5 Do Emulated Guardians Work?

5.1 Introduction

As the previous two chapters explained how the two Emulated Guardians – the EU's out-of-court dispute settlement bodies (ODSs) and Meta's Oversight Board – came about, this chapter now vets them against the efficacy criteria of any functional adjudicatory guardian of organizational power. These criteria build on the work of Peter Cane, one of common law's most eminent theorists of administrative law. Namely, the three criteria are, first, a set of stable, external, and legitimate rules; second, an authoritative guardian; and third an organizational culture that holistically effectuates rather than undermines the guardian's external feedback.[1]

The key message of this chapter is twofold. On the one hand, neither of the studied Emulated Guardians works or is indeed designed to work as a truly effective adjudicatory overseer of corporate power. The mandates of both bodies are weak and many normative and factual arrangements severely hamstring their ability to effectively hold social media companies accountable. This is, arguably, a common feature of emulative institutions, which quickly reproduce formal structures and institutions, without holistically reproducing the enabling factors – here: democratic

[1] It is worth noting that the elements of "authority" and "culture" represent a slight modification of what Cane originally described as "institution" and "practice." The sociolegal approach of this book required a subtle shift in emphasis, as Cane's original categories – rules, institutions, and practice – seem, to me, deeply interconnected (after all, what are rules and institutions without their practice?). Thus, I took the liberty of focusing on an institution's "authority" and the cultures of implementation and accountability. Peter Cane, *Controlling Administrative Power: An Historical Comparison* (Cambridge University Press 2016).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

rulemaking, the rule of law, judicial authority – enveloping the role model in its original context.

However, on the other hand, as flagged in the introduction, theorized in the second chapter and empirically traced in the case studies, both bodies (or their creators) *performatively* play with the symbolism and language of all these contextual factors. Albeit mostly discursive rather than substantively legal, those appropriations of adjudicative symbolism may incrementally position bodies like ODSs or the Oversight Board (or any other Emulated Guardian for that matter) in the public mindset as, somewhat, serious guardians of power. We already observe that the authority and global standing of the Oversight Board seems to stabilize. Likewise, as the very recent practical developments in the context of the first ODSs showed, ODSs might creatively overcome at least some of their serious shortcomings the case study uncovered. They might, as discussed in greater detail, be important experiments for incorporating large language models into decision-making architectures, offering valuable lessons for future automated justice systems. To repeat a catch phrase from the introduction: Emulated Guardians are not powerful because they are courts. Emulated Guardians are powerful because they *appear and act like* courts. Formal labels do not matter here – it is the idea, the insinuation, the performance that counts. This potentially enables Emulated Guardians to incrementally expand their normative bind vis-à-vis the organizations they shall control. Notably, such incremental construction of authority is not at all dubious. Rather, gradual gains in authority and strategic expansions of initially tight mandates are a common adjudicative strategy well studied for domestic courts – think about Marbury v Madison – as well as international courts and the European Court of Justice.[2] The source of that normativity, however, would not primarily be formal mandate but discursive pressure.

In other words, effectuating public law guarantees vis-à-vis private power through emulative arrangements works, ironically, largely *besides* formal public law. As already hinted at in the introduction, this book argues that as long as classic public law ideas like individual rights and adjudicative accountability over powerful organizations remain socioculturally appealing they may be powerful drivers of socially demanded rather than merely regulated change. Making those public law ideas

[2] cf. Chester Neal Tate, 'The Global Expansion of Judicial Power: The Judicialization of Politics' in Chester Neal Tate and Torbjörn Vallinder (eds.), *The Global Expansion of Judicial Power* (New York University Press 1995).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

work though requires a different – in many ways more rudimentary and rawer – approach than conventional command and control regulations offer. The ethereal but humongous and deeply institutionalized spheres of globalized, private, and profit-oriented power structures, seem to adopt public law ideas mainly out of self-interest and in contorted form, instead of bona fide implementation of classic public regulation. We will look deeper into this aspect in the final chapter.

Therefore, to come back to the concrete program of this chapter, we now dissect the ODSs and the Oversight Board's ability to exercise effective guardianship vis-à-vis the private, for-profit power structures they are tasked to control. Conducting such an analysis is necessarily a gradual, context-dependent rather than binary endeavor. I therefore refrain from clearly labeling either of the two Emulated Guardians studied here as wholly effective, entirely ineffective, legal, illegal, problematic, or promising. Such labels are simplifying and, I believe, beyond the point.

Absolute assessments are only snapshots anyways. ODSs are just becoming operational in the months the final words of this book were written. There will be a rich variety of different ODSs, national variations and, arguably, social media site-dependent practices. Likewise, the Oversight Board, social media platforms and the whole informational economy develop so fast that "definite" evaluations are not only quickly outdated but at odds with my whole observation that these phenomena are deeply performative and, thus, subject to constant change. Therefore, instead, this chapter offers three criteria – *rules*, *authority*, *culture* – to evaluate the efficacy of emulative guardianship over private, for-profit organizations in globalized commerce. The example we look at here is content moderation by social media companies. But, as argued in the introduction, I believe the core idea to emulate individual rights adjudication (or other public law classics) to control powerful organizations beyond the state is all but assured to reappear once we face the next "new" challenge, be it the governance of "artificial intelligence," bio-, geoengineering, space exploration, or something else.

Before we begin, a few words on the elements of an effective control regime. In essence, there are three criteria: rules, authority, and culture. The criterion of "rules" in this context refers to the normative material used by an Emulated Guardian to contain private power structures. "Authority" then denotes the power and efficacy of an Emulated Guardian as an institution and its normative output. Lastly, "culture" points to the various elusive but crucial contextual enabling factors, especially in an organization's sociology and culture. These are especially

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

a culture to implement, in good faith, external feedback and honestly commit to public law principles like rights and accountability.

