Who Should Regulate Extremist Content Online?

Alastair Reed and Adam Henschke

Abstract As liberal democracies grapple with the evolution of online political extremism, in addition to governments, social media and internet infrastructure companies have found themselves making more and more decisions about who gets to use their platforms, and what people say online. This raises the question that this paper explores, who should regulate extremist content online? In doing so the first part of the paper examines the evolution of the increasing role that social media and internet infrastructure companies have come to play in the regulating extremist content online, and the ethical challenges this presents. The second part of the paper explores three ethical challenges: i) the moral legitimacy of private actors, ii) the concentration of power in the hands of a few actors and iii) the lack of separation of powers in the content regulation process by private actors.

1 Framing the Problem

As liberal democracies grapple with the evolution of online political extremism, social media companies, and their supporting infrastructure, find themselves making more and more decisions about who gets to use their platforms, and what people say online. Moreover, the decision makers at these companies are increasingly uncomfortable with this power. In a public statement from 2018 Facebook CEO Mark Zuckerberg wrote "As I've thought about these content issues, I've increasingly come to believe that Facebook should not make so many important decisions about free expression and safety on our own" [94]. The CEO of Cloudflare, a company that supports the infrastructure of the internet, Matthew Prince, stated that as "the CEO of a major Internet infrastructure company", he could wake up in a bad mood and decide that "someone shouldn't be allowed on the Internet. No one should have that

A. Reed (✉)
Swansea University, Singleton Park, Swansea SA2 8PP, Wales, United Kingdom
e-mail: alastair.reed@swansea.ac.uk

A. Henschke
University of Twente, Enschede, Netherlands

A. Henschke et al. (eds.), *Counter-Terrorism, Ethics and Technology*,
Advanced Sciences and Technologies for Security Applications,
https://doi.org/10.1007/978-3-030-90221-6_11

power" [15, 63]. At the same time, governments and politicians question whether that Tech Companies are not doing enough to counter extremist content, hate speech, and misinformation online, and that these companies have a social responsibility to go further in their moderation of online content ([8, 53]; Bishop and Macdonald [81], 143).

The question of how and why social media companies and internet infrastructure came to find themselves the arbiters of online speech stems in part from counter-terrorism efforts, particularly the rise of the so-called Islamic State of Iraq and Syria (ISIS) in the mid 2010s. The issue of regulation of political extremists online is inseparable from the evolution of modern global terrorism. In this chapter, we follow the path of that recent history to highlight three ethical concerns arising from the responses to online political extremism. We then suggest that part of the problem with answering 'who should regulate extremist content online?' is that there are different aspects to how that content is being regulated. By reflecting on what sorts of institutions and services are being provided, we can suggest a more nuanced and collaborative approach to the regulation of online content.

Regulating extremist content online has faced two particular challenges. The first challenge is determining what extremist content *is*? [46]. Essentially what type of online content should be restricted and potentially removed online in the fight against extremism? This is not a straightforward task when we have no widely agreed upon definition of extremism or violent extremism. Too zealous an approach risks broadening the net too wide, and unnecessarily or disproportionately restricting individuals' rights and freedoms, whilst a narrower approach allows the free flowing of all but the most extreme of content. Where to draw the line is a controversial decision with no easy solutions [86]. The second is technical: how to *identify* extremist content online. Given the sheer scale and size of many platforms, looking for extremist content is like looking for the preverbal needle in a haystack.[1]

Though we recognise these points, this chapter looks at a less examined challenge of countering extremism propaganda online. Rather than questions of how this should be done, or what material is relevant, this chapter asks questions of *who* gets to make these decisions and *why*. As has been said, freedom of speech is as much about who gets to decide what is said as what is actually said [34].

2 The Status Quo: Regulation and Self-Regulation

The division of responsibilities between formal regulation and self-regulation varies across countries depending on local legal frameworks, regulators, and social norms. However, in most countries there is a legal framework that sets out what content is illegal and should be removed online, such as material from a proscribed terrorist organisation, hate speech, or child pornography. These laws then form the foundation

[1] Or, as others have noted, given that they are looking for data within data, it is more like looking for a needle in a pile of needles. See: [56].

on which social media companies base their content moderation on. However, the interpretation and enforcement of these laws are increasingly falling on social media companies themselves. As a result, these companies have developed large bureaucracies to monitor and regulate online speech on their platforms. As Jack Balkin argues, whilst in the past regulation of speech was targeted at the speakers, now governments' regulation is increasingly aimed at internet infrastructure (including social media companies), arguing that "in essence, nation-states attempt to get the privately owned infrastructure to do their work for them" ([5], 2015–16). This has led to a form private governance which now plays a central role in regulating extremist content online [5, 44]. In order to deal with the sheer volume of content uploaded daily the major social media companies have developed extensive technological solutions, often utilising artificial intelligence and machine learning,[2] to identify and remove offending content. In addition companies are increasingly collaborating through industry organisations such as the Global Internet Forum for Counter-Terrorism (GIFCT), for example a shared database of terrorist content [8, 26]. As the main platforms have developed their capabilities, they have become increasingly effective at identifying and removing offending content, as Bishop and MacDonald noted, in the case of the largest platforms—Facebook, YouTube and Twitter "referrals from users, law enforcement and governments are responsible for only a small minority of suspensions and take-downs; the vast majority of violations are detected by technology" ([8], 142).

The extent of responsibilities that legal frameworks place on social media companies varies across jurisdictions. In the United States, for example, first amendment free speech protections limit the legal scope of regulating extremist content online, compared to other liberal democracies in Europe or beyond. In Germany, the Network Enforcement Act (NetzDG) requires social media platforms to remove content that is "manifestly unlawful" within 24 h.[3] As Amélie Heldt notes, "[t]he obligation to remove unlawful content is in itself not problematic, but who gets to decide if user-generated content is "manifestly" unlawful? By delegating this task to social media platforms, the State has factually given the responsibility to decide upon the lawfulness of content to the reviewers in charge of content moderation" ([44], 342). Through seeking to make social media platforms more responsible for regulating the content on their platforms, governments have not only given these companies greater powers, but have effectively granted them the moral authority to develop, interpret, and enforce who says what online.

In practise, the regulation of content on a social media platform is governed by a wider set of community rules and guidelines, typically laid out in a platform's own term of service (ToS). As social media companies have evolved from conduits for hosting content to online communities to something more like traditional media,

[2] We note here that the use of artificial intelligence and machine learning in identifying and removing extremist content raises separate ethical challenges which are beyond the scope of this paper, but which the authors have addressed elsewhere. See: [46].

[3] The law applies to social media companies with more than two million users in Germany and allows up to seven days for cases that are less clear ([44], 341).

entertainment providers, and/or critical infrastructure, they have become increasingly involved in governing these communities by setting out in their ToS what content is or is not allowed on the platform. Importantly, as these ToS are largely self-generated, the content that is banned by a platform's ToS, can and regularly does go beyond what is merely unlawful. As platforms set out to govern their online communities, this now means not just upholding the law, but also writing the rules about what is acceptable within the community.

These ToS are often then used by law enforcement to request platforms to remove extremist content. Rather than proceeding down a more traditional legal route, with all the complexities that entails, law enforcement can request platforms to remove extremist content by flagging the content to the platforms, pointing out that it violates the platforms own ToS and therefore they should remove it on these grounds [58]. This is the basis on which the Europol's Internet Referral Unit (IRU) works on, which by itself has no enforcement powers. The IRU explains the implications of its referral process: "Thus the decision and removal of the referred terrorist content is taken by the concerned service provider under their own responsibility in reference to their terms of use" ([28], 6). In this situation, the content's removal is due to the material breaking the platform's own rules, and not because it was necessarily illegal.

