Matze's invocation of Parler's jury system in mainstream media interviews, which was used to further legitimate the platform's "constitutional" basis for self-governance. The idea of using juries to adjudicate and legitimate content moderation decisions in social media has been the subject of novel research (Fan and Zhang 2020). However, given the ideological skew of the website and its moderators (see Chapter 5) it is difficult to envisage a user getting an impartial "hearing". And in practice, the number of "jurors" was far too small to handle the reams of content reports as Parler's user base grew (see Chapter 7). In this way, the Community Jury exemplifies the overall efforts to discursively legitimate Parler's platform self-governance that I found in this analysis. The range of loose metaphorical invocations from American political

history may have bolstered the portrayal of Parler as the "free speech" platform, but this minimalist approach augured poorly for its platform self-governance in practice.

Summary

At first glance, and absent any further context, the publicly available evidence presented thus far might lead one to surmise that Parler's platform self-governance in 2020 was—while rather kooky, and undoubtedly US-centric—not altogether malign. The technical walkthrough found that Parler's app resembled its mainstream platform rivals with respect to most of its mediator characteristics such as menus and icons, yet it had some crucial differences: among them its colour palette, which hinted at its conservative skew; and a "personalisation" feature presented to users at sign-up, which encouraged them to follow predominantly right-wing figures, revealing the platform's imagined user. Meanwhile, in his regular media appearances throughout 2020, CEO John Matze drew on the keystone values of American democracy to promote and legitimate the platform's approach to self-governance. This rhetorical portrayal was reflected on the platform's website, which emphasised Parler's "heartland" credentials, and it reverberated through the platform's policy documents, which undergirded the platform's claimed commitment to justice. The overarching image that these discursive efforts sought to construct was that of Parler as a "populist platform", more sympathetic to the interests of "real" Americans deplatformed and silenced by the coastal "elites" of Silicon Valley—common tropes of the populist radical right (Berbrier 2000; Mudde 2016). In this way, my finding supports analysis by New America's characterisation of Parler as "a social media company whose very business model appeared to be predicated on the idea that it could grow its user base by promoting the discourses of the deplatformed and the disaffected among Trump's followers" (Rondeaux et al. 2023). In short, this was a platform designed and operated by, and for, a narrow band of self-described American "patriots". However, the weighty platform self-governance challenges that Parler faced in 2020 lay close to the surface of this smooth discursive portrayal. In the next chapter I turn from examining

Parler's platform self-governance as it was portrayed publicly to consider what it was like in practice, and assess the involvement of app store operators in the cross-platform governance of Parler.

Chapter 7: Parlous—the app store governance of Parler

Introduction

In this chapter, I turn from assessing Parler's discursive performance of its platform self-governance to consider the cross-platform governance of the platform in the same period, focusing on Apple and Google as app store operators (ASOs). I build on the conceptualisation of app store governance (ASG) developed in Chapters 3 and 4 to examine the responsibility that each ASO had accorded themselves with respect to hate speech.[112] Notwithstanding the opacity of the two companies' guidelines, I make a best effort to interpret and systematise them to assess how problematic the prevalence of hate speech on Parler would have proved during 2020. The empirical analysis of the sample of posts on Parler finds a high degree of hate speech, and hate speakers, on Parler, with no obvious longitudinal trends (7.1.1). Qualitative analysis reveals that the targets of hate speech on Parler were predominantly Muslims, immigrants, and women, especially female politicians. Violent and insurrectionist content was rife (7.1.2). The empirical findings are set against Apple's and Google's respective guidelines on hate speech, and it is argued that on the basis of their guidelines the two companies would have been warranted in removing Parler from app stores prior to January 2021 (7.2). Finally, the circumstances of Parler's eventual removal and reinstatement are discussed (7.3).

7.1 Parler in practice

In this section I report the results of a content analysis of the sample of posts made on Parler during 2020. I begin in 7.1.1 by reporting the top-level findings from the automated classification of the sample. I then turn in 7.1.2 to discuss findings from the manual analysis.

[112] For the sake of precision, in this chapter I use the app store review guidelines that were in force during 2020; thus, there may be some discrepancies between the wording here and elsewhere in the thesis, where I quote up-to-date app store review documentation. In practice, however, the guidelines remain broadly similar, and in most places identical, at the time of writing.

7.1.1 Hate speech on Parler: overarching findings

As detailed in Chapter 5, the automated component of the analysis involved obtaining measures of the likelihood of hate speech contained in the stratified random sample of 21,798 posts made on Parler during 2020, with classifications obtained from two automated systems, Rewire and Perspective API. Overall,[113] 3.3% of posts in the sample (572) were classified by the Rewire API as having a greater than 70% likelihood of being perceived as hateful. This result is strikingly similar to the level of hate speech measured by a different analysis of hate speech on Parler—despite differences in sampling and classification methods—which found that 3.29% of Parler posts contained hate (Israeli and Tsur 2022). For Perspective API, 5.1% (876) were classified as being likely to be perceived as toxic at the same threshold. While precise comparisons are not possible, these figures are also roughly comparable to the levels of hate speech detected on other fringe platforms. Zannettou et al. (2018) found that 5.4% of posts on Gab contained a hate word between 2016 and 2018—notably, a period in which Gab was removed from Google's Play Store specifically on the basis of hate speech on the platform (Coldewey 2017). It is however somewhat lower than the 12% of 4chan's /pol/ board that were found to contain at least one hate word (Hine et al. 2017). On a user basis, 3.8% (526) of users in the sample were classified as having made at least one post that met the Rewire API threshold.[114] For Perspective API, 5.7% of users (790) were classified as making at least one post likely to be perceived as toxic. This data is summarised in Table 6.

[113] During the app walkthrough I noticed that when a new user joined Parler, the platform would send a stock "greeting" post; I removed these posts from the classification analysis reported here. The post read, "I just joined Parler! Looking forward to meeting everyone here."

[114] By contrast to the previous comparison, this finding is notably lower than the figure found by Israeli and Tsur (2022), which found that 13% of Parler users shared at least one hateful post. This may reflect the longer period of their analysis, which began in August 2018. All else being equal, it may therefore be that though the overall level of hate on Parler remained fairly stable between 2018 and January 2021, the initially much smaller number of users meant that the proportion of users expressing hate was higher during this early period.

Table 6: Proportions of hate and toxicity detected in the Parler post sample

Sampled posts	19,237
Sampled posts excl. "new user" greeting	17,305
Share of posts scoring > 0.7	
Rewire: Hate	3.3%
Perspective: Toxicity	5.1%
Sampled users	13,936
Share of users with at least one post scoring > 0.7	
Rewire: Hate	3.8%
Perspective: Toxicity	5.7%

The proportion of posts that were classified as hateful (Rewire) and toxic (Perspective) for each month is visualised in Figure 18. As Figure 18 shows, the degree of hate and toxicity on Parler remained relatively constant over the course of 2020, with no clear trend despite the much higher volume of posts made in the latter part of the year. Overall, it does *not* appear that Parler became proportionately "more hateful" in 2020, though in absolute terms there was undoubtedly more hate expressed on the platform.

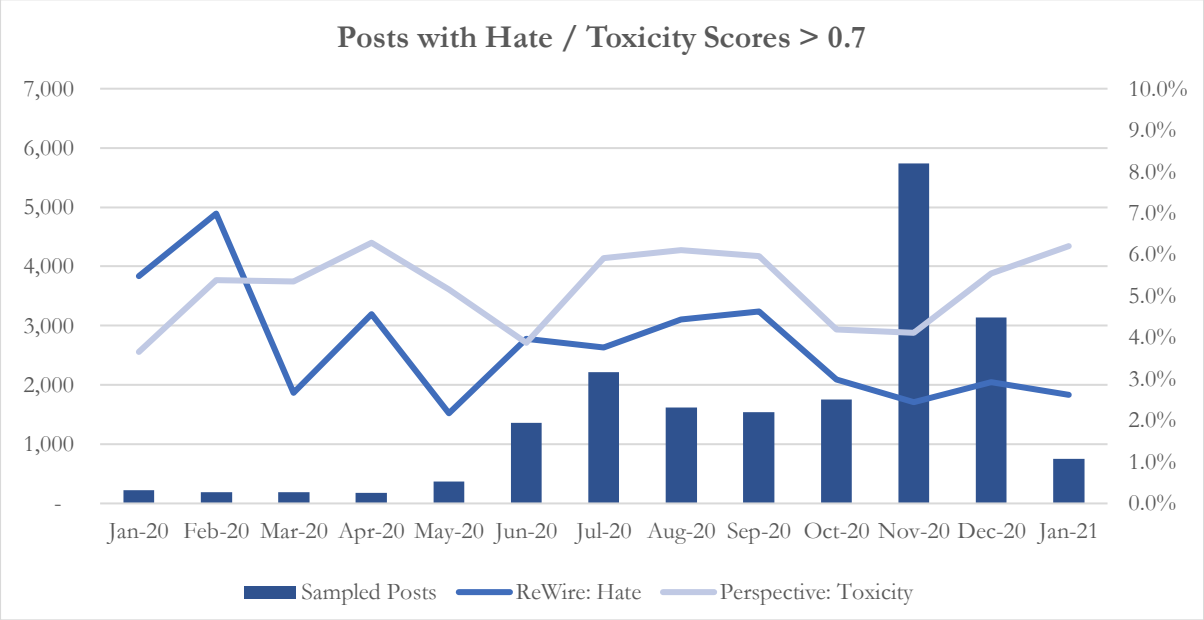


Figure 18: Volume of Parler posts in the sample overall (left axis), and proportion of posts which contained at least one hateful or toxic post in the sample (right axis).

Figure 19 shows the number of users who made at least one post classified as hateful (Rewire API) or toxic (Perspective API) post in each month. The close similarity between the two charts dispels the notion that hate on Parler was spread only by a small handful of users at any time during 2020, suggesting instead that it was a more systemic problem.

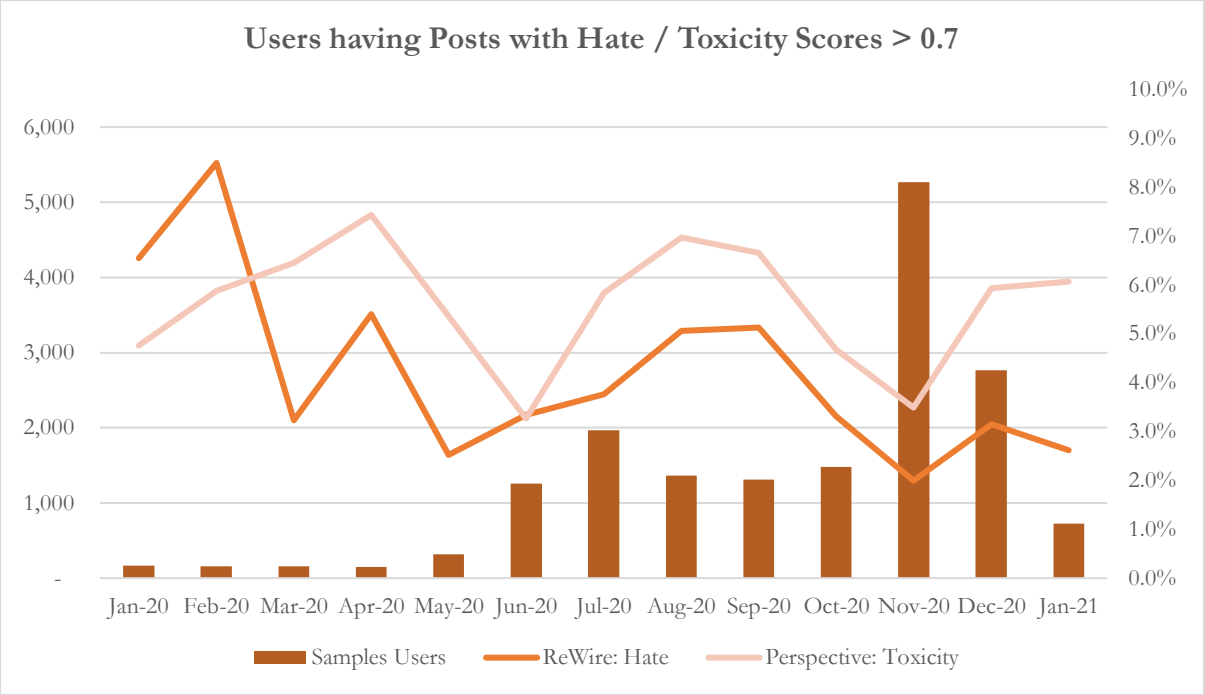


Figure 19: Volume of Parler users in the same sample (left axis) and proportion of users who made at least one hateful or toxic post in the sample (right axis).

7.1.2 Manual analysis

The automated classifications reported above provide a useful overview of the level of hate over time. Manual analysis, in turn, enables a more sensitive and nuanced understanding of the nature of hate speech on Parler during 2020. To restate, this analysis is based on the same sample of 19,237 posts that was classified automatically, and I analysed the posts chronologically in the order that they appeared online.

Several high-level characteristics about discourse on Parler were evident across the sample. First, Parler was political to a fault, and the perspectives expressed were almost exclusively right-wing. Almost every post contained a political comment or sentiment—so much so that, when working through the sample, I became surprised to chance upon a post on a non-political topic (such as a recipe), or a comment from a left-wing perspective (of which I counted four *overall*).

Moreover, even the discussion of many cultural phenomena were themselves the subject of political debate. For every cultural comment, such as a sports game ("come on Giants!"), there were substantially more posts making a comment about some politically salient aspect of culture, such as negative reactions to the decision by NBA basketball players to kneel in protest of the murder of George Floyd when the league resumed in July (Carroll 2020).

Second, although a large majority of posts on Parler (in English) were seemingly from right-wing American users, this right-wing user base was a relatively broad church (Newhouse 2020), at least compared with that of Parler's alt-rivals such as 8chan (Rieger et al. 2021). Despite the prevalence of truly sickening hate speech, as well as considerable conspiracism especially about the origins and effects of Covid-19, there was also a significant presence on the platform of somewhat more mainstream conservative voices.[115] This was clearest in the case of posts which drew on Christian religious messaging. Sometimes, Bible passages were quoted without elaboration or contextualisation, but more often, bespoke "prayers" and religiously inflected sentiments were offered with respect to the political issues of the day. For example, this could take the form of a prayer for the success of President Trump, or a prayer for police intervention to stop protests. Politics and religion most often combined in the case of abortion—a longstanding issue of controversy and contestation in American politics (Glenza 2021). In keeping with much of the rhetoric on Parler, discussion of abortion on the platform was extreme—with the procedure usually being framed as akin to "infanticide"—and was often accompanied by explicitly misogynistic hate speech. However, the salience of abortion on Parler, as well as gun ownership, another "signature issue" of the American right, suggests that the platform accommodated not

[115] Recent years have seen the Republican Party—including both its elected officials and its voters—shift significantly rightward, to embrace extreme stances on various issues. This is exemplified by the difference between the party's presidential candidates in 2008 and 2012, John McCain and Mitt Romney, to its standard-bearer since 2016, Donald Trump. Thus, a "mainstream" Republican today is far to the right compared with their counterpart in earlier decades (Bacon 2018).

just a radical, conspiracist fringe of the far-right, but also users with a relatively more mainstream set of right-wing values and interests.

Third, and related, posts on Parler were highly partisan, in the sense that users seemed to have a very clear sense who their perceived political enemies were. Democratic leaders and voters were commonly nicknamed "DemonRats" among similarly derogatory terms, and "communist", "socialist", and "Marxist" were frequent epithets applied to the policies and personalities of both mainstream Democratic leaders like Joe Biden and Kamala Harris, and to those of more radically left-wing figureheads like Bernie Sanders.

Taken together, these dynamics are important because they raise the possibility of the "transmission" of hate on Parler. This phenomenon was identified in the case of Gab by Gallacher and Bright (2021), who found a positive correlation between the amount of time users spent on the platform and how likely they were to spread hate, and a link between exposure to and expression of hate at the user level. This is especially significant on Parler, given its larger user base than Gab (Israeli and Tsur 2022) and the somewhat more ideologically heterogenous user base that this analysis has indicated. In short, it raises the prospect that highly politicised but initially less radicalised right-wing users may have been exposed to and influenced by expressly hateful sentiments on Parler, in an example of what Munn has called the alt-right "pipeline" (Munn 2019). Based on the iterative coding schema detailed in Chapter 5, my coding focused on identifying (i) the target(s) of hate speech and (ii) violent and insurrectionist rhetoric; the latter was not necessarily specifically targeted at an identity group but threatened political violence. I display the results of the content analysis in Figure 20. Once again it is important to remember that, although the sample is random, the underlying dataset from which it is drawn is not necessarily statistically representative of Parler during 2020.

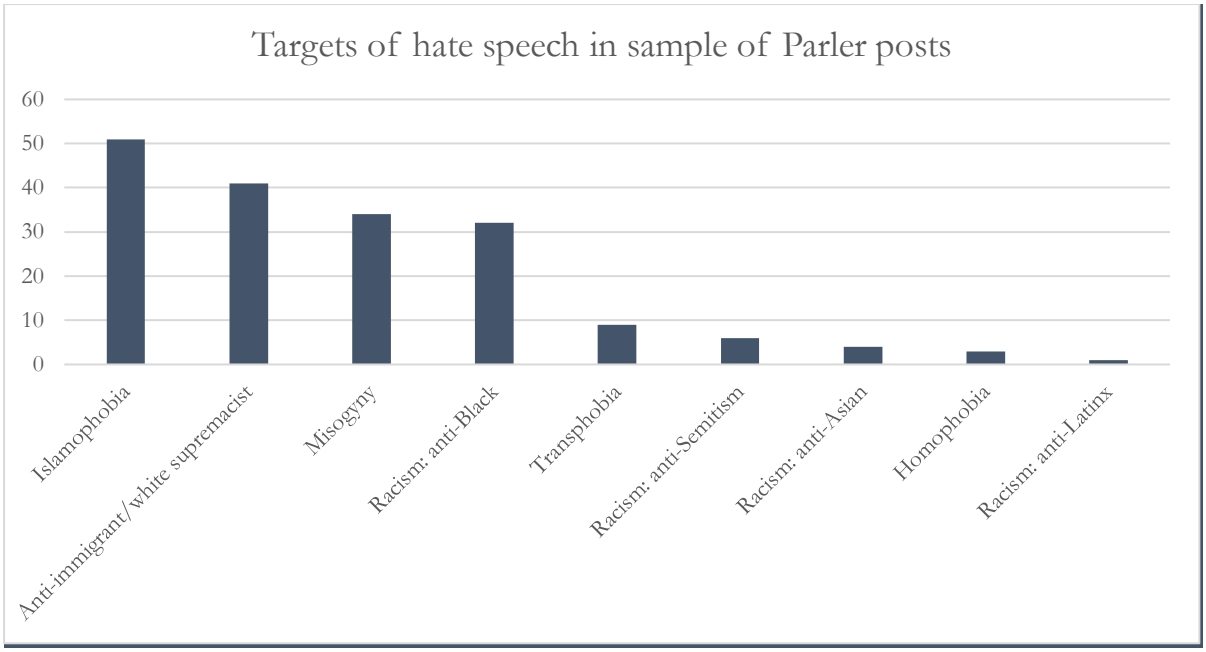


Figure 20: The identity groups most frequently targeted on Parler.

As the figure shows, the most common targets of hate that I identified in the sample were Muslims, immigrants (which also included general expressions of white supremacy and racial purity), women, and Black people. The high rates of Islamophobia and anti-immigrant sentiment are very similar to the same forms of hate speech found on Gab (Gallacher and Bright 2021), but compared to Gab, there appears to have been relatively less anti-Semitism on Parler.

Not only the degree but also the nature of Islamophobia in the sample is sobering. This included, most prominently, attacks on Representative Ilhan Omar, a well-known Muslim member of Congress, whose Somalian background was infamously targeted by President Trump on Twitter in 2019, when he told her and fellow opponents to "go back" to the countries they came from (Rogers and Fandos 2019). Similar, and worse, rhetoric was in evidence on Parler. Where it appeared, anti-immigrant sentiment was often coupled with expressions of white supremacy, or the notion of the west as civilisationally superior. Misogyny most often arose in the case of specific female politicians including Kamala Harris and Hillary Clinton, with users disparaging their character and suitability for office in extremely crude ways. Anti-Black racism spanned a continuum from explicitly racist expressions (including the N-word, which Matze had claimed Parler would assess on a case-by-case basis), to more subtle insinuations and stereotypes which it

is unlikely that word-based analysis would have detected (Rieger et al. 2021; Siapera, Moreo, and Zhou 2018).

Violent content was also strikingly prevalent on Parler. Much of it took the form of calling for the execution of political enemies, often in graphically descriptive terms. Much of it railed against liberals, Democrats and "Marxists", or the political establishment in general, and as such did not constitute hate speech in terms of specific identity categories. However, in several cases it did, including violent threats against specific politicians and groups like Black Lives Matter.

Finally, we can consider the extent to which Parler banned users. In total, only 21 (0.15%) of the 13,936 users in the sample were flagged as "banned" as of January 2021. Of the 526 users who made at least one hateful post in 2020, only two (0.38%) were among those banned.[116] This suggests a very low level of moderation of hate speech on Parler during 2020, which may be taken to reflect both principled reluctance and logistical difficulty. As we have seen, in terms of both its public rhetoric and rules, Parler professed to adopt a hands-off approach to hate speech—recall that hate speech was expressly *protected* on Parler for at least the first half of 2020. But it is important to also consider the logistical challenges Parler faced in moderating even the content that it was willing to tackle. As of late June 2020, Parler had only 200 volunteer moderators for 2.8 million users (Lerman 2020). As of January 2021, this had risen to 802 members of its "Community Jury", most of whom were unpaid, for its 13 million users (Thiel et al. 2021). Therefore, even if Parler had been more willing to moderate problematic content like hate speech, it would have been hamstrung by the sheer scale and logistical challenges that this complex task involved.

7.2 App store governance of hate speech on Parler

In Chapters 3 and 4 I discussed Apple and Google's processes and guidelines for reviewing (and if necessary, refusing or removing) apps and app updates submitted by third-party developers to

[116] In both cases, it is not possible to tell from the data whether the users in question had been suspended and then reinstated, or were permanently banned.

their app stores. I identified the "meta-moderation" approach that both companies take to governing user-generated hate speech, wherein moderation is delegated to platform operators in the first instance, while ultimate control rests with the companies. This implied a hypothetical threshold at which a platform could be considered sufficiently replete with hate speech to warrant its removal. In this section I draw together the rules that each app store had regarding hate speech with the empirical findings of the previous section to assess how app store operators governed hate speech on Parler, specifically in relation to this notion of a threshold for hate. I address each company's guidelines in turn before assessing both together.

7.2.1 Apple's App Store Review Guidelines

Four elements of Apple's App Store Review Guidelines (hereafter "ASRGs") in place during 2020[117] were pertinent to the prevalence of hate speech on Parler. These were the Introduction; "Objectionable Content" and "User Generated Content" under Safety (Section 1); and the Legal section (Section 5).

As noted, the Introduction to the ASRGs includes Apple's commitment to "strongly support all points of view being represented on the App Store, as long as the apps are respectful to users with differing opinions and the quality of the app experience is great," followed by the reference to Justice Potter Stewart's much-quoted aphorism that "I'll know it when I see it", for adjudicating "content or behaviour that we believe is over the line". Although the Introduction is the highest-level (least concrete) component of the ASRGs, the stated commitment here to "all points of view being represented" is significant. Crucially, the commitment is to support the representation of all points of view *on the App Store*, not *on any App Store app*; thus, there is no prohibition against specific platforms that may be highly ideologically skewed, as was the case with Parler.

[117] The version of the App Store Guidelines that I quote from in this discussion was archived by the Internet Archive on 30th June 2020, accessible at https://web.archive.org/web/20200630021420/http://developer.apple.com/app-store/review/guidelines/.

Section 1 of the guidelines, entitled "Safety", encompasses several forms of problematic content prohibited in apps. It begins by stating that:

> "When people install an app from the App Store, they want to feel confident that it's safe to do so—that the app doesn't contain upsetting or offensive content, won't damage their device, and isn't likely to cause physical harm from its use. We've outlined the major pitfalls below, but if you're looking to shock and offend people, the App Store isn't the right place for your app."

This preamble is designed to provide context for the more specific prohibitions that follow. Most significant here is the establishment of "offence" as possible grounds for the rejection of an app. This is reiterated in the first subsection of Safety, "1.1: Objectionable Content". The subsection starts with the warning that "apps should not include content that is offensive, insensitive, upsetting, intended to disgust, in exceptionally poor taste, or just plain creepy." Here "offence" reappears as part of a broader category of "insensitive, upsetting, and disgusting". The document goes on to outline six[118] specific examples of objectionable content, of which four are potentially relevant to the case study.[119] First, and most pertinently, the guidelines prohibit:

> "1.1.1: Defamatory, discriminatory, or mean-spirited content, including references or commentary about religion, race, sexual orientation, gender, national/ethnic origin, or other targeted groups, particularly if the app is likely to humiliate, intimidate, or harm a targeted individual or group. Professional political satirists and humorists are generally exempt from this requirement."

Without using the term "hate speech", this prohibition draws on very similar terms to those that are used in definitions of hate speech, including the definitions included in the community standards of mainstream social media platforms (see e.g., Meta 2023). As I argued in Chapter 4, hate speech is intricately bound up with one's affinity with and/or membership of a particular identity group. Moreover, the disallowing of apps "likely to humiliate, intimidate, or harm a

[118] A seventh example was added to the list after 2020, namely "Harmful concepts which capitalize or seek to profit on recent or current events, such as violent conflicts, terrorist attacks, and epidemics" (1.1.7). Though it is chronologically out of scope of this case study it will be interesting to consider the effects of this prohibition as part of ASG going forward.

[119] The fourth example of objectionable content was relevant to Parler, albeit outside the scope of hate speech. This was the prohibition on apps containing "overtly sexual or pornographic material" (1.1.4). Though pornographic material was not part of my text-based sample, contemporaneous media reports suggested that this was a problem for Parler's content moderation during 2020 (Timberg, Harwell, and Lerman 2020). The sixth example concerns apps that include "false information and features", such as trick/joke functionality (1.1.6). I am not aware of any indication Parler hosted or facilitated this form of problematic content at any point.

targeted individual or group" is notable for encompassing the various ways in which online hate speech inflicts harm: "humiliation" could be taken to include instantaneous emotional harm; and "intimidation" could be seen as related to both the longer-lasting effects of emotionally harmful hate speech (e.g., fear of leaving one's house), and the exclusion of diverse voices from the public sphere that can also result. In short, although the term "hate speech" is not employed in this guideline, it can be seen as encompassing it, in terms of the harms-based framework I presented in Chapter 4.