One concept I will frequently use as somewhat negative contrast to these three criteria is "ceremony." As a concept, ceremony originates from organization theory and denotes formal, external structures and organizational practices that accommodate public expectations but have little relation or effect on the internal actual practice of an organization.[3] In our example, a "ceremony" would be to establish a wholly ineffective but flashy and PR-savvy accountability forum. That would play well in overall discourse but change little in how platforms moderate down the road. Evelyn Douek called this "accountability theater."[4] In this sense, "ceremonies" and "accountability theater" is clearly negative. As shown in the second chapter, merely ceremonial emulations can be outrightly dangerous if they incrementally erode our shared sense of what kind of actor, what type of action, and what degree of rights protection and enforceability is even *possible* in globalized, transnational spheres. Remember that studies show, on the one hand, that privacy remains to be a cherished societal value, albeit, on the other hand, many grew to accept that privacy may be nice to have but is simply not enforceable. As one participant put it "this is capitalism."[5] If Emulated Guardians were only adjudicative ceremonies without practical benefit, they would, in the long run, erode our shared understanding of how far we *can* protect individual rights online.

Yet, at the same time, ceremonies, understood broadly as flashy but somewhat pointless habits are prevalent and arguably necessary elements of any organization and, arguably, society at large. Every formal organization, especially an adjudicator, features ceremonial structures and practices that are, at least partly, decoupled from their internal practices.[6] Judges wear robes and wigs, and rhetorically reaffirm their

[3] Seminal: John W Meyer and Brian Rowan, 'Institutionalized Organizations: Formal Structure as Myth and Ceremony' (1977) 83 American Journal of Sociology 340.

[4] Evelyn Douek, 'Content Moderation as Systems Thinking' (2022) 136 Harvard Law Review 526, 528, 533.

[5] Hagar Afriat and others, '"This Is Capitalism. It Is Not Illegal": Users' Attitudes toward Institutional Privacy Following the Cambridge Analytica Scandal' (2021) 37 The Information Society 115.

[6] Of course, ceremonial structures and practices in the sense of Meyer and Rowan are also practices in the sense adopted here. For Meyer and Rowan, the key distinction was not between formal structure and practice (although that also is a relevant distinction, as formal structures are especially "ripe" for emulation because they do not have to be necessarily backed by practice) but between internal practice and external structure, cf. Meyer and Rowan (n 3) 349 et seq.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

authority in phrases like "in the name of the people" and "all rise." These ceremonial aspects do not undermine the control of adjudicators or judges. Some of these aspects even reaffirm the adjudicator's controlling authority, performatively *making* the judge in the public's eyes. Thus, we see again, the deep ambivalence and transience of Emulated Guardians, which, in my eye, largely prohibit simplistic dismissals (especially as there does not really seem to be a politically feasible alternative to them).

5.2 Rules

We now proceed to the first element of a functional control regime of administrative(-like) power. That element is a set of stable, external, and democratic rules.[7] I argue that the two studied Emulated Guardians face three challenges in this regard. First, most of the rules they apply are embarrassingly undemocratic. Second, the Digital Services Act (DSA) seems to vindicate rather than restrict platform power as it implicitly declares platforms' unilateral, convoluted rules the main standard for free speech in Europe. This severely undermines the potential of ODSs to exercise meaningful guardianship vis-à-vis social media platforms in Europe. Third, as these phenomena are deeply ambivalent, I show how external, arguably legitimate rules like international human rights law, equally carry the potential to severely challenge core tenets of international law and governance. The example I present here is how the Oversight Board, a private body that shall control a private company, tweaks and twists human rights logics to argue that in "states of emergency" the *company* is under different policing duties vis-à-vis individuals *and sovereign states* (or to be precise, their social media accounts). One does not have to invoke an infamous German scholar of the first half of the twentieth century to detect the thinly veiled allusions to sovereignty and potentially paradigm shifting effects of such arguments for international law and governance.

5.2.1 Platform-Made Rules Are Undemocratic

Adjudication is third-party review. It only works if the allegedly controlled party cannot unilaterally amend the rules upon which its control is based. In other words, the whole point of individual rights adjudication is that it controls whether a powerful organization, that is, historically

[7] Cane (n 1) 2.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

the state, now, perhaps, social media platforms, abides by normative guarantees that are designed to limit their power. It is questionable to which extent Terms of Service and Community Guidelines, which are the most important sets of normative material here, are indeed designed to limit the power of platforms. As Gillespie put it, those community guidelines are "inevitably, provisional lines drawn in shifting sands, the residue of past tensions, and best efforts at workable compromises."[8]

In contrast, classic normative guarantees that are, at least in principle, designed to limit power are fundamental rights, international human rights law, or transnational or national regulation. To an extent, also normative material generated, in one form or the other, by the organization itself may self-restrict organizational power. After all, when courts review the legality of administrative action in national or EU contexts, they also look at statutory (and constitutional) law, which originates just as much from public power as the administrative agency does. However, the key factor waterproofing a rule as a functional basis for adjudicative control over power is the process of its genesis. In other words, while formal labels like public or private seem not decisive, what is decisive is the democratic viability of the rulemaking process. This is particularly important when the substantive matter is speech – a deeply sociocultural phenomenon. If an organization makes rules about online but nonetheless public discourse, those rules should tie back to the people they govern. In a public context, therefore, the rules that constrain power originate, in principle, from a democratic, transparent, and formalized process. Even though administrative agencies are as public as the legislator, the rules governing administrative agencies are not mostly made unilaterally by those agencies. And even those rules which are made unilaterally by administrative agencies only come about in somewhat democratized procedures. These procedures enable affected individuals to participate in the process of rulemaking.

None of this happens for platform rulemaking. Platform rules (meaning terms of service, community guidelines, and all the sublayer rules undergirding them) are unilaterally made and amended by the platforms at whim. One interviewee who advised Meta describes the process of making these rules as "basically a room full of US lawyers sits down and makes the rules after the latest crisis."[9] It is jarring that

[8] Tarleton Gillespie, *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media* (Yale University Press 2019) 71.

[9] Interview with Academic, Consultant to Facebook in 2018 (1 November 2021).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

many – including the European Union as Chapters 2 and 3 show – came to unquestioningly accept this rather brazen and unintuitive privatization of rulemaking power over absolute core issues of social, cultural, and political speech.

The ubiquitous counter argument is freedom of contract and property. The property argument is that social media companies may use their property – infrastructure, code, and so on – as they please. After all, platforms do own these infrastructures. The contract argument, pretty much up the same ally, holds that parties are free to agree to whatever they want and as users sign, or rather click, those agreements, who are we to weigh in? Both arguments are, frankly, ridiculous. Oil giants cannot dump crude oil in open waters, even though they own the crude oil. Even if they would indeed also own the waters, they would not be allowed to do so. Big Tabaco cannot legally tell us that cigarettes are healthy, and even if I were to legally buy the raw chemicals to produce LSD, I would not be allowed to do with them whatever I want. Property and contract are free only if their usage does not harm others. That is the whole point of living in a society and having laws. However, collective goods – such as each individual's personal construction of the self and society's foundational ability to communicate and address problems – must be governed by principles that extend beyond the pursuit of maximum profit by whoever was fastest to exploit a novel business model. Thus, clearly, reforming platform rulemaking is a question of political will.

