'Too much' and 'too little' content moderation: Internet governance as a case study in advanced liberal modes of government

Terry Flew, Professor of Digital Communication and Culture, The University of Sydney

ABSTRACT

The liberal art of government is forced to determine the precise extent to which and up to what point individual interest, that is to say, individual interests insofar as they are different and possibly opposed to one another, constitute a danger for the interest of all. The problem of security is the protection of the collective interest against individual interests. Conversely, individual interests have to be protected against everything that could be seen as an encroachment of the collective interest.

– Michel Foucault, The Birth of Biopolitics: Lectures at the Collège de France 1978-1979, p. 65.

In this paper I approach the question of content moderation on digital platforms as a case study in Foucauldian approaches to governmentality. The early history of the Internet was framed around a set of ideas that identified it as a "technology of freedom", unencumbered by the constraints associated with traditional media. Over time, a series of limitations have arisen with this optimistic framing of the open Internet, ranging from the commercial imperatives facing digital platforms that broker online interactions, the range of actual and potential harms associated with online content and user behaviour, and the limits of self-regulation.
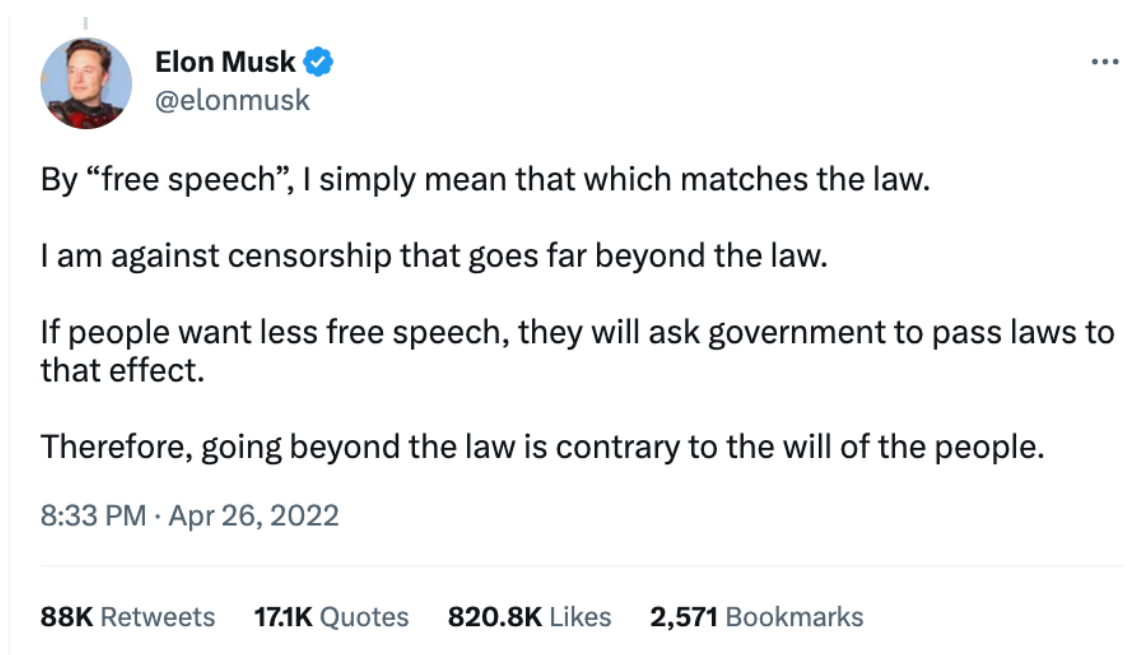
Such developments mean that we now live in an age where there are considerable expectations around tech companies being able to moderate digital content in the public interest. The policy and regulatory developments are, however, underpinned by a set of anxieties, ranging from the question of who should moderate and by what means, concerns about government interference with online spaces, and polarised political debates about "cancel culture" and the nefarious

influence of "Big Tech". Drawing upon Michel Foucault's lectures on neoliberal governmentality in The Birth of Biopolitics, I wish to consider what these debates tell us about the possibilities and prospects of governing online spaces so as to promote civic culture and standards of public discourse.

Introduction

When billionaire Elon Musk bought Twitter in October 2022, he laid off almost 70 per cent of the 7,500 staff employed by the company within a month. A very large number of those laid off were engaged in content moderation, or the division referred to within Twitter as Trust & Safety. Wanting to rebuild Twitter as a leaner company, where engineering values and principles were in the ascendancy, Musk has been innately suspicious of the Californian strand of left-liberalism which he and his followers see as being rife in Silicon Valley. The Trust & Safety Division seemed to be for Musk at the heart of what was wrong with Twitter: its left-leaning culture and values, its bureaucracy and – for an entrepreneur deeply rooted in the start-up ethos – its focus upon the wrong things, such as 'cultural safety' rather than business profitability.