On the basis of this guideline alone, the extent of hate on Parler throughout 2020 would seem from the outside to augur poorly for the platform's prospects of passing through Apple's review process. As I documented in Chapter 6, discriminatory and offensive posts targeting groups on the basis of religion, race, sexual orientation, gender, and national or ethnic origin were commonplace on Parler. Moreover, the phrase "references or commentary" makes clear that it is not just direct, second-person ("you") attacks that are in scope of Apple's prohibition, but also indirect, third-person ("they") attacks as well. In order words, the guideline extends beyond verbal assault or abuse aimed directly at fellow users, which is relevant since it seems that few Parler users were members of commonly targeted groups.

The second form of "objectionable content" cited by the ASRGs is less obviously relevant to hate speech, but implicitly strengthens the prohibition against it. This is the ban on:

> "Realistic portrayals of people or animals being killed, maimed, tortured, or abused, or content that encourages violence. "Enemies" within the context of a game cannot solely target a specific race, culture, real government, corporation, or any other real entity" (1.1.2).

This addresses another form of harm that can result from online hate speech, namely offline violence. Moreover, although this guideline is not outwardly related to violence against specific target groups, judging by the second sentence (and taken together with the previous stipulation), we may assume that the encouragement of violence against groups is not permitted in the App Store tout court, even beyond of gaming apps. As my analysis in Chapter 6 found, content

encouraging violence was commonplace on Parler throughout 2020, with rhetoric advocating for the execution of "traitors" especially prominent.

The third form of objectionable content that Apple prohibits is "depictions that encourage illegal or reckless use of weapons" (1.1.3). This rule is pertinent given the centrality of guns to the iconography and ideology of conservative America, emblemised by the staunch defence of the Second Amendment to the US Constitution. Gun culture was prominent on Parler, and users often touted their gun ownership when making violent threats to "defend" the country against perceived political opponents. For example, in my analysis I observed that a legislative package of gun control measures signed by the Governor of Virginia Ralph Northam in April 2020 (Stracqualursi 2020) was criticised fiercely on Parler. Sometimes, violent references to weaponry were linked to hate speech against target groups, with references to nooses unfortunately very common in this respect.

Finally, the prohibition on "inflammatory religious commentary or inaccurate or misleading quotations of religious texts" (1.1.5) strengthens the broader prohibition on hate speech (1.1.1) on the grounds of religious hatred specifically. Negative "commentary" on holy texts such as the Quran, and the extraction of decontextualised quotations from them, are often used to disparage the religion itself (Poole et al. 2021). In the subset of the sample that I coded as Islamophobic, I found a variety of myths and slurs directed at Islam, such as characterisations of it as satanic, which would fit within any reasonable definition of "inflammatory".

Taken together, the ASRGs pertaining to "objectionable content" in effect during 2020 represented a clear prohibition on the sort of hate speech that was prevalent on Parler, as well as on other forms of objectionable content that are adjacent to or overlapping with hate speech, including the reckless use of weapons which may foment offline violence, and inflammatory religious commentary. Therefore, judging solely on this subset of the guidelines, Parler was demonstrably in violation of these guidelines. It contained an undeniable degree of hate speech, as classified automatically and confirmed manually; moreover, the sample I analysed included examples of each form of objectionable content set out in this section of the ASRGs.

However, this first subset of the Section 1 guidelines on Objectionable Content (1.1) must be set against the second, which concerns apps containing User-Generated Content (1.2). This guideline is as follows:

> "Apps with user-generated content present particular challenges, ranging from intellectual property infringement to anonymous bullying. To prevent abuse, apps with user-generated content or social networking services must include: [120]
>
> 1 A method for filtering objectionable material from being posted to the app
> 2 A mechanism to report offensive content and timely responses to concerns
> 3 The ability to block abusive users from the service
> 4 Published contact information so users can easily reach you
>
> Apps with user-generated content or services that end up being used primarily for pornographic content, Chatroulette-style experiences, objectification of real people (e.g. "hot-or-not" voting), making physical threats, or bullying do not belong on the App Store and may be removed without notice. If your app includes user-generated content from a web-based service, it may display incidental mature "NSFW" content, provided that the content is hidden by default and only displayed when the user turns it on via your website."

Parler was certainly an app "with user-generated content or social networking services",[121] so this subset of the guidelines was applicable to Parler. However, it is less clear how the application of this guideline should be considered in relation to the prohibitions on objectionable content contained in 1.1. On one hand, it could be claimed that the guidelines in 1.2 supersede those in 1.1 entirely. This may be implicitly conveyed via the use of the similar terms "objectionable material" as used in the first bullet point and "offensive content" in the second: problematic content is clearly in scope of both 1.1 and 1.2, and thus, the more nuanced and specific rules regarding how this content should be handled by platforms are applicable. On the other hand, there is not a perfect overlap between the forms of content that are prohibited in 1.1 and those prohibited in 1.2. Some of the specific forms of problematic content mentioned in 1.2, such as "intellectual property infringement" and "anonymous bullying", are not mentioned in 1.1; nor are references to "Chatroulette-style experiences" and "objectification of real people".[122] These "particular

[120] I have enumerated the four requirements, which are bullet points in the ASRGs, for ease of discussion.
[121] In the remainder of this section, I use the term "social media platform" as shorthand for this phrase in Apple's guidelines.
[122] The reference to "hot-or-not" voting may be taken as a slight at the forerunner of the social media platform Facebook, "Facemash" which involved just such a system. Tensions between Apple and Meta persist.

challenges" are specific to social media platforms in the sense that they relate to interaction between individual users. I reflect further on the tension between these two sets of guidelines and their implications for app store governance of hate speech in 7.3.

The first of the four requirements contained in 1.2 mandates "a method for filtering objectionable material from being posted to the app". Interpreted literally, the requirement to stop objectionable material being posted in the first place would seem to necessitate that social media platforms in the App Store adopt ex ante content moderation in the form of, for example, matching with a set of pre-defined content filters; more flexibly, it may also involve systems to detect and remove objectionable content as soon as possible after posting. Effective systems of this sort typically rely on sophisticated algorithmic systems to work at scale (Gorwa et al. 2020), creating a significant burden for start-ups and smaller platforms. There is no evidence to suggest that Parler had a robust automated ex ante system in place (Thiel et al. 2021). Indeed, as we have seen in its Community Guidelines, the company acknowledged that it "relie[d] upon its users to report violations of these Guidelines" to the Community Jury for review (see Chapter 6).[123] Moreover, in terms of Parler's rhetorical portrayal of its platform self-governance, were it to have implemented an ex ante system this would have seriously undermined the company's claim to be a free speech utopia, where nothing is a priori "off limits". Consider that in its Community Guidelines, the company states: "In no case will Parler decide what content will be removed or filtered, or whose account will be removed, on the basis of the opinion expressed within the content at issue."[124] We can therefore conclude that Parler did not meet this specification during 2020.

Parler fares better, superficially, when assessed against Apple's second requirement for social media platforms, namely that platforms institute "a mechanism to report offensive content

[123] https://web.archive.org/web/20200908134955/https://legal.parler.com/documents/guidelines.pdf (italics mine)
[124] https://web.archive.org/web/20200908134955/https://legal.parler.com/documents/guidelines.pdf

and timely responses to concerns". This did exist on Parler: the platform's Community Jury procedure was based on reports of allegedly objectionable content from users. However, it is unlikely that any "responses" to reports were timely or effective. The Stanford Internet Observatory report on Parler identified an enormous backlog of cases for the overwhelmed community jury, estimating that Parler had only 802 moderators, the majority volunteers, for its 13 million user membership in January 2021. Furthermore, as active site participants, moderators *themselves* shared content that would have likely been removed from other platforms (Thiel et al. 2021).

The third requirement that platforms have an "ability to block abusive users from the service" is ambiguously worded, as it may either demand that platforms have this ability (a function more commonly referred to as banning), or alternatively that users themselves can block other, abusive users. If the former is true, Parler evidently had the ability to ban users—but it lacked the willingness to do so, given that only 0.15% of users in the sample were banned as of January 2021. On the user level, the platform did have a function allowing users to block each other. With respect to the fourth requirement, Parler did include an email address and mailing address in its Privacy Policy, which was linked to on its homepage, if users "have any questions, comments, or concerns".[125]

Finally, the Legal section (Section 5) of the ASRGs implores developers to ensure that their apps "comply with all legal requirements in any location where you make them available" and states that "apps that solicit, promote, or encourage criminal or clearly reckless behavior will be rejected". This provision can be seen as something of a catch-all, ensuring that legal content and activity is outlawed by definition—and, as a corollary, implying that all the above provisions on objectionable content are extra-legal. This section also contains various specific requirements

[125] Admittedly, this concerned Parler's data processing activities rather than its content moderation, but the fact that contact details including an email address and mailbox were accessible from its website probably satisfied this requirement.

regarding the safeguarding of user privacy (5.1) with respect to data collection and use, and intellectual property (5.2). Evidence of Parler's lax approach to data security had already emerged during 2020 (Hesse 2020; Taylor 2020) and this was further underlined after the Capitol riot and Parler's shuttering, when what Wired called an "absurdly basic bug let anyone grab all of Parler's data" (Greenberg 2021).

This analysis suggests that Parler fell short of meeting several of Apple's guidelines with respect to hate speech and its moderation. Before exploring what the apparent non-enforcement of these guidelines tells us about Apple's app store governance of hate speech, I discuss the equivalent rules that Google had in force at the time.

7.2.2 Google's Play Store Developer Policies

Google's guidelines for developers submitting apps and app updates to its Play Store, contained in its Play Developer Policy Center, adopt a nested structure similar to Apple's App Store guidelines.[126] The section of the Google Play Developer Policies that is relevant to hate speech on Parler is "Restricted Content". Each section features a question designed to summarise the purpose of the policy; for "Restricted Content", the question posed uses the by-now (painfully) familiar formulation: "does your app cross the line?". Like Apple's ASRGs, this section includes two relevant subsections: "Inappropriate Content"[127] and "User Generated Content".

Unlike Apple, in its guidelines on "Inappropriate Content", Google explicitly identifies "hate speech" as something that it prohibits. Under this heading, it states that:

> "We don't allow apps that promote violence, or incite hatred against individuals or groups based on race or ethnic origin, religion, disability, age, nationality, veteran status, sexual orientation, gender, gender identity, or any other characteristic that is associated with systemic discrimination or marginalization."

[126] https://web.archive.org/web/20200704134247/https://play.google.com/about/developer-content-policy/

[127] The version of Google's Play Developer Policies on Restricted Content that I quote from in this discussion was archived by the Internet Archive on August 3, 2020, and is accessible at https://web.archive.org/web/20200803163910/https://support.google.com/googleplay/android-developer/answer/9878810.

This description of hate speech fits neatly with the explication of the harms of hate speech presented in Chapter 4. It clearly captures the risk of offline violence that online hate speech can provoke, and by linking hate speech to the membership of groups that have experienced systematic discrimination or marginalisation, it acknowledges the role that historical inequities and asymmetric power relations play in shaping the persistent harm that hate speech can still inflict.

The guidelines provide several "examples of common violations" in this area. These include:

1. "Content or speech asserting that a protected group is inhuman, inferior or worthy of being hated.
2. Apps that contain hateful slurs, stereotypes, or theories about a protected group possessing negative characteristics (e.g. malicious, corrupt, evil, etc.), or explicitly or implicitly claims the group is a threat.
3. Content or speech trying to encourage others to believe that people should be hated or discriminated against because they are a member of a protected group.
4. Content which promotes hate symbols such as flags, symbols, insignias, paraphernalia or behaviors associated with hate groups."

This list represents an expansive range of prohibitions against different forms of hate speech. It reflects the common understanding that hate speech involves the expression of something negative and encompasses not only cases in which people may be targeted directly by hate speech, but also includes attempts to build hateful "coalitions" with like-minded individuals. And the inclusion of "hate symbols" captures the growing use of memes and other visual imagery to express hate in a subtle or non-verbal way.

I found each of these examples of hate speech in my sample of Parler posts. Many of the most egregious examples of hate speech I found involved the characterisation of identity groups as subhuman and civilisationally "backward" or "evil" (the latter was especially often asserted with respect to Muslims) and cast their mere presence as a threat to the mores of American society. However, because my sample was text-based, I did not encounter hate "symbols" as described in the fourth example. An open question to which I return below concerns the extent to which Parler was an app which "*promoted*" such material, as opposed to merely tolerating it.

Outside of the Hate Speech subsection itself, a second component of the Inappropriate Content subsection was also relevant to hate speech. Under the subheading of Violence, Google

states that "we don't allow apps that depict or facilitate gratuitous violence or other dangerous activities", including "graphic depictions or descriptions of realistic violence or violent threats to any person or animal." As we have seen, violent rhetoric, especially political violence, was also very prominent on Parler, some of which also constituted hate speech.

Much like Apple, Google had a separate subset of guidelines for apps that feature user-generated content, i.e., social media platforms.[128] UGC is defined as "content that users contribute to an app, and which is visible to or accessible by at least a subset of the app's users". This definition, while short, is notably more specific than Apple's in one important way. It posits, without specifying quantitatively, the notion of a threshold at which an app would qualify as featuring UGC and thus subject to additional requirements. This clause affords Google a lot of leeway, for instance to decide whether a messaging app that allows small group chats would be in scope. But for present purposes, the phrase "at least a subset of the app's users" indicates that even small, fringe platforms like Parler were in scope of these rules.

As with Apple, Google's guidelines for social media platforms contain several requirements with respect to user-generated content. The guidelines state that apps that allow it must:

> 1 "Require that users accept the app's terms of use and/or user policy before users can create or upload UGC;
> 2 define, in a manner consistent with the spirit of Google Play's Developer Program Policies, UGC that is objectionable, and prohibit that UGC via the app's terms of use and/or user policy;
> 3 implement robust, effective and ongoing UGC moderation, as is reasonable and consistent with the type(s) of UGC hosted by the app;
> 4 provide a user-friendly, in-app system for reporting and removal of objectionable UGC;
> 5 In the case of live-streaming apps, problematic UGC must be removed in as close to real-time as reasonably possible; and
> 6 remove or block abusive users who violate the app's terms of use and/or user policy
> 7 provide safeguards to prevent in-app monetization from encouraging objectionable user behavior."
>
> Apps whose primary purpose is featuring objectionable UGC will be removed from Google Play. Similarly, apps that end up being used primarily for hosting objectionable UGC, or that develop a

[128]https://web.archive.org/web/20200802054536/https://support.google.com/googleplay/android-developer/answer/9876937

> reputation among users of being a place where such content thrives, will also be removed from Google Play."

The fourth and sixth requirements are very similar to stipulations contained in Apple's guidelines, and thus the discussion of these in 7.2.1 applies equally here to Google. The fifth is out of scope as Parler did not have a live-streaming function. With respect to the seventh requirement, although Parler had a network for "influencers" to post sponsored content, it had no in-app monetization for ordinary users. The discussion therefore focuses on the first three requirements.

Parler required that users accept its Privacy Policy, Terms of Service, and Community Guidelines when signing up (see Chapter 5), thus meeting the first requirement. The second requirement superficially resembles Apple's stipulation that platforms must prevent "objectionable material" from being uploaded to their app. However, Google's rules go further by explicitly specifying that apps' definition of objectionable content and behaviour must comply with *Google's own rules* on such content, or at any rate "with the spirit of them", as outlined above. This represents an explicit avowal by Google of the "meta-moderation" role discussed in Part I. Furthermore, the third requirement, for platforms to "implement robust, effective and ongoing" content moderation, is more stringent in its specificity than Apple's equivalent. Taken together, with these three requirements Google establishes an indirect governance link not merely to the operators, but also to the users, of every social media platform in its app store.

One would expect prima facie that these requirements would have proved a problem for Parler in 2020. Not only did Parler contain hate speech at a roughly comparable level to that contained on fringe bedfellow Gab, the company's at-best sputtering attempts to moderate this sort of content on its platform were far from "robust, effective [or] ongoing". It is therefore instructive at this juncture to return to Google's removal of Gab in August 2017. Announcing that it had been removed, Gab cited a violation of Google's hate speech policy as the reason (Coldewey 2017). Google's statement explaining the decision, reported by TechCrunch, was as follows:

> "In order to be on the Play Store, social networking apps need to demonstrate a sufficient level of moderation, including for content that encourages violence and advocates hate against groups of

> people. This is a long-standing rule and clearly stated in our developer policies. Developers always have the opportunity to appeal a suspension and may have their apps reinstated if they've addressed the policy violations and are compliant with our Developer Program Policies" (reported in Coldewey 2017).

Google's decision regarding Gab[129] demonstrated that the company was willing to reject reasonably highly populated and functional social media platforms on the principal bases of the prevalence of hate speech and the insufficient moderation of it. And to a greater degree than Apple, Google's guidelines make clear that it expects platforms in its app store to follow at least the spirit, if not the letter, of its own definition and prohibition of content it deems objectionable, which includes hate speech.

It is clear from this analysis that Parler was evidently in violation of multiple provisions contained in both Apple's and Google's guidelines. So why was it not duly sanctioned? There is surely no "single reason" why Parler was not removed before January 2021, and thus this discussion is speculative; nonetheless, several factors may have played a role. One reason may be the discursive legitimation of Parler's platform self-governance identified in Chapter 6. Parler's focus on free speech, along with its invocation of other constitutional metaphors, shaped perception of the platform, even as its lived reality was very different from the utopian imagery the company used to characterise it. In doing so, Parler went much further than its fellow fringe platforms to "sanitise" its offering to mainstream conservative users. This portrayal may have helped dissuade sanction from app store operators prior to January 2021. Another factor is the lack of precision in app store guidelines with respect to hate speech and its relationship to rules on user-generated content. This ambiguity afforded Apple and Google considerable leeway to avoid sanctioning Parler during 2020—as a corollary to the swiftness with which they were able to act after the Capitol riot (see below). A third and final factor pertains to the political environment. The decision to remove a high-profile app—particularly one that was at or close to the top of their

[129] Gab was never allowed on to Apple's App Store, as it was rejected at the submission stage due to the pornography on the platform (Rodriguez 2016); evidently Apple knew it when it saw it.

own app store charts—would have brought greater attention to the powerful role that both companies play in app store governance and smartphone software. It is thus not surprising in retrospect that it took the explosive events at the Capitol, and the drastically different political environment that this created, to force the two companies to act. Crucially, however, this was not simply a case of Apple and Google "changing their mind when the facts change", to appropriate a quote attributed to John Maynard Keynes. My analysis shows that on the contrary, the "facts" of Parler were clear long before January 2021. Although the presidential election and subsequent controversy inevitably sparked greater foment on Parler, the platform had long been a hotbed of hate speech and insurrectionist content. The only change was the greater exposure to Parler that the Capitol riot brought upon it. In the final section, I turn to consider the circumstances surrounding Parler's removal and eventual reinstatement more closely.

7.3 Better late than (n)ever? Parler's removal from app stores

On 8th January 2021, Apple sent Parler a detailed letter, obtained by Buzzfeed News (Mac and Paczkowski 2021). The letter gave Parler 24 hours to submit a "moderation improvement plan" or face removal from the App Store. It began:

> "We require your immediate attention regarding serious App Store guideline violations that we have found with your app, Parler. We have received numerous complaints regarding objectionable content in your Parler service, accusations that the Parler app was used to plan, coordinate, and facilitate the illegal activities in Washington D.C. on January 6, 2021."

Apple went on to note that in its own investigation, it had "found that Parler is not effectively moderating and removing content that encourages illegal activity and poses a serious risk to the health and safety of users in direct violation of your own terms of service". This was accompanied both by a link to Parler's Elaboration on Guidelines document (see Chapter 6), and explicit references to two of Apple's own ASG guidelines. With respect to its Objectionable Content guideline, Apple claimed that "we found that your app includes content that some users may find upsetting, offensive, or otherwise objectionable. Specifically, we found direct threats of violence and calls to incite lawless action." With respect to User Generated Content guideline, Apple

claimed that "your app enables the display of user-generated content but does not have sufficient precautions in place to effectively manage objectionable content present in your app."

That Parler was in violation of Apple's rules as of January 2021 comes as no surprise given the findings presented in this chapter, which show that Parler had a significant amount of objectionable content including hate speech, and did little to moderate it, on the platform throughout 2020. However, it is unclear how much of a surprise it was to Apple—whose letter, by way of evidence, included links to tweets from apparently random Twitter users documenting complaints about problematic Parler posts, as well as a quote from Matze disclaiming responsibility for content on Parler and reiterating that "we're a neutral town square that just adheres to the law". Neither possibility casts Apple in a favourable light. It would be extraordinary to think that Apple could have been unaware of the level of objectionable content—not just "threats of violence", but also the documented plethora of hate speech—on the platform, nor of Parler's loud embrace of "free speech" as a basis for its platform self-governance. Recall that Parler had already topped Apple's own App Store charts in November; compared to mainstream rivals it was still a small platform, but hardly an unknown one. But nor does the other possibility—that Apple was well-aware of Parler's violations, yet only acted after they came into widespread view following an unprecedented attack at the heart of American democracy—reflect well on its enforcement of its own app store governance guidelines or its reckoning with their implications, given its stated commitment to keeping the app store safe.

Also on January 8th, Google announced, in a statement shared with reporters, that it was removing Parler from its Play Store (Peters 2021):

> "In order to protect user safety on Google Play, our longstanding policies require that apps displaying user-generated content have moderation policies and enforcement that removes egregious content like posts that incite violence. All developers agree to these terms and we have reminded Parler of this clear policy in recent months. We're aware of continued posting in the Parler app that seeks to incite ongoing violence in the U.S.. We recognize that there can be reasonable debate about content policies and that it can be difficult for apps to immediately remove all violative content, but for us to distribute an app through Google Play, we do require that apps implement robust moderation for egregious content. In light of this ongoing and urgent public

> safety threat, we are suspending the app's listings from the Play Store until it addresses these issues".

Justifying its removal of Parler, Google invoked several aspects of its guidelines discussed in 7.2. It notes the requirement for "robust" content moderation, which it claims that Parler lacked. The immediate context of the Capitol riot, which the statement characterises as an "ongoing and urgent public safety threat", is clear. Unlike Apple, the statement acknowledges communication between Google and Parler "in recent months" regarding egregious content, suggesting that there was an attempt to negotiate with Parler prior to the events of January 2021 which brought the platform into public focus.

Further information about both companies' decisions emerged in the following days. In an interview with Fox News anchor Chris Wallace on 15 January 2021, Apple CEO Tim Cook expanded on his company's decision to remove Parler. Cook told Wallace that "we looked at the incitement of violence that was on there and we don't consider that free speech and incitement of violence has an intersection" (Choudhary 2021). It is notable here that the "free speech" frame for Parler, discursively advanced by Matze throughout 2020, persisted beyond the platform's removal. And although, as we have seen, Fox News had been especially receptive to Matze's "free speech" discourse, mainstream news organisations such as the BBC also described Parler as a "'free speech' app" when reporting its removal, albeit putting "free speech" in speech marks (Anon 2021). Meanwhile, a report the following week in Business Insider shed light on Google's decision (Langley 2021). Parler's Chief Operating Officer Jeffrey Wernick confirmed the "reminders" that Parler had received from Google in the months leading up to its removal, and detailed a call with himself, Parler's policy officer Amy Peikoff, and Google employees. Wernick claimed that Google asked to postpone the call to a later date due to a technical issue, and that "it was clear that they didn't want to talk to us" (Langley 2021).

Parler was eventually restored to both app stores, and the explanations given for the reversal are also instructive. In May 2021, Parler returned to Apple's App Store following "months

of productive dialogue with Apple". The decision was announced by Apple in April 2021 in response to conservative lawmakers who had criticised its removal. Apple cited the "volume and types of prohibited content" in Parler as motivating its removal, but noted "substantial conversations with Parler in an effort to bring it into compliance with the Guidelines" (Powderly 2023). This included a claimed improvement to Parler's content moderation system, which Parler itself claimed incorporated algorithmic technology to detect violating posts (Lyons 2021b). This in effect created a two-tier system for content on Parler. Then, in September 2022, Parler was also restored to Google's Play Store, with little explanation provided by Google (Hatmaker 2022).

It is perhaps tempting to view Apple's and Google's removal of Parler in January 2021 as a case of "better late than never". However, its welcome parsimony notwithstanding, this formulation can be critiqued from two rather orthogonal angles. First, as the contemporary explanations and justifications make clear, Parler was not removed *because of* hate speech as such. Instead, the form of objectionable content that saw Parler ultimately sanctioned was described as "direct threats of violence and calls to incite lawless action" by Apple, and as "posts that incite violence" by Google. As my analysis found, in practice there was considerable overlap between posts on Parler expressing hate speech and posts calling for or instigating violence. This point of overlap can be seen as constituting what Susan Benesch (2012) has called "dangerous speech". However, as I detailed in Chapter 4, physical violence is not the only form of harm that hate speech can perpetrate; hate speech can also inflict emotional harm and discriminatory disadvantage on those directly targeted by it, as well as harm democratic society through the exclusion of diverse voices. This suggests that, despite the expansive wording of their app store guidelines, in practice Apple and Google are oriented towards preventing physical violence from hate speech above the other harms it can inflict. In this light, the removal of Parler in January 2021 may not necessarily be seen as even a belated assumption of the two app store operators' full responsibility they accord themselves to govern hate speech.

Second, and from an altogether different perspective, one can use the Parler case to interrogate the nature of the responsibility itself. The stewardship responsibility that both companies adopt reflects a balancing act between, on one side, exerting almost panoptical oversight of the content of any social media platform that seeks to enter their app store—raising significant privacy and free speech concerns—and on the other, tacitly permitting their smartphone software to be used by any community, for any purpose, that is not prohibited by law. It is this second line of critique that I take forward in the concluding chapter, wherein I explore alternative approaches to the app store governance of hate speech.

Summary

The second research question asks to what extent Apple and Google met their responsibility for governing hate speech in the case of Parler. Because of the meta-moderation model discussed in Chapter 3, this question had two components: how Parler performed its platform self-governance (RQ2a); and whether, as a result, Apple and Google ought to have acted to remove Parler sooner than they did (RQ2b). The analysis undertaken in the case study suggests that in the case of Parler, both companies fell short in meeting the responsibility that they accord themselves in their app store guidelines, raising wider questions about the consistency and efficacy of their app store governance. Parler was eventually removed from both app stores, but only amid a political crisis that directly implicated it, wherein the prevalence of hate speech was a secondary consideration to the level of violent conduct.