5.2.2 *The DSA Empowers Rather Than Controls Platforms*

The DSA does not remedy the democratic deficit of platform rulemaking. But that is not it. The DSA also stabilizes platforms' preference for internal, rather than external rules to "control" their behavior. The control of content moderation offered by ODSs in this model is, in practice, largely restricted to assessing whether platforms abide by their own rules, that is, the community standards.

Let me explain this. The DSA envisages platforms to base their moderation decisions on two sets of rules. These are, first, the rules made by the platform itself, and second, Member State and Union law. The first category is called "incompatibility" with terms of service; the second category is called "illegality." Illegal refers to any content that "is not in compliance with Union law or the law of any Member State which is in compliance with Union law, irrespective of the precise subject

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

matter or nature of that law," according to Article 2(h) of the DSA. These rules are external to platforms and should therefore provide a functional normative basis for adjudicatory control over platform enforcement. However, due to foreseeable practical impediments, this might not in fact be the case.

The DSA's definition of illegality is so broad that platforms might not be able to identify the relevant rules that should apply to them.[10] Similarly, some EU lawmakers interviewed for this present work assumed that only a minority of the moderation decisions will focus on "illegality," whereas possibly the majority of moderation decisions and ensuing disputes will relate to the platforms' terms of service.[11] That seems plausible. Platforms' enforcement of their own terms of service is a lot more seasoned, centralized, and established than the enforcement regimes required to take down any content that the DSA deems "illegal," especially in a pluralist and federated bloc like the EU.[12] Large social networks like Facebook or Instagram or other content platforms like YouTube moderate billions of pieces of content per month. I am skeptical that it is truly realistic for platforms to entirely overhaul their content moderation regimes to accommodate the specific interests of any given Member State, irrespective of its subject matter or nature."[13] The organizational effort to reflect the specifics of twenty-seven domestic legal systems in platforms' enforcement regimes far outweighs the benefits for platforms. Remember, some Member State legal frameworks arguably cover fewer platform users than the number of people living in a large city. In some cases, it might not even be a particularly big city – Malta, Luxembourg, and Cyprus together do not quite equal the number of inhabitants of Brooklyn. Therefore, from a legal realist perspective, it seems unlikely that platforms will substantively adapt their organizational structures to accommodate the normative specifics of "smaller" Member States.[14] As previous empirical work has uncovered evidence for, platforms favor a centralized system of

[10] Article 2(h) DSA.

[11] Interview with European Commission Staffer 1 (8 July 2022); Interview with European Commission Staffer 1 (29 October 2021); Interview with European Parliament, Staffer 2 (25 November 2022).

[12] Regarding the EU's categorization as federal system see Robert Schütze, *From Dual to Cooperative Federalism: The Changing Structure of European Law* (Oxford University Press 2009).

[13] For the quote see Article 2(h) DSA.

[14] cf. Lauren B Edelman, 'Legal Ambiguity and Symbolic Structures: Organizational Mediation of Civil Rights Law' (1992) 97 American Journal of Sociology 1531.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

rules they can apply as uniformly as possible.[15] Scale is a key component of platforms' business models.[16]

From an EU law perspective, that hints to an important finding. On the one hand, many of the DSA's key provisions read so vaguely that platform discretion and ceremonial adaptation may undercut any "Brussels Effect."[17] On the other hand and running counter to the logic of the "Brussels Effect" as described by Bradford, the DSA does indeed *not* establish a uniform, EU-wide substantive standard. Quite the contrary. When it comes to law enforcement, and intuitively so, if one thinks of the EU's federal structure, the DSA tries to enable Member States to muscle through their own standards but does not establish a common European one. Given European pluralism, that makes sense.

However, whether the DSA's uniform procedural framework suffices to motivate platforms to enforce, say, Maltese, Luxembourgian, or Cypriot law (and that of many other Member States) remains to be seen. They might enforce such laws when asked to in notice-and-takedown procedures; but proactively rather not. Eventually, "illegal" content, especially if illegal only as defined by nationally applicable legislation in some Member States, might not play a central role in platforms' moderation efforts. If such content is not deleted for its illegality but for its incompatibility with terms of service, however, ODSs will not adjudicate on the "illegality" of content.[18]

The other of these two categories, "incompatibility" with a platform's terms of service, imperfectly builds a control system on adjudicating

[15] Gillespie (n 8) 45 et seq. See also Kate Klonick, 'The New Governors: The People, Rules, and Processes Governing Online Speech' (2018) 131 Harvard Law Review 1598, 1631.

[16] cf. Stigler Center for the Study of the Economy and the State (University of Chicago), 'Stigler Committee on Digital Platforms Final Report' (2019) 34 et seq.

These are, incidentally, pertinent observations for debates about the EU's democratic deficit. In principle, one of the EU's major democratizing effects—despite other institutional flaws—has been to level the playing field between smaller and larger European states. However, federalizing enforcement and normative input vis-à-vis globally active, infrastructure-controlling, and vertically integrated digital corporations may undermine this empowering effect. Against that backdrop, the EU faces a difficult choice: while internal differentiation holds appeal from a legal pluralism perspective, the Union's strength in the global arena – whether in trade or digital regulation – derives precisely from acting as a cohesive bloc.

[17] Martin Husovec and Jennifer Urban, 'Will the DSA Have the Brussels Effect?' (*Verfassungsblog*, 21 February 2024) <https://perma.cc/DCX4-6FYK>; Moritz Schramm, 'Curb Your Enthusiasm: Why the Digital Services Act Might Not Have a "Brussels Effect"' (2025) European Journal of Legal Studies (Special Issue) 61.