Prior to acquiring Twitter, Musk described – on Twitter, of course, and to his 132 million followers – his vision of how speech rights should work on social media platforms:

Elon Musk @ @elonmusk

By "free speech", I simply mean that which matches the law.

I am against censorship that goes far beyond the law.

If people want less free speech, they will ask government to pass laws to that effect.

Therefore, going beyond the law is contrary to the will of the people.

8:33 PM · Apr 26, 2022

88K Retweets   17.1K Quotes   820.8K Likes   2,571 Bookmarks

Musk moved quickly to enact such 'free speech absolutism' shortly after he acquired Twitter. As well as laying off a large number of staff and contractors involved with content moderation, he reinstated a series of high-profile accounts that had previously been blocked from the site, including Andrew Tate, Jordan Peterson, Kanye West (Ye) and, most famously, Donald Trump (Ecarma, 2022). Finally, it appeared that conservatives in the U.S. had the owner of a large social media platform they had long been seeking, having complained for years about 'Big Tech' censorship and the online 'cancel culture' they associated with Silicon Valley (Napoli, 2021). But the moment of 'free speech absolutism' ended quite abruptly on the platform. Ye was deplatformed after remarks deemed unacceptably anti-Semitic (actually pro-Hitler), and professional conspiracy theorist Alex Jones was not permitted to return to the site. Outlining the new content moderation policies, Musk tweeted:

← **Thread**

Elon Musk ✔
@elonmusk

New Twitter policy is freedom of speech, but not freedom of reach.

Negative/hate tweets will be max deboosted & demonetized, so no ads or other revenue to Twitter.

You won't find the tweet unless you specifically seek it out, which is no different from rest of Internet.

6:31 PM · Nov 18, 2022

41.1K Retweets   **20.4K** Quotes   **405.9K** Likes   **3,786** Bookmarks

Vanity Fair writer Caleb Ecarma has described Twitter's new content moderation policy as being 'seemingly wholly dependent on what side of the bed Musk wakes up on or how users respond to random polls' (Ecarma, 2022). It is certainly a long way away from the policies of other platform companies, most notably Meta, who have created an infrastructure of third-party adjudication on the company's content moderation decisions through the Oversight Board (formerly Facebook Oversight Board), which I will discuss later in this paper. Musk's constant public utterances on this topic on a web site that he now owns are idiosyncratic by the standards of contemporary business leaders, and his decision-making methods are attention seeking and odd, with the plebiscite of Twitter users to vote on whether Trump's account should be reinstated on the site being a case in point.
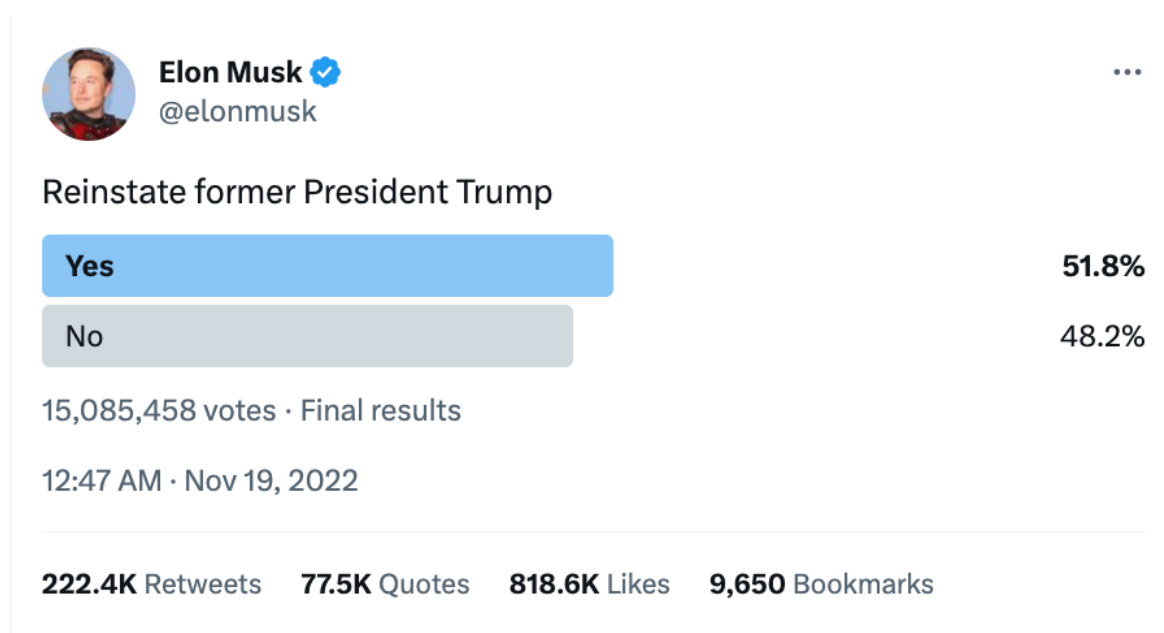
While Musk's approach to content moderation questions is distinctive, the issues being faced are not necessarily new. The global Internet of the 2020s differs from that of the 1990s in that communication and interaction among users now largely takes place through a relatively small number of digital platforms, which filter and moderate content on the basis of a variety of goals to keep a balance between a range of stakeholders, that include not only users, but advertisers,