In this conclusion to the case study, I turn from assessing how Apple and Google performed the governance responsibility that they accord to themselves as app store operators, to consider the broader societal implications of the fact that Apple and Google have this role at all. The retention of Parler in app stores through 2020—despite the platform's apparent contravention of Apple's and Google's guidelines regarding hate speech—illuminates several of the conceptual limitations of app store governance that I first indicated in Chapter 3. Three are most consequential

and form the basis for the remainder of the discussion. These are Apple's and Google's duopolistic dominance over the smartphone market, and the powerful position this gives them in app store governance; the substantive ambiguity of their app store guidelines; and the procedural arbitrariness and unaccountability of the application of these guidelines to specific cases.

First, when Apple and Google both removed Parler from their respective app stores within hours of one another, it served as a rare but striking reminder of how much control the companies exert over their smartphone platforms. In his interview with Fox News following Parler's removal, Tim Cook claimed that "we obviously don't control what's on the internet, but we've never viewed that our platform should be a simple replication of the internet. We have rules and regulations, and we just ask that people abide by those" (Choudhary 2021). Yet this downplays the control that Apple (and Google) exert over their smartphone software, as well as the two companies' shared dominance over the smartphone market that their duopoly enables. Even if they did not directly coordinate their decision in this case (as they did in the development of contact tracing apps: Sharon 2020), the two companies' shared dominance means that harmful apps can be withdrawn from the app stores of a large proportion of the world's smartphones in an instant. There are times when such a governance action may be justified, and indeed Parler may have been one of those cases. Yet Apple and Google cannot be relied upon to be sufficiently "enlightened" in any, let alone every, given case (Benedek and Kettemann 2020:146)—whether to adjudicate competing claims regarding free expression and public safety as in this case, regarding the demands of data privacy and public health surveillance as in the case of Covid-19 contact tracing (Morley et al. 2020), or in myriad other controversies yet to arise. And as my analysis has found, a reasoned decision about Parler could, in any case, have been made sooner, with respect to the companies' own rules, given the extent of hate speech on the platform through 2020. As such, the Parler case hardly inspires confidence in how well Apple and Google carry out the "stewardship" role that they rhetorically accord themselves.

The second limitation springs from the first. The case study set out to explore the apparent contradiction between the two companies' guidelines on objectionable content on one hand and user-generated content on the other, as these are applied to social media platforms. In the case of Parler, both companies seem to have taken a hands-off approach until it was too late. The case study has thus exposed the strategic ambiguity at the heart of the meta-moderation role that Apple and Google adopt with respect to social media platforms in their app stores, which can be discursively reconfigured depending on the companies' interests in any given case.

The third limitation concerns the lack of procedural oversight over the decisions that the two companies make regarding apps in their app stores. App store governance has long been, and remains, a black box—and the Parler case highlights the social costs of this procedural impenetrability. Hitherto, Apple and Google have remained largely immune to demands akin to those placed on social media platforms to make their governance decisions and processes, if not participatory, then at least visible, as Meta has moved some way towards with its Oversight Board. At least in my own experience in this case study, this made investigating their decisions feel somewhat like runology. Moreover, Parler was a high-profile platform with 13 million users, among them leading US politicians. This may have obliged Apple and Google to provide a higher degree of explanation than is the case for other app removals.

In Part III, I build on the findings and insights of the case study to explore potential policy responses to the limitations of app store governance that the case study has illuminated.

Part III: Implications, reflections, and conclusions

Chapter 8: Policy implications and impacts

Introduction

In this concluding chapter, I draw the thesis to a close by using insights from the case study to inform a broader policy discussion about the responsibility of app store operators to govern hate speech and how this may be enhanced and enforced by recent and forthcoming regulatory changes. To restate, the case study elaborated three limitations of the app store governance of hate speech that were first identified in Part I, namely the dominance of Apple and Google, ambiguity about the application of different guidelines on hate speech on social media, and the lack of procedural oversight of app store governance decision-making. In this concluding chapter I turn attention to the third and final research question, which is as follows:

RQ3: What additional considerations should shape the app store governance of hate speech on social media?

First, a brief synopsis of the argument so far is provided (8.1). I then use lessons from the case of Parler's removal from app stores to engage with the normative implications of the ambiguity and opacity of app store governance. App store operators hold a curious position with respect to online speech governance: they can pressure platforms to conform with their view of acceptable content moderation, and sanction those that fail to toe the line, yet there is no clear form of redress for individuals indirectly affected by the effects of this standard-setting. In this way, answering Keller's pointed (2019) question of "who do you sue?" in cases of content removal is even more complicated to answer when the actor responsible for the removal is not a platform but an app store (8.2). I then explore new legislation that may affect app store governance. Apple's dominance over its iOS platform may soon wane, with new legislation forcing it to open its smartphone software to allow the use of third-party app stores, among other legislative provisions that may impact on ASG. I draw the discussion together to address RQ3. I find that proposed legislation, especially the EU's Digital Markets Act and Digital Services Act, in Europe, seem likely

to address many of the limitations of ASG identified in the case study, yet they leave some questions unresolved (8.3). Finally, directions for future research are identified, the limitations of the study are considered, and closing remarks are offered (8.4).

8.1 Synopsis

In this case study I studied Parler: a fringe social media platform which, during 2020, was both the agent and the subject of platform governance. Parler adopted and sought to discursively legitimate what it claimed was a constitutionally inspired approach to platform self-governance, to attract users disillusioned with the restrictions on harmful content adopted by larger rival platforms. Parler thus permitted a range of problematic material, including hate speech, that is typically prohibited by other platforms, and adopted a content moderation regime that was demonstrably inadequate to the scale of its 13 million users.

Because it had a smartphone app, Parler was the subject to the rules and procedures of the smartphone app stores operated by Apple and Google, as a form of cross-platform governance. As self-appointed stewards of their respective app stores, Apple and Google accord themselves a responsibility to constrain the extent and reach of objectionable material, including hate speech, in platforms that their app stores host. Their guidelines pertained both to the extent of hate speech on Parler, and to the efficacy of Parler's content moderation regime. The eventual decision by Apple and Google to suspend Parler from their app stores in January 2021 was a rare, significant demonstration of their app store governance power. It seems to have been justified in the narrow sense of Apple's and Google's app store review guidelines; in fact, in terms of their own guidelines, the two platforms may even have acted too late, in view of the level of hate speech that persisted on the platform long before January 2021 and Parler's refusal to meaningfully address it.

However, the decision to remove Parler also raises weighty questions about the nature, scope, and legitimacy of app store governance. As the self-styled stewards of their respective app stores, Apple and Google justify the control that they exert over app store governance on the

grounds of ensuring the safety and wellbeing of smartphone users. But the case of Parler suggests that this model is flawed. Very little of the app store review process is public, besides the vague guidelines that both companies provide; this opacity affords the two companies significant leeway for flexible interpretation and implementation. Given the widespread adoption of smartphones and the significant role that they play in modern life—enabling everything from the quotidian to the revolutionary—the lack of public oversight over app store guidelines, how apps are pressured to conform to them, and the removal of those that refuse, seems to be an oversight in itself.

8.2 Who do you sue now? The platform power of app store operators

The case study highlighted the shortcomings of both Parler's self-governance, and of the cross-platform governance efforts of app store operators Apple and Google, with respect to the hate speech that prevailed on Parler throughout 2020. In view of the harms of hate speech identified in Chapter 4, this shared governance failure raises questions about the harm that fringe platforms like Parler can inflict at multiple levels if left unchecked. Yet the episode also complicates the application of the question posed by Daphne Keller in 2019, which asks who users can sue when a platform removes their content, to cases in which an app store operator removes a platform wholesale.

Apple's removal of Parler from the App Store meant that the platform could not be used as an app on iOS devices, while Google's removal of Parler from its Play Store made it significantly more difficult for users to find, download, and use Parler as an app on their Android devices. Having assessed the justifiability of Parler's removal in terms of these companies' own rules in Chapter 7, we can turn to consider its implications in terms of the rights of users affected by these and similar decisions. These two considerations are not easily separable. In *Who Do You Sue?*, Daphne Keller observes that "speakers in the United States have few or no legal rights when [social media] platforms take down their posts," and that, in cases where platforms remove content, even when platforms may appear to be using their own discretion, removal decisions "are often

profoundly influenced by governments" (Keller 2019:2; see also Leerssen 2015). This "influence" takes two forms: (i) initiatives like the agreement between leading social media platforms and the European Commission to remove hate speech, as well as laws like NetzDG through which states oblige platforms to "do their work for them" (Balkin 2018:2016); and (ii) a more subtle form of influence that Keller calls platforms' "anticipatory obedience" to governments' wishes, which "spares governments the need to enact actual laws" (2019:2). At least in the US, where courts have been very reluctant to force platforms to restore content (the situation is more complicated in other jurisdictions such as Germany: Kettemann and Tiedeke 2020), this leaves those affected by content removals with little redress. Therefore, while states can and in some cases have delegated responsibilities to private platform companies as speech governors, these platform companies do not have the traditional responsibilities and oversight that attach to state governance of speech (Keller 2019; Klonick and Kadri 2019). The "triangular" free speech framework, consisting of states, platforms, and individuals, is thus more complex than traditional "old-school" speech regulation (Balkin 2004, 2018; Griffin 2022). At least in the American judicial context, as it stands, legal tools to challenge both "state and private platform power … are underdeveloped—and in some cases may simply not be possible to devise" (Keller 2019:27).

The situation becomes more complicated still in the case of cross-platform governance, wherein one platform company, such as Apple or Google, delegates content moderation to the platform companies that they host in their app stores, as a necessary condition for their inclusion and retention. This raises the rather complex prospect of an implicit "double delegation": *from* states *to* app store operators, and then *from* app store operators *to* social media platforms.[130] A further complicating factor is the limited set of concrete governance tools that app store operators have at their disposal with respect to violative platforms that they host. The main tool ASOs have is the ability to remove an app in its entirety from their app stores; as Hofmann et al. observe,

[130] And this is not to mention even further delegation within the architecture of a single platform, such as to moderators of specific groups or communities: Gillett and Suzor 2022; Hasinoff and Schneider 2022).

more "targeted intervention against individual illegal content is not possible" (2022:3). That said, ASOs do have some "softer" tools at their disposal. They can try to negotiate with platforms, as they did with Parler (though the process seems to have broken down prior to Parler's removal), and they can downrank apps in their app store search and "top charts" architecture (though this would not have prevented Parler users navigating from a direct link, e.g., on its website or another platform, to reach the platform's page in the app stores).

The nature of ASG—which we may summarise as "speak vaguely and carry a big stick"—implicates other rights, in addition to freedom of expression. Undoubtedly, removing a platform such as Parler wholesale adversely affects the free speech rights of its users, insofar as they can no longer use a particular smartphone app to communicate on a particular social media platform. But as they host only the app, not the content, app stores cannot remove content from the internet as such. Nor can they prevent individuals from "speaking their mind" on other websites and platforms—indeed they not know who these individual users are. This highlights a curious paradox that is analogous to that which arises from the delegation of moderation to social media platforms by states observed by Keller. When platforms remove content at the (implicit or explicit) behest of states, litigators "may see no room to object and no one to sue when private companies silence online speakers" (2019:27). In the same way, if a platform removes content, or even bans a user, to toe the line of an app store operator's (vague) guidelines, then the user in question may not have an obvious path to redress. It seems that such a case arose with Telegram's banning of the channels of the conspiracy theorist Attila Hildmann in 2021, "presumably due to pressure from the app stores" (Hofmann et al. 2022:3). This case (in which a platform was forced to act based on the threat of removal) taken together with the case of Parler (in which a non-compliant platform was removed) underscores the "extensive freedom" that ASOs have typically enjoyed (Hofmann et al. 2022:9) in deciding when to deem that a social media platform has failed in its delegated moderation role. This may, however, be set to change, in at least some jurisdictions.

8.3 The impact of new legislation on app store governance

Several laws that are either under consideration or that have recently been passed may come to affect the dynamics of app store governance as they applied in the case of Parler (Cowls, Morley, and Floridi 2023; Hofmann et al. 2022; Kollnig and Shadbolt 2023). Most prominently, this includes the European Union's Digital Markets Act (DMA) and Digital Services Act (DSA), which were announced in December 2020 as part of a package designed to make "Europe fit for the digital age" by creating "new online rules for platforms" (Turillazzi et al. 2022). The two laws address different aspects of platform power, and both hold implications for app store governance. In view of the so-called "Brussels effect" (Bradford 2020), which sees global corporations comply with European regulation not just in Europe but wherever their services are offered, it is reasonable to think that the DMA and DSA package may encourage wider changes in app store governance tout court. Indeed, it seems that Apple is preparing for the new laws by allowing sideloading in iOS 17, its next iPhone software release (Gurman 2022). Here I assess the expected impact of each law on app store governance.

8.3.1 The Digital Markets Act

The DMA is designed to ensure that what it calls "gatekeeper platforms … behave in a fair way online" (European Commission 2023a). To do this, it enumerates a set of "core platform services" (CPS), which includes "online intermediation services" (European Parliament legislative resolution on the Digital Markets Act, 2022 (hereafter "DMA"); art 2.2.a) such as "software application stores" (DMA art. 2.12). It then establishes the new category of "gatekeeper", which may be applied to operators of core platform services if they meet a set of criteria. This includes having a sizeable impact on the EU market and an entrenched and durable intermediation position, such that they control, through their CPS(s), a gateway between other businesses (e.g., app developers) and users (e.g., device owners). Based on these criteria, it seems clear that Apple and Google as companies, and specifically their app stores as core platform services, fall under the regulation

(Hofmann et al. 2022).[131] As a result, several provisions will apply to app stores, including a prohibition on app store operators preventing users from subscribing to services or purchasing content for their device from outside the app store (DMA art. 5.c) and a requirement to apply fair and non-discriminatory conditions of access to app stores for developers (DMA art. 6.1.k).

The provision of the DMA that will have the most impact on ASG is the provision that will force Apple and Google to "allow the installation and effective use of third-party software applications or software application stores using, or interoperating with, operating systems of that gatekeeper and allow these software applications or software application stores to be accessed by means other than the core platform services of that gatekeeper" (DMA art. 6.1.c). This would, in effect, amount to the requirement that the practice of sideloading apps be permitted on smartphone platforms, which would force a change in Apple's smartphone ecosystem. However, it is important to note two caveats. First, this is contained in the so-called "grey list" of "obligations … susceptible of being further specified"—likely, given Apple and Google's size and dominance, on the basis of bilateral dialogue with the Commission (Grünwald et al. 2022). Second, the article also notes that "the gatekeeper shall not be prevented from taking proportionate measures to ensure that third-party [apps and app stores] do not endanger the integrity of the hardware or operating system provided by the gatekeeper" (DMA art. 6.1.c). These caveats initially cast some doubt on whether sideloading would be enforced, but a European Commission spokesperson confirmed that under the DMA users would be "able to opt for alternative sources of apps on [their] smartphone" and that "the DMA would allow a smartphone owner to also opt for other *safe* app stores" (Porter and Vincent 2022; italics added).

Sideloading would mark a major blow to Apple's control over its smartphone ecosystem, but the example of its rival Google, which already permits sideloading on Android devices, is instructive. As Hofmann et al (2022) observe, the market share of Google's Play Store remains

[131] The provisions of the DMA apply to the specific CPS for which a company has dominance, not to all the company's platforms.

very high, and Google has taken a forceful approach to removing third-party app stores that it designates, often without evidence, as unsafe (see Chapter 3). Nonetheless, the DMA—alongside the proposed Open App Markets Act in the US, which would also enforce sideloading (Cowls, Morley, and Floridi 2023)—would transform app store governance. New app store operators could feasibly adopt different (tighter or looser) standards for platforms that they host. It is at least plausible that an alternative app store could adopt laxer standards for hate speech and other objectionable content such that platforms like Gab or Parler could be allowed. However, reports claim that Apple is considering mandating certain security requirements even for apps outside the app store, and even charging a fee for verifying these apps (Gurman 2022). This suggests that Apple is searching for ways to maintain the large "safety premium" that it currently attaches to its current in-app-store review processes.

8.3.2 The Digital Services Act

The DSA, by contrast, focuses on "ensuring a safe and accountable online environment" (European Commission 2023b). Unlike the DMA's dichotomic "gatekeeper" assignation, the DSA establishes a series of cumulative obligations that scale based on a platform operator's role, size, and impact in the online ecosystem. There are four levels of obligation, which apply in ascending order to what are called intermediary services (e.g., domain name registrars), hosting services, online platforms, and very large online platforms. App stores are explicitly identified as constituting "online platform"—the second most onerous designation. However, the "very large" designation is applicable to any "online platform" that reaches more than 45 million people in Europe on a monthly basis (European Parliament legislative resolution on the Digital Services Act, 2022; (hereafter "DSA"); art 33.1); therefore, it can be assumed that both Apple's App Store and Google's Play Store will receive this most burdensome designation (Hofmann et al. 2022; Kollnig and Shadbolt 2023). The DSA contains a range of measures aimed at greater transparency, oversight and safety concerning the operation of online platforms. It includes obligations for platform operators of all sizes to remove illegal content when they are notified that they are hosting

it (DSA art. 9); creates a procedure for reporting content that platform operators remove, whether on the basis of illegality or due to a platform's own terms of service (DSA art. 15); creates a procedure for handling complaints about such removals (DSA art. 20); and creates a process for out-of-court dispute settlement on such complaints (DSA art. 21). In view of the limitations of ASG that I have identified in this thesis, the most impactful provisions specifically for app store governance relate to transparency and reporting requirements, and systemic risk assessment.

The DSA seems to promise much regarding the transparency and reporting obligations that ASOs will face. Articles 15 and 24 together create a regime for "transparency reporting obligations" which will oblige ASOs to provide information about measures taken against content, which should be distinguished between content removed for illegality versus that removed owing to violation of the ASO's own guidelines. At a high level this will systematise ASG, obliging reports of the sort that mainstream platform operators like Meta have long provided. In the specific case of Parler, this would have standardised the format in which Apple and Google reported publicly, and to Parler, the reasons why the platform was removed.

In addition, the DSA obliges platforms to set up a process through which they can be notified for "specific items of information that [an outside party] considers to be illegal content". What constitutes illegal content at the level of an app remains unclear (Cowls et al. 2023), but at a minimum, it seems that Apple and Google will be obliged to receive and consider complaints about illegal content on apps. This in turn raises the potential for abuse and dispute resolution. The DSA contains safeguards against "manifestly unfounded complaints" about illegal content (DSA art 23). It also obliges providers to establish an "effective complaint-handling system" to complain about app removals. Though the app store operators already have such systems, third-party developers note that they do not resolve issues effectively (Hofmann et al. 2022). Taken together these processes will go some way to forcing open the black box of app store governance.

Furthermore, the DSA obliges very large online platforms to assess whether there are any systematic risks "stemming from the design or functioning of their service and its related systems

… or from the use made of their services" (DSA art 34). In response to such risks, platform operators are encouraged to consider adapting the "design, features or functions of their services", their "terms and conditions and their enforcement", and their "content moderation processes" (DSA art 35). This latter point measure is explicitly taken to include, "where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation." This obliges ASOs to consider and seek to mitigate illegal hate speech was present in apps that their app stores host (though note that this does not cover legal but potentially harmful material).

Finally, ASOs are obliged to "provide access to data to vetted researchers … for the sole purpose of conducting research that contributions to the detection, identification and understanding of systemic risks" (DSA art 40). This includes "any actual or foreseeable negative effects for the exercise of fundamental rights" in what Kollnig and Shadbolt (2023:23) have called a "very broad definition" of actual and potential risks. It may be hoped that this opens up ASG to academic scrutiny, which, as my own efforts here have shown, is currently hamstrung by a lack of available data access. However, uncertainty and oversights remain (Leerssen, Heldt, and Kettemann 2023).

It will be interesting to see how the provisions contained in the DMA regarding sideloading interact with those in the DSA just discussed. Apple and Google clearly fall within the scope of the DSA given their size—but this may not be the case for smaller third-party providers as new entrants into app store governance, who may therefore not be subject to the DSA's steepest obligations. Counterintuitively, therefore the two acts taken together may expand the scope of ASG to encompass other actors, but not accord all the new responsibilities to these new actors.

8.3.3 Other efforts

Similar if smaller-scale efforts related to app stores are under consideration in other jurisdictions, such as Australia, whose Competition and Consumer Commission has mooted "upfront rules and

regulations" to open up Apple and Google's smartphone ecosystems (J. Taylor 2021), and South Korea, whose Telecommunications Business Act was amended in 2021 to force Apple and Google to allow alternative payment systems (Clark and Porter 2021). In addition, the Open App Markets Act (OAMA) remains under consideration by the United States Senate, having passed through the Senate Judiciary Committee with a rare bipartisan 20–2 vote during the last congressional session, in February 2022 (MacCarthy 2022). Compared with the EU DMA and DSA package, the OAMA aims solely at Apple and Google's control over their respective app stores and the economic consequences of this dominance for consumers and developers. The current version of the bill defines an app store as "a publicly available website, software application, or other electronic service that distributes apps from third-party developers to users of a computer, a mobile device, or any other general purpose computing device" (Blumenthal 2021:2.2), and introduces several constraints on their management. As with the EU's DMA, the OAMA is generally economic in nature. It contains several concrete constraints on Apple and Google's conduct as app store operators, including prohibitions on requiring developers to use their in-app payment systems, and penalizing developers for offering lower prices elsewhere or for making these offers known (OAMA sec 3.a). It also imposes new obligations on Apple and Google, including allowing users to "install third-party apps or app stores through means other than its app store" (OAMA sec. 3.d.2), thus, like the DMA, enforcing sideloading. However, since the OAMA is a standalone bill not part of a broader legislative package, it does not offer any of the provisions in the EU's companion Digital *Services* Act which may introduce new obligations not merely on how apps are found and bought, but also how they are governed.

Despite some areas that are yet to be specified (Cowls et al. 2023; Hofmann et al. 2022), it is clear that the changes to ASG wrought by the European DMA/DSA package are significant. Specifically, they address several of the limitations of ASG identified in Chapters 3 and 4 and explicated in the case study. The DMA (and OAMA in the US, if passed) will oblige Apple to allow

new app store operators, loosening the company's stranglehold on its smartphone software. The DSA will oblige ASOs to make available considerably more information about how they undertake app store governance, at the level of specific decisions, through regular high-level reports, and even to vetted researchers for the examination of systemic risks. It will also help level the playing field between ASOs and app developers in the case of disputes and encourage ASOs to optimise for the expeditious removal of illegal hate speech in Europe. Together, it seems that these laws will have meaningful impact on the key problems with app store governance that the thesis has identified: the duopolistic dominance of Apple and Google over app store governance; the ambiguity regarding how the companies apply rules on objectionable content versus user-generated content on platforms; and the broader lack of procedural oversight of and outside involvement in app store governance decision-making.

However, despite the progress made by these laws, some challenges may go unresolved, or could even be exacerbated. First, it remains to be seen whether, in its forced abandonment of the "hard" technical approach to preventing third party app stores, Apple will simply turn to Google's "soft" approach, which uses software signals and settings to strongly dissuade rather than prevent users from using alternative app stores, and removes third party app stores on "safety" grounds, as in the case of Aptoide. The DMA may empower regulators to take further steps against Google's (and if necessary, Apple's) practices in this regard. Second, the arrival of new actors into the domain of app store governance, including some who may not be subject to the new European rules on account of their size, may fragment app store governance. Just as the past few years have seen a rise of fully-fledged platforms designed for the "alt-right", such as Parler, it is not beyond the realm of possibility that similar actors will now be empowered to create their own app *stores*, with greater tolerance for objectionable material than Apple and Google.

8.4 Reflections, limitations, and directions for future research

The contributions of the thesis can be contextualised and situated through wider reflections, acknowledgements of the limitations of the research, and identifications of directions for future research in the area.

As the recency of the references—both scholarly and journalistic—cited in this thesis attest, many of the topics covered are the subject of growing interest and attention. At the legislative level, it is heartening to see the adoption of legislation that may force improvement in several of the limitations of ASG that my thesis has identified. In view of the "Brussels effect" (Bradford 2020), the European legislation may be expected to do for digital platforms what the GDPR did for data protection—not just impose concrete obligations on companies, but spark awareness and interest in the broader issues that these laws seek to tackle. Legislative interest in app store governance is a welcome if belated addition to more established debates over how platform companies govern the content and activity of their own users.

The increased interest in the topics I cover in this thesis was also sparked by developments in the social media landscape during the writing of this thesis, such as Donald Trump's launch of his Truth Social platform in 2021 (Cowls 2022) and Elon Musk's purchase of Twitter in 2022.[132] There has also been growing attention paid to what might be called "general-ish" platforms: those that ostensibly cater to a wide audience, but which have found popularity among specific communities, some of which have proved problematic in the view of authorities. This includes Telegram, which Germany recently mooted banning outright (Meaker 2022) and which removed certain conspiracist channels seemingly in response to app store pressure (Hofmann et al. 2022). It also includes video game communication platform Discord, which at the time of writing faces pressure in the US after a large leak of classified military documents was released on the site (Ganguly 2023). Other decentralised platforms like Mastodon have been touted as more reliable

[132] For a moment in 2022, it even seemed that Kanye West would buy Parler itself, which would have completed a curious trifecta of rich, politically controversial figures launching or purchasing "boutique" social media platforms and catering them to specific political audiences or viewpoints.

and sustainable alternatives to Twitter (Nicholas 2023). This flourishing of alternative platforms to longstanding incumbents like Facebook and Twitter will only bring greater scrutiny to the stack providers, including app store operators, who provide essential infrastructure to them.

However, from another vantage point, it could be argued that this shift in focus to stack governance is a case of "old wine in new bottles". Whether it is the operators of social media platforms or the providers of tech stack infrastructure who are making decisions about online speech, in both instances, it is private companies that are making governance decisions that affect society at large. Private platform companies can hardly be relied upon, by themselves, to make such socially weighty decisions when their incentives may lead in directions that are orthogonal, if not entirely antagonistic, to the public interest. Yet neither does the idea of vesting governments with greater control over online speech have much appeal at a time when even the world's great democracies are falling victim to populist and nativist politics. In short, the central challenge at the heart of online speech governance—who gets to decide?—remains frustratingly difficult to answer.