This brings us to the question of the paper—*who* has the moral authority to make such decisions about what remains online? The role of regulating extremist content online has increasingly shifted from public authorities to private companies, and in the process, these companies now have significant capacity to decide what is allowed online. This brings into focus both the outsourcing to private companies to interpret, and enforce existing legislation, and also the grey area between content that is removed because it breaks a country's laws, and content that is not illegal but removed by private companies because it breaks their own terms of service. Both of these challenges raise questions about the legitimacy of private companies in playing these roles: as we discuss later, in liberal democracies we generally recognise that the state has some moral authority to make such decisions, but it is less clear if private companies have equivalent moral authority.

In response to public and political pressure, platforms can feel compelled to take further actions, and as a consequence update their ToS to ban from their platform a wider array of extremist content. This wider spectrum of what is perceived as extremist content has continued to evolve as perceptions and understandings of how extremists use the internet have developed. For example, Facebook had long banned white supremacy from its platform as 'hateful' content, however it was not until Facebook updated its rules after the Christchurch attacks, that it also included white nationalism and white separatism [30, 71]. As noted at the start of this chapter, this is a place that these CEOs and private companies neither wanted, nor indeed, expected to be in. Interestingly, this challenge has its genesis in the largely uncontroversial decision that terrorist groups like ISIS should not have free run of social media. Terrorism and the means used to combat it are casting a long shadow over the recent evolution of information and communication technologies.

3 Terrorism as a Driver for Deplatforming: From ISIS to Political Extremists

The rise of ISIS and its exploitation of propaganda via social media forced social media companies to overhaul their content moderation responses to tackle this new threat. After initial challenges, the major platforms such as Facebook and Twitter were largely successful in driving ISIS supporters off their platforms ([19], 108). However, in comparison to the challenges to come, the situation with ISIS was unique. First, due to their extreme violence and barbaric practices ISIS were almost universally condemned and a general consensus that ISIS and their propaganda should be confronted. Secondly, ISIS material is usually easily identifiable and clearly branded by the group, making its detection much simpler ([19], 108–9).

As Facebook's approaches to content removal developed, so did the list of organisations and individuals that it banned from its platform. However, the individuals and groups banned by Facebook were not just terrorists, but also were increasingly those that operate in the fringes of more familiar political beliefs [35]. In particular with the rise of new far-right movements the line between extremist and mainstream politics became increasingly blurred.

In March 2018 Facebook took the step of banning the British far-right group "Britain First" from the platform, removing its official home page and the pages of its leaders Paul Golding and Jayda Fransen [29, 49]. Stating that "[w]e do not do this lightly, but they have repeatedly posted content designed to incite animosity and hatred against minority groups, which disqualifies the Pages from our service" [29]. Whilst a fringe political organisation, the group had a large social media presence, with more than 1.8 million followers and 2 million likes on its Facebook page, more than double the amount of the Labour party (the mainstream party with the most likes) [49, 61]. The following year Facebook went further designating the group and its leaders under its new definition of 'dangerous groups and individuals', along with a number of other far-right organisations and individuals [89]. This designation also extended the ban to "[p]osts and other content which expresses praise or support for these figures and groups".[4] Although Britain First had ceased to be a political party a few months before the original ban, it is believed that the policy would apply to the proscribed individuals if they ran for or assumed political office in the future [47].

Following a 2018 ban from Facebook, Britain First launched legal action against the company for 'political discrimination' with the group's leader Paul Golding stating: "For too long now social networks have censored certain political viewpoints and thus interfered with the political process" [62].[5] After Facebook's removal of ads supportive of Britain First posted by a third organisation in January 2019,[6] the group accused the company of "political gerrymandering" [90]. A similar position was taken in January 2019 by Saoradh, a political party representing dissident Irish

[4] Facebook statement quoted in: [47].

[5] The group ultimately decided not to continue with the court case the following year [27].

[6] The Facebook adverts were bought by a page called 'Political Gamers TV' supporting a petition by Britain First to halt the reconstruction of a Mosque See: [40, 90].

republicans after Facebook removed its pages from its platform [4, 54]. As it sought a court order to re-instate its pages, Saoradh's lawyers explained: "Facebook has now taken to remove what they deem to be unacceptable political messages, that sets a very, very dangerous precedent and it's an attack, a deliberate attack, on the freedom of expression …Therefore our clients have no alternative but to seek injunctive relief to compel Facebook to uphold what is a very, very basic principle, the right to a political opinion and the right to expression" [54].

Putting aside the nature of Britain First's and Saoradh political views and the content of their material (which—given liberal democracy's commitment to political pluralism and free speech—some may find objectionable), the point the groups were making was clear; by choosing to de-platform them, Facebook was interfering in the political process. By exercising their power over which groups can espouse their political views on Facebook, the company had enormous power over what gets said and by whom. Whilst Britain First and Saoradh were on the fringes of mainstream politics, Facebook has also taken steps to ban organisations, which have 'one foot in the political mainstream', such as the Greek far-right political party 'Golden Dawn'[7] which faced a ban by the platform despite having elected members in both the national and European parliament [35, 79].

In September 2019, Facebook removed the account page of the Italian far-right group CasaPound from its platform, along with the pages of its representatives and supporters, on the grounds that they violated Facebook's Terms of Service by containing hate speech and content that amounted to incitement of violence [41]. In the court case that followed, Facebook was ordered to re-activate CasaPound's account page, with the court "setting a penalty of €800.00 for each day of violation" [41]. In the court's ruling it noted, among other points, that "[T]he exclusion of the applicants from Facebook is in contrast with the right to pluralism… eliminating or strongly compressing the possibility for association… to express its political messages." [72]. This ruling coheres with the view that constraints on free public expression of political beliefs is antithetical to liberal democratic commitments to free and pluralistic societies.

These cases highlight the complex intersection of competing rights that need to be balanced against each other. On the one hand, even in liberal democracies, it is legitimate to restrict content that constitutes hate speech and incitement of violence (Henschke Forthcoming). On the other hand, however, liberal democracies define themselves in part by reference to the right to free speech and political pluralism.[8] This brings us back to the motivating question: who has the authority to decide what extremist content is and what the appropriate responses should be? Should decisions that potentially impact on the public sphere be made by private companies? The near universal agreement that ISIS should be deplatformed has led us to the situation where (fringe) political parties are losing the capacity for public expression. And, while we might agree that the content, and political beliefs of Golden Dawn and the

[7] In October 2020 the leadership of Golden Dawn were convicted of running a criminal organization [6].

[8] For more on this, see: [75, 77, 82, 85].

like are not only objectionable, and perhaps dangerous for democracies, we are still left with the issue of whether social media companies are the right institutions to make decisions about who gets to speak in the new public squares.

These questions where thrown into sharp relief with the 2020 US presidential campaign and the series of events that ultimately led to US President Donald Trump's suspension from many social media platforms. In the run up to the 2020 US presidential elections saw social media platforms revising and updating their policies and terms of service [7]. Twitter in particular, employed extensive measures noting that on Twitter users find "real-time political conversation, resources, and breaking news. And an essential part of our service is taking action on content that attempts to manipulate, disrupt, or cause confusion about civic processes" [83]. As part of this approach Twitter began fact-checking the President's tweets, placing some behind warnings and labelling others as manipulated media or misleading [17, 84].

In the aftermath of the January 6th storming of the Capitol Building, Twitter, Facebook and most major platforms took steps to suspend or ban the President from their platforms [11, 48]. The decision was criticised by many, whilst others supported the actions taken [48]. The then U.S. Secretary of State Mike Pompeo tweeted "Silencing speech is dangerous. It's un-American. Sadly, this isn't a new tactic of the Left" [57]. At the same time, others highlighted the importance of moderating social media content to prevent misinformation, hate speech, and incitement of violence online, but still voiced unease at the process by which the President was suspended from social media.[9] The German Chancellor Angela Merkel, highlighted the importance of free opinion, and that while noting "[t]his fundamental right can be intervened in, but according to the law and within the framework defined by legislators—not according to a decision by the management of social media platforms" [2].

Putting aside any questions over the grounds for the decisions to suspend Trump from social media or questions over whether this decision was biased or politically motivated or not, we need again to ask, who gets to decide this, and why?