[18] If at all, ODSs will adjudicate the "illegality" mainly based on appeals launched by so-called "trusted flaggers" (people who report other users, see esp. Article 22 DSA) but those will arguably be way fewer cases than by people whose content was deleted for violating a platform's community standards.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

internal rules. These platform-made rules, however, cannot be an appropriate normative yardstick for controlling platforms because platforms have so much leeway in devising and amending them. The best outcome that such control could bring is what one adviser to the Commission called an "arbitrariness check" – that is, monitoring whether platforms follow their own dysfunctional rules.[19]

Intuitively, one might say that for controlling private power such scanty control could suffice. Moreover, indeed, stringent enforcement of platform-made rules would signify progress compared to the situation we have now. However, the root of the problem is not only platforms' enforcement of their rules but how these rules have come about. Terms of service (and the cascading community guidelines) are neither procedurally nor normatively linked to the individuals they govern. As argued above, terms of service are deeply undemocratic. They mirror what a group of corporate lawyers and product managers at social media companies deem good policy (for the company, not necessarily the public writ large). Yet, those rules code and normalize a very particular and presumably rather corporate understanding of global speech norms.[20] This undermines the EU's own goal to reign in on digital corporations. Therefore, the rules determining free speech in Europe are unlikely to be any more "European" than those determining online speech in Indonesia, Argentina, or South Africa. Overall, we will have (or already have) a global standard. That standard is not set by the EU but by private companies. And with the DSA, the EU valorizes rather than democratizes this corporate standard.

5.2.3 *From Flawed Design to Better Practice*

That being said, not every design flaw ends up staining later practice. It seems very well plausible that at least some ODSs overcome those in-built design flaws and develop ODSs into more than they are designed to be. The developments around ADROIT, User Rights, and the Appeals Centre – three ODSs fresh off the shelf – offer a shimmer of entrepreneurship and institutional innovation. As some of these new ODSs are run by people coming from content moderation, rather than from a consumer protection side, we may observe strategic sensitivity that external and concise rules are important, and that solid adjudicative argumentation and cautious institution-building

[19] Interview with Academic, Consultant to the European Commission (29 June 2022).

[20] This reminds of the "coding" of corporate law through global elites of corporate lawyers described in Katharina Pistor, *The Code of Capital: How the Law Creates Wealth and Inequality* (Princeton University Press 2019).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

are arguably necessary to make an ODS work. At the same time, as ADROIT indicated, also the consumer protection ADR has its place in this new ecosystem.

Similarly, the Oversight Board quickly overcame some of its in-built design flaws. For example, the Oversight Board's charter establishes that the Board's main review function is to "determine" whether content enforcement decisions are "consistent with Facebook's content policies and values."[21] In contrast, the role for human rights law, which would have been an external and legitimate basis for review, could be read as merely decorative. The Board should only "pay particular attention to the *impact* of removing content in light of human rights norms protecting free expression" but, at least according to the wording of the charter, the Board should not use international human rights law as a basis for its review of Meta's enforcement decisions.[22] Judging from its normative fundamentals, the Board was thus largely restricted to applying Meta's own, internally devised rules. However, for meaningful adjudicatory review, Facebook's content policies and values are woefully inadequate. In the Board's own words, they are often "inconsistent" and "lack clarity" while the way they are handled by the company remains "vague" and "uncertain."[23] Further, Meta may (and does) unilaterally amend these policies and values, thereby rendering any adjudicatory control based on them largely ceremonial.[24] Therefore, if we look at the Oversight Board's institutional design, we must conclude that this design could be viewed as largely ceremonial.

However, in practice, the Oversight Board overcame this in-built limitation of only being able to work with Meta's internal rules. As argued in Chapter 4, the Board does in fact draw upon international human rights law to derive normative standards for Meta's actions.[25] The Board regularly assesses Meta's internal decision-making, communication with individuals, and the quality of the rules Meta applies in terms of whether these adhere to international human rights law. Admittedly, the Board's normative assessments often focus on establishing standards and not on the kind of extensive reasoning that occurs in domestic courts.[26] In other words, the

[21] Charter Article 2 Section 2.
[22] Charter Article 2 Section 2 (emphasis added); see also in detail at 4.4.2.1.
[23] For the former see Breast Cancer symptoms and nudity [2021] Oversight Board 2020-004-IG-UA. For the latter see Former President Trump's suspension [2021] Oversight Board 2021-001-FB-FBR. at section 8.3.I.
[24] See at 4.1.2.2 and 4.5.4.
[25] See at 4.5.
[26] Ibid.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

Board itself expounded an external normative framework beyond Meta's control. It arguably did so out of self-interest. Adjudicating based on, and in the language of, human rights is a legitimacy-affording asset for the Board vis-à-vis the company (and other organized powers like states or the EU). And it seems to be working. Some commentators are already comparing the Board to a UN human rights treaty body and even point to a "human rights future."[27] For the Board, recourse to human rights is a means to build legitimacy incrementally. Hence, despite the normative straitjacket imposed on it by the charter, the Board seems to be emulating what adjudicators have arguably done for centuries: reinterpreting its role more broadly than initially foreseen.

Does all this mean that the Board or experiments like it will single-handedly save us? Surely not. But it shows that dynamics of contestation, voice, and accountability, of reform and critique follow power structures we too often view only through the limiting lens of private law and economics.

5.2.4 *Why Human Rights Law Is Ambivalent Too*

Broadly and boldly interpreting its role as human rights guardian also comes with ambivalence. For example, its turn to international human rights law might legitimize the Board vis-à-vis the company but simultaneously somewhat shields it from national regulation. This is, from Meta's perspective, of course a good thing. However, from a democratic perspective one may rightly question the legitimacy of claims of a private actor which basically ties itself to public international law – on very convenient terms, as none of this public international law is formally binding for it – instead of the usually more strict and (perhaps) concise regulatory demands of, say, the European Union.

The ambivalent role of public international law as grounds for private adjudicative review for private power becomes particularly vivid whenever actual international actors are involved. By now, there are several Oversight Board cases concerning, in one way or the other, government speech. There is, of course, the famous Trump decision. But there are more. For example, in the context of ethnic unrest in the Tigray region in Ethiopia, the Ethiopian government ushered several statements through

[27] Laurence R Helfer and Molly K Land, 'The Facebook Oversight Board's Human Rights Future' (2023) 44 Cardozo Law Review 2223; see also Nele Meier and Angelo Jr Golia, 'The Emerging Normative System of Meta's Oversight Board: An Introduction' MPIL Research Paper Series No. 2022-09 41.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

its social media accounts that could easily be described as inflammatory. In fact, inflammatory statements occur regularly on social media profiles of government officials, as we saw, for example, also in Cambodia, Myanmar, and the United States. So, the Board faced the somewhat fundamental question how it and the company should deal with government communication that could be seen as at least mediately entailing the risk of further violence and real-world harm.