regulators, professional content providers (e.g. media companies), and many others. As Julie Cohen has observed, the "network of networks" is … a network of platforms; Internet access and use are intermediated from beginning to end' (Cohen, 2017, p. 143). A critical stakeholder in these relations is advertisers, who continually need to be persuaded to place content on these sites as the basis of the platforms' continued viability. Indeed, Elon Musk's biggest challenge at Twitter may be less the loss of site users as the failure to attract advertisers to the revised site: it is estimated that advertising revenues on Twitter are down by 40% of what they were prior to Musk's takeover (Milmo, 2023). With rising inflation and interest rate increases threatening consumer demand globally, the current economic environment is one that is challenging for social media companies generally, and particularly challenging for a heavily indebted company that pursues content moderation policies that generate uncertainty.

Musk's challenges with developing a content moderation that strikes an appropriate balance between so-called 'free speech absolutism', questions of where lines are drawn between 'acceptable' and 'unacceptable' forms of public speech, and how to run a contemporary social media business with multiple and diverse stakeholders and inherently antagonistic interests, provides a backdrop around which we can make a few connections. One is the relationship between classic debates about free speech rights and social obligations in liberal societies, and the different ways in which competing norms and principles have been balanced over time. I will draw upon the late work of Michel Foucault on liberal and neoliberal modes of governmentality – albeit with some qualifications about the term 'neoliberal' – and how they have mapped onto different governmental practices towards content regulation, particularly the shifting balance between censorship and classification of media content.

A second question arises from the platformised Internet, and the extent to which familiar challenges emerge around content regulation, but questions of responsibility shift from public actors to private actors. I will consider the Oversight Board developed by Meta as a case study in what I have elsewhere termed quasi-self-regulation (Flew, 2021), whereby corporations create their own quasi-judicial structures to establish regulatory regimes to which they possess accountability, but which are outside of the jurisdictional purview of nation-states. Finally, I will connect the rise of digital platform companies to the resurgence of debates about 'stakeholder capitalism' in recent years and consider the scope and limitations of such conceptions of reconciling private and public interests through regulatory regimes.

Too much and too little freedom? Michel Foucault and the rise of classificatory regimes

While Michel Foucault did not live to see the global popularisation of the Internet, I would hazard the observation that he would have found the balances that are involved in Internet governance align with the propositions he developed around liberal modes of governmentality. In his 1978-79 lectures at the College de France, published in English in 2008 as *The Birth of Biopolitics* (Foucault, 2008), Foucault distinguishes between a history of liberalism that focuses upon foundational concepts (sovereignty, rule of law, human rights, constitutions, social contract etc.), and one that focuses upon practices of government and how they evolve over time. From this perspective, he identifies the critical insight of political economy as it emerged in 18th century Europe as setting up internal rationales of government whereby the question of whether governments are governing 'too much' or 'too little' becomes, not a question of abstract rights, but rather a series of empirical questions where 'success or failure, rather than legitimacy or illegitimacy, now become the criteria of governmental action'

(Foucault, 2008, p. 16). Political economy thus became central to liberalism as a new regime of truth alternative to that of *raison d'Etat* (reason of state), where 'a government is never sufficiently aware that it always risks governing too much, or a government never knows too well how to govern just enough' (Foucault, 2008, p. 17). The importance of political economy is that it establishes the market as the 'site of truth' of liberal government, setting limits to state action and promoting 'frugal government' in terms of the relationship between instruments and goals (Foucault, 2008, pp. 28, 30).

Within Foucault's account of liberal government, freedom emerges not as a set of constitutionally enshrined rights, but rather as both an object of and a constraint upon governmental practice:

> Freedom is never anything other … than an actual relation between governors and governed, a relation in which the measure of the "too little" existing freedom is given by the "even more" freedom demanded (Foucault, 2008, p. 63).