4 A Deeper Cut: De-Platforming the Platforms

The examples above have focussed on the control that social media platforms have on the content posted on their sites, and also over which individuals and organisations can post content on their platforms. Another type of online de-platforming that has recently emerged focuses not on removing individuals or organisations from a given platform, but literally removing the platform itself from the internet. Online platforms rely on a whole host of auxiliary services to be able to exist online. If these internet infrastructure service providers decide to remove their services, it can prevent the

[9] We note here that the governor of the US state of Florida has signed a bill to ban this sort of deplatforming of political actors [42].

platform itself from operating or operating effectively online.[10] The platforms can effectively be de-platformed.

In the wake of the white supremacist violence at a rally in the US city of Charlottesville in 2017, there was a rise in popular and political pressure for private companies to take further action against extremist material online [14, 33, 64]. Whilst the initial focus had been on social media companies to better regulate the content posted on their platforms, a new front opened up, "[t]hat front lies deeper within the web's infrastructure, in the realm of web hosts, domain registrars, and various other web services. The companies that provide the back-end services of the web have historically resisted pressure to police the behavior of sites that use them and have mostly avoided the spotlight in controversies over online speech" [64]. This marked a change at the level at which extremist content was 'regulated' on the internet, and highlights a wider debate about the role of internet infrastructure companies, that support the workings of the internet, and whether they should remain content neutral. In this view internet infrastructure companies are seen as the plumbing of the internet and should not be making decisions about content ([5, 10]; Balkin [39], 2038).

The first major change came in the wake of Charlottesville, when the domain registry service 'Go-Daddy' cancelled the registration of the neo-nazi site Daily Stormer. Daily Stormer briefly transferred to Google domains before being cancelled by the provider, which also banned it from YouTube, relegating the website to the dark web [9, 21, 63, 73]. As said, Cloudflare, a service that provides online infrastructure support, including protection from distributed denial of service (DDOS) attacks, soon followed suit [63, 69]. Whilst believing they made the right choice, the CEO Prince expressed unease about the power that Cloudflare could exert [15, 63]. Prince further highlighted that due to the ease at which online attacks could be orchestrated on websites, websites need the services of a network like Cloudflare. Otherwise, they would be at risk of being kicked off-line by anyone that they offend by their content, in practise allowing a form of vigilante justice to police the internet.[11] Going on to note the growing dependence on a few giant networks to provide these services, he argued that soon being online may mean relying on the services of a "company with a giant network like Cloudflare, Google, Microsoft, Facebook, Amazon, or Alibaba" highlighting that Cloudflare by itself already handles 10% of all internet requests [69].

Following the El Paso mass shooting in 2019, Cloudflare decided to take similar action terminating its contract with controversial online platform 8chan,[12] seen by

[10] This includes a spectrum of companies that provide services such as web hosting, domain name registries, security (i.e. DDoS protection), online payment processing, among other services. For more see: ([5, 39]).

[11] Prince notes that the initial demands for Cloudflare to terminate their contract with Daily Stormer came from hackers that wanted Cloudflare to '[g]et out of the way' so they could knock it off line with a DDoS attack [69].

[12] 8chan went offline after Cloudflare and other web infrastructure companies refused to provide it with the services in needed to remain online. However, it remerged 3 months later online as 8kun [16]; For more on 8chan's struggle to stay online see: [18].

many as an online haven for far-right and other extremist views ([20], 12–14, [93]). Prince noted that the El Paso shooter had apparently been inspired by 8chan and had posted a screed to the platform before the attack [70]. Furthermore, this he noted was not an isolated incident, highlighting similar activity earlier that year before both the Christchurch attack on two Mosques and the Poway synagogue attack by lone shooters. Similarly, Prince noted his unease at the arbitrary power the company had, writing "Cloudflare is not a government. While we've been successful as a company, that does not give us the political legitimacy to make determinations on what content is good and bad. Nor should it" [70].

These ethical challenges where highlighted again in the wake of the January 6th storming of the US Capitol Building in a series of events that forced 'free speech' social network Parler off-line. Positioning itself as a free speech alternative to social media platforms such as Twitter and Facebook, Parler had been one of the fastest growing apps in the preceding months. In the wake of the 2020 presidential election, as platforms like Twitter and Facebook clamped down on misinformation about who had won the elections, millions of conservatives migrated to alternative platforms such as Parler [51].

In the aftermath of the events of January 6th, and former US President Trump's perceived role in inciting violence numerous social media platforms including Twitter and Facebook took actions suspending the President's account [11, 48]. As the President and many of his followers sought to migrate to Parler as an alternative platform something unexpected happened. Apple and then Google suspended Parler from their App stores, for not taking sufficient action to police posts made on the Platform. This significantly limited Parler's ability to gain new followers. Shortly afterwards, Amazon Web Services terminated its contract with the platform for repeated violations of Amazon's rules in effect taking the platform offline [60]. Parler's Chief Executive Johan Matze, accused the tech giants of a '"coordinated effort" to "completely remove free speech off the internet"'[60].

We should not see the actions against Parler in isolation. They were part of wider reactions by private companies in the aftermath of Jan 6th to try to moderate far-right content, as well as mis/disinformation and conspiracy theories on their platforms. However, whilst companies such as Facebook and Twitter took action against content on their own platforms, what makes the case of Parler different, was that Apple, Google and Amazon, took actions against another platform for the content it hosted and a perceived failure to take sufficient action against extremist speech and actions.

Setting aside for one moment questions about the nature of the content on Parler, and whether Parler had or had not taken sufficient action, we have a situation where in effect, a small group of private companies through their actions de-platformed a social media platform over the content it hosted. Given the concentration of power, these companies' decisions were not just whether to keep Parler as client, but whether Parler could or should remain on the internet. Again, this raises questions about the legitimacy of these platforms to make such far-reaching decisions.

So, what we have seen in just five years is a slippery slope in action. Originally prompted by widespread agreement that the terrorist group ISIS was using social media in ways that connected directly to their violent and extreme actions and beliefs,

we find ourselves in a situation where a then sitting President of the US has been removed from the most popular social media platforms, and now some platforms themselves are finding it increasingly hard to remain online. However, a slippery slope is defined not just by the fact that we have moved from one condition to another one, but that that subsequent condition is one of significant ethical concern [78, 92]. For clarity's sake, reference to a slippery slope is frequently considered to either be a weak argument, or a criticism of a particular argument. This is not our view here—some slippery slopes are of legitimate concern, but one has to be able to show that a given slide from one state to another is occurring, and that the outcome is one that is morally problematic.

The first half of this chapter has shown that we have slid from banning ISIS to deplatforming platforms. The second half of this chapter looks at the ethical concerns with such a slide. In particular, the ethical question that this chapter is now concerned with is whether the private companies ought to be restricting what people say online.

5 Ethical Challenges

Again, free speech debates are not so much about what is said, but about who has the authority to decide what is said [34]. In this section we examine three ethical challenges presented by the role private companies play in policing extremist content online. The first is the question of whether these private companies are legitimate actors. The second is the implications of the concentration of power into the hands of a few private companies has on their role of policing the internet. Finally, we look at questions about the lack of separation of power in regulating content online, where private companies become judge, jury and executioner.

As context, let us recall that these issues have largely evolved because governments were reluctant to make decisions about people's political beliefs and their right to public communication. In liberal democracies, censorship, where the state makes the determination on who gets to say what and where, is typically limited to public communications that are highly offensive, are likely to induce or incite significant dangerous or illegal activity, or that occur in a context of significant and long running discrimination (van Mill [85]; Henschke Forthcoming). One significant reason for this disinclination for governments to decide who gets to say what and where, is that interference in free speech is frequently seen as a marker of authoritarianism. "The right to free speech is hardly in tension with democracy; it is a precondition for it" ([82], 121). When considering the centrality of information and communication technologies to modern life, and the deep integration of social media into people's personal and social lives, we ought to be ethically and politically concerned if governments started making decisions about who says what online.