It is important to acknowledge the limitations of this thesis. However interesting and important, my case study covered the governance of just one platform, Parler, and thus the applicability of the findings I obtained are limited in the extent of their generalisability to other cases. Relatedly, my positionality as a researcher must be acknowledged, in terms of the topic I chose, the data I gathered, the analyses I conducted, and findings I derived. This latter point is especially important in the context of interpreting and classifying hate speech, where one's (lack of) lived experience of discrimination and marginalisation inevitably affects how they experience exposure to hate speech, even at a distance from the circumstances of its original expression. This may thus limit the internal validity of the findings as they relate to hate speech on Parler. Finally, the case study inevitably suffered from missing data. As a retrospective study, I faced challenges in gathering data to reconstruct as accurate as possible a portrait of Parler during 2020. The key actors in the study, such as app store operators, are notoriously difficult to gain information from

(Kollnig and Shadbolt 2023), thus limiting me in some cases to gathering and interpreting found data.

Each of these epistemological limitations—generalisability, internal validity, and missingness—are real and important. In mitigation I would only humbly note that each limitation is rooted in epistemological conceptions and structural conditions that go far beyond the scope of this work—and that these are conditions that I hope this work might play a small part in countering. Focusing on Parler as just one platform may have limited the generalisability of the study in terms of app store governance, but as I hope the case study demonstrated, single platforms (large or small) are important areas of study in their own right. Attending to hate speech was limited by my positionality, but part of my intention in doing so was partly to highlight the (methodological and political) challenge of determining what is and is not hate speech for the purposes of platform governance. And the lack of available information I encountered about how app store governance decisions are made has served only to sharpen my critique of the ambiguity of app store governance guidelines and the opacity regarding their implementation, informing my recommendations in this regard.

Several avenues for potential future research in this space can be identified. One direction involves rights other than freedom of expression that are impacted by platform governance in the expanded form that I have conceptualised it here. For example, a focus on freedom of assembly and association as complementary rights to freedom of expression may together serve to better defend online speech. In the case of the removal of platforms from app stores, it is not clear that it is individuals' freedom of speech that has necessarily been infringed, as much as it is their freedom to assemble in a particular (online) space, and to associate with like-minded others. The removal of a platform does not prevent individuals from "speaking their mind" elsewhere online, but it can disrupt or even destroy the space in which they have chosen to participate and the relationships they have founded. As with any right, freedom of assembly and association are not absolute; and in fact, in many rights documents, the right to freely assembly is (in contrast to

freedom of expression) limited by the condition of "peaceability" (Inazu 2013). This may provide a useful precedent for the authorities who determine the limits of this right; the relevance of this to Parler's decidedly unpeaceful assembly is clear.

Second, the entrance of new governors into the space of app store governance is likely to create fruitful avenues for research. Whether these new governors share Apple's and Google's app store governance approach will be interesting to witness, especially with respect to whether different app store governance approaches serve as a competitive advantage to these new entrants. For example, would a platform—for the sake of example, Donald Trump's Truth Social—removed from a mainstream app store find a safe harbour elsewhere, in an alternative app store? More generally, it will be interesting to examine the effects of the loosening control of the rules of app store governance as well as the revenue that singular app stores provide.

Third, the release of new and different digital devices may further transform app store governance. Companies like Meta, Sony, and Microsoft have already released (expensive) augmented/virtual reality (AR/VR) devices, and Apple is readying its own. If these devices achieve mainstream adoption this may cement a new class of app store operators with control over the distribution of apps. While the smartphone market is now quite mature, the introduction of new kinds of digital devices and the distinct affordances they bring may exacerbate digital divides affecting participation on future social media platforms.

Finally, the apparent fracturing of the social media landscape on geopolitical and ideological lines is one of the most interesting developments yet at the research front of platform governance. At the time of writing, several western countries are considering following India in banning TikTok (Brandom and Christopher 2023), which, if enforced, would be an extraordinary statement of the power of states over social media and demonstrate that the expression of such power is not limited to authoritarian states. Furthermore, ideological splintering, both in terms of out-and-out ideological platforms like Truth Social, and what I have called "general-ish" platforms

like Telegram, will also be worth tracing, especially with respect to how these platforms seek to legitimate their self-governance approach.

8.5 Closing remarks

This thesis considered the overarching question of how app stores should responsibly govern hate speech on social media platforms. Part I found that as app store operators, Apple and Google accord themselves a responsibility to govern hate speech as the "stewards" of their respective app stores. However, the case study on Parler in Part II highlighted the limitations of this responsibility in terms of the substantive and procedural opacity of app store governance. The policy analysis in Part III found that new legislation in Europe (potentially, via the Brussels effect, influencing other jurisdictions) will oblige greater openness and oversight over how this important form of cross-platform governance is undertaken.

One difficult question that this thesis has highlighted but not resolved concerns the question of who gets to decide what can and cannot be said online. As the Parler case demonstrates, fundamentally political decisions are often made by the private companies that provide infrastructure to social media platforms. As such, the thesis captured a snapshot in time when a set of stack operators collectively made a significant intervention in online speech which had the effect of taking a social media platform offline. As forthcoming legislation opens up the market for app stores and necessitates more transparency on the part of incumbent operators, it is possible that this decision-making power may begin to fade, in turn supplanted by other actors operating other "hidden levers" (DeNardis 2012). In this context, the characterisation of internet governance as a "difficult horse to catch" (Ziewitz and Pentzold 2014:306) seems especially accurate, and an appropriate note on which to end this thesis. As digital technology, society, and politics continue to co-evolve and reshape one another, the only prediction that seems safe to make is that the question of power will be pivotal to future academic investigations in this domain; a closing conviction to state is that it ought to be.

References

AAP. 2019. '4chan, Liveleak Blocked by Australian Internet Providers'. Retrieved 2 April 2023 (https://www.9news.com.au/national/telcos-block-access-to-4chan-liveleak/128de87c-0fba-4bac-9e14-d3a1cfbbc6d4).

Abbott, Kenneth W., and Duncan Snidal. 2009. 'Strengthening International Regulation through Transmittal New Governance: Overcoming the Orchestration Deficit'. *Vand. J. Transnat'l L.* 42:501.

Action Aid. 2020. '$2.8bn "Tax Gap" Exposed by ActionAid Research Reveals Tip of the Iceberg of "Big Tech's Big Tax Bill" in the Global South'. *ActionAid International*. Retrieved 5 May 2021 (https://actionaid.org/news/2020/28bn-tax-gap-exposed-actionaid-research-reveals-tip-iceberg-big-techs-big-tax-bill-global).

ADL. 2022. 'Online Hate and Harassment: The American Experience 2022 | ADL'. Retrieved 31 March 2023 (https://www.adl.org/resources/report/online-hate-and-harassment-american-experience-2022).

Agarwal, Saharsh, Uttara M. Ananthakrishnan, and Catherine E. Tucker. 2022. 'Deplatforming and the Control of Misinformation: Evidence from Parler'.

Aliapoulios, Max, Emmi Bevensee, Jeremy Blackburn, Barry Bradlyn, Emiliano De Cristofaro, Gianluca Stringhini, and Savvas Zannettou. 2021. 'A Large Open Dataset from the Parler Social Network'. Pp. 943–51 in *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 15.

Aliapoulios, Max, Emmi Bevensee, Jeremy Blackburn, Emiliano De Cristofaro, Gianluca Stringhini, and Savvas Zannettou. 2021. 'An Early Look at the Parler Online Social Network'. *ArXiv:2101.03820 [Physics]*.

Alkiviadou, Natalie. 2019. 'Hate Speech on Social Media Networks: Towards a Regulatory Framework?' *Information & Communications Technology Law* 28(1):19–35.

Alphabet. 2023. 'Alphabet SEC Filing'. Retrieved 30 March 2023 (https://abc.xyz/investor/static/pdf/20230203_alphabet_10K.pdf?cache=5ae4398).

Anderson, Benedict. 2006. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Verso books.

Anderson, Kathryn Freeman. 2013. 'Diagnosing Discrimination: Stress from Perceived Racism and the Mental and Physical Health Effects*'. *Sociological Inquiry* 83(1):55–81. doi: 10.1111/j.1475-682X.2012.00433.x.

Anon. 2021. 'Google Suspends "free Speech" App Parler'. *BBC News*, January 9.

Ansell, Chris, and Alison Gash. 2008. 'Collaborative Governance in Theory and Practice'. *Journal of Public Administration Research and Theory* 18(4):543–71.

Apperley, Thomas, and Jussi Parikka. 2018. 'Platform Studies' Epistemic Threshold'. *Games and Culture* 13(4):349–69.

Apple. 2008. 'IPhone to Support Third-Party Web 2.0 Applications'. Retrieved 10 February 2022 (https://web.archive.org/web/20081215230338/http://www.apple.com/pr/library/2007/06/11iphone.html).

Apple. 2021. *Building a Trusted Ecosystem for Millions of Apps: A Threat Analysis of Sideloading*.

Apple. 2022. 'App Store Review Guidelines'. *Apple Developer*. Retrieved 14 March 2023 (https://developer.apple.com/app-store/review/guidelines/).

Apple. 2023. 'Apple Reports First Quarter Results'. *Apple Newsroom*. Retrieved 30 March 2023 (https://www.apple.com/newsroom/2023/02/apple-reports-first-quarter-results/).

Arun, Chinmayi. 2021. *Facebook's Faces. SSRN Scholarly Paper*. ID 3805210. Rochester, NY: Social Science Research Network. doi: 10.2139/ssrn.3805210.

Aswad, Evelyn. 2016. 'The Role of US Technology Companies as Enforcers of Europe's New Internet Hate Speech Ban'. *HRLR Online* 1:1.

Awan, Imran, and Irene Zempi. 2016. 'The Affinity between Online and Offline Anti-Muslim Hate Crime: Dynamics and Impacts'. *Aggression and Violent Behavior* 27:1–8.

Bacon, Perry. 2018. 'The Republican Party Has Changed Dramatically Since George H.W. Bush Ran It'. *FiveThirtyEight*. Retrieved 10 April 2023 (https://fivethirtyeight.com/features/the-republican-party-has-changed-dramatically-since-george-h-w-bush-ran-it/).

Baines, Annalise, Muhammad Ittefaq, and Mauryne Abwao. 2021. '# Scamdemic,# Plandemic, Or# Scaredemic: What Parler Social Media Platform Tells Us about COVID-19 Vaccine'. *Vaccines* 9(5):421.

Balkin, Jack M. 2004. 'Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society'. *New York University Law Review* 79:1.

Balkin, Jack M. 2014. 'Old School/New School Speech Regulation'.

Balkin, Jack M. 2017. 'Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation'. *UCDL Rev.* 51:1149.

Balkin, Jack M. 2018. 'Free Speech Is a Triangle'. *Colum. L. Rev.* 118:2011.

Barlow, John Perry. 2016. 'A Declaration of the Independence of Cyberspace'. *Electronic Frontier Foundation*. Retrieved 9 August 2019 (https://www.eff.org/cyberspace-independence).

Barrett, Paul M., and J. Grant Sims. 2021. *False Accusation: The Unfounded Claim That Social Media Companies Censor Conservatives*. Technical Report. NYU Stern Center for Business and Human Rights. https ….

Bauer, Johannes M., and Tiago S. Prado. 2020. *Digital Platforms and Innovation: Lessons for Innovation Policy and Regulation. SSRN Scholarly Paper*. ID 3749670. Rochester, NY: Social Science Research Network. doi: 10.2139/ssrn.3749670.

BBC. 2020. 'Trump Covid Post Deleted by Facebook and Hidden by Twitter'. *BBC News*, October 6.

Beckert, Jens. 2016. *Imagined Futures: Fictional Expectations and Capitalist Dynamics*. Harvard University Press.

Beetham, David. 2013. *The Legitimation of Power*. Macmillan International Higher Education.

Belk, Russell, and Gülnur Tumbat. 2005. 'The Cult of Macintosh'. *Consumption Markets & Culture* 8(3):205–17.

Ben-David, Anat, and Adam Amram. 2018. 'The Internet Archive and the Socio-Technical Construction of Historical Facts'. *Internet Histories* 2(1–2):179–201. doi: 10.1080/24701475.2018.1455412.

Benedek, Wolfgang, and Matthias C. Kettemann. 2020. *Freedom of Expression and the Internet: Updated and Revised 2nd Edition*. Council of Europe.

Benesch, Susan. 2012. 'Dangerous Speech: A Proposal to Prevent Group Violence'. *Voices That Poison: Dangerous Speech Project*.

Benesch, Susan. 2020. 'But Facebook's Not a Country: How to Interpret Human Rights Law for Social Media Companies'. *JREG Bulletin* 38:86.

Benikova, Darina, Michael Wojatzki, and Torsten Zesch. 2018. 'What Does This Imply? Examining the Impact of Implicitness on the Perception of Hate Speech'. Pp. 171–79 in *Language Technologies for the Challenges of the Digital Age*, *Lecture Notes in Computer Science*, edited by G. Rehm and T. Declerck. Cham: Springer International Publishing.

Bennett, Dalton, Emma Brown, Atthar Mirza, Sarah Cahlan, Joyce Sohyun Lee, Meg Kelly, Elyse Samuels, and Jon Swaine. 2021. '41 Minutes of Fear: A Video Timeline from inside the Capitol Siege'. *Washington Post*, January 16.

Benveniste, Alexis, and Kaya Yurieff. 2020. 'Meet Rebekah Mercer, the Deep-Pocketed Co-Founder of Parler, a Controversial Conservative Social Network'. *CNN*. Retrieved 10 May 2022 (https://www.cnn.com/2020/11/15/media/rebekah-mercer-parler/index.html).

Berbrier, Mitch. 2000. 'The Victim Ideology of White Supremacists and White Separatists in the United States'. *Sociological Focus* 33(2):175–91.

Beres, Damon. 2021. 'Apple's Walled Garden Is Becoming Harder to Escape'. *Slate*, June 8.

Berry, Robert James. 2006. *Environmental Stewardship*. A&C Black.

Bietti, Elettra. 2020. 'From Ethics Washing to Ethics Bashing: A View on Tech Ethics from within Moral Philosophy'. Pp. 210–19 in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, *FAT* '20*. New York, NY, USA: Association for Computing Machinery.

Bilewicz, Micha\l, and Wiktor Soral. 2020. 'Hate Speech Epidemic. The Dynamic Effects of Derogatory Language on Intergroup Relations and Political Radicalization'. *Political Psychology* 41:3–33.

Binder, Matt. 2023. 'Conservative Social Media Platform Parler Acquired and Then Immediately Shut down by New Owner'. *Mashable*. Retrieved 18 April 2023 (https://mashable.com/article/parler-acquired-shut-down).

Binns, Reuben, Michael Veale, Max Van Kleek, and Nigel Shadbolt. 2017. 'Like Trainer, like Bot? Inheritance of Bias in Algorithmic Content Moderation'. Pp. 405–15 in *International Conference on Social Informatics*. Springer.

Black, Julia. 2008. 'Constructing and Contesting Legitimacy and Accountability in Polycentric Regulatory Regimes'. *Regulation & Governance* 2(2):137–64.

Blake, Aaron. 2022. 'Analysis | "Let's Have Trial by Combat": How Trump and Allies Egged on the Violent Scenes Wednesday'. *Washington Post*, January 31.

Blake, Andrew. 2020. 'Parler CEO Says Up-and-Coming Social Media Service Will Not Ban Users for Hate Speech'. *Washington Times*. Retrieved 23 March 2023 (https://www.washingtontimes.com/news/2020/aug/5/parler-ceo-says-up-and-coming-social-media-service/).

Blank, Grant. 2017. 'The Digital Divide among Twitter Users and Its Implications for Social Research'. *Social Science Computer Review* 35(6):679–97.

Bloch-Wehba, Hannah. 2019. 'Global Platform Governance: Private Power in the Shadow of the State'. *SMU Law Review* 72(1):27.

Blumenthal, Richard. 2021. *S.2710 - Open App Markets Act*.

Bodansky, Daniel. 1999. 'The Legitimacy of International Governance: A Coming Challenge for International Environmental Law?' *American Journal of International Law* 596–624.

Bogost, Ian, and Nick Montfort. 2009. 'Platform Studies: Frequently Questioned Answers'.

Bohn, Dieter. 2021. 'Apple Isn't Just a Walled Garden, It's a Carrier'. *The Verge*. Retrieved 15 February 2023 (https://www.theverge.com/2021/6/7/22521476/apple-walled-garden-carrier-app-store-innovation).

Bolinger, Renée Jorgensen. 2017. 'The Pragmatics of Slurs'. *Noûs* 51(3):439–62.

Boyd, Kenneth. 2020. 'Parler and the Problems of a "Free Speech" Social Network'.

Bradford, Anu. 2020. *The Brussels Effect: How the European Union Rules the World*. Oxford University Press.

Brandom, Russell, and Nilesh Christopher. 2023. 'What the U.S. Can Learn from India's TikTok Ban'. *Rest of World*. Retrieved 15 April 2023 (https://restofworld.org/2023/america-india-tiktok-ban/).

Brodkin, Jon. 2018. 'EU: Google Illegally Used Android to Dominate Search, Must Pay $5B Fine'. *Ars Technica*. Retrieved 11 February 2022 (https://arstechnica.com/tech-policy/2018/07/eu-google-illegally-used-android-to-dominate-search-must-pay-5b-fine/).

Bronstein, Carolyn. 2021. 'Deplatforming Sexual Speech in the Age of FOSTA/SESTA'. *Porn Studies* 8(4):367–80. doi: 10.1080/23268743.2021.1993972.

Brown, Abram. 2020. 'Parler's Founder Explains Why He Built Trump's New Favorite Social Media App'. *Forbes*. Retrieved 26 May 2022

(https://www.forbes.com/sites/abrambrown/2020/06/27/parlers-founder-explains-why-he-built-trumps-new-favorite-social-media-app/).

Brown, Alexander. 2018. 'What Is so Special about Online (as Compared to Offline) Hate Speech?' *Ethnicities* 18(3):297–326.

Brügger, Niels. 2018. 'Web History and Social Media'. *The SAGE Handbook of Social Media* 196–212.

Brügger, Niels, Gerard Goggin, Ian Milligan, and Valérie Schafer. 2017. 'Introduction: Internet Histories'. *Internet Histories* 1(1–2):1–7. doi: 10.1080/24701475.2017.1317128.

Brustein, Joshua. 2020. 'Social Media App Parler, a GOP Darling, Isn't Catching On'. *Bloomberg.Com*, July 15.

Buchanan, Allen, and Robert O. Keohane. 2006. 'The Legitimacy of Global Governance Institutions'. *Ethics & International Affairs* 20(4):405–37.

Bucher, Taina. 2012. 'A Technicity of Attention: How Software'makes Sense". *Culture Machine* 13.

Bucher, Taina, and Anne Helmond. 2018. 'The Affordances of Social Media Platforms'. *The SAGE Handbook of Social Media* 1:233–54.

Buntain, Cody, Martin Innes, Tamar Mitts, and Jacob Shapiro. 2023. 'Cross-Platform Reactions to the Post-January 6 Deplatforming'. *Journal of Quantitative Description: Digital Media* 3.

Busch, Christoph. 2022. 'Regulating the Expanding Content Moderation Universe: A European Perspective on Infrastructure Moderation'. *UCLA JL & Tech.* 27:32.

Büthe, Tim, and Walter Mattli. 2013. *The New Global Rulers: The Privatization of Regulation in the World Economy*. Princeton University Press.

Butler, Judith. 1997. *Excitable Speech: A Politics of the Performative*. routledge.

Cadwalladr, Carole, and Emma Graham-Harrison. 2018. 'Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach'. *The Guardian* 17:22.

Cameron, Dell, and Dhruv Mehrotra. 2021. 'Leaked Parler Data Points to Users at Police Stations, U.S. Military Bases'. *Gizmodo*. Retrieved 6 April 2023 (https://gizmodo.com/leaked-parler-data-points-to-users-at-police-stations-1846059897).

Campbell, Donald T., and Julian C. Stanley. 2015. *Experimental and Quasi-Experimental Designs for Research*. Ravenio books.

Carroll, Roy. 2020. 'NBA Players Protest Racial Injustice as League Returns to Action'. *Reuters*, July 30.

Cath, Corinne. 2021. 'The Technology We Choose to Create: Human Rights Advocacy in the Internet Engineering Task Force'. *Telecommunications Policy* 45(6):102144. doi: 10.1016/j.telpol.2021.102144.

Cath, Corinne, and Jenna Ruddock. 2022. 'One Year After the Storming of the US Capitol, What Have We Learned About Content Moderation Through Internet Infrastructure?' *Tech Policy*

Press. Retrieved 23 February 2022 (https://techpolicy.press/one-year-after-the-storming-of-the-us-capitol-what-have-we-learned-about-content-moderation-through-internet-infrastructure/).

Catipon, Regina, and Onur Savas. 2021. 'A Different Story: How Conservative Narratives Diverge Between Twitter and Parler.' Pp. 15–22 in *Text2Story@ ECIR*.

Celeste, Edoardo. 2019. 'Terms of Service and Bills of Rights: New Mechanisms of Constitutionalisation in the Social Media Environment?' *International Review of Law, Computers & Technology* 33(2):122–38. doi: 10.1080/13600869.2018.1475898.

Cepollaro, B., and D. Zeman. 2020. 'Non-Derogatory Uses of Slurs [Special Issue]'. *Grazer Philosophische Studien* 97.

Chang, Yu-Wei, Ping-Yu Hsu, Shih-Hsiang Huang, and Jiahe Chen. 2019. 'Determinants of Switching Intention to Cloud Computing in Large Enterprises'. *Data Technologies and Applications* 54(1):16–33. doi: 10.1108/DTA-12-2018-0104.

Chen, Ashley Gold, Shawna. 2021. 'Google Suspends Parler from App Store after Deadly U.S. Capitol Violence'. *Axios*. Retrieved 30 January 2023 (https://www.axios.com/2021/01/09/capitol-mob-parler-google-ban).

Choudhary, Vidhi. 2021. 'Apple CEO Tim Cook Says Parler Suspension and Free Speech Unrelated'. *TheStreet*. Retrieved 13 April 2023 (https://www.thestreet.com/investing/apple-chief-tim-cook-says-parler-suspension-and-free-speech-unrelated).

Christou, George, and Seamus Simpson. 2009. 'New Governance, the Internet, and Country Code Top-Level Domains in Europe'. *Governance* 22(4):599–624.

Citron, Danielle Keats. 2014. *Hate Crimes in Cyberspace*. Harvard University Press.

Citron, Danielle Keats, and Helen Norton. 2011. 'Intermediaries and Hate Speech: Fostering Digital Citizenship for Our Information Age'. *BUL Rev.* 91:1435.

Clark, Mitchell, and Jon Porter. 2021. 'Apple and Google Must Allow Developers to Use Other Payment Systems, New Korean Law Declares'. *The Verge*. Retrieved 23 September 2022 (https://www.theverge.com/2021/8/31/22643800/apple-google-south-korea-app-store-payment-legislation-passes).

Coldewey, Devin. 2017. 'Alt-Social Network Gab Booted from Google Play Store for Hate Speech'. *TechCrunch*. Retrieved 12 April 2023 (https://techcrunch.com/2017/08/17/alt-social-network-gab-booted-from-google-play-store-for-hate-speech/).

Cooper, Charles. 2012. 'Apple, Google Remove Hezbollah TV App'. *CNET*. Retrieved 13 July 2022 (https://www.cnet.com/tech/tech-industry/apple-google-remove-hezbollah-tv-app/).

Courty, Audrey. 2020. 'Parler: What You Need to Know about the "free Speech" Twitter Alternative'. *The Conversation*. Retrieved 26 May 2022 (http://theconversation.com/parler-what-you-need-to-know-about-the-free-speech-twitter-alternative-142268).

Cowls, Josh. 2022. 'Trump Central? App Stores as a New Front in the Platform Governance of Donald Trump'. Retrieved 29 March 2023 (https://www.oii.ox.ac.uk/news-events/news/trump-central-app-stores-as-a-new-front-in-the-platform-governance-of-donald-trump).

Cowls, Josh, Philipp Darius, Dominiquo Santistevan, and Moritz Schramm. 2021. 'Constitutional Metaphors: Facebook's "Supreme Court" and Platform Legitimation'.

Cowls, Josh, Philipp Darius, Dominiquo Santistevan, and Moritz Schramm. 2022. 'Constitutional Metaphors: Facebook's "Supreme Court" and the Legitimation of Platform Governance'. *New Media & Society* 14614448221085560. doi: 10.1177/14614448221085559.

Cowls, Josh, and Cindy Ma. 2023. 'The Internet's Town Square? A Critical Discourse Analysis of Parler's "Past Imaginary"'. *Social Media + Society* 9(2):20563051231181100. doi: 10.1177/20563051231181099.

Cowls, Josh, and Jessica Morley. 2022. 'App Store Governance: The Implications and Limitations of Duopolistic Dominance'. Pp. 75–92 in *The 2021 Yearbook of the Digital Ethics Lab*. Springer.

Cowls, Josh, Jessica Morley, and Luciano Floridi. 2023. 'App Store Governance: Implications, Limitations, and Regulatory Responses'. *Telecommunications Policy* 47(1):102460.

Cowls, Josh, and Ralph Schroeder. 2018. 'Tweeting All the Way to the White House'. *Trump and the Media* 151–57.

Cowls, Josh, and Bertie Vidgen. Under Review. *Platform Governance Terminologies: Online Hate Speech*.

Crenshaw, Kimberle. 1990. 'Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color'. *Stan. L. Rev.* 43:1241.

Culliford, Elizabeth, and Katie Paul. 2019. 'Unhappy with Twitter, Thousands of Saudis Join pro-Trump Social Network Parler'. *Reuters*, June 13.

Culpepper, Pepper D., and Kathleen Thelen. 2020. 'Are We All Amazon Primed? Consumers and the Politics of Platform Power'. *Comparative Political Studies* 53(2):288–318. doi: 10.1177/0010414019852687.

Dagoula, Chrysi. 2019. 'Mapping Political Discussions on Twitter: Where the Elites Remain Elites'. *Media and Communication* 7(1):225–34.

Daniels, Jessie. 2009. *Cyber Racism: White Supremacy Online and the New Attack on Civil Rights*. Rowman & Littlefield Publishers.

Daniels, Jessie. 2013. 'Race and Racism in Internet Studies: A Review and Critique'. *New Media & Society* 15(5):695–719.

Daniels, Jessie. 2018. 'The Algorithmic Rise of the "Alt-Right"'. *Contexts* 17(1):60–65. doi: 10.1177/1536504218766547.

Dasgupta, Nabarun. 2021. 'Countering Hesitancy and Misinformation on Side Effects to Complete the Course of COVID Vaccination'. *Journal of Patient Experience* 8:23743735211067310.

Dave, Paresh. 2022. 'PayPal Shuts down Its Services in Russia Citing Ukraine Aggression'. *Reuters*, March 5.

Deibert, Ronald, ed. 2008. *Access Denied: The Practice and Policy of Global Internet Filtering*. Cambridge, Mass: MIT Press.