Some of the "doctrines" developed in this context read like a direct challenge to core tenets of international law. Let us, for example, have a look at the Alleged Crimes in Raya Kobo case. In Alleged Crimes in Raya Kobo the Board had to assess whether Meta was right to delete content posted by a private individual in the context of the Ethiopian conflict with Tigrayans. One critical factor in this case was that, although that specific piece of content had not been posted by the Ethiopian government, that same government been posting similar or even more concerning things. Thus, discussing Meta's right to remove this specific piece of content implicitly discussed whether Meta was allowed to moderate similar content posted by the Ethiopian government. For example, only several weeks before the decision was published, Meta removed a post by the Ethiopian prime minister urging civilians to take up arms and "bury" Tigrayan rebels.[28] It should be noted that UN experts sent to Ethiopia have repeatedly highlighted the prevalence of hate speech and its contributory role in atrocities.[29]

The question the Board then faced was how to legitimize Meta's role as a "speech" moderator in potentially life-threatening circumstances.[30] The

[28] 'Facebook Deletes Ethiopia PM's Post That Urged Citizens to "Bury" Rebels' (*BBC*, 3 November 2021) <https://perma.cc/7VAB-XABF>.

[29] See e.g. Human Rights Council, 'Report of the International Commission of Human Rights Experts on Ethiopia, A/HRC/51/46' (2022) para 23, and explicitly referencing the 2014 UN Framework for Analysis of Atrocity Crimes 127.

[30] Regarding the situation on the ground, the Board assessed: "the content could contribute to ethnic hatred. Ethnic conflict situations call for heightened scrutiny over how people should report on and discuss human rights violations committed by parties to the conflict. These considerations apply to Facebook posts, which can be used to spread unverifiable rumors at great speed. The Board notes that some Ethiopian government officials have instigated or spread hate speech targeting Tigrayans . . . On its own, an unverifiable rumor may not directly and immediately cause harm. However, when such content appears on an important, influential and popular social media platform during an ongoing conflict, the risk and likelihood of harm become more pronounced. . . . Furthermore, cumulative impact can amount to causation through a 'gradual build-up of effect,' as happened in Rwanda where calls to genocide were repeated (see the Nahimana case, Case No. ICTR-99-52-T, paras 436, 478, and 484-485)." Alleged Crimes in Raya Kobo [2021] Oversight Board 2021-014-FB-UA 8.3 III.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

Ethiopian government's involvement in communicating hate speech further complicates the matter. Some of the communication the company would have to suppress, and already had suppressed,[31] directly or indirectly stemmed from the (democratically elected) Ethiopian government. And yet serious allegations of human rights violations were being directed at that self-same government.[32]

Therefore, one of the main questions was how to balance, on the one hand, Ethiopian sovereignty and, on the other, the potentially lethal effects of (government-propagated) hate speech in that context.[33]

The human rights rules the Board applied to this situation go to the heart of public law and the notion of sovereignty. The Board framed this whole predicament as a balancing question between, on the one hand, freedom of expression, which in this case would protect the communication of the government on Meta's infrastructures, and on the other hand, Ethiopian sovereignty. Remarkably though, the Board held the company's commitments to free expression would and should be "derogate[ed] in times of *public emergency*."[34] It would be, unsurprisingly, the company which ultimately decides then what counts as state of emergency. In its decision, the Board reasoned:

[31] See above n 28.

[32] UN Office of Genocide Prevention and the Responsibility to Protect, 'Statement by Alice Wairimu Nderitu, United Nations Special Adviser on the Prevention of Genocide, on the Continued Deterioration of the Situation in Ethiopia' (2021) <https://perma.cc/ZX3W-V4V4>; Office of the High Commissioner for Human Rights, 'Ethiopia: Bachelet Urges End to "Reckless" War as Tigray Conflict Escalates (Press Release)' (2021) <https://perma.cc/L2BY-GWPX>.

[33] Reminiscent questions surfaced when the Board had to decide whether Meta could permanently ban then US President Donald Trump from its platforms. Prior to his ban, the former President used Meta's services in ways plausibly contributing to the violent attack on the US Capitol on 6 January 2021. The Board noted that "[Trump's] posts were made during a dynamic and fluid period of ongoing violence. There was a clear immediate risk of harm to life, electoral integrity, and political participation. The violence at the Capitol started within an hour of the rally organized through the use of Facebook and other social media. . . . Analyzing these factors, the Board concludes that . . . Facebook's imposition of account-level restrictions on January 6 and the extension of those restrictions on January 7 was necessary and proportionate," Former President Trump's suspension (n 23).

Even though the Board did not directly invoke the "public emergency" argument in its Trump decision, functionally, its reasoning is similar. Because of the real-world harm that followed the President's remarks (several people died at the Capitol riot, not to speak of the lasting damage to American democracy), the company restricted official communication on political issues – even if uttered by a democratically elected head of state.

[34] Alleged Crimes in Raya Kobo (n 30) 8.3 (emphasis added).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

The Board finds that in a country where there is an ongoing armed conflict and an assessed inability of governmental institutions to meet their human rights obligations under international law, Meta may restrict freedom of expression that it otherwise would not (see ICCPR Article 4 on derogations in times of *public emergencies*).[35]

In simple terms this means that Meta, a private, for-profit company declares a state of emergency vis-à-vis a government and may then restrict the communication of said government. And the core argument for this is the ICCPR. This is, from a conventional perspective a pretty remarkable argument. The traditional "who does what" in international law is that states ought to protect human rights and may, under very narrow circumstances, declare emergencies. In contrast, private corporations typically have little to do with neither, especially the latter.

From a human rights perspective, some might find the Board's expansive reasoning logical. For decades, scholars and activists have demanded stricter human rights responsibilities from nonstate actors, and especially private companies.[36] Indeed, as this book seeks to emphasize, thinking about the power of some corporations discursively triggers and theoretically warrants conceptualization in public law terms. It does so, incidentally, not because these companies should be likened to states, but because public law is fundamentally about anchoring organized power to the public will and interest – regardless of the historically contingent nature of that power.[37] Using international human rights law instead of Meta's own rules makes the Board appear less ceremonial and underscores the fact that the Board takes its control function seriously. This is, from a purely functional viewpoint, a positive sign. Who would not agree that powerful organizations should strive to mitigate the social and political perils that may arise from "improper" use of their products and services, especially in scenarios of possible ethnic cleansing or even genocide. It would be absurd to assert that private governance actors

[35] Ibid 8.3 III (emphasis added).