Whilst in this chapter we highlight some of the ethical challenges of private companies having the power to control content online, it is worth reflecting what happens when this power rest with governments. We have often seen with more authoritarian governments, the blocking of access to social media sites in the face of

criticism and/or in the wake of anti-government demonstrations.[13] So, there is a case for the state to be minimally involved in these decisions. However, as we will show, simply expecting Tech Companies to fill this void poses significant challenges.

5.1 Moral Legitimacy of Private Actors

The first question that arises is: are private companies legitimate actors to make such decisions? The challenge here is to determine where private companies derive their 'moral authority' from to be able to ban content from their platform. In the case of governments an argument based on the social contract could be made, in that "the first duty of government is the security of its people" [50]. The idea is that members of the public cede certain rights to the government and in return the government has a responsibility to provide security.[14] And removing extremist content from social media is a part of this responsibility. However, private companies are not the government. So where do they get their moral legitimacy from?

We suggest that the moral legitimacy of private actors is derived from social license and responsibility. On social licence, the idea is that companies have a 'social licence' "as a means of pursuing new relationships between industry and communities to reflect public values and ensure community support for projects" ([1], 3). These "societal concerns oblige large corporations to act more "responsibly" … Companies—and their operations—must increasingly satisfy not only the conditions of their formal licences, but also the concerns of host communities and broader society… Hence, it is commonly contended that companies need a "social licence" in addition to their legal and regulatory obligations" ([66], 341). Much like the social contract that allows a government to make decisions about the lives of its citizens, the social licence afforded private companies is something that society, or at least relevant members in that society, grant to that company. "Typically, an operation's social licence is theorised as comprising ongoing acceptance or approval from the local community" ([66], 344). Their moral legitimacy comes in part from the agreement of society that they can continue to exist, in order to provide goods and services and so on.

Parallel to this is the notion of responsibility, whereby the company can be held responsible for the outcomes of their decisions. If a company is shown to be causing undue harm to the environment, or to people, they may be held responsible for that. Similarly, if they are shown to *act responsibly*, by admitting to, and responding to negative impacts of their practices, then the community may see them as earning legitimacy to operate. "Moral legitimacy can be achieved by engaging with affected

[13] For example see: [24, 55].

[14] We note here that this is opens a larger discussion of the ethics of state use of power, legitimacy, and the tension between state-sponsored censorship and the responsibility of the state to provide security to its citizens. There is not space in this chapter to cover those questions, however. For see more on this see: [22, 23, 76].

persons and groups and by finding solutions and compromises with them in order to overcome dissent" ([25], 679). The point here is that the ethical legitimacy is not simply earned, but must be attended to and maintained. Arguably, when the threats posed by ISIS' online activities became more apparent, the social media companies had to act in a way that showed that they were responsible—they saw the role that their services and products played in advancing ISIS' activities, and acted in ways that showed, or at least purported to show, that they cared about the negative impacts that their services were allowing. This brings to the surface a deeper point. The question of who makes decisions, leads to a second question of *how* they make the decisions. The moral legitimacy of the actor in part depends on how they make decisions.

This does not mean that the answers are necessarily clear or easy. For instance, like any issues of representation, who counts as society? What happens if one significant sector of society deeply believe that a political actor needs to be deplatformed, while another significant sector of society deeply believe that that political actor represents their views, and so needs to retain their right to public communication? Moreover, how do we know when society has withdrawn that social license? The purpose of this chapter is not to offer answers to these questions, but instead it is to first show that in the vacuum left by governments reluctant to censor political extremists, Tech Companies are the pivotal actors. Second, we hope to show the contours of where their moral authority to make relevant decisions might come from. Finally, we are pointing out particular emerging questions which would form the basis for discussions moving forward.

5.2 Concentration of Power

In the online world, power is concentrated into the hands of a few giant private tech companies (Fernandez [32]; Kang [37, 52], 2). In terms of social media, the field is dominated by a few big players Facebook (including Instagram and Whatsapp), YouTube and Twitter (Statista n.d.; [13], 88–92).[15] For web infrastructure companies the picture is less clear as there is a much wider diversity of small and large companies across the plethora of infrastructure services. However, as the examples with Daily Stormer, 8chan and Parler show, there are limited options, with the technical ability and capacity, to keep platforms online at scale. This concentration of power into a few hands, as we argue below, should necessitate a higher level of scrutiny and new obligations on the decision makers. Through executing their power to decide who is or is not allowed on their platforms, a few private companies are in practice deciding who can have a voice online.

If there was a much larger plurality of social media platforms, then the banned individual or group could simply move to another platform. Having a plurality of

[15] Further, Evelyn Douek has argued that the increased collaboration of platforms through industry bodies to confront extremist content can "augment the power of already powerful actors by allowing them to decide standards for smaller players" [26].

the press means that a small amount of decision makers do not get to determine what is or is not the news, or whose ideas are allowed in political debate. With this concentration of power of both social media, and the supporting infrastructure, liberal democracies are at risk of significantly narrowing the set of people who get to make the decisions about the public communication of political ideas. Note also that this is a global phenomenon—decisions by Facebook, Twitter, Google etc., impact not just the national discourse, but discussions everywhere. The point here is that, in liberal democracies committed to political pluralism, we need not just multiple avenues for public expression of ideas, but for these avenues to encompass a range of views.

The concentration of power with a small set of companies means that a small group of tech executives get to decide who can and who cannot have a voice online. Such power we argue, and its far-reaching implications on public debate, necessitates that these decisions should face a far higher level of scrutiny than they currently do. One of the main concerns about government censorship is that the concentration of power necessitated by government allows for a very small number of people to make decisions that impact a large number of people. The evolution of the internet, the rise of a small set of companies to dominate the social media space, and the dependence of them and other smaller companies on an increasingly small number of service providers means that we are seeing a similar concentration of power in the hands of a few key actors. To be clear, we are not necessarily disagreeing with decisions to deplatform ISIS or other political extremists. Rather, our point is around those with decision making power in these private companies have power that is significantly disproportionate to those who are affected by their decisions. To explain the concern here, in liberal democracies, the authority of the state lies in people having the ability to vote in elections to bring about a peaceful transitions of power. Private companies have none of these, and given the market dominance of a small number of companies, people have very limited ability to choose other platforms. And in 4 years there is no election to decide who the next Twitter should be.

A final aspect to this concentration of power is that many of these companies, whether the public facing or the supporting infrastructure, are effectively US based. The issue here is the values and standards that are being developed, interpreted, and applied, have not just a developed world view on things, but will typically have an English speaking and American framing and foundation to these issues ([44], 340). Moreover, the social and legal factors that will influence the decisions about whether to protect or deplatform a speaker or company will be heavily US based. If, for instance, there are significant protests in Fiji about particular content, or particular views being deplatformed, is this going to have the same effect as significant protests in the US? Moreover, in line with the point above, if people are dissatisfied with particular laws or ToS in the US, they can seek to change those laws through political processes. But if people in Fiji are dissatisfied, then what processes are available to them to change US laws or ToS? The point here is that, not only is there a concentration of decision making power, but also a concentration of social power.

The overall point here is that, given the way that social media and supporting services have evolved, we are facing an issue now that a very limited number of

people have the capacity to decide who gets deplatformed, and who doesn't. Again, the question we need to ask is, is this concentration of power ethically, socially, and politically justifiable? We suggest here that the answers to this are going to require not just a consideration of why these decision makers have this moral authority, if at all, but will also require reflection on what sorts of institutions these private companies are (see below), as well as further reflection on the best ways to ensure decision making occurs in a way that is responsive to a range of stakeholders' views and concerns, a point we return to at the end of the chapter.