Deibert, Ronald, John Palfrey, Rafal Rohozinski, and Jonathan Zittrain. 2008. *Access Denied: The Practice and Policy of Global Internet Filtering*. The MIT Press.

Delgado, Richard. 1982. 'Words That Wound: A Tort Action for Racial Insults, Epithets, and Name-Calling'. *Harv. CR-CLL Rev.* 17:133.

DeNardis, Laura. 2012. 'Hidden Levers of Internet Control'. *Information, Communication & Society* 15(5):720–38. doi: 10.1080/1369118X.2012.659199.

DeNardis, Laura, and Samantha Bradshaw. 2022. 'Internet Infrastructure as an Emerging Terrain of Disinformation'. *Centre for International Governance Innovation*. Retrieved 22 July 2022 (https://www.cigionline.org/articles/internet-infrastructure-as-an-emerging-terrain-of-disinformation/).

DeNardis, Laura, and Andrea M. Hackl. 2015. 'Internet Governance by Social Media Platforms'. *Telecommunications Policy* 39(9):761–70.

DeNardis, Laura, and Francesca Musiani. 2014. 'Governance by Infrastructure: Introduction,"The Turn to Infrastructure in Internet Governance". *The Turn to Infrastructure in Internet Governance'(September 15, 2014)*.

Dickter, Cheryl L. 2012. 'Confronting Hate: Heterosexuals' Responses to Anti-Gay Comments'. *Journal of Homosexuality* 59(8):1113–30.

Dieter, Michael, Carolin Gerlitz, Anne Helmond, Nathaniel Tkacz, Fernando N. van der Vlist, and Esther Weltevrede. 2019. 'Multi-Situated App Studies: Methods and Propositions'. *Social Media + Society* 5(2):2056305119846486. doi: 10.1177/2056305119846486.

van Dijck, J., T. Poell, and M. De Waal. 2018. *The Platform Society: Public Values in a Connective World*. Oxford, New York: Oxford University Press.

Dillet, Romain. 2022. 'Apple to Charge 27% Fee for Dutch Dating Apps Using Alternative Payment Options'. *TechCrunch*. Retrieved 30 March 2023 (https://techcrunch.com/2022/02/04/apple-to-charge-27-fee-for-dutch-dating-apps-using-alternative-payment-options/).

DiStefano, Mark, Alex Spence, and Ryan Mac. 2019. 'Pro-Trump Activists Are Boosting A Twitter App Used By Banned Personalities And It Appears To Have Already Stalled'. *BuzzFeed*. Retrieved 3 April 2023 (https://www.buzzfeed.com/markdistefano/pro-trump-activists-are-boosting-a-twitter-app-for-banned).

Ditch the Label. 2021. 'Uncovered: Online Hate Speech in the Covid Era'. *Ditch the Label*. Retrieved 16 May 2022 (https://www.ditchthelabel.org/research-papers/hate-speech-report-2021/).

Dolgov, Anna. 2015. 'Paypal Blocks Russian Account Linked to Nemtsov Report on Ukraine'. *The Moscow Times*. Retrieved 9 June 2022

(https://www.themoscowtimes.com/2015/05/15/paypal-blocks-russian-account-linked-to-nemtsov-report-on-ukraine-a46583).

donk_enby. 2020. 'Parler Tricks'.

Donovan, Joan. 2019. 'Navigating the Tech Stack: When, Where and How Should We Moderate Content?' *Centre for International Governance Innovation*. Retrieved 9 February 2022 (https://www.cigionline.org/articles/navigating-tech-stack-when-where-and-how-should-we-moderate-content/).

Donovan, Joan, Becca Lewis, and Brian Friedberg. 2019. 'Parallel Ports: Sociotechnical Change from the Alt-Right to Alt-Tech'.

Dori-Hacohen, Shiri, Keen Sung, Jengyu Chou, and Julian Lustig-Gonzalez. 2021. 'Restoring Healthy Online Discourse by Detecting and Reducing Controversy, Misinformation, and Toxicity Online'. Pp. 2627–28 in *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*.

Douek, Evelyn. 2019. *Facebook's 'Oversight Board:' Move Fast with Stable Infrastructure and Humility*. *SSRN Scholarly Paper*. ID 3365358. Rochester, NY: Social Science Research Network.

Douek, Evelyn. 2022. 'Content Moderation as Systems Thinking'. *Harvard Law Review*.

Drezner, Daniel W. 2004. 'The Global Governance of the Internet: Bringing the State Back In'. *Political Science Quarterly* 119(3):477–98.

Dubois, Elizabeth, and Devin Gaffney. 2014. 'The Multiple Facets of Influence'. *American Behavioral Scientist* 58:1260–77. doi: 10.1177/0002764214527088.

Duguay, Stefanie, Jean Burgess, and Nicolas Suzor. 2020. 'Queer Women's Experiences of Patchwork Platform Governance on Tinder, Instagram, and Vine'. *Convergence* 26(2):237–52. doi: 10.1177/1354856518781530.

Dutton, William H., and Malcolm Peltu. 2007. 'The Emerging Internet Governance Mosaic: Connecting the Pieces'. *Information Polity* 12(1–2):63–81.

Dwoskin, Elizabeth, and Rachel Lerman. 2020. '"Stop the Steal" Supporters, Restrained by Facebook, Turn to Parler to Peddle False Election Claims'. *Washington Post*, November 13.

Economides, Nicholas, and Ioannis Lianos. 2010. 'A Critical Appraisal of Remedies in the EU Microsoft Cases'. *Colum. Bus. L. Rev.* 346.

Edelman, Benjamin, and Damien Geradin. 2016. 'Android and Competition Law: Exploring and Assessing Google's Practices in Mobile'. *European Competition Journal* 12(2–3):159–94. doi: 10.1080/17441056.2016.1254483.

End Violence Against Women Coalition, and Glitch UK. 2020. *The Ripple Effect: COVID-19 and the Epidemic of Online Abuse*.

Epic Games. 2021. 'Epic Games Files EU Antitrust Complaint Against Apple'. Retrieved 14 March 2023 (https://www.epicgames.com/site/en-US/news/epic-games-files-eu-antitrust-complaint-against-apple).

Epstein, Dmitry. 2013. 'The Making of Institutions of Information Governance: The Case of the Internet Governance Forum'. *Journal of Information Technology* 28(2):137–49. doi: 10.1057/jit.2013.8.

Epstein, Dmitry, Christian Katzenbach, and Francesca Musiani. 2016. 'Doing Internet Governance: Practices, Controversies, Infrastructures, and Institutions'. *Internet Policy Review* 5(3).

Esser, Arne C. 2021. 'How Does the Language of Corpora from Radicalized Communities Discovered on Parler Compare to Online Conversations on Twitter Regarding the 2021 Capitol Riots and Election Fraud?' Master's Thesis.

European Commission. 2018. 'Antitrust: Commission Fines Google €4.34 Billion for Abuse of Dominance Regarding Android Devices'. *European Commission*. Retrieved 11 February 2022 (https://ec.europa.eu/commission/presscorner/detail/en/IP_18_4581).

European Commission. 2021. 'The Commission Proposes to Extend the List of 'EU Crimes''. *European Commission - European Commission*. Retrieved 13 June 2022 (https://ec.europa.eu/commission/presscorner/detail/en/ip_21_6561).

European Commission. 2023a. 'The Digital Markets Act: Ensuring Fair and Open Digital Markets'. Retrieved 16 April 2023 (https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets_en).

European Commission. 2023b. 'The Digital Services Act: Ensuring a Safe and Accountable Online Environment'. Retrieved 16 April 2023 (https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment_en).

European Commission. 2023c. 'The Digital Services Act Package | Shaping Europe's Digital Future'. Retrieved 14 March 2023 (https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package).

European Parliament. 2022a. *European Parliament Legislative Resolution of 5 July 2022 on the Proposal for a Regulation of the European Parliament and of the Council on Contestable and Fair Markets in the Digital Sector (Digital Markets Act)*.

European Parliament. 2022b. *Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and Amending Directive 2000/31/EC (Digital Services Act) (Text with EEA Relevance)*. Vol. 277.

Fan, Jenny, and Amy X. Zhang. 2020. 'Digital Juries: A Civics-Oriented Approach to Platform Governance'. Pp. 1–14 in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Honolulu HI USA: ACM.

Farkas, Johan, Jannick Schou, and Christina Neumayer. 2018. 'Cloaked Facebook Pages: Exploring Fake Islamist Propaganda in Social Media'. *New Media & Society* 20(5):1850–67.

Federal Ministry of Justice. 2021. 'German Criminal Code (Strafgesetzbuch – StGB)'. Retrieved 31 March 2023 (https://www.gesetze-im-internet.de/englisch_stgb/englisch_stgb.html#p1241).

Fedoruk, Benjamin, Harrison Nelson, Russell Frost, and Kai Fucile Ladouceur. 2021. 'The Plebeian Algorithm: A Democratic Approach to Censorship and Moderation'. *JMIR Formative Research* 5(12):e32427.

Fink, Christina. 2018. 'DANGEROUS SPEECH, ANTI-MUSLIM VIOLENCE, AND FACEBOOK IN MYANMAR'. *Journal of International Affairs* 71(1.5):43–52.

Flew, Terry, and Fiona R. Martin. 2022. *Digital Platform Regulation: Global Perspectives on Internet Governance*. Springer Nature.

Flick, Uwe. 2017. 'The SAGE Handbook of Qualitative Data Collection'. *The SAGE Handbook of Qualitative Data Collection* 1–736.

Floridi, Luciano. 2020. 'The Fight for Digital Sovereignty: What It Is, and Why It Matters, Especially for the EU'. *Philosophy & Technology* 33(3):369–78. doi: 10.1007/s13347-020-00423-6.

Floridi, Luciano. 2021a. 'The End of an Era: From Self-Regulation to Hard Law for the Digital Industry'. *Philosophy & Technology* 34(4):619–22. doi: 10.1007/s13347-021-00493-0.

Floridi, Luciano. 2021b. 'Trump, Parler, and Regulating the Infosphere as Our Commons'. *Philosophy & Technology* 34(1):1–5. doi: 10.1007/s13347-021-00446-7.

Flyvbjerg, Bent. 2006. 'Five Misunderstandings about Case-Study Research'. *Qualitative Inquiry* 12(2):219–45.

Foer, Franklin. 2017. 'Facebook's War on Free Will'. *The Guardian*. Retrieved 5 May 2021 (http://www.theguardian.com/technology/2017/sep/19/facebooks-war-on-free-will).

for Countering Digital Hate. 2021. 'Euro '21 Final, Racism and Instagram'. *Center for Countering Digital Hate | CCDH*. Retrieved 31 March 2023 (https://counterhate.com/research/euro-21-final-racism-and-instagram/).

Fordham, Evie. 2020. 'Parler CEO Says Social Media Companies Are Behaving like Publishers by Censoring'. *Yahoo*. Retrieved 17 March 2023 (https://www.yahoo.com/now/parler-ceo-says-social-media-145832887.html).

Fortuna, Paula, and Sérgio Nunes. 2018. 'A Survey on Automatic Detection of Hate Speech in Text'. *ACM Computing Surveys (CSUR)* 51(4):1–30.

Fox, Eleanor M. 2019. 'Platforms, Power, and the Antitrust Challenge: A Modest Proposal to Narrow the U.S.-Europe Divide'. *Nebraska Law Review* 98:297.

Fraser, Nancy. 1990. 'Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy'. *Social Text* (25/26):56–80.

Fuchs, Christian, ed. 2012. *Internet and Surveillance: The Challenges of Web 2.0 and Social Media*. New York: Routledge.

Fuchs, Christian. 2021. *Digital Capitalism: Media, Communication and Society Volume Three*. Routledge.

Gagliardone, Iginio, Danit Gal, Thiago Alves, and Martinez Sainz Martinez. 2015. *Countering Online Hate Speech*. UNESCO Publishing.

Gallacher, John, and Jonathan Bright. 2021. 'Hate Contagion: Measuring the Spread and Trajectory of Hate on Social Media'.

Ganesh, Bharath. 2018. 'The Ungovernability of Digital Hate Culture'. *JIA SIPA*. Retrieved 1 April 2023 (https://jia.sipa.columbia.edu/ungovernability-digital-hate-culture).

Ganesh, Maya Indira. 2020. 'The Ironies of Autonomy'. *Humanities and Social Sciences Communications* 7(1):1–10. doi: 10.1057/s41599-020-00646-0.

Ganguly, Manisha. 2023. 'Pentagon Leaks: How Discord Video Game Chat Platform Landed in the Spotlight'. *The Guardian*, April 14.

Garud, Raghu, Henri A. Schildt, and Theresa K. Lant. 2014. 'Entrepreneurial Storytelling, Future Expectations, and the Paradox of Legitimacy'. *Organization Science* 25(5):1479–92.

Geertz, Clifford. 1996. *After the Fact: Two Countries, Four Decades, One Anthropologist*. Cambridge, MA, USA: Harvard University Press.

George, Alexander L., and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. MIT Press.

van Geuns, Suzanne, and Corinne Cath-Speth. 2020. 'How Hate Speech Reveals the Invisible Politics of Internet Infrastructure'. *Brookings*. Retrieved 29 July 2022 (https://www.brookings.edu/techstream/how-hate-speech-reveals-the-invisible-politics-of-internet-infrastructure/).

Ghaffary, Shirin. 2020. 'Twitter Has Finally Started Fact-Checking Trump'. *Vox*. Retrieved 19 May 2022 (https://www.vox.com/recode/2020/5/26/21271210/twitter-fact-check-trump-tweets-mail-voting-fraud-rigged-election-misleading-statements).

Gillespie, Tarleton. 2010. 'The Politics of "Platforms"'. *New Media & Society* 12(3):347–64. doi: 10.1177/1461444809342738.

Gillespie, Tarleton. 2015. 'Platforms Intervene'. *Social Media + Society* 1(1):2056305115580479. doi: 10.1177/2056305115580479.

Gillespie, Tarleton. 2017. 'Governance of and by Platforms'. *SAGE Handbook of Social Media* 254–78.

Gillespie, Tarleton. 2018. *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. Yale University Press.

Gillespie, Tarleton. 2023. 'The Fact of Content Moderation; Or, Let's Not Solve the Platforms' Problems for Them'. *Media and Communication* 11(2).

Gillespie, Tarleton, Patricia Aufderheide, Elinor Carmi, Ysabel Gerrard, Robert Gorwa, Ariadna Matamoros-Fernández, Sarah T. Roberts, Aram Sinnreich, and Sarah Myers West. 2020. 'Expanding the Debate about Content Moderation: Scholarly Research Agendas for the Coming Policy Debates'. *Internet Policy Review* 9(4).

Gillett, Rosalie, and Nicolas Suzor. 2022. 'Incels on Reddit: A Study in Social Norms and Decentralised Moderation'. *First Monday*.

Glenza, Jessica. 2021. '"Historical Accident": How Abortion Came to Focus White, Evangelical Anger'. *The Guardian*, December 5.

GoDaddy. 2021. 'GoDaddy Statement Regarding AR15.Com'. Retrieved 9 June 2022 (https://aboutus.godaddy.net/newsroom/press-releases/press-release-details/2021/GoDaddy-Statement-Regarding-AR15.com/default.aspx).

Goertz, Gary, and James Mahoney. 2012. *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences*. Princeton University Press.

Goldsmith, Jack L. 1998. 'Against Cyberanarchy'. *The University of Chicago Law Review* 65(4):1199–1250.

Goldsmith, Jack, and Tim Wu. 2006. *Who Controls the Internet?: Illusions of a Borderless World*. Oxford University Press.

Gonzalez, Oriana, and Danielle Alberti. 2022. 'The Political Leanings of the Supreme Court Justices'. *Axios*. Retrieved 21 April 2023 (https://www.axios.com/2019/06/01/supreme-court-justices-ideology).

Google. 2023a. 'Developer Policy Center'. Retrieved 14 March 2023 (https://play.google.com/about/developer-content-policy/).

Google. 2023b. 'Hate Speech Policy - YouTube Help'. Retrieved 2 April 2023 (https://support.google.com/youtube/answer/2801939?hl=en&ref_topic=9282436).

Google. 2023c. 'Inappropriate Content - Play Console Help'. Retrieved 15 March 2023 (https://support.google.com/googleplay/android-developer/answer/9878810).

Gorrell, Genevieve, Mark A. Greenwood, Ian Roberts, Diana Maynard, and Kalina Bontcheva. 2018. 'Twits, Twats and Twaddle: Trends in Online Abuse towards Uk Politicians'. in *Twelfth international AAAI conference on web and social media*.

Gorwa, Robert. 2019. 'What Is Platform Governance?' *Information, Communication & Society* 22(6):854–71. doi: 10.1080/1369118X.2019.1573914.

Gorwa, Robert. 2021. 'Elections, Institutions, and the Regulatory Politics of Platform Governance: The Case of the German NetzDG'. *Telecommunications Policy* 45(6):102145.

Gorwa, Robert, Reuben Binns, and Christian Katzenbach. 2020. 'Algorithmic Content Moderation: Technical and Political Challenges in the Automation of Platform Governance'. *Big Data & Society* 7(1):2053951719897945. doi: 10.1177/2053951719897945.

Government of Canada. 2023. 'Consolidated Federal Laws of Canada, Criminal Code'. Retrieved 31 March 2023 (https://laws-lois.justice.gc.ca/eng/acts/c-46/).

Grant, Maria J., and Andrew Booth. 2009. 'A Typology of Reviews: An Analysis of 14 Review Types and Associated Methodologies'. *Health Information & Libraries Journal* 26(2):91–108. doi: 10.1111/j.1471-1842.2009.00848.x.

Grant, Nico. 2022. 'Parler Returns to Google Play Store'. *The New York Times*, September 2.

Graz, Jean-Christophe, and Andreas Nölke. 2007. *Transnational Private Governance and Its Limits*. Routledge.

Greenberg, Andy. 2021. 'An Absurdly Basic Bug Let Anyone Grab All of Parler's Data'. *Wired*, January 12.

Greenhalgh, Spencer, Daniel G. Krutka, and Shannon M. Oltmann. 2021. 'Gab, Parler, and (Mis) Educational Technologies: Reconsidering Informal Learning on Social Media Platforms'. *The Journal of Applied Instructional Design* 10(3).

Griffin, Rachel. 2022. 'New School Speech Regulation as a Regulatory Strategy against Hate Speech on Social Media: The Case of Germany's NetzDG'. *Telecommunications Policy* 46(9):102411. doi: 10.1016/j.telpol.2022.102411.

Grimmelmann, James. 2015. 'The Virtues of Moderation'. *Yale JL & Tech.* 17:42.

Groeger, Lena V., J. Kao, A. Shaw, M. Syed, and M. Eliahou. 2021. 'What Parler Saw during the Attack on the Capitol. ProPublica'.

Grünwald, Andreas, Christoph Nüssing, Theresa Oehm, and Dominik Arncken. 2022. 'The EU Digital Markets Act – The Holy Grail of Big Tech Regulation?' *Morrison Foerster*. Retrieved 22 September 2022 (https://www.mofo.com/resources/insights/null).

Gupta, Priya S. 2019. 'The Fleeting, Unhappy Affair of Amazon HQ2 and New York City'. *Transnational Legal Theory* 10(1):97–122. doi: 10.1080/20414005.2019.1609785.

Gurman. 2020. 'Apple Weighs Letting Users Switch Default IPhone Apps to Rivals'. *Bloomberg.Com*, February 20.

Gurman, Mark. 2022. 'Apple to Allow Outside App Stores in Overhaul Spurred by EU Laws'. *Bloomberg.Com*, December 13.

Guterres, António. 2019. 'The Wildfire of Hate Speech | United Nations Secretary-General'. Retrieved 3 September 2021 (https://www.un.org/sg/en/content/sg/articles/2019-06-18/the-wildfire-of-hate-speech).

Guthey, Eric, Timothy Clark, and Brad Jackson. 2009. *Demystifying Business Celebrity*. Routledge.

Hagey, Keach, and Jeff Horwitz. 2021. 'Parler, a Platform Favored by Trump Fans, Struggles for Survival'. *WSJ*. Retrieved 15 November 2022 (https://www.wsj.com/articles/parler-struggles-survival-amazon-lawsuit-trump-fans-11610414745).

Haggart, Blayne, and Clara Iglesias Keller. 2021. 'Democratic Legitimacy in Global Platform Governance'. *Telecommunications Policy* 45(6):102152. doi: 10.1016/j.telpol.2021.102152.

Haggart, Blayne, Courtney Radsch, Chris Riley, Natasha Tusikov, Heidi Tworek, and Samuel Woolley. 2022. 'As Elon Musk Upends Twitter, Whither the Digital Town Square?' *Centre for International Governance Innovation*. Retrieved 3 April 2023 (https://www.cigionline.org/articles/as-elon-musk-upends-twitter-whither-the-digital-town-square/).

Hall, Madison. 2023. 'At Least 476 Rioters Have Pleaded Guilty for Their Role in the Capitol Insurrection so Far. This Table Shows Them All.' *Insider*, February 16.

Hancock, Ange-Marie. 2007. 'Intersectionality as a Normative and Empirical Paradigm'. *Politics & Gender* 3(2):248–54.

Hardwick, Tim. 2022. 'EU Provisionally Agrees on Law That Would Force Apple to Allow Alternative App Stores, Sideloading, and IMessage Interoperability'. *MacRumors*. Retrieved 14 March 2023 (https://www.macrumors.com/2022/03/25/eu-provisionally-agrees-dma-law-apple-app-store/).

Harsin, Jayson. 2015. 'Regimes of Posttruth, Postpolitics, and Attention Economies'. *Communication, Culture & Critique* 8(2):327–33.

Hartzell, Stephanie L. 2018. 'Alt-White: Conceptualizing the 'Alt-Right'as a Rhetorical Bridge between White Nationalism and Mainstream Public Discourse'. *Journal of Contemporary Rhetoric* 8(1/2):6–25.

Hasinoff, Amy A., and Nathan Schneider. 2022. 'From Scalability to Subsidiarity in Addressing Online Harm'. *Social Media + Society* 8(3):20563051221126040. doi: 10.1177/20563051221126041.

Hassoun, Nicole. 2012. *Globalization and Global Justice: Shrinking Distance, Expanding Obligations*. Cambridge University Press.

Hatmaker, Taylor. 2020. 'Parler Tops App Store Rankings Following Biden's Win'. *TechCrunch*. Retrieved 21 March 2022 (https://social.techcrunch.com/2020/11/09/parler-app-store-facebook/).

Hatmaker, Taylor. 2022. 'Google Lets the Parler App Back into the Play Store'. *TechCrunch*. Retrieved 21 April 2023 (https://techcrunch.com/2022/09/02/parler-google-play-store-return/).

Haupt, Joachim. 2021. 'Facebook Futures: Mark Zuckerberg's Discursive Construction of a Better World'. *New Media & Society* 23(2):237–57. doi: 10.1177/1461444820929315.

Hawley, George. 2018. *Making Sense of the Alt-Right*. Columbia University Press.

Helberger, Natali, Jo Pierson, and Thomas Poell. 2018. 'Governing Online Platforms: From Contested to Cooperative Responsibility'. *The Information Society* 34(1):1–14. doi: 10.1080/01972243.2017.1391913.

Heldt, Amélie Pia. 2019. 'Reading between the Lines and the Numbers: An Analysis of the First NetzDG Reports'. *Internet Policy Review* 8(2).

Helmond, Anne. 2015. 'The Platformization of the Web: Making Web Data Platform Ready'. *Social Media+ Society* 1(2):2056305115603080.

Helmore, Edward. 2022. '"Don" of a New Era: The Rise of Peter Thiel as a US Rightwing Power Player'. *The Guardian*, May 30.

Hemmer, Nicole. 2016. *Messengers of the Right: Conservative Media and the Transformation of American Politics*. University of Pennsylvania Press.

Hern, Alex. 2021. 'Apple and Facebook at Odds over Privacy Move That Will Hit Online Ads'. *The Guardian*, January 28.

Hesse. 2020. 'Parler Wasn't Hacked, but That Doesn't Mean It's Safe to Use'. *Lifehacker*. Retrieved 11 April 2023 (https://lifehacker.com/parler-wasnt-hacked-but-that-doesnt-mean-its-safe-to-u-1845757409).

Hestres, Luis. 2013. 'App Neutrality: Apple's App Store and Freedom of Expression Online'. *Hestres, LE (2013). App Neutrality: Apple's App Store and Freedom of Expression Online. International Journal of Communication* 7:1265–80.

Hill, Clara. 2021. 'Apple Removes Anti-Vaxx Dating App from App Store'. *The Independent*, August 2.

Hine, Gabriel, Jeremiah Onaolapo, Emiliano De Cristofaro, Nicolas Kourtellis, Ilias Leontiadis, Riginos Samaras, Gianluca Stringhini, and Jeremy Blackburn. 2017. 'Kek, Cucks, and God Emperor Trump: A Measurement Study of 4chan's Politically Incorrect Forum and Its Effects on the Web'. Pp. 92–101 in *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 11.

Hitkul, Avinash Prabhu, Dipanwita Guhathakurta, Jivitesh jain, Mallika Subramanian, Manvith Reddy, Shradha Sehgal, Tanvi Karandikar, Amogh Gulati, Udit Arora, Rajiv Ratn Shah, and Ponnurangam Kumaraguru. 2021. *Capitol (Pat)Riots: A Comparative Study of Twitter and Parler*. arXiv:2101.06914. arXiv. doi: 10.48550/arXiv.2101.06914.

Hoffmann, Anna Lauren, Nicholas Proferes, and Michael Zimmer. 2018a. '"Making the World More Open and Connected": Mark Zuckerberg and the Discursive Construction of Facebook and Its Users'. *New Media & Society* 20(1):199–218. doi: 10.1177/1461444816660784.

Hoffmann, Anna Lauren, Nicholas Proferes, and Michael Zimmer. 2018b. '"Making the World More Open and Connected": Mark Zuckerberg and the Discursive Construction of Facebook and Its Users'. *New Media & Society* 20(1):199–218. doi: 10.1177/1461444816660784.

Hofmann, Jeanette. 2020. '12 The Multistakeholder Concept as Narrative: A Discourse Analytical Approach'. *Researching Internet Governance: Methods, Frameworks, Futures* 249.

Hofmann, Jeanette, Christian Katzenbach, and Kirsten Gollatz. 2017. 'Between Coordination and Regulation: Finding the Governance in Internet Governance'. *New Media & Society* 19(9):1406–23.