From this account of liberalism, freedom becomes something that is not innate in humanity or a universal right, but rather something that is produced, consumed, managed and organised by government, Moreover, governmental practice is at risk of damaging freedom through its actions: 'the liberalism we can describe as the art of government formed in the eighteenth century, entails at its heart a productive/destructive relationship with freedom' (Foucault, 2008, p. 64). In particular, the counterpoint of freedom is security, and 'the problem of security is the protection of the collective interest against individual interests. Conversely, individual interests have to be protected against everything that could be seen as an encroachment of the collective interest' (Foucault, 2008), p. 65). Foucault proposes that, ultimately:

The problems of what I shall call the economy of power peculiar to liberalism are internally sustained … by this interplay of freedom and security (Foucault, 2008, p. 65).

I have argued elsewhere (Flew, 2018)that Foucault's governmental conception of the relationship between freedom and security – or constraint – provides a useful foundation for understanding how the classification of media and other forms of content has evolved from the 1960s onwards. In particular, the proposition that 'governmental reason does not divide subjects between an absolutely reserved dimension of freedom and another dimension of submission that is either consented to or opposed', but rather the division is … in the very domain of governmental practice itself, between the operations that can be carried out and those that cannot, between what to do and the means to use on the one hand, and what not to do on the other (Foucault, 2008, pp. 11-12), applies quite well to contemporary media content classification regimes.

The legal case *R. v Penguin Books Ltd. (1959)*, generally known as the 'Lady Chatterley's Lover' case, indicated that content regimes based upon the exercise of the juridical power of the state, through the Chief Censor or equivalent, to prohibit content without consideration of context or avenues for appeal would prove to be untenable. In finding against claims that the novel had the capacity to 'deprave or corrupt', the case introduced several critical elements into censorship law, which remain in most countries to this day, including: the need to place any individual element of a work in the context of the work as a whole; the artistic, literary or educational merit of the work; and considerations of audience, both in terms of the 'reasonable person', and distinctions within that category, particularly those between adults and children. Subsequent legislation that reformed censorship guidelines incorporated these and other

criteria into a revised framework. In Australia, the National Classification Code enacted in 1972, following an amended *Customs Act 1971*, articulated three guiding principles for content classification:

1. Adults should be able to read, see and hear what they want.

2. Minors should be protected from material likely to harm or disturb them.

3. Everyone should be protected from exposure to unsolicited material which they find offensive.

Within such governmental frameworks, the shift from censorship to classification can be seen, not as a once-and-for-all shift from restrictions to freedom, but rather as facilitating a series of rules, laws, principles and guidelines that enable 'governmentalisation' of the regulatory space for media content, by minimizing the amount of actual banning of materials (and with it, the familiar paradox of censorship, which is that banning something only serves to make it more attractive), but at the same time rendering the scope to apply the tools of media content classification more widely. As Nicole Moore has observed, evaluation of regimes of censorship revolve less upon questions of the freedom or otherwise of the human subject, and instead come to be associated with 'an attempt to identify the limits and effects of regulatory power as such' (Moore, 2014, p.63), and as applied across a much wider array of social, cultural and communicative spaces.

Two further elements can be noted with such a governmental regime of classification. One is that the move to restrict censorship to be an exception involves establishing such regulation as a normal component of public administration, taking it out of the realm of the law, the police, and the courts. As Ian Hunter has observed, with regards to education, this entails establishing

the government bureau as an administrative centre driven by 'the deployment of techniques of quantitative calculation and of procedural decision-making in a particular domain' (Hunter, 1994, p. 151). This in turn requires the bureaucratisation of classification decision-making, characterised by 'the separation of person and office' (Gorski, 2005, p. 269), and where the classification agency became what Max Weber (1978, pp. 957-959) termed a particular form of 'office', staffed by 'bureaucrats … "personally" committed to the ethos and purposes of their distinctive office even though that ethos lies outside of their own personal (i.e. individual) moral predilections or principles' (du Gay, 2009, p. 150). The personal comportment expected of media classification officers requires, among other things, 'the pragmatic rejection of principled politics' in making decisions, the 'capacity to detach governmental decisions from personal loyalties and … passions' (Hunter, 1994, pp. 151, 155), and the 'construction of a buffer between civic comportment and personal principles' (du Gay, 2009, p. 152).

The other point to be noted is that, with principles and morality being downplayed as the drivers of classification decisions, the question of precedent becomes increasingly important. With vague concepts such as 'community standards', and the expectations of a 'reasonable person', providing shifting and unreliable benchmarks for deliberation and judgement, we see the rise of what is known in legal terms as case-based reasoning, or the concept of *casuistry* (from the Latin *casus*). The concept of casuistry can be found in Catholic moral thought, associated with the work of Jesuit theologians who grappled with the need to assess traditional morality in light of changing reality, as well as the moral challenges arising from the private confession of sins and questions of conscience. It placed an emphasis upon the detailed documentation of how general principles were applied in particular cases, providing guidance to those involved in the training of priests in pastoral pedagogy, based upon analysis of practical examples.