5.3 Lack of Separation of Power

The third area of concern is the lack of separation of powers. This builds on the point made above, that an important part of question of who makes the decisions, is how they make decisions. In terms of online content regulation, as the whole process is in effect carried out by the private companies themselves, private companies are in practice acting as prosecutor, judge, jury and executioner.[16] The platforms, through their ToS, determine the rules for governing the content allowed on their platforms, with the interpretation and enforcement of these rules at the companies' discretion [44]. The sanctions invoked in the case of breaking the terms of service range from removing content, to suspending accounts to banning individuals or organizations from the platform are similarly decided by the platform. And finally, appeals processes are run by the platforms themselves. As a result, this private governance of online speech raises questions of transparency and due process ([5], 2031). In short questions about how they make decisions, and if they make them in the appropriate way.

The argument here is not of any impropriety on behalf of the private companies in enforcing their terms of service,[17] rather that separation of powers is a well-established safeguard in liberal democracies against potential abuses of power. In most liberal democracies there is some version of separation of powers in government, between the legislature, judiciary and executive.[18] The fact that private companies currently decide the rules of what is allowed to be said online, enforce these rules, and run the appeals processes concentrates even more power in their hands and removes potential safeguards against abuse or bias. The ethical issues here obviously arise in situations where that concentrated power is abused—if a tech company capriciously decides what the rules of online activity are, who says what, and/or how any appeals are run, we have a system that lacks the basic pre-conditions of justice. There is a further issue of bias. We find expression of this in discussions of criminal

[16] Balkin highlights this as the problem of 'collateral censorship' that emerges from private governance of online regulation ([5], 2031).

[17] It should be noted that many tech companies have been taking steps to increase the transparency of their content moderation practices, and to include clear appeals processes.

[18] For more on this, see for example: [67, 87, 88].

justice and the separation of powers. Christopher Wellman notes: "The crucial point is that things would deteriorate into a horribly dangerous mess if each individual were personally responsible for punishing those who wronged her. The explanation for this has been laid out plainly by social contract theorists… victims who personally mete out the punishment are more likely to punish the innocent and over-punish the guilty" ([91], 428). Wellman's point is that the ends of justice are better met when there is a separation of powers. We suggest that a similar principle likely arises here—the decisions about who says what online are better made if there is some effective separation between those who write the rules, enforce them, and then act to adjudicate disputes about their enforcement.

However, it should be noted that these concerns are not lost on the companies. Facebook for example has recently created an independent 'Oversight Board' as it believed "that it shouldn't be making so many decisions about speech and online safety on its own" [31]. The board's role is to "review a select number of highly emblematic cases and determine if decisions were made in accordance with Facebook's stated values and policies" [31]. The decisions made by the board will then be binding for Facebook to implement (unless it breaks the law) [31]. This board had its first major test following Donald Trump's statements about the January 6th insurrection at the US Capitol Building, and decided that—due to the risks he posed to public safety, the former US President would remain banned from Facebook until 2023 [74]. Heldt has argued that through the establishing of an independent oversight board and the publication of the guidelines by which its moderators interpret the rules in Facebook's ToS, it has created "structures and procedures similar to administrative law" ([44], 354). Whilst Facebook has taken steps to add in elements of independence and separation of power into their bureaucracies of private governance, questions remain about the legitimacy of institutions like Facebook's Oversight Board, when compared to the separation of powers within a liberal democracy.

So, now we can see that Facebook is attempting to develop a set of processes that divest the company of particular decisions about who gets deplatformed, when, why, and for how long. But this raises further questions—first, how independent is this Oversight Board? Whoever gets to decide the make-up of the board has significant power to influence the direction that any future decisions are made. Second, even if the board itself acts fairly independently, what happens when a decision by the board is likely to have significant economic costs? How does Facebook management adjudicate between the board's decisions and the economic advice? Finally, this gives Facebook a significant advantage over other would-be social media companies. How can a smaller company, much less a start-up have the capitol to setup this quasi-legal infrastructure from the outset that now society deems important? This brings us back to the issue of concentration of power, discussed earlier. Again, our point is not to offer answers to these questions, but to map out the complexity faced when deciding to deplatform political actors. Where it was once relatively easy to make a decision to remove ISIS' material, we are now faced with a highly complex space.

6 Different Institutions, Different Ethical Responsibilities

We have discussed the notion of private companies having a social licence to operate, and that they may lose their moral legitimacy if they do not act responsibility. A further point about legitimacy derives from what type of institution we are discussing.[19] In this section we examine whether we can see social media companies, and the companies that provide the supporting infrastructure, simply as private companies, or as news media companies, or are they instead public infrastructure? Different answers to these questions suggest different ethical responsibilities. Our suggestion here is that not only can we see different social media and related companies differently, we ought actually to see them differently depending on the service that they provide.

If we consider social media and web infrastructure companies to simply be private companies, according to a view like that of shareholder theory,[20] their obligations are restricted to following the law and providing the best returns to shareholders [38]. If seen as purely private companies, they would have on the one hand no obligation to remove or restrict extremist content online beyond what is purely illegal. In short, unless the individual or group is proscribed and/or the material breaks relevant laws such as incitement to violence or hate crimes, the company would have no obligation to remove or restrict access to the content. On the other hand, as a private company, private companies are free to set their own terms of service, deciding both what can be said on their platforms and by whom. Hence, they have no obligation to provide access to their platform to everyone. Thus, seeing them simply as private companies, as long as no laws are broken, then they have no particular responsibility to regulate what is online. However, at the same time, just as any other private company can refuse to offer a product or service to a client, then these companies are free to regulate online content as they will.

Instead, we could consider social media platforms as *media* companies, with all the editorial oversight requirements and responsibility for the content published that this requires. In which case social media platforms might be legally liable for all of the content posted on their platforms. In contrast to the current situation in which platforms are not held liable for the content of their users, a legacy of the 'safe harbor' provisions in section 230 of the US 1996 Communications Act.[21] These provisions have underpinned much of the evolution of social media platforms, and we note here that change in such a position would challenge the concept of social media as we currently know it, which has been built on the premise that they are not responsible for the content of their users' posts.

Alternatively, it could be argued that private companies are providing a public good, and should be seen as a provider of a *public infrastructure utility*, like a water or electricity company. This different conception prompts us to consider that their

[19] For more on this, see: [59].

[20] We note here that a theory like stakeholder theory might take a different view, that the tech companies have a broader set of commitments that includes people beyond shareholders. For more on stakeholder theory see: [68].

[21] For a wider discussion on the debate see: [36].

moral responsibilities might be different than a normal public company. "If they are, instead, more like public infrastructure—like that of a road system or energy system—then they may have to constrain their responsibility to shareholders and profits by reference to public safety and extremist content that poses a public safety threat would likely be justifiably disrupted" [46]. In providing a public good such institutions have a responsibility to the *safety* of their users. In the case of social media companies and web infrastructure companies this responsibility could be seen to include keeping their users safe from extremist content online. This responsibility for safety offers the justification for their content moderation.

However, seeing private companies as public utilities likely generates wider obligations. Public utilities are normally heavily regulated and required to provide equal access to their services to everyone. For example, water or electricity utilities are usually required to provide their services to all members of the public and not discriminate in their choice of customers.[22] In this case if we see private companies as public utilities, this likely places wider obligations on them to provide equal access to their platforms or services.

This leads to another question, should companies that provide different types of service be seen as different types of institution, and hence have different responsibilities and obligations? For example, should social media companies such as Facebook and YouTube, which host, organise and promote users' content, be seen differently to internet infrastructure companies such as Cloudflare which support the back-end of the web?[23] This point highlights the ongoing debate about the appropriate level at which content on the internet be should regulated. However, this not an argument for the latter to have no responsibility to regulate. Rather, it is to say that all should have some level of responsibility for regulating the content that their services support. For example, enforcing action against material from proscribed terrorist groups, or other illegal content such as child pornography. However, should we expect the same regulation of content by social media companies, as by content delivery networks, web hosts or domain registrars? [64]. And if so, is this best understood by seeing then as different types of institutions, entailing different levels of responsibilities and obligations?[24]

For example Balkin argues, "[d]ifferent parts of the internet infrastructure[25] should have different responsibilities to protect freedom of speech online" ([5], 2037). He sets out three groups of companies with different responsibilities: Basic Internet Services (including hosting services, telecommunications services, domain name

[22] For a wider discussion on whether internet infrastructure companies should be seen as delivering a public good see: [39].