[36] For an overview over the debate see Andrew Clapham, *Human Rights Obligations of Non-State Actors* (Oxford University Press 2006); from the UN context see esp. John Ruggie, 'Report of the Special Representative of the Secretary-General on the Issue of Human Rights and Transnational Corporations and Other Business Enterprises, John Ruggie' (United Nations, Human Rights Council, 17th session 2011) A/HRC/17/31; and Astrid Sanders, 'The Impact of the "Ruggie Framework" and the "United Nations Guiding Principles on Business and Human Rights" on Transnational Human Rights Litigation' in Jena Martin and Karen E Bravo (eds.), *The Business and Human Rights Landscape* (Cambridge University Press 2015).

[37] See in this context also n 334 on sovereignty in the previous chapter at 4.5.5.1.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

like Meta should stay on the sidelines and thereby become complicit in government actions like inciting racial hatred or thwarting an election.

5.2.5 *Private Checks on Public Power?*

However, a bitter aftertaste remains. In front of these scenarios, we develop a better sense of how fundamentally Emulated Guardians may transform global governance. One does not need to invoke a certain German scholar to stress the implications of granting a private corporation the authority to decide what constitutes a "public emergency." Allowing Meta to declare a "public emergency," would allow the company to suspend its own commitment to respecting free expression in a specific context. It alleviates the company to what could be described as ultimately "sovereign" position, whereas individuals as much as states are subjects to the platform's unilateral emergency powers. In short, the Oversight Board's invoking of a "state of public emergency" shows how platforms incrementally become private *checks* on public power.

In principle, it is not so remarkable that private companies seek to sway and "discipline" public actors in situations of atrocities or other human rights violations. There were, for example, company boycotts of Apartheid-era South Africa or the Russian Federation after its full-scale invasion of Ukraine in 2022. Also, some American companies pledged to fulfill emission-reduction goals even though the Trump government pulled out of the Paris Accord. Yet, the situation described here appears different as it directly emulates a legal and quite direct, formalized method of disciplining public power. When online platforms restrict government communication, governments lose an important channel for communicating and conducting their policies. This can be viewed as a de facto, corporate, private check on public power, which is a remarkable development. In fact, one interviewed Board member mentioned – as already briefly discussed in the previous chapter – that they were "acutely aware of the implications of statements like that of a public emergency."[38]

So far, the Board has exercised this check in line with "liberal" expectations. It asserted that governments should not encourage ethnic hatred and/or impede the peaceful transfer of power. From a normative perspective, the Board arguably could hardly have decided otherwise. The Board's normative reasoning has perhaps even helped prevent future harm.

[38] Interview with Oversight Board Member (15 November 2022). For the discussion in the previous chapter see at 4.6.3.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

However, this specific Emulated Guardian, the Oversight Board, is still a private body without a democratic mandate; Meta even more so. Both lack meaningful procedural or participatory connections to the people they are claiming to protect. Hence, Emulated Guardians risk delegating decisions to actors lacking the (legitimate) procedures, norms, and institutions originally tasked with making such decisions – that is, international law and governments.

However, would private complicity in alleged government terror be the better solution? Hardly so. Are despotic governments more legitimate decision-makers than private organizations, especially when the latter are nominally accountable to an Emulated Guardian? Hard to say. Do we have any guarantee that private power structures align themselves meaningfully with public law ideals or are they themselves – arguably hinging on the political preferences of their owner – prone to authoritarian tendencies? Properly answering those questions requires thoroughly deconstructing notions of public and private as well as the presumed higher normative legitimacy of ideals of public power which, this self-awareness seems warranted, is a presumption much more common in Europe than the United States. Would it be even desirable if state-oriented notions of international law block novel, adjudicative actors operating outside the conventional scope and mechanisms of international law?

The empowerment the Board derives from international human rights law vis-à-vis the company also, paradoxically, encapsulates the Board and, by extension, the company from regulatory intervention. Visualized in the classic pyramid of legal hierarchy, the Board tries to elevate the problems Meta faces to the international-law level, thereby decoupling Meta's governance from domestic or supranational regulation. This feeds into the narrative of successful self-regulation, pushing the company's administration-like practices (discursively) out of the reach of regulators. The Board's coy labeling of itself as a "non-judicial grievance mechanism" thus belies the reality that Meta retains decision-making power over what constitutes a "public emergency."[39] Such decision-making powers are not trivial or necessarily business-related but historically linked to the very heart of sovereignty.

There are no easy, and hardly definitive answers to these questions. All point to a rather fundamental reshuffle of power, form, and expectation. What seems for certain though is, to reiterate, that public law ideas – like

[39] Claimed COVID cure [2021] Oversight Board 2020-006-FB-FBR at 8.3; see also Meta, Corporate Human Rights Policy at 03, accessible via <https://perma.cc/EK3G-FC7E>.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

individual rights adjudication, judicial review of administrative power, fundamental rights, and so on – increasingly escape the classic forms and mechanisms of public law. Yet, public law ideas, for their sheer sociocultural entrenchment, appear to remain a *leitmotif* of future discourse and practice shaping future governance institutions. Those future governance institutions, however, may build on past experiences with public law without being public legal actors themselves. This is a challenging, and at least from a scholarly perspective, vastly exciting prospect. We will discuss this outlook in more detail in the remaining pages of this book.

5.3 Authority

The second criterion determining the might of the two Emulated Guardians studied here is their adjudicative authority. Are the EU's out-of-court dispute settlement bodies and the Oversight Board authoritative enough to steer private governance operations into a more normatively desirable direction? In other words, can they exercise meaningful control vis-à-vis the corporate behemoths of our time? Can they – and, by extensions, Emulated Guardians as a class of actors – remedy rights violations and thereby advance justice for individuals? Do nonbinding decisions and recommendations foster accountability or merely sugar-coat unchanged practices?