Casuistry has returned to significance in recent years in debates in fields such as bioethics (Jonson & Toulmin, 1988), where either/or moral judgements are often difficult, and where such case-based reasoning can 'provide a dialectical form of exchange between, on the one hand, what appears to be the facts of a particular case and, on the other, one or more generally accepted moral principles which appeared to be relevant to this case' (Mahoney, 2000, p. 98). The merit of such an approach as an alternative to deductive reasoning based upon general moral principles is said to lie in its ability to identify commonalities between new cases and those that have appeared previously, enabling the latter to provide precedents that can assist in identifying new moral principles based upon a particular class of cases. In relation to classification, casuistic reasoning may emerge in a context where conflicting moral principles are in play, such as freedom of expression and protection of children. It also reflects the need for guidance on particular cases being provided to some degree by past decisions, in the absence of an overarching set of ethical principles that can be drawn upon.

The 'platformised Internet' and new classificatory regimes

The rise of the Internet has obviously disrupted such classificatory regimes. The "golden era" of people bearing clipboards and closely scrutinising *Last Tango in Paris* or its equivalents to identify problematic scenes barely registers in an era where digital content is ubiquitous, fast changing, and increasingly uploaded by users themselves rather than accredited media professionals. It is estimated that 3.7 million new videos are uploaded to YouTube every day, or about 150,000 videos per minute, of an average length of 4.4 minutes (Hayes, 2023). If we multiply those activity across a range of social media and video sharing platforms, and acknowledge that all operate on *ex post* rather than *ex ante* principles – that is, content can only be reviewed after it appears on the platform, as distinct from media gatekeeping practices,

where content is typically screened before it reaches the public – then the task of media classifiers appears to be an impossible one. The 2011 Review of the Australian National Classification Scheme undertaken by the Australian Law Reform Commission, which I chaired, concluded that Australia's *Classification Act 1991* was 'an analogue piece of legislation in a digital world' (Australian Law Reform Commission, 2012, p. 11). The ALRC Review concluded that, at the very least, a contemporary classification system that could work in an era of digital media convergence required a shift from across-the-board classification to a focus on potentially problematic content, a platform-neutral approach to regulation that recognised both continuities and differences between online content and traditional media, and a co-regulatory framework that empowered industry to undertake its own classificatory practices, underpinned by genuine enforcement criteria for non-compliance (Australian Law Reform Commission, 2012).

At the same time, we need to be cautious about presuming that because delivery technologies have changed, the societal issues that underpinned regulatory action have disappeared. A key idea in the early evolution of the Internet was that, as it was fundamentally different to traditional media, a 'policy of freedom' was required for such 'technologies of freedom' (de Sola Pool, 1983). In retrospect, we can see in such moves, enshrined in US legislation such as Section 230 of the *Communications Decency Act 1996*, an ideological gambit which ensured that laws and policies related to the Internet would need to be crafted anew and not involve adaptations of existing legislation, that a presumption towards non-intervention on the part of governments would be the norm, and that US values based around freedom of expression and the First Amendment would be hegemonic in international institutions developed for Internet governance (Carr, 2016; Flew, 2021). [1]

At any rate, what has taken place with the evolution of the Internet has been what has been referred to as the *platformisation of the Internet* (Flew, 2021; Helberger et al., 2018; Helmond, 2015; Kretschmer et al., 2021; Napoli, 2019). The concentration of control over key functions of the digital economy by a relatively small number of global technology corporations, and the communications power associated with being *de facto* gatekeepers of much of the world's online content (Balkin, 2018; Castells, 2009) has had paradoxical consequences. While it has concentrated immense power in the hands of these tech giants, it has also rendered decision-making around content moderation and related governance practices more visible, and more capable of being impacted upon, than the open internet (Flew & Su, 2022).