[23] This is an argument made by Cloudflare's Mathew Prince [70]. Suzanne van Geuns and Corinne Cath-Speth, put forward an argument that web infrastructure companies like Cloudflare should really be seen as traffic controllers, highlighting the choices that these companies make over the flow of traffic on the internet [39]. However, we also note that the distinction is increasingly blurred with some big tech companies including both social media platforms and web infrastructure services.

[24] We note here that there is a wider debate here about the neutrally of the inner workings of the internet which is beyond the scope of this chapter. For a brief discussion see: [10].

[25] Balkin includes social media companies in his definition of internet infrastructure.

services, caching and defense services), Payment Services and Content Curators (including social media and search engines). Basic Internet Services (such as Cloudflare) and Payment services he argues should not regulate content, while content curators like social media companies have different responsibilities ([5], 2038). Whilst other would argue that basic internet service companies should still have some responsibility to regulate content (as noted above), it might still be reasonable to expect that these responsibilities are different to and less extensive to those of social media companies. If we see companies as having different levels of responsibility depending on the service they provide, this maybe best understood as seeing them as different types of institutions which determines their ethical responsibilities and their legitimacy to take action to regulate.

7 Conclusion: Is Co-Regulation a Solution?

We have argued in this paper that the role that social media companies and web infrastructure companies play in regulating extremist content online, raises significant ethical questions that warrant further attention. Similarly, we have noted above that placing responsibility for online regulation and its enforcement solely within the remit of governments also causes significant ethical dilemmas, in particular with authoritarian states. One approach to resolving these challenges is the idea of co-regulation. In the current situation, internet regulation is partially covered by both public authorities and also self-regulation by social media and web infrastructure companies. The ethical dilemmas highlighted in this chapter fall within the grey area between public authorities' regulation and companies' self-regulation, one solution is to move towards a more dynamic regulatory environment between both actors in which regulatory frameworks are made in a consensual manner. However, as Natali Helberger, Jo Pierson and Thomas Poell have argued "the realization of core public values in these sectors should be the result of the dynamic interaction between platforms, users, and public institutions" ([43], 10).

A first benefit of co-regulation is that it can help reduce the problems of granting authority to one single powerful sort of actor. "These decisions, which deeply affect public safety, the character of public communication, and freedom of expression, should not be left to governments, or to individual platforms and their users. As history shows over and over again, unilateral government regulation of public communication tends to sit in tension with freedom of speech. Furthermore, since social media corporations are primarily driven by commercial interests, they cannot be trusted to always act in the interest of the public good either" ([43], 8). Ideally, co-regulation is a way of not just balancing different interests, but ensuring that each set of actors limit the other set's interests.

The benefits of a co-regulation approach are that it avoids either effectively letting the state be the sole decision maker about political speech, or the challenges of simply

having companies engage in self-regulation [65]. Whilst the latter suffers from questions of accountability and legitimacy [12] outlined above, public authority regulation not only runs the risk of the state being the arbiter of political speech, but often suffers from being slow and unresponsive and lacking the necessary technological know-how [3]. Through developing a process by which the different actors can work more dynamically together, maybe we are able to go some way to resolving the ethical challenges above. Co-regulation between governments and private companies may increase the democratic accountability of who regulates the internet, and in part answer some of the questions of the moral legitimacy of actors. Furthermore, the closer interaction of governments and private companies, may also provide an avenue to address the challenges of concentration of power and separation of powers. That said, we recognise that such public/private cooperation may not only leave the ethical issues we have identified unresolved, but close collaboration between government and private interests may increase the problems identified. It is beyond the scope of this chapter to specify what co-regulation involves; our point here is that if we're talking about how to deal with extremist material online, co-regulation is an area that has ethical significance. All that said, we consider that this is the opening to a discussion on how best to regulate politically extreme content online, rather than a solution. Co-regulation offers one part of that solution, but needs further discussion.

At the heart of content regulation is balancing of rights, between protecting users from extremist material and upholding freedom speech and free public communication. The moral legitimacy of private companies to restrict rights in the pursuit of safety of its users is dependent on what type of institution they are. Furthermore, the obligations of private companies to protect rights of freedom of speech and communication also depends on what type of institution they are. Hence answering this question is fundamental to determining the role they should play in content regulation. While we have travelled far from shutting down ISIS, it is clear that the need to limit terrorist content online has effectively been the start of discussions about who gets to regulate content online, and these discussions are likely to go on for a long time.

References

1. Aitken M, Toreini E, Carmichael P, Coopamootoo K, Elliott K, van Moorsel A (2020) Establishing a social licence for financial technology: reflections on the role of the private sector in pursuing ethical data practices. Big Data Soc 7(1):2053951720908892. https://doi.org/10.1177/2053951720908892
2. AP News (2021) Germany's Merkel: trump's twitter eviction 'problematic.' AP NEWS. 11 Jan 2021, sec. Donald Trump. https://apnews.com/article/merkel-trump-twitter-problematic-dc9732268493a8ac337e03159f0dc1c9
3. Ayres I, Braithwaite J (1992) Responsive regulation: transcending the deregulation debate. Oxford University Press, USA
4. Bain M (2019) Facebook sued by Dissident Group Saoradh over removed pages. Belfasttelegraph. 31 Jan 2019. https://www.belfasttelegraph.co.uk/news/northern-ireland/facebook-sued-by-dissident-group-saoradh-over-removed-pages-37767707.html

5. Balkin JM (2018) Free speech is a triangle. Columbia Law Rev 118:2011–2056
6. BBC (2020) Greece golden dawn: Neo-Nazi leaders guilty of running crime gang. BBC News. 7 Oct 2020, sec. Europe. https://www.bbc.com/news/world-europe-54433396
7. Bergengruen V (2020) 'The devil will be in the details.' how social media platforms are bracing for election chaos. Time. 23 Sept 2020. https://time.com/5892347/social-media-platforms-bracing-for-election/
8. Bishop P, Macdonald S (2019) Terrorist content and the social media ecosystem: the role of regulation. In: Digital Jihad: online communication and violent extremism, pp 135–152. https://cronfa.swan.ac.uk/Record/cronfa52902
9. Brandom R (2017a) Google says it will ban Neo-Nazi site after domain name switch. The Verge. 14 Aug 2017. https://www.theverge.com/2017/8/14/16145064/google-daily-stormer-ban-neo-nazi-registrar-godaddy
10. Brandom R (2017b) Charlottesville is reshaping the fight against online hate—the verge. The Verge. 15 Aug 2017. https://www.theverge.com/2017/8/15/16151740/charlottesville-daily-stormer-ban-neo-nazi-facebook-censorship
11. Byers D (2021) How Facebook and Twitter decided to take down trump's accounts. NBC News. 14 Jan 2021. https://www.nbcnews.com/tech/tech-news/how-facebook-twitter-decided-take-down-trump-s-accounts-n1254317
12. Campbell AJ (1999) Self-regulation and the media. Federal Commun Law J 51(3). https://www.repository.law.indiana.edu/fclj/vol51/iss3/11/
13. Cicilline DN (2020) Investigation of competition in digital markets: subcommittee on antitrust commercial and administrative law of the committee on the judiciary, 449
14. CNBC (2017) Internet firms flex muscle to exile white supremacists. CNBC. 16 Aug 2017, sec. Technology. https://www.cnbc.com/2017/08/16/internet-firms-flex-muscle-to-exile-white-supremacists.html
15. Conger K (2017) Cloudflare CEO on terminating service to Neo-Nazi site: 'the daily stormer are assholes.' Gizmodo. 16 Aug 2017. https://gizmodo.com/cloudflare-ceo-on-terminating-service-to-neo-nazi-site-1797915295
16. Conger K (2019) It's back: 8chan returns online. The New York Times. 4 Nov 2019, sec. Technology. https://www.nytimes.com/2019/11/04/technology/8chan-returns-8kun.html
17. Conger K (2020) Twitter has labeled 38% of trump's tweets since tuesday. The New York Times. 5 Nov 2020, sec. Technology. https://www.nytimes.com/2020/11/05/technology/donald-trump-twitter.html
18. Conger K and Popper N (2019) Behind the scenes, 8chan scrambles to get back online. The New York Times. 5 Aug 2019, sec. Technology. https://www.nytimes.com/2019/08/05/technology/8chan-website-online.html
19. Conway M (2020) Routing the extreme right. RUSI J 165(1):108–113. https://doi.org/10.1080/03071847.2020.1727157
20. Conway M, Macnair L, Scrivens R (2019) Right-wing extremists' persistent online presence: history and contemporary trends. ICCT Policy Brief 24
21. Cox J (2017) After shutdown, daily stormer users are moving to a dark web version of site. 15 Aug 2017. https://www.vice.com/en/article/evvxvz/white-supremacist-website-daily-stormer-goes-offline
22. Cudd A (2012) Contractarianism. Stanford Encyclopedia of Philosophy/Winter 2013 Edition. https://plato.stanford.edu/archives/win2013/entries/contractarianism/
23. D'Agostino F, Gaus G, Thrasher J (2011) Contemporary approaches to the social contract. Stanford Encyclopedia of Philosophy/Winter 2012 Edition. https://plato.stanford.edu/archives/win2012/entries/contractarianism-contemporary/
24. Deahl D (2018) Iran has banned telegram after claiming the app encourages 'armed uprisings. The Verge. 1 May 2018. https://www.theverge.com/2018/5/1/17306792/telegram-banned-iran-encrypted-messaging-app-russia
25. Demuijnck G, Fasterling B (2016) The social license to operate. J Bus Ethics 136(4):675–685. https://doi.org/10.1007/s10551-015-2976-7