As seen in the case studies, both Emulated Guardians have several means at their disposal to exert control and advance accountability. However, in comparison to traditional public law adjudication, these mechanisms are embryonic. ODSs' decisions are expressly nonbinding.[40] The Board's decisions "bind" Meta, but the Board makes so few decisions that any normative influence it has resides mainly in its nonbinding recommendations.[41] To put it bluntly, the authority of Emulated Guardians to formally adjudicate is minuscule. Yet, neither is formal mandate the only source of authority, nor are formal legal rules and decisions the only well of normativity. Broadly understood, the authority of accountability forums like Emulated Guardians depends on their efficacy – in which way whatsoever – to curb corporate misconduct and advance individual freedom vis-à-vis the dominant power structures of the digital sphere. The fact that Emulated Guardians, to an extent, mirror the judicial process does not necessarily mean that they are limited to the

[40] See at 3.6 and 3.7.
[41] See at 4.6.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

same tools an, anyways idealized, understanding of courts would imagine, that is, communicate exclusively through binding decisions. Rather, Emulated Guardians offer various forms of external feedback and, inevitably, turn to society as main driver to force digital corporations to implement that feedback. If your feedback is not *legally* binding, you may still orchestrate enough public pressure to nonetheless lend it the weight of thousands or perhaps even millions of disgruntled individuals.

Both types of Emulated Guardians, the emerging rich variety of ODSs as well the Oversight Board, feature mechanisms that seem to buttress their nonbinding normative pronouncements vis-à-vis the private governance structures they control.[42] These could be costs, public pressure, and dialogue-based engagement.

We will focus here on the latter two elements. At the time of writing, we simply lack empirical data to assess the efficacy of the DSA's litigation fees in the context of Article 21; even more so the disciplinary efficacy of Commission fines under the DSA.

In contrast, we have plenty of real-world examples of attempts to rein in on Meta through dialogue and public pressure. Most examples are not necessarily promising. For example, whistleblower Frances Haugen, a former high-ranking Facebook executive, testified to Congress that

> The company's leadership knows ways to make Facebook and Instagram safer and won't make the necessary changes because they have put their immense profits before people. . . . [I]f there is one thing that I hope everyone takes away from these disclosures it is that Facebook chooses profit over safety every day.[43]

So, given such dire in-house perspectives, how persuasive and effective is Emulated Guardians' feedback to corporate actors really? As discussed in the third and fourth chapter, the Oversight Board and ODSs differ in many respects. The Oversight Board is a relatively small, unitary, and tightly run organization. It has a clear focus on structural deficits within Meta. It decides relatively few cases per year and its main normative output are not its decisions but its nonbinding recommendations, which it seeks to buttress by naming and shaming efforts and strategic communication. In contrast, ODSs will be very diverse. The four ODSs certified at the point of this writing all are vastly different.[44] They will reflect the diverse jurisdictional cultures from which they originate. Being a business themselves, ODSs must decide as many cases as possible.

[42] See at 3.7 and 4.6.
[43] 'Statement of Frances Haugen' <https://perma.cc/3QN6-JMY4>.
[44] See in detail at 3.8.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

They are not designed to orchestrate extensive naming and shaming campaigns or engage in other forms of strategic communication. Rather, the key factor behind Article 21 DSA is mass. The EU envisages simply so many cases that platforms, at some point, simply "have to" comply. The crux is, however, whether enough users will indeed play along as it is unclear whether the prospect of nonbinding procedures with unclear effects indeed motivates enough people to ramp up pressure vis-à-vis platforms.

Therefore, this section investigates two things. First, we assess whether Meta indeed changes its practice based on the Oversight Board's feedback. Second, we question, from a more theoretical angle, whether the European social media users are likely to be altruistic enough to effectuate ODSs, which, lacking binding decision-making powers, might not be able to remedy all those individual grievances.

5.3.1 Does Meta Care for the Oversight Board?

So, does Meta take the Board seriously enough to adopt its recommendations, accept its normative guidance, and ultimately subject itself to the Board's emulation of public law adjudication?

Referring to "key audiences," the Oversight Board's decisions have been examined by stakeholders and legal academics and, as we have seen, Meta has verifiably followed some of its recommendations. Clearly, Meta's lack of transparency as to how far it really applies with the Board's recommendations is still a problem. One interviewee at the Board, for example, said they thought many people at Meta "do not care about the Board."[45] The shifts in corporate policy and personnel in early 2025, aimed at aligning Meta with the second Trump administration – including the outright denunciation of previous content moderation as "censorship" – do not inspire much confidence that Meta's rank and file will show greater regard for the Board in the near future. A particularly dispiriting aspect of this policy shift was, arguably, the manner in which it unfolded. When Zuckerberg decided to ban then-US President Trump from Meta following his use of the platform during the storming of the Capitol on January 6, 2021, he quickly engaged with the Oversight Board to outline the rules for removing high-profile figures. Yet now, as Meta scales back enforcement and revises its policies, he appears to have bypassed the Board entirely. Moreover, given the Board's own assessment that the company had only verifiably fully put in place fifty-seven out of 288 of

[45] Interview with Oversight Board Staffer 6 (23 November 2022).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

its recommendations by November 2024, we cannot talk of meaningful accountability yet.[46] However, the fifty-seven recommendations the Board deemed *had* been verifiably fully implemented are a significant improvement compared to the sixteen recommendations in that category in the first quarter of 2022.[47]

Figure 5.1 shows another trend worth monitoring. That is the decline in appeal numbers, which have steadily decreased since early 2022. The implications of this decline are unclear. On the one hand, it could either indicate user disillusionment with the Oversight Board – perhaps due to the near-zero chances of their cases being adjudicated or Meta's inconsistent implementation of the Board's recommendations. Assuming the average user is less aware of the developments surrounding the Board than the somewhat niche group of people studying it, the former explanation seems more plausible.