In the context of the platformised Internet, digital platform companies find themselves implicated in incidents, or what Ananny and Gillespie term public shocks, that 'suddenly highlight a platform's infrastructural qualities and call it to account for its public implications' (Ananny and Gillespie, 2017, p. 2). There have been a variety of these, but perhaps the incident with the most significant and lasting consequences for perceptions of platforms was the Cambridge Analytica scandal, where whistleblower Christopher Wylie revealed the directly enabling role played by Facebook in enabling illegal data harvesting of 87 million user accounts for a variety of political campaigns, under the misleading pretext of online quizzes. When this story was broken by *Guardian* journalist Carole Cadwalladr in 2018, the apology of Facebook CEO Mark Zuckerberg, and the associated commitment to 'do better next time', was met with incredulity around the globe. Academic and commentator Zeynep Tufecki referred to the Zuckerberg 'apology tour' where he has apologies 14 times in the 14-year history of Facebook as a public company for data breaches, always with a promise to do better next time. Politicians and legislators around the world were sharpening their critique of the platforms and the failure of self-regulation, whether through the House of Commons committees on hate

speech and disinformation in the UK (House of Commons, 2019; House of Commons Home Affairs Committee, 2017), an increasingly frustrated US Congress demanding accountability from the tech giants, or other equivalent committees in Canada, Australia and elsewhere. The waring of Californian Democrat Senator Dianne Feinstein in 2017 to tech companies that 'I don't think you get it … You created these platforms, and they are being misused. And you have to be the ones to do something about it – or we will' (quoted in Flew, et al., 2019) was coming to fruition. The need to public action that went beyond managing such matters in house was apparent, as the threat of direct state regulation of digital platforms was becoming real.

The Oversight Board and private global governance

The most interesting response has been that of the Facebook Oversight Board (now the Oversight Board). Recognising the problems associated with the company having unilateral decision-making powers over content on its own platform, but being reluctant to cede governance powers to an external regulator, the Oversight Board was announced in May 2020 and commenced activities in 2021, with 40 high-profile members from throughout the world, and established with a governance structure, financial arrangements, and charter that sought to secure its structural independence from Facebook (Clegg, 2020; Douek, 2019). The processes through which the Oversight Board makes content decisions involve a series of steps, from referral of a case and selection of cases by the Oversight Board, deliberation and recommendations that are then presented to Facebook for implementation, which are in turn intended to set precedents for subsequent and similar cases. Factors that inform this process relate to the significance, or real-world impact of the content; whether it involves issues that are severe, large-scale, and important for public discourse; and whether decisions on the content raise wider questions about current policy or its enforcement, including uncertainly,

competing values, and disagreement about how it relates to Facebook's underlying policy or policies (Harris, 2019).

The Oversight Board, not surprisingly, has its critics. A group calling itself the Real Facebook Oversight Board, headed by Guardian investigative journalist Carole Cadwalladr, have alleged that it is a 'Potemkin Village' and corporate window-dressing with no substantive impact on the company's business practices. Others have wondered about whether a singular approach to speech rights and content moderation can adequately serve the diversity of the world's polities and communities. But my interest in the Oversight Board relates to the form of governance innovation it entails. This is not industry self-regulation in its traditional form, as the Oversight Board has been structured in such a way to enable its operations to be financially independent of Meta as its sponsoring entity. It also functions as a quasi-judicial entity, creating a body of case law that can form the basis for future decisions, along the lines of casuistry as discussed earlier. It is not a lobbying group for Meta or for the tech sector, and the standing of the participants is intended to dispel this concern, including as it does a former Danish Prime Minister, a former Guardian editor and various international human rights and legal experts. It is, however, established as an explicit alternative to rule-based regulation and nation-state regulation, making the case for 'soft law' and corporate social responsibility as alternative ways of addressing matters of public concern surrounding online speech to direct regulation by state agencies.

David Morar has described the Oversight Board as a 'single-company private governance institution' (Morar, 2021). It occupies a vacuum left on the one hand by the inadequacy of existing governance mechanisms for content moderation decisions, such as Terms of Service guidelines, and on the other by the hesitancy of nation-states in setting new rules for digital and

social media platforms around content rules. This is most apparent in the United States, where the First Amendment to the US Constitution sets a barrier to new forms of legislation governing speech, as it is contestable in the courts. It exhibits the characteristics of self-regulation that can enable action in the face of political gridlock, such as responsiveness, flexibility, greater compliance by the affected company, and informed and targeted intervention towards particular kinds of speech (Ayun, 2020). If the Oversight Board does prove to operate in a transparent manner that ensures due process and equal treatment and enables past decisions to set genuine precedent for ongoing content moderation activities, it may set up a model for other tech companies, and for governments keen to establish co-regulatory regimes with these companies, as the European Union's Digital Services Act envisages. Drawing upon what Golia and Teubner term *societal constitutionalism* (Golia & Teubner, 2021), Marta Maroni has identified the Oversight Board as a case study in 'how a private organisation develops its own set or understanding of fundamental rights, coupled with stronger administrative procedures to reinforce its organisation's position, legitimacy and autonomy' (Maroni, 2019). In doing so, she follows Golia and Teubner in observing that the emergence of such 'self-contained regimes' can be nonetheless 'intertwined with the creation of substantive rules in special fields of law, and the production of procedural norms' (Maroni, 2019).