26. Douek E (2020) The rise of content cartels. Knight First Amendment Institute, The Tech Giants, Monopoly Power, and Public Discourse. February. https://knightcolumbia.org/content/the-rise-of-content-cartels
27. Erwin A (2019) Britain first ends Facebook challenge over Northern Ireland page ban. Belfast-telegraph. 15 May 15. https://www.belfasttelegraph.co.uk/news/northern-ireland/britain-first-ends-facebook-challenge-over-northern-ireland-page-ban-38114561.html
28. EUROPOL (2020) EU IRU transparency report 2019. Europol. 13 Oct 2020. https://www.europol.europa.eu/publications-documents/eu-iru-transparency-report-2019
29. Facebook (2018) Taking action against Britain first. Taking action against Britain first (blog). 14 Mar 2018. https://about.fb.com/news/h/taking-action-against-britain-first/
30. Facebook (2019) Standing against hate. About Facebook (blog). 27 Mar 2019. https://about.fb.com/news/2019/03/standing-against-hate/
31. Facebook Oversight Board. n.d. Oversight boardIndependent judgement. Transparency. Legitimacy. Facebook Oversight Board. Accessed 12 June 2021. https://oversightboard.com/
32. Fernandez R, Adriaans I, Klinge TJ, Hendrikse R (2021) How big tech is becoming the government. SOMO. 5 Feb 2021. https://www.somo.nl/how-big-tech-is-becoming-the-government/
33. Finkle J, Rodriguez S (2017) Tech companies in the crosshairs on white supremacy and free speech. Reuters. 14 Aug 2017. https://www.reuters.com/article/us-virginia-protests-godaddy-idUSKCN1AU0CV
34. Fish S (1994) There's no such thing as free speech: and it's a good thing, too. Oxford University Press, New York. https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=273279
35. Fisher M (2018) Inside Facebook's secret rulebook for global political speech—The New York Times. The New York Times. 27 Dec 2018. https://www.nytimes.com/2018/12/27/world/facebook-moderators.html
36. Flew T, Martin F, Suzor N (2019) Internet regulation as media policy: rethinking the question of digital communication platform governance. J Digital Media Policy 10(1):33–50. https://doi.org/10.1386/jdmp.10.1.33_1
37. Floridi L (2021) Trump, Parler, and regulating the Infosphere as our commons. Philosophy and Technology, March. https://doi.org/10.1007/s13347-021-00446-7
38. Friedman M (1972) The social responsibility of business. The New York Times Magazine, 13 Sept. In: Friedman M (ed) An economist's protest: columns on political economy. Thomas Horton & Daughters, Glen Ridge, NJ, pp 177–84
39. Geuns SV, Cath-Speth C (2020) How hate speech reveals the invisible politics of internet infrastructure. Brookings (blog). 20 Aug 2020. https://www.brookings.edu/techstream/how-hate-speech-reveals-the-invisible-politics-of-internet-infrastructure/
40. Ghosh S (2019) Facebook cleared political ads for a far-right group it banned just 8 months ago. Business Insider Nederland. 8 Jan 2019. https://www.businessinsider.nl/far-right-group-britain-first-facebook-ads-mosque-ban-2019-1/
41. Global Freedom of Expression (2020) Facebook v. CasaPound. Global Freedom of Expression. http://globalfreedomofexpression.columbia.edu/cases/casapound-v-facebook/
42. Godwin C (2021) Florida governor signs bill to ban big tech 'Deplatforming.' BBC News. 24 May 2021, sec. Technology. https://www.bbc.com/news/technology-56952435
43. Helberger N, Pierson J, Poell T (2018) Governing online platforms: from contested to cooperative responsibility. Inf Soc 34(1):1–14. https://doi.org/10.1080/01972243.2017.1391913
44. Heldt A (2019) Let's meet halfway: sharing new responsibilities in a digital age. J Inf Policy 9:336–369. https://doi.org/10.5325/jinfopoli.9.2019.0336
45. Henschke A (Forthcoming) Free speech, free public communication and counterterrorism. In: Feltes J, Henschke A, Miller S, Elgar E (eds) Counter-terrorism: the ethical issues
46. Henschke A, Reed A (2021) Toward an ethical framework for countering extremist propaganda online. Stud Conflict Terrorism. https://doi.org/10.1080/1057610X.2020.1866744