Hofmann, Vincent, Christina Dinar, Matthias C. Kettemann, Lena Hinrichs, Julius Böke, and Max Gradulewski. 2022. 'The Power of App Stores and Their Normative Orders'. in *How Platforms Respond to Human Rights Conflicts Online. Best Practices in Weighing Rights and Obligations in Hybrid Online Orders*, edited by M. C. Kettemann. Hamburg: Verlag Hans-Bredow-Institut.

Hollister, Sean. 2021. 'Apple's $64 Billion-a-Year App Store Isn't Catching the Most Egregious Scams'. *The Verge*. Retrieved 10 February 2022 (https://www.theverge.com/2021/4/21/22385859/apple-app-store-scams-fraud-review-enforcement-top-grossing-kosta-eleftheriou).

Hope Not Hate. 2020. *A Better Web*.

Horta Ribeiro, Manoel, Homa Hosseinmardi, Robert West, and Duncan J. Watts. 2023. 'Deplatforming Did Not Decrease Parler Users' Activity on Fringe Social Media'. *PNAS Nexus* 2(3):pgad035. doi: 10.1093/pnasnexus/pgad035.

Horwitz, Jeff, and Keach Hagey. 2020. 'Parler Makes Play for Conservatives Mad at Facebook, Twitter'. *Wall Street Journal*, November 14.

Hosseinmardi, Homa, Amir Ghasemian, Aaron Clauset, Markus Mobius, David M. Rothschild, and Duncan J. Watts. 2021. 'Examining the Consumption of Radical Content on YouTube'. *Proceedings of the National Academy of Sciences* 118(32):e2101967118.

Howard, Jeffrey W. 2019. 'Dangerous Speech'. *Philosophy & Public Affairs* 47(2):208–54.

Howard, Philip N. 2020. *Lie Machines: How to Save Democracy from Troll Armies, Deceitful Robots, Junk News Operations, and Political Operatives*. Yale University Press.

Hsieh, Hsiu-Fang, and Sarah E. Shannon. 2005. 'Three Approaches to Qualitative Content Analysis'. *Qualitative Health Research* 15(9):1277–88.

Hurrell, Andrew, and Terry Macdonald. 2012. 'Global Public Power: The Subject of Principles of Global Political Legitimacy'. *Critical Review of International Social and Political Philosophy* 15(5):553–71. doi: 10.1080/13698230.2012.727305.

Hutchinson, Jade, Amarnath Amarasingam, Ryan Scrivens, and Brian Ballsun-Stanton. 2021. 'Mobilizing Extremism Online: Comparing Australian and Canadian Right-Wing Extremist Groups on Facebook'. *Behavioral Sciences of Terrorism and Political Aggression* 1–31.

Hylton, Keith N. 2019. 'Digital Platforms and Antitrust Law'. *Nebraska Law Review* 98:272.

Imhoff, Roland, and Julia Recker. 2012. 'Differentiating Islamophobia: Introducing a New Scale to Measure Islamoprejudice and Secular Islam Critique'. *Political Psychology* 33(6):811–24.

Inazu, John. 2013. 'Virtual Assembly'. *Cornell Law Review* 98(5):1093.

Ingham-Barrow, Isobel. 2018. 'More than Words: Approaching a Definition of Islamophobia'. *Muslim Engagement and Development (MEND)*.

Isaac, Mike, and Kellen Browning. 2020. 'Fact-Checked on Facebook and Twitter, Conservatives Switch Their Apps'. *The New York Times*, November 11.

Israeli, Abraham, and Oren Tsur. 2022. 'Going Extreme: Comparative Analysis of Hate Speech in Parler and Gab'. *ArXiv Preprint ArXiv:2201.11770*.

Jakubik, Johannes, Michael Vössing, Dominik Bär, Nicolas Pröllochs, and Stefan Feuerriegel. 2022. 'Online Emotions During the Storming of the US Capitol: Evidence from the Social Media Network Parler'. *ArXiv Preprint ArXiv:2204.04245*.

Jasanoff, Sheila, and Sang-Hyun Kim. 2009. 'Containing the Atom: Sociotechnical Imaginaries and Nuclear Power in the United States and South Korea'. *Minerva* 47(2):119.

Jereza, Rae. 2022. '"I'm Not This Person": Racism, Content Moderators, and Protecting and Denying Voice Online'. *New Media & Society* 14614448221122224. doi: 10.1177/14614448221122224.

Johnson, David R., and David Post. 1996. 'Law and Borders: The Rise of Law in Cyberspace'. *Stanford Law Review* 1367–1402.

Johnson, Hailey, Karl Volk, Robert Serafin, Cinthya Grajeda, and Ibrahim Baggili. 2022. 'Alt-Tech Social Forensics: Forensic Analysis of Alternative Social Networking Applications'. *Forensic Science International: Digital Investigation* 42:301406.

Kan, Michael. 2018. 'PayPal Bans Infowars Over Hate Speech'. *PCMag UK*. Retrieved 9 June 2022 (https://uk.pcmag.com/software/117545/paypal-bans-infowars-over-hate-speech).

Katzenbach, Christian, and Stefan Larsson. 2017. 'Imagining the Digital Society – Metaphors from the Past and Present – Digital Society Blog'. *HIIG*. Retrieved 28 March 2023 (https://www.hiig.de/en/imagining-the-digital-society-metaphors-from-the-past-and-present/).

Katzenbach, Christian, and Lena Ulbricht. 2019. 'Algorithmic Governance'. *Internet Policy Review* 8(4):1–18. doi: 10.14763/2019.4.1424.

Kaye, D. 2018. *A Human Rights Approach to Platform Content Regulation*. Report of the UN Special Rapporteur on the Promotion and Protection of the ….

Kaye, David. 2017. 'Mandate of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression'.

Keller, Daphne. 2019. 'Who Do You Sue?' *Hoover Institution*. Retrieved 22 October 2019 (https://www.hoover.org/research/who-do-you-sue).

Kelly, Makena. 2019. 'Cloudflare to Revoke 8chan's Service, Opening the Fringe Website up for DDoS Attacks'. *The Verge*. Retrieved 10 February 2022 (https://www.theverge.com/2019/8/4/20754310/cloudflare-8chan-fredrick-brennan-ddos-attack).

Kelly, Makena, and Zoe Schiffer. 2020. '"Stop the Steal" Spreads across the Internet after Infecting Facebook'. *The Verge*. Retrieved 3 April 2023 (https://www.theverge.com/2020/11/6/21552752/facebook-stopthesteal-ban-twitter-parler-discord-trump).

Kettemann, Matthias C. 2020. *The Normative Order of the Internet: A Theory of Rule and Regulation Online*. Oxford University Press.

Kettemann, Matthias C., and Wolfgang Schulz. 2020. 'Setting Rules for 2.7 Billion: A (First) Look into Facebook's Norm-Making System; Results of a Pilot Study'.

Kettemann, Matthias C., and Anna Sophia Tiedeke. 2020. 'Back up: Can Users Sue Platforms to Reinstate Deleted Content?' *Internet Policy Review* 9(2):1–20.

Khalid, A. 2022. 'Apple Faces €5.5 Billion Lawsuit from Netherlands over Its App Store'. *Engadget*. Retrieved 29 April 2022 (https://www.engadget.com/apple-faces-class-action-lawsuit-from-netherlands-over-its-app-store-001610098.html).

Kiela, Douwe, Hamed Firooz, Aravind Mohan, Vedanuj Goswami, Amanpreet Singh, Pratik Ringshia, and Davide Testuggine. 2020. 'The Hateful Memes Challenge: Detecting Hate Speech in Multimodal Memes'. *Advances in Neural Information Processing Systems* 33:2611–24.

King, Gary, Robert O. Keohane, and Sidney Verba. 2021. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton university press.

Kirk, Hannah Rose, Bertram Vidgen, Paul Röttger, Tristan Thrush, and Scott A. Hale. 2022. 'Hatemoji: A Test Suite and Adversarially-Generated Dataset for Benchmarking and Detecting Emoji-Based Hate'.

Kirtz, Jaime Lee, and Zeerak Talat. 2023. 'Futures for Research on Hate Speech in Online Social Media Platforms'. P. 467 in *Challenges and perspectives of hate speech research*.

Klein, Hans. 2002. 'ICANN and Internet Governance: Leveraging Technical Coordination to Realize Global Public Policy'. *The Information Society* 18(3):193–207.

Klonick, Kate. 2017. 'The New Governors: The People, Rules, and Processes Governing Online Speech'. *Harvard Law Review* 131:1598.

Klonick, Kate. 2021. 'Inside the Making of Facebook's Supreme Court'. *New Yorker*, February 12.

Klonick, Kate, and Thomas Kadri. 2019. *Facebook v. Sullivan: Public Figures and Newsworthiness in Online Speech*. *SSRN Scholarly Paper*. ID 3332530. Rochester, NY: Social Science Research Network. doi: 10.2139/ssrn.3332530.

Kohlbacher, Florian. 2006. 'The Use of Qualitative Content Analysis in Case Study Research'. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 7(1). doi: 10.17169/fqs-7.1.75.

Kohler-Koch, Beate, and Berthold Rittberger. 2006. 'The Governance Turn in EU Studies'. *J. Common Mkt. Stud.* 44:27.

Kollnig, Konrad, and Nigel Shadbolt. 2023. 'Ready for the EU Digital Services Act? How Decisions by Apple and by Google Impede App Privacy'.

Kosseff, Jeff. 2019. *The Twenty-Six Words That Created the Internet*. Cornell University Press.

Kuhn, Timothy. 2008. 'A Communicative Theory of the Firm: Developing an Alternative Perspective on Intra-Organizational Power and Stakeholder Relationships'. *Organization Studies* 29(8–9):1227–54.

Kurbalija, Jovan. 2016. *An Introduction to Internet Governance*. Diplo Foundation.

Kutner, Max. 2016. 'Meet Robert Mercer, Billionaire Benefactor of Breitbart'. *Newsweek*. Retrieved 3 April 2023 (https://www.newsweek.com/2016/12/02/robert-mercer-trump-donor-bannon-pac-523366.html).

Lada, Akos, Meihong Wang, and Tak Yan. 2021. 'How Does News Feed Predict What You Want to See?' *Tech at Meta*. Retrieved 2 April 2023 (https://tech.facebook.com/engineering/2021/1/news-feed-ranking/).

Lamerichs, Nicolle, Dennis Nguyen, Mari Carmen Puerta Melguizo, Radmila Radojevic, and Anna Lange-Böhmer. 2018. 'Elite Male Bodies: The Circulation of Alt-Right Memes and the Framing of Politicians on Social Media'. *Participations* 15(1):180–206.

Lane, Julie. 2019. 'Cultivating Distrust of the Mainstream Media'. *News on the Right: Studying Conservative News Cultures* 157–73.

Langley, Hugh. 2021. 'How Google Finally Decided to Remove Parler after Months of Flagging the App's Harmful Content'. *Business Insider*. Retrieved 13 April 2023 (https://www.businessinsider.com/insiders-talk-tensions-between-google-and-parler-2021-1).

Lapidot-Lefler, Noam, and Azy Barak. 2012. 'Effects of Anonymity, Invisibility, and Lack of Eye-Contact on Toxic Online Disinhibition'. *Computers in Human Behavior* 28(2):434–43. doi: 10.1016/j.chb.2011.10.014.

Ledwich, Mark, and Anna Zaitsev. 2019. 'Algorithmic Extremism: Examining YouTube's Rabbit Hole of Radicalization'. *ArXiv Preprint ArXiv:1912.11211*.

Lee, Timothy B. 2017. 'Cloudflare's CEO Has a Plan to Never Censor Hate Speech Again'. *Ars Technica*. Retrieved 9 February 2022 (https://arstechnica.com/tech-policy/2017/12/cloudflares-ceo-has-a-plan-to-never-censor-hate-speech-again/).

Leerssen, Paddy, Amélie Heldt, and Matthias C. Kettemann. 2023. 'Scraping By? Europe's Law and Policy on Social Media Research Access'. P. 405 in *Challenges and perspectives of hate speech research*.

Leerssen, Patrick. 2015. 'Cut Out By The Middle Man: The Free Speech Implications Of Social Network Blocking and Banning In The EU'. *JIPITEC* 6(2).

Leetaru, Kalev. 2019. 'Facebook Was A Democracy 2009-2012 But We Didn't Vote So It Turned Into A Dictatorship'. *Forbes*. Retrieved 23 July 2021 (https://www.forbes.com/sites/kalevleetaru/2019/04/13/facebook-was-a-democracy-2009-2012-but-we-didnt-vote-so-it-turned-into-a-dictatorship/).

Lehtiniemi, Tuukka, and Minna Ruckenstein. 2018. 'The Social Imaginaries of Data Activism'. *Big Data & Society* 6(1):2053951718821146.

Lerman, Rachel. 2020. 'The Conservative Alternative to Twitter Wants to Be a Place for Free Speech for All. It Turns out, Rules Still Apply.' *Washington Post*, November 23.

Lessig, Lawrence. 1999. *Code: And Other Laws of Cyberspace*. Basic Books.

Levy, Ari. 2020. 'Trump Fans Are Flocking to the Social Media App Parler — Its CEO Is Begging Liberals to Join Them'. *CNBC*. Retrieved 19 May 2022 (https://www.cnbc.com/2020/06/27/parler-ceo-wants-liberal-to-join-the-pro-trump-crowd-on-the-app.html).

Light, Ben, Jean Burgess, and Stefanie Duguay. 2018. 'The Walkthrough Method: An Approach to the Study of Apps'. *New Media & Society* 20(3):881–900.

Lima, Cristiano. 2020. 'Cruz Joins Alternative Social Media Site Parler in Swipe at Big Tech Platforms'. *POLITICO*. Retrieved 19 May 2022 (https://www.politico.com/news/2020/06/25/ted-cruz-joins-parler-339811).

Lin, Johnny. 2020. 'How to Make $80,000 Per Month on the Apple App Store'. *Transparency Matters*. Retrieved 10 February 2022 (https://blog.lockdownprivacy.com/2020/11/25/how-to-make-80000.html).

Livingstone, Sonia, and Ranjana Das. 2013. 'Interpretation/Reception'.

Lomas, Natasha. 2019. 'Aptoide, a Play Store Rival, Cries Antitrust Foul over Google Hiding Its App'. *TechCrunch*. Retrieved 14 April 2023 (https://techcrunch.com/2019/06/04/aptoide-a-play-store-rival-cries-antitrust-foul-over-google-hiding-its-app/).

Lothian-McLean, Moya. 2019. 'Mumsnet Users Declare Flora Margarine Boycott after Brand Cuts Ties with Site over Claims of Transphobia'. *The Independent*. Retrieved 23 June 2022 (https://www.independent.co.uk/life-style/mumsnet-transphobic-flora-upfield-boycott-a9155016.html).

Lounsbury, Michael, and Mary Ann Glynn. 2001. 'Cultural Entrepreneurship: Stories, Legitimacy, and the Acquisition of Resources'. *Strategic Management Journal* 22(6–7):545–64.

Loveluck, Laura, and Ghalia al-Alwani. 2018. 'Saudi Electronic Army Floods Twitter with Insults and Mistruths after Khashoggi's Disappearance'. *Washington Post*, October 19.

Lynn Jr, Laurence E., Carolyn J. Heinrich, and Carolyn J. Hill. 2001. *Improving Governance: A New Logic for Empirical Research*. Georgetown University Press.

Lyons, Kim. 2020. 'Twitter Removes Tweets by Brazil, Venezuela Presidents for Violating COVID-19 Content Rules'. *The Verge*. Retrieved 17 March 2023 (https://www.theverge.com/2020/3/30/21199845/twitter-tweets-brazil-venezuela-presidents-covid-19-coronavirus-jair-bolsonaro-maduro).

Lyons, Kim. 2021a. 'Amazon Is Kicking Parler off Its Web Hosting Service'. *The Verge*. Retrieved 14 March 2023 (https://www.theverge.com/2021/1/9/22222637/amazon-workers-aws-stop-hosting-services-parler-capitol-violence).

Lyons, Kim. 2021b. 'Parler Returns to Apple App Store with Some Content Excluded'. *The Verge*. Retrieved 21 March 2022 (https://www.theverge.com/2021/5/17/22440005/parler-apple-app-store-return-amazon-google-capitol).

M. Otala, Jacqueline, Gillian Kurtic, Isabella Grasso, Yu Liu, Jeanna Matthews, and Golshan Madraki. 2021. 'Political Polarization and Platform Migration: A Study of Parler and Twitter Usage by United States of America Congress Members'. Pp. 224–31 in *Companion Proceedings of the Web Conference 2021*.

Ma, Cindy. 2021. 'What Is the "Lite" in "Alt-Lite?" The Discourse of White Vulnerability and Dominance among YouTube's Reactionaries'. *Social Media+ Society* 7(3):20563051211036384.

Mac, Ryan, and Rosie Gray. 2021. 'Parler Wanted Donald Trump On Its Site. Trump's Company Wanted A Stake.' *BuzzFeed News*. Retrieved 3 April 2023 (https://www.buzzfeednews.com/article/ryanmac/trump-parler-ownership).

Mac, Ryan, and John Paczkowski. 2021. 'Apple Has Threatened To Ban Parler From The App Store'. *BuzzFeed News*. Retrieved 21 March 2022 (https://www.buzzfeednews.com/article/ryanmac/apple-threatens-ban-parler).

MacCarthy, Mark. 2022. 'The Open App Markets Bill Moves out of the Senate Judiciary Committee'. *Brookings*. Retrieved 21 March 2022 (https://www.brookings.edu/blog/techtank/2022/03/10/the-open-app-markets-bill-moves-out-of-the-senate-judiciary-committee/).

Machkovech, Sam. 2015. '8chan-Hosted Content Disappears from Google Searches [Updated]'. *Ars Technica*. Retrieved 24 February 2021 (https://arstechnica.com/tech-policy/2015/08/8chan-hosted-content-disappears-from-google-searches/).

Mager, Astrid, and Christian Katzenbach. 2021. 'Future Imaginaries in the Making and Governing of Digital Technology: Multiple, Contested, Commodified'. *New Media & Society* 23(2):223–36. doi: 10.1177/1461444820929321.

Mahoney, James, and Gary Goertz. 2006. 'A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research'. *Political Analysis* 14(3):227–49.

Mann, Monique, Peta Mitchell, Marcus Foth, and Irina Anastasiu. 2020. '#BlockSidewalk to Barcelona: Technological Sovereignty and the Social License to Operate Smart Cities'. *Journal of the Association for Information Science and Technology* 71(9):1103–15. doi: https://doi.org/10.1002/asi.24387.

Margetts, H., B. Vidgen, and S. A. Hale. 2019. 'Online Hate and the Nasty Election'. *UK Election Analysis* 16–17.

Marwick, Alice E. 2013. *Status Update: Celebrity, Publicity, and Branding in the Social Media Age*. Yale University Press.

Marwick, Alice E., and Danah Boyd. 2011. 'I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience'. *New Media & Society* 13(1):114–33.

Marwick, Alice E., and Ross Miller. 2014. *Online Harassment, Defamation, and Hateful Speech: A Primer of the Legal Landscape. SSRN Scholarly Paper*. ID 2447904. Rochester, NY: Social Science Research Network.

Massanari, Adrienne. 2017. '#Gamergate and The Fappening: How Reddit's Algorithm, Governance, and Culture Support Toxic Technocultures'. *New Media & Society* 19(3):329–46. doi: 10.1177/1461444815608807.

Matamoros-Fernández, Ariadna, and Johan Farkas. 2021. 'Racism, Hate Speech, and Social Media: A Systematic Review and Critique'. *Television & New Media* 22(2):205–24. doi: 10.1177/1527476420982230.

Matias, J. Nathan. 2019. 'The Civic Labor of Volunteer Moderators Online'. *Social Media + Society* 5(2):2056305119836778. doi: 10.1177/2056305119836778.

Matsuda, Mari J. 1993. *Words That Wound: Critical Race Theory, Assaultive Speech, And The First Amendment*. 1st edition. Boulder, Colo: Westview Press.

Matsuda, Mari, Charles Lawrence, Richard Delgado, and Kimberlé Crenshaw. 1993. 'Words That Wound: Critical Race Theory, Assaultive Speech, and The First Amendment'. *Westview Press*.

Mayntz, Renate. 2003. 'New Challenges to Governance Theory'. in *Governance as Social and Political Communication*. Manchester University Press.

McMillan Cottom, Tressie. 2020. 'Where Platform Capitalism and Racial Capitalism Meet: The Sociology of Race and Racism in the Digital Society'. *Sociology of Race and Ethnicity* 6(4):441–49.

Meaker, Morgan. 2022. 'Germany Has Picked a Fight With Telegram'. *Wired UK*, February 3.

Merriam Webster. 2023. 'Definition of Stewardship'. Retrieved 14 March 2023 (https://www.merriam-webster.com/dictionary/stewardship).

Meta. 2023. 'Facebook Community Standards: Hate Speech'. Retrieved 15 March 2023 (https://transparency.fb.com/en-gb/policies/community-standards/hate-speech/).

Meyer, Ilan H. 2003. 'Prejudice, Social Stress, and Mental Health in Lesbian, Gay, and Bisexual Populations: Conceptual Issues and Research Evidence'. *Psychological Bulletin* 129(5):674–97. doi: 10.1037/0033-2909.129.5.674.

Mihalcik, Carrie. 2019. 'Apple Pulls HKmap.Live App Used in Hong Kong Protests'. *CNET*. Retrieved 20 April 2023 (https://www.cnet.com/tech/mobile/apple-pulls-hkmap-live-app-used-in-hong-kong-protests/).

Milano, Silvia, Mariarosaria Taddeo, and Luciano Floridi. 2019. 'Recommender Systems and Their Ethical Challenges'. *Available at SSRN 3378581*.

Milano, Silvia, Mariarosaria Taddeo, and Luciano Floridi. 2020. 'Recommender Systems and Their Ethical Challenges'. *AI & SOCIETY*. doi: 10.1007/s00146-020-00950-y.

Miles, Tom. 2018. 'U.N. Investigators Cite Facebook Role in Myanmar Crisis'. *Reuters*, March 12.

Miller, Edie. 2018. 'Why Is British Media so Transphobic?' *The Outline*. Retrieved 22 June 2022 (https://theoutline.com/post/6536/british-feminists-media-transphobic).

Miller, Ryan A. 2017. '"My Voice Is Definitely Strongest in Online Communities": Students Using Social Media for Queer and Disability Identity-Making'. *Journal of College Student Development* 58(4):509–25. doi: 10.1353/csd.2017.0040.

Miškolci, Jozef, Lucia Kováčová, and Edita Rigová. 2020. 'Countering Hate Speech on Facebook: The Case of the Roma Minority in Slovakia'. *Social Science Computer Review* 38(2):128–46. doi: 10.1177/0894439318791786.

Moore, Martin, and Damian Tambini. 2018. *Digital Dominance: The Power of Google, Amazon, Facebook, and Apple*. Oxford University Press.

Morar, David. 2018. *Perspectives on Internet Governance: (Why) Does Internet Governance Matter? SSRN Scholarly Paper*. ID 3173935. Rochester, NY: Social Science Research Network. doi: 10.2139/ssrn.3173935.

Moreno, Johan. 2021. 'Google Estimated To Be Paying $15 Billion To Remain Default Search Engine On Safari'. *Forbes*. Retrieved 30 March 2023 (https://www.forbes.com/sites/johanmoreno/2021/08/27/google-estimated-to-be-paying-15-billion-to-remain-default-search-engine-on-safari/).

Morley, Jessica, Josh Cowls, Mariarosaria Taddeo, and Luciano Floridi. 2020. 'Ethical Guidelines for COVID-19 Tracing Apps'. *Nature*.

Morozov, Evgeny. 2011. *The Net Delusion: How Not to Liberate the World*. Penguin UK.

Mudde, Cas. 2016. *The Populist Radical Right: A Reader*. Taylor & Francis.

Müller, Karsten, and Carlo Schwarz. 2021. 'Fanning the Flames of Hate: Social Media and Hate Crime'. *Journal of the European Economic Association* 19(4):2131–67.

Munn, Luke. 2019. 'Alt-Right Pipeline: Individual Journeys to Extremism Online'. *First Monday*. doi: 10.5210/fm.v24i6.10108.

Munn, Luke. 2021. 'More than a Mob: Parler as Preparatory Media for the U.S. Capitol Storming'. *First Monday*. doi: 10.5210/fm.v26i3.11574.

Musiani, Francesca. 2015. 'Practice, Plurality, Performativity, and Plumbing: Internet Governance Research Meets Science and Technology Studies'. *Science, Technology, & Human Values* 40(2):272–86.

Musiani, Francesca. 2020. 'Science and Technology Studies Approaches to Internet Governance: Controversies and Infrastructures as Internet Politics'. *Researching Internet Governance: Methods, Frameworks, Futures* 85–104.

Musiani, Francesca, Derrick L. Cogburn, Laura DeNardis, and Nanette S. Levinson. 2016. *The Turn to Infrastructure in Internet Governance*. Springer.

Myers West, Sarah. 2018. 'Censored, Suspended, Shadowbanned: User Interpretations of Content Moderation on Social Media Platforms'. *New Media & Society* 20(11):4366–83. doi: 10.1177/1461444818773059.

Nagel, Thomas. 2005. 'The Problem of Global Justice'. *Philosophy & Public Affairs* 33(2):113–47.

Nash, Victoria. 2019. 'Revise and Resubmit? Reviewing the 2019 Online Harms White Paper'. *Journal of Media Law* 0(0):1–10. doi: 10.1080/17577632.2019.1666475.

Nast, Condé. 2017. 'We Need to Talk About Digital Blackface in GIFs'. *Teen Vogue*. Retrieved 7 February 2022 (https://www.teenvogue.com/story/digital-blackface-reaction-gifs).

Natale, Simone, Paolo Bory, and Gabriele Balbi. 2019. 'The Rise of Corporational Determinism: Digital Media Corporations and Narratives of Media Change'. *Critical Studies in Media Communication* 36(4):323–38. doi: 10.1080/15295036.2019.1632469.

Naylor, Brian. 2021. 'Read Trump's Jan. 6 Speech, A Key Part Of Impeachment Trial'. *NPR*, February 10.

Newhouse, Alex. 2020. 'Parler Is Bringing Together Mainstream Conservatives, Anti-Semites and White Supremacists as the Social Media Platform Attracts Millions of Trump Supporters'.

The Conversation. Retrieved 3 April 2023 (http://theconversation.com/parler-is-bringing-together-mainstream-conservatives-anti-semites-and-white-supremacists-as-the-social-media-platform-attracts-millions-of-trump-supporters-150439).