What then should we make of the Oversight Board as a case study in what Golia and Teubner term 'constitutionality beyond the state' (Golia & Teubner, 2021, p. 360)? While it is early to make definitive judgements, four issues have been raised by analysts that require critical scrutiny:

1. Mandate: while the ongoing concerns about content moderation on social media have created the conditions for action, Facebook/Meta has done so in a manner that is

unilateral and based upon its own statement of values and community guidelines, without reference to rules and standards derived from an elected legislature, or a Supreme Court or another source that carries democratic legitimacy.

2. Legal foundation: while the Oversight Board Charter references other sources of law, such as international human rights law, its decisions are to be based upon the company's own values, content policies and community standards guidelines, and draw upon principles, such as 'authenticity', that have no legal foundation other than the company's own guidelines.

3. Accountability: the Oversight Board is ultimately accountable to the company rather than to recognisable democratic institutions, and there is a reliance upon their 'ethics of office' as a guarantee of due process and legal accountability.

4. Effectiveness: as the Oversight Board hears about 20-30 case per year, this constitutes a fraction – estimated to be less than 0.004% - of appeals received about content moderation decisions (Gopal, 2021).

Moreover, as Lakshmi Gopal points out, the Oversight Board 'is not able to comment on the algorithms that Facebook uses to organize and display user content or the balance that Facebook sets between user engagement and community safety' (Gopal, 2021). Its responses to the 'Facebook Files' released by whistleblower Frances Haugen were essentially procedural ones; it could not comment on the underlying critique of how the social media company draws upon controversy to drive engagement. Gopal has argued that 'Facebook seems to have left its most fundamental speech-related business decisions completely beyond the reach of the Oversight Board, leaving those paid experts … with only its most downstream, politicized, and public-facing controversies on disputed user-generated content' (Gopal, 2021). For these and other reasons, public interest advocates such as Rys Farthing and Dhakshayini Sooriyakumaran

have argued that light-touch self-regulation is the wrong path for governments to allow for social media giants, and that formal, 'black letter' law is required, combined with meaningful sanctions for non-compliance, and that measures such as the Oversight Board have not adequately addressed their systemic compliance issues around matters of serious public concern (Farthing & Sooriyakumaran, 2021).

Conclusion

Drawing upon the Foucauldian framework outlined at the beginning of this presentation, we can see the problem of censorship, as it was debated from the 1960s onwards, as one of 'too much government'. Moreover, it was the kind of government that exercised sovereign power over the content choices of its citizens: you cannot read *Lady Chatterley's Lover*, watch *Last Tango in Paris* etc. The move towards a classificatory regime that added a range of qualifying factors into such decisions, around age, location, context, artistic merit, and so on, can be seen as 'cutting the King's head off', to use an expression from Michel Foucault. Content regulation became, through classificatory regimes, a technology of advanced liberal government.

With new initiatives in the age of the platformised Internet such as the Oversight Board, we see content governance increasingly both globalised and privatised. Finding the right terminology for such initiatives is challenging; it is self-regulation, but not in the forms with which we are familiar. A term such as third-party quasi-self-regulation may capture some of the intent, as does private global governance. But it is also an exercise in Meta as a company seeking to derive competitive advantage from content moderation as a service, as well as addressing its own challenges in terms of misinformation and trust in its platform.

The concept of 'societal constitutionalism' has been offered recently, although I suspect this underestimates the extent to which global platform companies such as Meta continue to rely upon the tacit and at times explicit support of nation-state governments, as seen with current lobbying to restrict access to Chinese social media platforms such as TikTok, and the likelihood of national legislatures developing their own speech rules for social media platforms that supercede such in-house arrangements. The blurring of public and private forms of jurisdiction would certainly have made sense to Foucault. In *The Birth of Biopolitics*, he warned of 'state-phobia' and indicated that he avoided a theory of the state as one would 'avoid an indigestible meal' (Foucault, 2008, p. 77). He instead proposed that 'the state is nothing else but the mobile effect of a regime of multiple governmentalities' (p. 77). Rather than 'too much' or 'too little' government, contemporary liberalism in digital societies may instead be grappling with 'too much' or 'too little' governance, and the legitimacy, transparency, and accountability of the hybrid regimes of public and private actors responsible for overseeing such institutional arrangements.

References Cited

Australian Law Reform Commission. (2012). *Classification—Content Regulation and Convergent Media* [Final Report].