47. Hern (2019) Facebook bans far-right groups including BNP, EDL and Britain first. The Guardian. 18 Apr 2019, sec. Technology. http://www.theguardian.com/technology/2019/apr/18/facebook-bans-far-right-groups-including-bnp-edl-and-britain-first
48. Hern (2021) Opinion divided over trump's ban from social media. The Guardian. 11 Jan 2021. http://www.theguardian.com/us-news/2021/jan/11/opinion-divided-over-trump-being-banned-from-social-media
49. Hern A, Rawlinson K (2018) Facebook bans Britain first and its leaders. The Guardian. 14 Mar 2018, sec. World news. http://www.theguardian.com/world/2018/mar/14/facebook-bans-britain-first-and-its-leaders
50. Heyman S (1991) The first duty of government: protection, liberty and the fourteenth amendment. Duke Law J 41(3):507–571
51. Isaac M, Browning K (2020) Fact-checked on Facebook and Twitter, conservatives switch their apps. The New York Times. 11 Nov 2020, sec. Technology. https://www.nytimes.com/2020/11/11/technology/parler-rumble-newsmax.html
52. Kang C, McCabe D (2020) House lawmakers condemn big tech's 'monopoly power' and urge their breakups. The New York Times. 6 Oct 2020, sec. Technology. https://www.nytimes.com/2020/10/06/technology/congress-big-tech-monopoly-power.html
53. Kayali L, Braun E (2020) France pushes tougher eu rules for social media in wake of terror attack. POLITICO. 26 Oct 2020. https://www.politico.eu/article/france-renews-social-media-regulation-push-at-eu-level-in-wake-of-terror-attack/
54. Kearney V (2019) Facebook: dissident republicans Saoradh take legal action. BBC News. 30 Jan 2019, sec. Northern Ireland. https://www.bbc.com/news/uk-northern-ireland-47043375
55. Letsch C, Rushe D (2014) Turkey blocks YouTube amid 'national security' concerns. 28 Mar 2014, sec. World news. http://www.theguardian.com/world/2014/mar/27/google-youtube-ban-turkey-erdogan
56. Loadenthal M (2015) Introduction: like finding a needle in a pile of needles: political violence and the perils of a brave new digital world. Critical Stud Terrorism 8(3):456–465. https://doi.org/10.1080/17539153.2015.1094266
57. Lonas L (2021) Pompeo, Cruz and other trump allies condemn Twitter's ban on president. The Hill. 9 Jan 2021. https://thehill.com/policy/technology/533486-pompeo-cruz-and-other-trump-allies-condemn-twitters-ban-on-president
58. Macdonald S, Staniforth A (2021) The tech industry and the regulation of online terrorist content: what do law enforcement think? Hedayah (blog). 16 Jan 2021. https://www.hedayahcenter.org/media-center/latest-news/blog-post-the-tech-industry-and-the-regulation-of-online-terrorist-content-what-do-law-enforcement-think/
59. Miller S (2010) The moral foundations of social institutions: a philosophical study. Cambridge University Press
60. Nicas J, Alba D (2021) Amazon, Apple and Google cut off Parler, an app that drew trump supporters—The New York Times. The New York Times. 9 Jan 2021. https://www.nytimes.com/2021/01/09/technology/apple-google-parler.html
61. Nouri L, Lorenzo-Dus N, Watkin A-L (2021) Impacts of radical right groups' movements across social media platforms—a case study of changes to Britain first's visual strategy in its removal from Facebook to gab. Stud Confl Terrorism:1–27.https://doi.org/10.1080/1057610X.2020.1866737
62. O'Neill L (2018) Britain first using northern Ireland laws to sue Facebook over censorship claims. Belfast Telegraph. 3 Oct 2018. https://www.belfasttelegraph.co.uk/news/uk/britain-first-using-northern-ireland-laws-to-sue-facebook-over-censorship-claims-37377224.html
63. Oremus W (2017a) Cloudflare CEO Matthew prince is right: we can't count on him to police internet. Slate Magazine. 8 2017. https://slate.com/technology/2017/08/cloudflare-ceo-matthew-prince-is-right-we-can-t-count-on-him-to-police-online-speech.html
64. Oremus W (2017b) GoDaddy joins the resistance. Slate Magazine. 16 Aug 2017. https://slate.com/technology/2017/08/the-one-big-problem-with-godaddy-dropping-the-daily-stormer.html

65. Palzer C (2003) Self-monitoring v. Self-regulation v. Co-regulation. In: Closs W, Nikoltchev S (eds) Co-regulation of the media in Europe. European Audiovisual Observatory, pp 29–31
66. Parsons R, Moffat K (2014) Constructing the meaning of social licence. Soc Epistemol 28(3–4):340–363. https://doi.org/10.1080/02691728.2014.922645
67. Persson T, Roland G, Tabellini G (1997) Separation of powers and political accountability. Q J Econ 112(4):1163–1202
68. Phillips R, Edward Freeman R, Wicks AC (2003) What stakeholder theory is not. Bus Ethics Q 13(4):479–502. https://doi.org/10.5840/beq200313434
69. Prince M (2017) Why we terminated daily stormer. The Cloudflare Blog (blog). 17 Aug 2017. https://blog.cloudflare.com/why-we-terminated-daily-stormer/
70. Prince M (2019) Terminating service for 8Chan. The Cloudflare Blog (blog). 5 Aug 2019. https://blog.cloudflare.com/terminating-service-for-8chan/
71. Reuters S (2019a) Facebook bans white nationalism, white separatism on its platforms. Reuters. 27 Mar 2019. https://www.reuters.com/article/us-facebook-hatespeech-idUSKCN1R81ZH
72. Reuters S (2019b) Italian judge orders Facebook to reopen neo-fascist group's account. Reuters. 12 Dec 2019. https://www.reuters.com/article/uk-italy-facebook-neofascists-idUKKBN1YG2AR
73. Robertson A (2017) Neo-Nazi site moves to dark web after GoDaddy and Google bans. The Verge. 15 Aug 2017. https://www.theverge.com/2017/8/15/16150668/daily-stormer-alt-right-dark-web-site-godaddy-google-ban
74. Rodriguez S (2021) Facebook says Donald Trump to remain banned for two years, effective from January 7. CNBC. 4 June 2021, sec. Technology. https://www.cnbc.com/2021/06/04/facebook-says-donald-trump-to-remain-banned-from-platform-for-2-years-effective-from-jan-7.html
75. Sadurski W (1999) Freedom of speech and its limits. Law and Philosophy Library. Springer Netherlands. https://doi.org/10.1007/978-94-010-9342-2
76. Scanlon T (2000) What we owe to each other. Belknap Press of Harvard University Press, Cambridge, MA
77. Schauer F (1982) Free speech: a philosophical enquiry. Cambridge University Press, Cambridge
78. Schauer F (1985) Slippery slopes. Harv Law Rev 99(2):361–383. https://doi.org/10.2307/1341127
79. Siapera E, Veikou M (2016) The digital golden dawn: emergence of a nationalist-racist digital mainstream. In: The digital transformation of the public sphere: conflict, migration, crisis and culture in digital networks, pp 35–59
80. Statista. n.d. Most used social media 2021. Statista. Accessed 12 June 2021. https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/
81. Stewart H, Elgot J (2018) May calls on social media giants to do more to tackle terrorism. The Guardian. 24 Jan 2018, sec. Business. http://www.theguardian.com/business/2018/jan/24/theresa-may-calls-on-social-media-giants-to-do-more-to-tackle-terrorism
82. Sunstein CR (1993) Democracy and the problem of free speech. Free Press
83. Twitter (2020) The 2020 US elections and Twitter|Twitter Help. https://help.twitter.com/en/using-twitter/us-elections
84. Tyko K (2020) Trump Tweet about problems with mail-in ballots quickly labeled by twitter as misleading. USA TODAY. 26 Oct 2020. https://www.usatoday.com/story/tech/2020/10/26/donald-trump-twitter-mail-ballots-election-tweet-misleading/6049734002/
85. van Mill D (2017) Freedom of speech. In: Zalta EN (ed) The stanford encyclopaedia of philosophy. https://plato.stanford.edu/archives/win2013/entries/freedom-speech/
86. van der Vegt I, Gill P, Macdonald S, Kleinberg B (2019) Shedding light on terrorist and extremist content removal. Global Research Network on Terrorism and Technology, paper no. 3. https://rusi.org/publication/other-publications/shedding-light-terrorist-and-extremist-content-removal
87. Vibert F (2007) The rise of the unelected: democracy and the new separation of powers. Cambridge University Press

88. Vile MJC (2012) Constitutionalism and the separation of powers. Liberty Fund, Indianapolis. https://muse.jhu.edu/book/21621
89. Vincent J (2019) Facebook Bans UK's biggest far-right organizations, including EDL, BNP, and Britain first. The Verge. 18 Apr 2019
90. Wakefield J (2019) Facebook takes down Britain first ads. BBC News. 7 Jan 2019, sec. Technology. https://www.bbc.com/news/technology-46746601
91. Wellman CH (2009) Rights and state punishment. J Philos 106(8):419–439
92. Williams B (1985) Which slopes are slippery. In: Lockwood M (ed) Moral dilemmas in modern medicine. Oxford University Press, Oxford, pp 126–137
93. Wong JC (2019) 8chan: the far-right website linked to the rise in hate crimes. The Guardian. 5 Aug 2019, sec. Technology. http://www.theguardian.com/technology/2019/aug/04/mass-shootings-el-paso-texas-dayton-ohio-8chan-far-right-website
94. Zuckerberg M (2018) A blueprint for content governance and enforcement. https://m.facebook.com/nt/screen/?params=%7B%22note_id%22%3A751449002072082%7D&path=%2Fnotes%2Fnote%2F&refsrc=http%3A%2F%2Fwww.google.com%2F&_rdr