Ng, Lynnette Hui Xian, Iain Cruickshank, and Kathleen M. Carley. 2021. 'Coordinating Narratives and the Capitol Riots on Parler'. *ArXiv Preprint ArXiv:2109.00945*.

Ng, Lynnette Hui Xian, Iain J. Cruickshank, and Kathleen M. Carley. 2022. 'Cross-Platform Information Spread during the January 6th Capitol Riots'. *Social Network Analysis and Mining* 12(1):133.

Nicas, Jack. 2019. 'Apple Removes App That Helps Hong Kong Protesters Track the Police'. *The New York Times*, October 10.

Nicas, Jack, Raymond Zhong, and Daisuke Wakabayashi. 2021. 'Censorship, Surveillance and Profits: A Hard Bargain for Apple in China'. *The New York Times*, May 17.

Nicholas, Josh. 2023. 'Elon Musk Drove More than a Million People to Mastodon – but Many Aren't Sticking around | Mastodon | The Guardian'. *The Guardian*. Retrieved 16 April 2023 (https://www.theguardian.com/news/datablog/2023/jan/08/elon-musk-drove-more-than-a-million-people-to-mastodon-but-many-arent-sticking-around).

Nicolás, Elena Sánchez. 2021. '"Big Five" Tech Giants Spent €19m Lobbying EU in 2020'. *EUobserver*. Retrieved 13 April 2021 (https://euobserver.com/science/151072).

Nieborg, David B., Chris J. Young, and Daniel Joseph. 2020. 'App Imperialism: The Political Economy of the Canadian App Store'. *Social Media + Society* 6(2):2056305120933293. doi: 10.1177/2056305120933293.

Noble, Safiya Umoja. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.

Nockleby, John T. 2000. 'Hate Speech'. *Encyclopedia of the American Constitution* 3(2):1277–79.

Noor, Nahiyan Bin, Niloofar Yousefi, Billy Spann, and Nitin Agarwal. 2023. 'Comparing Toxicity Across Social Media Platforms for COVID-19 Discourse'.

Norgaard, Julia R., and Harold Walbert. 2023. 'Group Polarization?: An Analysis of Parler Data in the Wake of the Capitol Riot'. *Journal of Entrepreneurship and Public Policy* ahead-of-print(ahead-of-print). doi: 10.1108/JEPP-08-2022-0087.

Ntim, Zac. 2018. 'Mumsnet Brings in Tougher Forum Rules after Transgender Row'. *The Guardian*, June 13.

Open Handset Alliance. 2007. 'Industry Leaders Announce Open Platform for Mobile Devices | Open Handset Alliance'. Retrieved 11 February 2022 (http://www.openhandsetalliance.com/press_110507.html).

Ortiz, Stephanie M. 2020. 'Trolling as a Collective Form of Harassment: An Inductive Study of How Online Users Understand Trolling'. *Social Media + Society* 6(2):2056305120928512. doi: 10.1177/2056305120928512.

Oxford Analytica. 2021a. 'Apple's Parler Move Will Feed US Debate on Section 230'. *Emerald Expert Briefings* (oxan-es).

Oxford Analytica. 2021b. 'Parler Case Ups Scrutiny of IT Services Providers'. *Emerald Expert Briefings* (oxan-es).

Oxford Analytica. 2021c. 'Parler-Amazon Rift Fuels Politicisation of IT Services'. *Emerald Expert Briefings* (oxan-db).

Paczkowski, John, and Ryan Mac. 2021. 'Amazon Will Suspend Hosting For Pro-Trump Social Network Parler'. *BuzzFeed News*. Retrieved 17 March 2023 (https://www.buzzfeednews.com/article/johnpaczkowski/amazon-parler-aws).

Padovani, Claudia, Francesca Musiani, and Elena Pavan. 2010. 'Investigating Evolving Discourses on Human Rights in the Digital Age'. *International Communication Gazette* 72(4–5):359–78.

Page, Matthew J., Joanne E. McKenzie, Patrick M. Bossuyt, Isabelle Boutron, Tammy C. Hoffmann, Cynthia D. Mulrow, Larissa Shamseer, Jennifer M. Tetzlaff, Elie A. Akl, and Sue E. Brennan. 2021. 'The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews'. *International Journal of Surgery* 88:105906.

Parekh, Bhikhu. 2006. 'Hate Speech: Is There a Case for Banning?' *Public Policy Research* 12(4):213–23. doi: 10.1111/j.1070-3535.2005.00405.x.

Pariser, Eli. 2022. 'Musk's Twitter Will Not Be the Town Square the World Needs'. *Wired*, October 28.

Park, Jaihyun, JungHwan Yang, Katherine Bunsold, and Amada Tolbert. 2022. 'Cross-Platform Analysis of Twitter and Parler during the 2020 US Presidential Election'.

Parliament of Australia. 2019. 'Criminal Code Amendment (Sharing of Abhorrent Violent Material) Bill 2019'. Retrieved 2 April 2023 (https://www.aph.gov.au/Parliamentary_Business/Bills_Legislation/Bills_Search_Results/Result?bId=s1201).

Paul, Kari. 2019. 'A Brutal Year: How the "techlash" Caught up with Facebook, Google and Amazon'. *The Guardian*, December 28.

Peikoff, Amy L. 2008. 'Beyond Reductionism: Reconsidering the Right to Privacy'. *New York University Journal of Law & Liberty* 3:1.

Perez, Sarah. 2020. 'In Antitrust Hearing, Zuckerberg Admits Facebook Has Copied Its Competition'. *TechCrunch*. Retrieved 5 May 2021 (https://social.techcrunch.com/2020/07/29/in-antitrust-hearing-zuckerberg-admits-facebook-has-copied-its-competition/).

Perez, Sarah. 2021. 'Apple Confirms It Will Allow Parler to Return to App Store'. *TechCrunch*. Retrieved 14 March 2023 (https://techcrunch.com/2021/04/19/apple-confirms-it-will-allow-parler-to-return-to-app-store/).

Perrigo, Billy. 2022. 'Inside Facebook's African Sweatshop'. *Time*. Retrieved 2 April 2023 (https://time.com/6147458/facebook-africa-content-moderation-employee-treatment/).

Persily, Nathaniel. 2017. 'Can Democracy Survive the Internet?' *Journal of Democracy* 28(2):63–76. doi: 10.1353/jod.2017.0025.

Peter, Fabienne. 2017. 'Political Legitimacy'. in *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta. Metaphysics Research Lab, Stanford University.

Peters, Jay, and Kim Lyons. 2021. 'Apple Removes Parler from the App Store'. *The Verge*. Retrieved 21 September 2022 (https://www.theverge.com/2021/1/9/22221730/apple-removes-suspends-bans-parler-app-store).

Phillips, Whitney. 2015. *This Is Why We Can't Have Nice Things: Mapping the Relationship Between Online Trolling and Mainstream Culture*. MIT Press.

Pieroni, Emma, Peter Jachim, Nathaniel Jachim, and Filipo Sharevski. 2021. 'Parlermonium: A Data-Driven Ux Design Evaluation of the Parler Platform'. *ArXiv Preprint ArXiv:2106.00163*.

Pierre, Jon. 2000. *Debating Governance: Authority, Steering, and Democracy*. OUP Oxford.

Pierre, Jon, and B. Guy Peters. 2020. *Governance, Politics and the State*. Bloomsbury Publishing.

Plantin, Jean-Christophe, Carl Lagoze, Paul N. Edwards, and Christian Sandvig. 2018. 'Infrastructure Studies Meet Platform Studies in the Age of Google and Facebook'. *New Media & Society* 20(1):293–310.

Plantin, Jean-Christophe, and Aswin Punathambekar. 2019. 'Digital Media Infrastructures: Pipes, Platforms, and Politics'. *Media, Culture & Society* 41(2):163–74. doi: 10.1177/0163443718818376.

Poell, Thomas, David Nieborg, and José Van Dijck. 2019. 'Platformisation'. *Internet Policy Review* 8(4):1–13.

Pohle, Julia, and Thorsten Thiel. 2020. 'Digital Sovereignty'. *Internet Policy Review* 9(4).

Poole, Elizabeth, Eva Haifa Giraud, and Ed de Quincey. 2021. 'Tactical Interventions in Online Hate Speech: The Case Of# StopIslam'. *New Media & Society* 23(6):1415–42.

Porter, Jon. 2018. 'Tumblr Was Removed from Apple's App Store over Child Pornography Issues'. *The Verge*. Retrieved 14 March 2023 (https://www.theverge.com/2018/11/20/18104366/tumblr-ios-app-child-pornography-removed-from-app-store).

Porter, Jon, and James Vincent. 2022. 'Apple Would Be Forced to Allow Sideloading and Third-Party App Stores under New EU Law'. *The Verge*. Retrieved 29 March 2022 (https://www.theverge.com/2022/3/25/22996248/apple-sideloading-apps-store-third-party-eu-dma-requirement).

Poulton, Michael S. 2005. 'Organizational Storytelling, Ethics and Morality: How Stories Frame Limits of Behavior in Organizations'. *EJBO-Electronic Journal of Business Ethics and Organization Studies*.

Powderly, Timothy. 2023.

Price, Rob. 2017. 'Google's App Store Has Banned Gab — a Social Network Popular with the Far-Right — for "Hate Speech"'. *Business Insider*. Retrieved 14 March 2023 (https://www.businessinsider.com/google-app-store-gab-ban-hate-speech-2017-8).

Puddu, ANGELICA MARIA. 2021. 'A Tale of Two OSNs: Online Political Debate around Capitol Hill Assault on Twitter and Parler'.

Rahman, K. Sabeel. 2017. 'The New Utilities: Private Power, Social Infrastructure, and the Revival of the Public Utility Concept'. *Cardozo L. Rev.* 39:1621.

Rahman, K. Sabeel, and Kathleen Thelen. 2019. 'The Rise of the Platform Business Model and the Transformation of Twenty-First-Century Capitalism'. *Politics & Society* 47(2):177–204.

Ravazzi, Chiara, Francesco Malandrino, and Fabrizio Dabbene. 2022. 'Towards Proactive Moderation of Malicious Content via Bot Detection in Fringe Social Networks'. *IEEE Control Systems Letters* 6:2960–65.

Reichelmann, Ashley, James Hawdon, Matt Costello, John Ryan, Catherine Blaya, Vicente Llorent, Atte Oksanen, Pekka Räsänen, and Izabela Zych. 2021. 'Hate Knows No Boundaries: Online Hate in Six Nations'. *Deviant Behavior* 42(9):1100–1111.

Reuters. 2018. 'Aptoide Wins Court Battle against Google in Landmark Case'. *Reuters*, October 22.

Reuters. 2022. 'Dutch Regulator Levies 10th Fine on Apple in Dating App Row, Assessing New Proposal'. *Reuters*, March 28.

Rider, Karina, and David Murakami Wood. 2019. 'Condemned to Connection? Network Communitarianism in Mark Zuckerberg's "Facebook Manifesto"'. *New Media & Society* 21(3):639–54. doi: 10.1177/1461444818804772.

Rieger, Diana, Anna Sophie Kümpel, Maximilian Wich, Toni Kiening, and Georg Groh. 2021. 'Assessing the Extent and Types of Hate Speech in Fringe Communities: A Case Study of Alt-Right Communities on 8chan, 4chan, and Reddit'. *Social Media + Society* 7(4):20563051211052904. doi: 10.1177/20563051211052906.

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

Robertson, Adi. 2018. 'Facebook Used to Be a Democracy — but Nobody Voted'. *The Verge*. Retrieved 11 March 2021 (https://www.theverge.com/2018/4/5/17176834/mark-zuckerberg-facebook-democracy-governance-vote-failure).

Robertson, Adi. 2021. 'Judge Orders Apple to Allow External Payment Options for App Store by December 9th, Denying Stay'. *The Verge*. Retrieved 29 April 2022 (https://www.theverge.com/2021/11/9/22773082/epic-apple-fortnite-lawsuit-ruling-injunction-stay-app-store-anti-steering-rules).

Roccas, Sonia, and Marilynn B. Brewer. 2002. 'Social Identity Complexity'. *Personality and Social Psychology Review* 6(2):88–106.

Rodriguez, Salvador. 2016. 'Gab, the Alt-Right's Favorite Social Network, Gets Rejections From Apple, Twitter'. *Inc.Com*. Retrieved 12 April 2023 (https://www.inc.com/salvador-rodriguez/gab-apple-twitter.html).

Rogers, Katie, and Nicholas Fandos. 2019. 'Trump Tells Congresswomen to "Go Back" to the Countries They Came From'. *The New York Times*, July 14.

Rogers, Richard. 2020. 'Deplatforming: Following Extreme Internet Celebrities to Telegram and Alternative Social Media'. *European Journal of Communication* 35(3):213–29.

Romano, Aja. 2017. 'Neo-Nazi Site Daily Stormer Resurfaces with Russian Domain Following Google and GoDaddy Bans'. *Vox*. Retrieved 10 February 2022 (https://www.vox.com/culture/2017/8/16/16156210/daily-stormer-russian-domain-godaddy-google-ban).

Rondeaux, Candace, Ben Dalton, Cuong Nguyen, Michael Simeone, Thomas Taylor, and Shawn Walker. 2023. 'Parler and the Road to the Capitol Attack'. *New America*. Retrieved 20 April 2023 (http://newamerica.org/future-frontlines/reports/parler-and-the-road-to-the-capitol-attack/).

Roose, Kevin. 2019. 'Why Banning 8chan Was So Hard for Cloudflare: "No One Should Have That Power"'. *The New York Times*, August 5.

Rose, Gillian. 2022. 'Visual Methodologies: An Introduction to Researching with Visual Materials'. *Visual Methodologies* 1–100.

Rosenau, James N. 2009. 'Governance in the Twenty-First Century'. Pp. 7–40 in *Palgrave Advances in Global Governance*. Springer.

Rosenau, James N., Ernst-Otto Czempiel, and Steve Smith. 1992. *Governance Without Government: Order and Change in World Politics*. Cambridge University Press.

Röttger, Paul, Bertie Vidgen, Dong Nguyen, Zeerak Waseem, Helen Margetts, and Janet Pierrehumbert. 2021. 'HateCheck: Functional Tests for Hate Speech Detection Models'. Pp. 41–58 in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Online: Association for Computational Linguistics.

Ruddock, Jenna, and Justin Sherman. 2021. 'Widening the Lens on Content Moderation'. *Joint PIJIP/TLS Research Paper Series*.

Ruediger, Marco Aurélio, Amaro Grassi, Victor Piaia, Danilo Carvalho, and Polyana Barboza. 2021. *The Global Far Right: Brazil Establishes Own Ecosystem on Parler and Mimics American Far-Right Wing*.

Ryan, Celine. 2018. 'Conservatives Burned By Twitter Flock To New Platform'. *The Daily Caller*. Retrieved 18 May 2022 (https://dailycaller.com/2018/12/12/conservatives-twitter-alternative-platform-parler-news/).

Salminen, Joni, Fabio Veronesi, Hind Almerekhi, Soon-Gvo Jung, and Bernard J. Jansen. 2018. 'Online Hate Interpretation Varies by Country, but More by Individual: A Statistical Analysis Using Crowdsourced Ratings'. Pp. 88–94 in *2018 fifth international conference on social networks analysis, management and security (snams)*. IEEE.

Samples, John. 2018. 'Fighting Words and Free Speech | Cato at Liberty Blog'. *Cato Institute*. Retrieved 23 March 2023 (https://www.cato.org/blog/fighting-words-free-speech).

Schafer, Valérie, Francesca Musiani, and Marguerite Borelli. 2016. 'Web Archiving, Governance and STS'. *French Journal of Media Research* 6:1–23.

Schmidt, Vivien A. 2018. 'Rethinking EU Governance: From 'Old'to 'New'Approaches to Who Steers Integration'. *JCMS: Journal of Common Market Studies* 56(7):1544–61.

Schneider, Volker. 2004. 'State Theory, Governance and the Logic of Regulation and Administrative Control'. *Governance in Europe* 25–41.

Schouten, Fredreka. 2018. 'Who Are Mega-Donors Bob and Rebekah Mercer, and Why Are They Influential?' *USA TODAY*. Retrieved 3 April 2023 (https://www.usatoday.com/story/news/politics/2017/03/07/who-are-trump-donors-bob-and-rebekah-mercer/98812284/).

Schreckinger, Ben. 2019. 'Amid Censorship Fears, Trump Campaign "checking out" Alternative Social Network'. *POLITICO*. Retrieved 18 May 2022 (https://www.politico.com/story/2019/05/28/trump-campaign-twitter-1345357).

Schultz, Marisa. 2020. 'Facebook, Twitter Censors Trump "at Their Peril," GOP Official Says'. *Fox News*. Retrieved 3 April 2023 (https://www.foxnews.com/politics/facebook-twitter-censors-trump-at-their-peril-gop-official).

Schulz, Wolfgang. 2018. 'Regulating Intermediaries to Protect Privacy Online–the Case of the German NetzDG'. *Personality and Data Protection Rights on the Internet, Forthcoming*.

Scrivens, Ryan, and Amarnath Amarasingam. 2020. 'Haters Gonna "like": Exploring Canadian Far-Right Extremism on Facebook'. *Digital Extremisms: Readings in Violence, Radicalisation and Extremism in the Online Space* 63–89.

Seering, Joseph, Tony Wang, Jina Yoon, and Geoff Kaufman. 2019. 'Moderator Engagement and Community Development in the Age of Algorithms'. *New Media & Society* 21(7):1417–43. doi: 10.1177/1461444818821316.

Sellars, Andrew. 2016. 'Defining Hate Speech'. *Berkman Klein Center Research Publication* (2016–20):16–48.

Sharon, Tamar. 2020. 'Blind-Sided by Privacy? Digital Contact Tracing, the Apple/Google API and Big Tech's Newfound Role as Global Health Policy Makers'. *Ethics and Information Technology*. doi: 10.1007/s10676-020-09547-x.

Shephard, Alex. 2023. 'What Is Elon Musk Building?' *The New Republic*, March 21.

Shieber, Jonathan. 2021. 'Parler Jumps to No. 1 on App Store after Facebook and Twitter Ban Trump | TechCrunch'. Retrieved 20 May 2022 (https://techcrunch.com/2021/01/09/parler-jumps-to-no-1-on-app-store-after-facebook-and-twitter-bans/).

Siapera, Eugenia, Elena Moreo, and Jiang Zhou. 2018. *Hate Track: Tracking and Monitoring Online Racist Speech*. Irish Human Rights and Equality Commission.

Siegel, Alexandra A. 2020. 'Online Hate Speech'. *Social Media and Democracy: The State of the Field, Prospects for Reform* 56–88.

Siegler, MG. 2010. 'Steve Jobs Reiterates: "Folks Who Want Porn Can Buy an Android Phone"'. *TechCrunch*. Retrieved 10 February 2022 (https://social.techcrunch.com/2010/04/19/steve-jobs-android-porn/).

Simpson, Robert Mark. 2013. 'Dignity, Harm, and Hate Speech'. *Law and Philosophy* 32(6):701–28.

Sipka, Andrea, Aniko Hannak, and Aleksandra Urman. 2021. 'Comparing the Language of QAnon-Related Content on Parler, Gab, and Twitter'. *ArXiv Preprint ArXiv:2111.11118*.

Smith, Adam. 2020. 'Parler Reveals Strange Things It Bans as People Suspended from Twitter Head to Rival App'. *The Independent*. Retrieved 21 March 2022 (https://www.independent.co.uk/tech/parler-app-rules-ban-social-media-twitter-ceo-a9595606.html).

Soral, Wiktor, Micha\l Bilewicz, and Miko\laj Winiewski. 2018. 'Exposure to Hate Speech Increases Prejudice through Desensitization'. *Aggressive Behavior* 44(2):136–46.

Srnicek, Nick. 2017. *Platform Capitalism*. John Wiley & Sons.

Stake, Robert E. 1995. *The Art of Case Study Research*. SAGE.

Stanford Internet Observatory. 2021. 'Parler's First 13 Million Users'. Retrieved 20 May 2022 (https://cyber.fsi.stanford.edu/io/news/sio-parler-contours).

Sterling, Toby. 2022. 'Dutch Watchdog Fines Apple $5.7 Million Again in App Store Dispute'. *Reuters*, February 7.

Stoker, Gerry. 1998. 'Governance as Theory: Five Propositions'. *International Social Science Journal* 50(155):17–28.

Stracqualursi, Veronica. 2020. 'Virginia Governor Signs Background Checks, "red Flag" and Other Gun Control Bills into Law | CNN Politics'. *CNN*. Retrieved 11 April 2023 (https://www.cnn.com/2020/04/10/politics/ralph-northam-signs-gun-bills/index.html).

Strange, Susan. 1996. *The Retreat of the State: The Diffusion of Power in the World Economy*. Cambridge University Press.

Streeter, Thomas. 2015. 'Steve Jobs, Romantic Individualism, and the Desire for Good Capitalism'. *International Journal of Communication* 9:19.

Strossen, Nadine. 2018. *Hate: Why We Should Resist It with Free Speech, Not Censorship*. Oxford University Press.

Suchman, Mark C. 1995. 'Managing Legitimacy: Strategic and Institutional Approaches'. *Academy of Management Review* 20(3):571–610.

Sullivan, Mark. 2020. 'I Joined Parler, the Right-Wing Echo Chamber's New Favorite Alt-Twitter'. *Fast Company*. Retrieved 26 May 2022 (https://www.fastcompany.com/90522049/i-joined-parler-the-right-wing-echo-chambers-new-favorite-alt-twitter).

Sumner, L. W. 2013. 'Review of The Harm in Hate Speech'. *Law and Philosophy* 32(2/3):377–83.

Suzor, Nicolas. 2018. 'Digital Constitutionalism: Using the Rule of Law to Evaluate the Legitimacy of Governance by Platforms'. *Social Media+ Society* 4(3):2056305118787812.

Suzor, Nicolas, and Rosalie Gillett. 2022. 'Self-Regulation and Discretion'. Pp. 259–79 in *Digital Platform Regulation: Global Perspectives on Internet Governance*. Springer International Publishing Cham.

Suzor, Nicolas, Tess Van Geelen, and Sarah Myers West. 2018. 'Evaluating the Legitimacy of Platform Governance: A Review of Research and a Shared Research Agenda'. *International Communication Gazette* 80(4):385–400. doi: 10.1177/1748048518757142.

Swenson, Kyle. 2019. 'Rebekah Mercer, the Billionaire Backer of Bannon and Trump, Chooses Sides'. *Washington Post*, January 5.

Swindells, Katherine, and George Corbineau. 2021. 'Revealed: The Army of Big Tech Lobbyists Targeting Capitol Hill'. *New Statesman*. Retrieved 5 May 2021 (https://www.newstatesman.com/business/companies/2021/02/revealed-army-big-tech-lobbyists-targeting-capitol-hill).

Tachaiya, Jakapun, Arman Irani, Kevin M. Esterling, and Michalis Faloutsos. 2021. 'SentiStance: Quantifying the Intertwined Changes of Sentiment and Stance in Response to an Event in Online Forums'. Pp. 361–68 in *Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*.

Tallberg, Jonas, Karin Bäckstrand, and Jan Aart Scholte. 2018. *Legitimacy in Global Governance: Sources, Processes, and Consequences*. Oxford University Press.

Taylor, Frank M. 2020. 'Parler Isn't Safe (Parlez-Vous Danger?) – Frank M Taylor'. *Frank M Taylor*. Retrieved 27 May 2022 (https://blog.frankmtaylor.com/2020/11/25/parler-isnt-safe-parlez-vous-danger/).

Taylor, Josh. 2021. 'Australian Watchdog Considers Regulating Apple and Google to Boost App Store Competition'. *The Guardian*, August 18.

Taylor, Josh. 2022. '"It's Not That Hard": Does Kicking Kiwi Farms off the Internet Prove Tech Firms Can Act against Hate Speech?' *The Guardian*, September 19.

Taylor, Linnet. 2021. 'Public Actors Without Public Values: Legitimacy, Domination and the Regulation of the Technology Sector'. *Philosophy & Technology*. doi: 10.1007/s13347-020-00441-4.

Tell MAMA. 2017. *Beyond the Incident — Outcomes for Victims of Anti-Muslim Prejudices*.

Ten Oever, Niels, Stefania Milan, and Davide Beraldo. 2020. '10 Studying Discourse in Internet Governance through Mailing-List Analysis'. *Researching Internet Governance: Methods, Frameworks, Futures* 231.

The Law Commission. 2014. *Hate Crime: Should the Current Offences Be Extended ?*

The Law Commission. 2020. *Hate Crime Laws: A Consultation Paper*.

Thiel, David, Renée DiResta, and Shelby Grossman. 2021. *Contours and Controversies of Parler*. Technical report, Stanford University.

Timberg, Craig. 2020. 'Parler and Gab, Two Conservative Social Media Sites, Keep Alleged Russian Disinformation up, despite Report'. *Washington Post*, October 7.

Timberg, Craig, and Drew Harwell. 2021. 'Pro-Trump Forums Erupt with Violent Threats Ahead of Wednesday's Rally against the 2020 Election'. *Washington Post*, January 10.

Timberg, Craig, Drew Harwell, and Rachel Lerman. 2020. 'Parler's Got a Porn Problem: Adult Businesses Target pro-Trump Social Network'. *Washington Post*, December 3.

Timberg, Craig, and Isaac Stanley-Becker. 2020. 'QAnon Learns to Survive — and Even Thrive — after Silicon Valley's Crackdown'. *Washington Post*, October 28.

Torene, Spencer, Andrew Follmann, Thomas Teague, Peter Chang, and Blake Howald. 2022. 'Automated Hashtag Hierarchy Generation Using Community Detection and the Shannon Diversity Index, with Applications to Twitter and Parler'. *International Journal of Semantic Computing* 16(04):473–96.

Trujillo, Milo, Maurício Gruppi, Cody Buntain, and Benjamin D. Horne. 2020. 'What Is BitChute? Characterizing the "Free Speech" Alternative to YouTube'.

Tufekci, Zeynep. 2018. 'YouTube, the Great Radicalizer'. *The New York Times*, June 8.

Turillazzi, Aina, Federico Casolari, Mariarosaria Taddeo, and Luciano Floridi. 2022. *The Digital Services Act: An Analysis of Its Ethical, Legal, and Social Implications*. *SSRN Scholarly Paper*. ID 4007389. Rochester, NY: Social Science Research Network. doi: 10.2139/ssrn.4007389.

Tusikov, Natasha. 2019a. *Chokepoints: Global Private Regulation on the Internet*. University of California Press.

Tusikov, Natasha. 2019b. 'Defunding Hate: PayPal's Regulation of Hate Groups'. *Surveillance & Society* 17(1/2):46–53.