Ayun, C. (2020, May 6). The Facebook Oversight Board: An Experiment in Self-Regulation. *Just Security*. https://www.justsecurity.org/70021/the-facebook-oversight-board-an-experiment-in-self-regulation/

Balkin, J. M. (2018). Free Speech is a Triangle. *Columbia Law Review*, *118*(7), 2011–2056.

Carr, M. (2016). *U.S. Power and the Internet in international Relations: The Irony of the Information Age*. Palgrave Macmillan.

Castells, M. (2009). *Communication Power*. Oxford University Press.

Cohen, J. E. (2017). Law for the Platform Economy. *U.C. Davis Law Review*, *51*(1), 133–204.

de Sola Pool, I. (1983). *Technologies of Freedom*. Harvard University Press.

Ecarma, C. (2022, November 21). We are officially in the Elon Musk era of content moderation. *Vanity Fair*. https://www.vanityfair.com/news/2022/11/elon-musk-twitter-content-moderation

Farthing, R., & Sooriyakumaran, D. (2021). Why the Era of Big Tech Self-Regulation Must End. *Australian Quarterly*, *92*(4), 3–10.

Flew, T. (2018). Regulation beyond Government: Weber, Foucault, and the liberal governance of media content. In *Youth, Technology, Governance, Experience: Adults Understanding Young People* (pp. 48–64). Routledge.

Flew, T. (2021). *Regulating Platforms*. Polity Press.

Flew, T., Martin, F., & Suzor, N. P. (2019). Internet Regulation as Media Policy: Rethinking the question of digital communication platform governance. *Journal of Digital Media and Policy*, *10*(1), 33–50. https://doi.org/10.1386/jdtv.10.1.33_1

Flew, T., & Su, C. (2022). *Mapping International Enquiries into the Power of Digital Platforms* (CREATe Working Paper 2022/2). UK Copyright and Creative Economy Centre.

Foucault, M. (2008). *The Birth of Biopolitics: Lectures at the College de France 1978-79*. Palgrave Macmillan.

Golia, A., & Teubner, G. (2021). Societal Constitutionalism: Background, Theory, Debates. *ICL Journal*, *15*(4), 357–411.

Gopal, L. (2021, October 12). Facebook's Oversight Board & the Rule of Law: The Importance of Being Earnest. *American Bar Association Business Law Today*.

https://businesslawtoday.org/2021/10/facebooks-oversight-board-the-rule-of-law-the-importance-of-being-earnest/

Hayes, A. (2023, February 23). YouTube Stats: Everything You Need to Know In 2023! *Wyzowl*. https://www.wyzowl.com/youtube-stats/

Helberger, N., Pierson, J., & Poell, T. (2018). Governing online platforms: From contested to cooperative responsibility. *The Information Society*, *34*(1), 1–14. https://doi.org/10.1080/01972243.2017.1391913

Helmond, A. (2015). The Platformization of the Web: Making Web Data Platform Ready. *Social Media + Society*, *July-November*, 1–11.

House of Commons. (2019). *Disinformation and fake news': Final report* (p. 111). https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/1791.pdf

House of Commons Home Affairs Committee. (2017). *Hate crime: Abuse, hate and extremism online* (No. HC609). House of Commons.

Kretschmer, M., Furgal, U., & Schlesinger, P. (2021). *The emergence of platform regulation in the UK: an empirical-legal study* (CREATe Working Paper 2021/6). AHRC Creative Industries Policy & Evidence Centre/University of Glasgow.

Maroni, M. (2019, October 17). Some reflections on the announced Facebook Oversight Board. *EUI Centre for Media Pluralism and Media Freedom*. https://cmpf.eui.eu/some-reflections-on-the-announced-facebook-oversight-board/

Milmo, D. (2023, January 18). Twitter hit by 40% revenue drop amid ad squeeze, say reports. *The Guardian*. https://www.theguardian.com/technology/2023/jan/18/twitter-revenue-drop-advertising-squeeze-elon-musk#:~:text=Advertising%20is%20the%20main%20source,bn%20(%C2%A335bn)%20takeover.

Moore, N. (2014). Censorship Is. *Australian Humanities Review*, *54*, 45–65.

Morar, D. (2021, February 23). Facebook's Oversight Board makes an imperfect case for private governance. *Brookings Tech Tank*. https://www.brookings.edu/blog/techtank/2021/02/23/facebooks-oversight-board-makes-an-imperfect-case-for-private-governance/

Napoli, P. (2019). *Social Media and the Public Interest: Media Regulation in the Disinformation Age*. Columbia University Press.

Napoli, P. (2021). The Symbolic Uses of Platforms: The Politics of Platform Governance in the U.S. *Journal of Digital Media & Policy*, *13*(1).

---

[1] Platform companies as media companies debate: Napoli & Caplan; Flew et. al., Winseck