Tusikov, Natasha. 2021. 'Censoring Sex: Payment Platforms' Regulation of Sexual Expression'. Pp. 63–79 in *Media and Law: Between Free Speech and Censorship*. Vol. 26. Emerald Publishing Limited.

Twitter Comms [@TwitterComms]. 2020. 'We Have Placed a Public Interest Notice on This Tweet from @realdonaldtrump. Https://T.Co/6RHX56G2zt'. *Twitter*. Retrieved 19 May 2022 (https://twitter.com/TwitterComms/status/1266267446979129345).

Twomey, Aisling. 2017. 'A Civil Society Perspective on Anti-Traveller and Anti-Roma Hate: Connecting Online to on the Street'. Pp. 355–66 in *Critical Perspectives on Hate Crime*. Springer.

Tworek, Heidi, and Paddy Leerssen. 2019. 'An Analysis of Germany's NetzDG Law'. *Transatlantic Working Group* 11.

United Nations. 1966. 'International Covenant on Civil and Political Rights'. *OHCHR*. Retrieved 31 March 2023 (https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights).

United Nations. 2019. *UNITED NATIONS STRATEGY AND PLAN OF ACTION ON HATE SPEECH*.

US Federal Communications Commission. 2019. 'Obscene, Indecent and Profane Broadcasts'.

US Federal Communications Commission. 2022. 'The FCC and Speech'. Retrieved 3 April 2023 (https://www.fcc.gov/consumers/guides/fcc-and-speech).

Vaara, Eero, and Janne Tienari. 2011. 'On the Narrative Construction of Multinational Corporations: An Antenarrative Analysis of Legitimation and Resistance in a Cross-Border Merger'. *Organization Science* 22(2):370–90.

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

Vallance, Chris, and Shiona McCallum. 2022. 'Online Safety Bill: Plan to Make Big Tech Remove Harmful Content Axed'. *BBC News*, November 28.

Van Dijck, José, Tim de Winkel, and Mirko Tobias Schäfer. 2021. 'Deplatformization and the Governance of the Platform Ecosystem'. *New Media & Society* 14614448211045662. doi: 10.1177/14614448211045662.

Van Eeten, Michel JG, and Milton Mueller. 2013. 'Where Is the Governance in Internet Governance?' *New Media & Society* 15(5):720–36.

Van Leeuwen, Theo. 2007. 'Legitimation in Discourse and Communication'. *Discourse & Communication* 1(1):91–112. doi: 10.1177/1750481307071986.

Vendelø, Morten Thanning. 1998. 'Narrating Corporate Reputation: Becoming Legitimate through Storytelling'. *International Studies of Management & Organization* 28(3):120–37.

Vidgen, Bertie, Tristan Thrush, Zeerak Waseem, and Douwe Kiela. 2021. 'Learning from the Worst: Dynamically Generated Datasets to Improve Online Hate Detection'. Pp. 1667–82 in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Online: Association for Computational Linguistics.

Vogel, David. 2008. 'Private Global Business Regulation'. *Annu. Rev. Polit. Sci.* 11:261–82.

Vogels, Emily A., Andrew Perrin, and Monica Anderson. 2020. 'Most Americans Think Social Media Sites Censor Political Viewpoints'. *Pew Research Center: Internet, Science & Tech*. Retrieved 7 March 2023 (https://www.pewresearch.org/internet/2020/08/19/most-americans-think-social-media-sites-censor-political-viewpoints/).

Volokh, Eugene. 2018. 'Then-Candidate Trump's Speech at Rally Wasn't Constitutionally Unprotected "Incitement" of Violence Against Protesters'. *Reason.Com*. Retrieved 23 March 2023 (https://reason.com/volokh/2018/09/11/then-candidate-trumps-speech-at-rally-wa/).

Wachs, Sebastian, and Michelle F. Wright. 2018. 'Associations between Bystanders and Perpetrators of Online Hate: The Moderating Role of Toxic Online Disinhibition'. *International Journal of Environmental Research and Public Health* 15(9):2030.

Wagner, Ben. 2018. 'Ethics as an Escape from Regulation: From Ethics-Washing to Ethics-Shopping'. *Being Profiling. Cogitas Ergo Sum* 1–7.

Wakefield, Jane. 2020. 'Google Starts Appeal against £2bn Shopping Fine'. *BBC News*, February 12.

Waldherr, Annie, Stephanie Geise, and Christian Katzenbach. 2019. 'Computational Communication Science| Because Technology Matters: Theorizing Interdependencies in Computational Communication Science With Actor–Network Theory'. *International Journal of Communication* 13(0):21.

Waldron, Jeremy. 2012. *The Harm in Hate Speech*. Harvard University Press.

Walter, Dawn. 2021. 'Why Infrastructure Power Is as Important as Platform Governance'. *Response-Ability Summit*. Retrieved 1 November 2022 (https://response-ability.tech/why-infrastructure-power-is-as-important-as-platform-governance/).

Ward, Ethan. 2021. 'Parlez-Vous Le Hate?: Examining Topics and Hate Speech in the Alternative Social Network Parler'. Master's Thesis, University of Waterloo.

Waseem, Zeerak. 2016. 'Are You a Racist or Am I Seeing Things? Annotator Influence on Hate Speech Detection on Twitter'. Pp. 138–42 in *Proceedings of the First Workshop on NLP and Computational Social Science*. Austin, Texas: Association for Computational Linguistics.

Waseem, Zeerak, and Dirk Hovy. 2016. 'Hateful Symbols or Hateful People? Predictive Features for Hate Speech Detection on Twitter'. Pp. 88–93 in *Proceedings of the NAACL Student Research Workshop*. San Diego, California: Association for Computational Linguistics.

Weatherbed, Jess. 2023. 'Parler Goes Offline after Being Sold to a New Owner'. *The Verge*. Retrieved 15 April 2023 (https://www.theverge.com/2023/4/14/23683207/parler-shutting-down-relaunch-aquisition-announcement-conservative-social-media).

Weber, Anne. 2009. 'Manual on Hate Speech'. *Council of Europe*. Retrieved 7 February 2022 (https://book.coe.int/en/human-rights-and-democracy/4198-pdf-manual-on-hate-speech.html).

Weber, Max. 1994. *Weber: Political Writings*. Cambridge University Press.

Weise, Karen, and Paul Mozur. 2021. 'LinkedIn to Shut Down Service in China, Citing "Challenging" Environment'. *The New York Times*, October 14.

Welch, Chris. 2018. 'Netflix Stops Offering In-App Subscriptions for New and Returning Customers on IOS'. *The Verge*. Retrieved 14 March 2023 (https://www.theverge.com/2018/12/28/18159373/netflix-in-app-subscriptions-iphone-ipad-ios-apple).

Wells, Georgia, Jeff Horwitz, and Aruna Viswanatha. 2020. 'Facebook CEO Mark Zuckerberg Stoked Washington's Fears About TikTok'. *Wall Street Journal*, August 23.

Westergren, Martin. 2016. 'The Political Legitimacy of Global Governance Institutions: A Justice-Based Account'. PhD Thesis, Department of Political Science, Stockholm University.

Whine, Michael. 2003. 'The Far Right on the Internet'. Pp. 219–38 in *The governance of cyberspace*. Routledge.

White, Aoife. 2021. 'Apple May Face Antitrust Complaint as EU Steps Up Spotify Probe'. *Bloomberg.Com*, March 4.

White, Wilson. 2022. 'Google and U.S. Developers Find Agreement over Google Play Store'. *Google*. Retrieved 14 March 2023 (https://blog.google/outreach-initiatives/public-policy/google-and-us-developers-find-agreement-over-google-play-store/).

Williams, Matthew L., and Pete Burnap. 2016. 'Cyberhate on Social Media in the Aftermath of Woolwich: A Case Study in Computational Criminology and Big Data'. *The British Journal of Criminology* 56(2):211–38. doi: 10.1093/bjc/azv059.

Williams, Matthew L., Pete Burnap, Amir Javed, Han Liu, and Sefa Ozalp. 2020. 'Hate in the Machine: Anti-Black and Anti-Muslim Social Media Posts as Predictors of Offline Racially and Religiously Aggravated Crime'. *The British Journal of Criminology* 60(1):93–117.

Williams, Robertson. 2018a. 'Permitted Speech'. *Parler [The Internet Archive]*. Retrieved 17 March 2023 (https://web.archive.org/web/20180709191600/https://home.parler.com/test-2/).

Williams, Robertson. 2018b. 'The Thing About Passwords – Parler'. *Parler [The Internet Archive]*. Retrieved 3 April 2023 (https://web.archive.org/web/20180709191430/https://home.parler.com/test-1/).

Windrum, Paul. 2004. 'Leveraging Technological Externalities in Complex Technologies: Microsoft's Exploitation of Standards in the Browser Wars'. *Research Policy* 33(3):385–94. doi: 10.1016/j.respol.2003.09.002.

Winter, Aaron. 2019. 'Online Hate: From the Far-Right to the 'Alt-Right'and from the Margins to the Mainstream'. Pp. 39–63 in *Online othering*. Springer.

Wong, Julia Carrie. 2019. 'Facebook to Be Fined $5bn for Cambridge Analytica Privacy Violations – Reports'. *The Guardian*, July 12.

Wood, Alex J., Mark Graham, Vili Lehdonvirta, and Isis Hjorth. 2019. 'Networked but Commodified: The (Dis) Embeddedness of Digital Labour in the Gig Economy'. *Sociology* 53(5):931–50.

Woods, Lorna. 2020. 'Coronavirus and Harm in Broadcast Content | Media@LSE'. Retrieved 7 February 2022 (https://blogs.lse.ac.uk/medialse/2020/12/17/coronavirus-and-harm-in-broadcast-content/).

Wypych, Michał, and Michał Bilewicz. 2022. 'Psychological Toll of Hate Speech: The Role of Acculturation Stress in the Effects of Exposure to Ethnic Slurs on Mental Health among Ukrainian Immigrants in Poland'. *Cultural Diversity and Ethnic Minority Psychology* No Pagination Specified-No Pagination Specified. doi: 10.1037/cdp0000522.

Yang, Zachary, Anne Imouza, Kellin Pelrine, Sacha Lévy, Jiewen Liu, Gabrielle Desrosiers-Brisebois, Jean-François Godbout, André Blais, and Reihaneh Rabbany. 2021. 'Online Partisan Polarization of COVID-19'. Pp. 893–901 in *2021 International Conference on Data Mining Workshops (ICDMW)*. IEEE.

Yin, Robert K. 2018. *Case Study Research and Applications: Design and Methods*. Sixth edition. Los Angeles: Sage.

Zannettou, Savvas, Barry Bradlyn, Emiliano De Cristofaro, Haewoon Kwak, Michael Sirivianos, Gianluca Stringini, and Jeremy Blackburn. 2018. 'What Is Gab: A Bastion of Free Speech

or an Alt-Right Echo Chamber'. Pp. 1007–14 in *Companion Proceedings of the The Web Conference 2018*.

Zarsky, Tal Z. 2014. 'Social Justice, Social Norms and the Governance of Social Media'. *Pace L. Rev.* 35:154.

Zeng, Jing, and D. Bondy Valdovinos Kaye. 2022. 'From Content Moderation to Visibility Moderation: A Case Study of Platform Governance on TikTok'. *Policy & Internet* 14(1):79–95.

Zeng, Jing, and Mike S. Schäfer. 2021. 'Conceptualizing "Dark Platforms". Covid-19-Related Conspiracy Theories on 8kun and Gab'. *Digital Journalism* 9(9):1321–43. doi: 10.1080/21670811.2021.1938165.

Ziems, Caleb, Bing He, Sandeep Soni, and Srijan Kumar. 2020. 'Racism Is a Virus: Anti-Asian Hate and Counterhate in Social Media during the Covid-19 Crisis'. *ArXiv Preprint ArXiv:2005.12423*.

Ziewitz, Malte, and Christian Pentzold. 2014. 'In Search of Internet Governance: Performing Order in Digitally Networked Environments'. *New Media & Society* 16(2):306–22.

Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for the Future at the New Frontier of Power*. Profile Books.

Zuckerman, Ethan, and Chand Rajendra-Nicolucci. 2021. 'Deplatforming Our Way to the Alt-Tech Ecosystem'. *Toward a Better Internet*. Retrieved 21 September 2022 (http://knightcolumbia.org/blog/deplatforming-our-way-to-the-alt-tech-ecosystem).

Appendices

Appendix 1: Own work cited

Journal articles

Cowls, J. and Ma, C. (2023). The internet's town square? A critical discourse analysis of Parler's "past imaginary". *Social Media + Society.*

Cowls, J. and Vidgen, B. (Under review). Platform Governance Terminologies: Online Hate Speech.

Cowls, J., Morley, J., and Floridi, L. (2023). App store governance: implications, limitations, and regulatory responses. *Telecommunications Policy* 41(1).

Cowls, J.*, Darius, P.*, Santistevan, D*. and Schramm, M.* (2022). Constitutional metaphors: Facebook's "supreme court" and the legitimation of platform governance. *New Media & Society.* (*All authors contributed equally.)

Morley, J., Cowls, J., Taddeo, M., & Floridi, L. (2020). Ethical guidelines for COVID-19 tracing apps. *Nature*, *582*(7810), 29-31.

Conference presentations

Au, Y., Cowls, J., and Ma, C. (2023, May). When Platforms are Gatekeepers: Exploring Metaphors of Gates and their Guardians. Selected for presentation at International Communication Association 2023 preconference on History of Digital Metaphors, Toronto, Canada.

Cowls, J., Darius, P., Santistevan, D, and Schramm, M. (2021). Constitutional Metaphors: Facebook's 'Supreme Court' and Platform Legitimation. Presented at the First Annual Conference of The Platform Governance Research Network, March 2021, virtual.

Cowls, J. (2019). The Silicon Cage: "Legitimate" governance 100 years after Weber. Presented at Data Power Conference, September 2019, Bremen, Germany.

Book chapters and blog posts

Cowls, J. and Morley, J. (2022). App store governance: the implications and limitations of duopolistic dominance.

Cowls J. (2022). Trump Central? App stores as a new front in the platform governance of Donald Trump. Oxford Internet Institute. https://www.oii.ox.ac.uk/news-events/news/trump-central-app-stores-as-a-new-front-in-the-platform-governance-of-donald-trump/

Appendix 2: List of articles found in systematised thematic review

Agarwal, Saharsh, Uttara M. Ananthakrishnan, and Catherine E. Tucker. 2022. 'Deplatforming and the Control of Misinformation: Evidence from Parler'.

Aliapoulios, Max, Emmi Bevensee, Jeremy Blackburn, Barry Bradlyn, Emiliano De Cristofaro, Gianluca Stringhini, and Savvas Zannettou. 2021a. 'A Large Open Dataset from the Parler Social Network'. Pp. 943–51 in *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 15.

Aliapoulios, Max, Emmi Bevensee, Jeremy Blackburn, Barry Bradlyn, Emiliano De Cristofaro, Gianluca Stringhini, and Savvas Zannettou. 2021b. 'An Early Look at the Parler Online Social Network'. *ArXiv Preprint ArXiv:2101.03820*.

Baines, Annalise, Muhammad Ittefaq, and Mauryne Abwao. 2021. '# Scamdemic,# Plandemic, Or# Scaredemic: What Parler Social Media Platform Tells Us about COVID-19 Vaccine'. *Vaccines* 9(5):421.

Bär, Dominik, Nicolas Pröllochs, and Stefan Feuerriegel. 2023. 'New Threats to Society from Free-Speech Social Media Platforms'.

Buntain, Cody, Martin Innes, Tamar Mitts, and Jacob Shapiro. 2023. 'Cross-Platform Reactions to the Post-January 6 Deplatforming'. *Journal of Quantitative Description: Digital Media* 3.

Busch, Christoph. 2022. 'Regulating the Expanding Content Moderation Universe: A European Perspective on Infrastructure Moderation'. *UCLA JL & Tech.* 27:32.

Catipon, Regina, and Onur Savas. 2021. 'A Different Story: How Conservative Narratives Diverge Between Twitter and Parler.' Pp. 15–22 in *Text2Story@ ECIR*.

Dasgupta, Nabarun. 2021. 'Countering Hesitancy and Misinformation on Side Effects to Complete the Course of COVID Vaccination'. *Journal of Patient Experience* 8:23743735211067310.

Dori-Hacohen, Shiri, Keen Sung, Jengyu Chou, and Julian Lustig-Gonzalez. 2021. 'Restoring Healthy Online Discourse by Detecting and Reducing Controversy, Misinformation, and Toxicity Online'. Pp. 2627–28 in *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*.

Esser, Arne C. 2021. 'How Does the Language of Corpora from Radicalized Communities Discovered on Parler Compare to Online Conversations on Twitter Regarding the 2021 Capitol Riots and Election Fraud?' Master's Thesis.

Fedoruk, Benjamin, Harrison Nelson, Russell Frost, and Kai Fucile Ladouceur. 2021. 'The Plebeian Algorithm: A Democratic Approach to Censorship and Moderation'. *JMIR Formative Research* 5(12):e32427.

Greenhalgh, Spencer, Daniel G. Krutka, and Shannon M. Oltmann. 2021. 'Gab, Parler, and (Mis) Educational Technologies: Reconsidering Informal Learning on Social Media Platforms'. *The Journal of Applied Instructional Design* 10(3).

Groeger, Lena V., J. Kao, A. Shaw, M. Syed, and M. Eliahou. 2021. 'What Parler Saw during the Attack on the Capitol. ProPublica'.

Guhathakurta, Dipanwita, Mallika Subramanian, Manvith Reddy, Shradha Sehgal, Tanvi Karandikar, Amogh Gulati, Udit Arora, Rajiv Ratn Shah, and Ponnurangam Kumaraguru. 2021. 'Capitol (Pat) Riots: A Comparative Study of Twitter and Parler'. *ArXiv Preprint ArXiv:2101.06914*.

Hitkul, Avinash Prabhu, Dipanwita Guhathakurta, Jivitesh jain, Mallika Subramanian, Manvith Reddy, Shradha Sehgal, Tanvi Karandikar, Amogh Gulati, Udit Arora, Rajiv Ratn Shah, and Ponnurangam Kumaraguru. 2021. *Capitol (Pat)Riots: A Comparative Study of Twitter and Parler*. arXiv:2101.06914. arXiv. doi: 10.48550/arXiv.2101.06914.

Horta Ribeiro, Manoel, Homa Hosseinmardi, Robert West, and Duncan J. Watts. 2023. 'Deplatforming Did Not Decrease Parler Users' Activity on Fringe Social Media'. *PNAS Nexus* 2(3):pgad035. doi: 10.1093/pnasnexus/pgad035.

Israeli, Abraham, and Oren Tsur. 2022. 'Going Extreme: Comparative Analysis of Hate Speech in Parler and Gab'. *ArXiv Preprint ArXiv:2201.11770*.

Jakubik, Johannes, Michael Vössing, Dominik Bär, Nicolas Pröllochs, and Stefan Feuerriegel. 2022. 'Online Emotions During the Storming of the US Capitol: Evidence from the Social Media Network Parler'. *ArXiv Preprint ArXiv:2204.04245*.

Janke, Sebastian. n.d. 'Analysis and Contextualization of the Emotions and the Narrative Created on Parler through Social Media Posts in the Days before the Storm on the U.S. Capitol on the 6th of January 2021'.

Johnson, Hailey, Karl Volk, Robert Serafin, Cinthya Grajeda, and Ibrahim Baggili. 2022. 'Alt-Tech Social Forensics: Forensic Analysis of Alternative Social Networking Applications'. *Forensic Science International: Digital Investigation* 42:301406.

Karell, Daniel, Andrew Linke, Edward Holland, and Edward Hendrickson. 2023. '"Born for a Storm": Hard-Right Social Media and Civil Unrest'. *American Sociological Review* 00031224231156190.

M. Otala, Jacqueline, Gillian Kurtic, Isabella Grasso, Yu Liu, Jeanna Matthews, and Golshan Madraki. 2021. 'Political Polarization and Platform Migration: A Study of Parler and Twitter Usage by United States of America Congress Members'. Pp. 224–31 in *Companion Proceedings of the Web Conference 2021*.

Munn, Luke. 2021. 'More than a Mob: Parler as Preparatory Media for the US Capitol Storming'. *First Monday*.

Ng, Lynnette Hui Xian, Iain Cruickshank, and Kathleen M. Carley. 2021. 'Coordinating Narratives and the Capitol Riots on Parler'. *ArXiv Preprint ArXiv:2109.00945*.

Ng, Lynnette Hui Xian, Iain J. Cruickshank, and Kathleen M. Carley. 2022a. 'Coordinating Narratives Framework for Cross-Platform Analysis in the 2021 US Capitol Riots'. *Computational and Mathematical Organization Theory* 1–17.

Ng, Lynnette Hui Xian, Iain J. Cruickshank, and Kathleen M. Carley. 2022b. 'Cross-Platform Information Spread during the January 6th Capitol Riots'. *Social Network Analysis and Mining* 12(1):133.

Noor, Nahiyan Bin, Niloofar Yousefi, Billy Spann, and Nitin Agarwal. 2023. 'Comparing Toxicity Across Social Media Platforms for COVID-19 Discourse'.

Norgaard, Julia R., and Harold Walbert. 2023. 'Group Polarization?: An Analysis of Parler Data in the Wake of the Capitol Riot'. *Journal of Entrepreneurship and Public Policy*.

Oxford Analytica. 2021a. 'Apple's Parler Move Will Feed US Debate on Section 230'. *Emerald Expert Briefings* (oxan-es).

Oxford Analytica. 2021b. 'Parler Case Ups Scrutiny of IT Services Providers'. *Emerald Expert Briefings* (oxan-es).

Oxford Analytica. 2021c. 'Parler-Amazon Rift Fuels Politicisation of IT Services'. *Emerald Expert Briefings* (oxan-db).

Park, Jaihyun, JungHwan Yang, Katherine Bunsold, and Amada Tolbert. 2022. 'Cross-Platform Analysis of Twitter and Parler during the 2020 US Presidential Election'.

Pieroni, Emma, Peter Jachim, Nathaniel Jachim, and Filipo Sharevski. 2021. 'Parlermonium: A Data-Driven Ux Design Evaluation of the Parler Platform'. *ArXiv Preprint ArXiv:2106.00163*.

Ravazzi, Chiara, Francesco Malandrino, and Fabrizio Dabbene. 2022. 'Towards Proactive Moderation of Malicious Content via Bot Detection in Fringe Social Networks'. *IEEE Control Systems Letters* 6:2960–65.

Ruediger, Marco Aurélio, Amaro Grassi, Victor Piaia, Danilo Carvalho, and Polyana Barboza. 2021. 'The Global Far Right: Brazil Establishes Own Ecosystem on Parler and Mimics American Far-Right Wing'.

Sipka, Andrea, Aniko Hannak, and Aleksandra Urman. 2021. 'Comparing the Language of QAnon-Related Content on Parler, Gab, and Twitter'. *ArXiv Preprint ArXiv:2111.11118*.

Tachaiya, Jakapun, Arman Irani, Kevin M. Esterling, and Michalis Faloutsos. 2021. 'SentiStance: Quantifying the Intertwined Changes of Sentiment and Stance in Response to an Event in Online Forums'. Pp. 361–68 in *Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*.

Thiel, David, Renée DiResta, and Shelby Grossman. 2021. *Contours and Controversies of Parler*. Technical report, Stanford University.

Torene, Spencer, Andrew Follmann, Thomas Teague, Peter Chang, and Blake Howald. 2022. 'Automated Hashtag Hierarchy Generation Using Community Detection and the Shannon Diversity Index, with Applications to Twitter and Parler'. *International Journal of Semantic Computing* 16(04):473–96.

VanDyke, David. n.d. 'Coded Data: Tracking Discursive Trends in the January 6 Parler Data'. Pp. 72–94 in *The Capitol Riots*. Routledge.

Ward, Ethan. 2021. 'Parlez-Vous Le Hate?: Examining Topics and Hate Speech in the Alternative Social Network Parler'. Master's Thesis, University of Waterloo.

Yang, Zachary, Anne Imouza, Kellin Pelrine, Sacha Lévy, Jiewen Liu, Gabrielle Desrosiers-Brisebois, Jean-François Godbout, André Blais, and Reihaneh Rabbany. 2021. 'Online Partisan Polarization of COVID-19'. Pp. 893–901 in *2021 International Conference on Data Mining Workshops (ICDMW)*. IEEE.

Appendix 3: Example of coding in mainstream media corpus

To demonstrate how I coded articles in the mainstream media corpus, here is one such passage, where paragraphs have been numbered for ease of discussion:

> [1] Parler, the new social media platform, has added about a million new users over the past two weeks, said John Matze, CEO of Parler, attributing the "mass migration" to "frustration with censorship" on Twitter. He offered his remarks on SiriusXM's Breitbart News Sunday with host Joel Pollak.
>
> [2] Parler describes itself as "a non-biased free speech driven entity" which "accepts your right to express your thoughts, opinions and ideals online."
>
> [3] Parler was the "number one [app] for news" across mobile platforms in recent days, Matze said. "We've seen an influx of probably almost a million users in week and a half to two weeks."
>
> [4] Parler added "a couple hundred thousand people overnight" after Twitter's recent banning of British conservative commentator Katie Hopkins, said Matze. Twitter banned Hopkins for "hateful conduct." The company also banned Carpe Donktum," a meme maker who mocks the left.
>
> [5] Matze considered, "It looks like a lot of people are using their Twitter handles and Twitter accounts to drive people to Parler because of their frustration with the censorship on that platform."

Each paragraph was considered, and coded, separately. First, direct quotes by Matze in ¶1 and 3-5, as well as the self-description of Parler in ¶2, were coded accordingly, as *Direct quote: Matze* and *Direct quote: Corporate communications* respectively. Next, the passage containing Matze's reference to Twitter, an *External entity: Rival platforms/Silicon Valley*, was coded as such. Lastly, various codes were applied from the *Legitimating themes* code group. Matze's references to "frustration with censorship" in ¶1 and 5 were coded as pertaining to *Freedom of expression*.[133] Finally, ¶2 was coded as concerning both *Freedom of expression* and *tTansparency, impartiality and equality* given the use of "non-biased". As such, the majority of passages in this excerpt received more than one code.

[133] Close reading of the passage reveals that the quote in ¶5 was simply a longer version of the same quote extracted in ¶1—a common journalistic practice. This may engender concern that repetitions such as this may skew the results reported below. However, those results report the number of articles in the dataset containing (at least) one use of a given code, rather than the total number of times a code was used.

Appendix 4: Queries sent to APIs for hate speech classification

Please see supplementary material (Appendices 4.1 and 4.2) submitted separately.