On the other hand, the decline might reflect a normalization of user interest, as users gain a clearer understanding of the Board's purpose and the initial hype – especially following the Trump decision, which likely drove the mid-2021 to early-2022 spike – gradually subsides. If so, this is not necessarily a bad thing. Fewer appeals could, in fact, simplify the

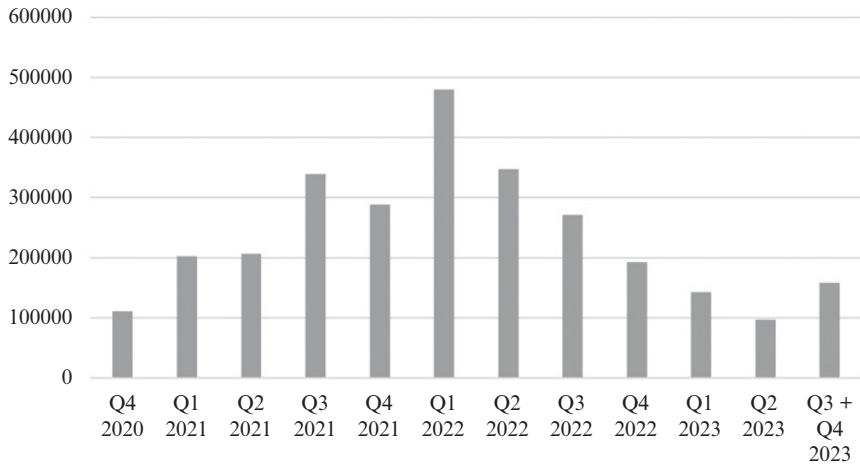


Figure 5.1 User appeals to the Oversight Board (2020–2023)

[46] Oversight Board 'Recommendation Tracker' as of November 2024, accessible via <https://perma.cc/4M6Y-WX7Q>.

[47] Transparency Report Q 1 2022, p 21, accessible via <https://perma.cc/G4YL-S68T>.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

process of selecting cases, a task Oversight Board staffers jokingly refer to as "fishing mode."

Nonetheless, the Oversight Board's unique strength has always been its ability to generate highly targeted discourses. As it stands, it remains the only Emulated Guardian with the research capabilities, public relations influence, and academic and policy networks necessary to rally global public support for specific policy reforms at Meta. To sustain this role, a dramatic drop in appeal numbers would be problematic. That said, even if the Board received a low six-figure number of cases annually, its members would still have plenty of material to work with.

To me, the key impression is that the Board quickly identified and began establishing the mechanisms that would be necessary to assume a (partially) controlling role over Meta – it refers to external legitimate rules, dialogically engages with the company and drums up and takes on public discourse around specific issues that it then channels into recommendations. Thanks to the Board's legal experts and its overall adjudicative perception, the Board has been able to do what many other adjudicators had done before it: engage in the slow, incremental construction of authority vis-à-vis the power structure it controls. The "strength" of adjudicatory authority escapes simple scales or indicators but must be assessed on a context-sensitive, case-by-case basis. For the Oversight Board, that context could be better – especially regarding Meta's presumably underdeveloped "accountability culture" (see below at Section 5.4).

However, so far, it could also be a whole lot worse. Entirely ceremonial corporate legitimacy stunts usually do not cost several hundred million dollars and hire substantial amounts of serious people, many of whom elite-educated lawyers. Of course, a big check and a few Ivy League lawyers do not make a potent guardian. But roughly 100 highly qualified, motivated people, who operate largely independently in their day-to-day, and who are backed by more than a quarter billion dollars for several years while under intense global scrutiny to hold Meta accountable, indeed sound like the ingredients to potentially shake things up. To compare, the annual budget of the International Court of Justice amounts to roughly 30 million dollars these days, that of the European Court of Justice, which employs more than 2,300 people to almost 500 million euros, and the European Court of Human Rights has roughly 75 million euros to spend annually.[48]

[48] International Court of Justice, 'Report of the International Court of Justice, 1 August 2023- 31 July 2024, A/79/4' para 300 <https://perma.cc/PX8M-5HJK>; European Court of Justice, 'Annual Report 2023' 16 <https://perma.cc/H7R9-ZKUD>; European Court of Human Rights, 'ECHR Budget 2024 (Press Release)' <https://perma.cc/6685-KQNP>.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

Clearly, from Meta's perspective, there is a ceremonial element to the Board. But from the Board itself, whose day-to-day operations seem to function (with some exceptions) largely independent from Meta,[49] neither this research nor other reporting indicated any serious penchant for timidity when it comes to Meta. How this would unfold if Meta were indeed to scale back large parts of its content moderation remains uncertain. Either way, for various reasons outlined in this and the next chapter Emulated Guardians like the Oversight Board must be seen as a first step on a long journey toward comprehensive protection of individual freedoms and the public interest in a world of global governance that seems increasingly dominated by vastly powerful corporations controlled by a handful of men. Whether the Oversight Board and Meta will ultimately be the key players in the broader effort to reclaim public accountability for private, unilateral power structures like digital corporations remains to be seen.

The fact that the Board is – and will remain – bereft of any public and democratic mandate makes it unusual in the otherwise crowded arena of globally active, individual rights-focused adjudicators. International adjudicators in the traditional sense of the word originate from treaties or other expressions of sovereign will.[50] Private international adjudicators can also come into being by the consent of two disputing parties, for the purpose of dealing with contractual disputes – not the administrification of private governance. Other specialized adjudicators like the Court of Arbitration for Sport (CAS) or the dispute settlement system at the International Corporation for Assigned Names and Numbers (ICANN) also are only a semi-good fit as neither of them customarily deal with controlling the rulemaking and rule enforcement of massive, administration-like operations like Meta's private governance.[51]

Whether the Board succeeds in combining public pressure with dialogue-based oversight remains to be seen – and will largely depend on Meta executives' internal approach to the Board. The Board's most realistic path forward may be to take up activists' ideas and demands for reform. In this way, Board members' reasoning, decisions, and recommendations may become a valuable normative point of reference and

[49] See at 4.3.3.

[50] For a theoretical overview highlighting also other sources of normative legitimacy than state consent see Armin von Bogdandy and Ingo Venzke, *In wessen Namen? Internationale Gerichte in Zeiten globalen Regierens* (Suhrkamp 2014).

[51] For an overview see John G White, 'ICANN's Uniform Domain Name Dispute Resolution Policy in Action' (2001) 16 Berkeley Technology Law Journal 229.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

"legitimacy asset" for factions within the company.[52] These factions could use Board recommendations to force through governance changes that they would otherwise not be able to.[53] Referring to Eleanor Westney's study of Japan's emulations in the Meiji era, we recall that a "foreign organizational model may be used to legitimate changes that are desired by the innovators, but which require justification by the appeal to an outside model."[54] Using the Oversight Board and its decisions as an external point of reference may hence be a legitimizing factor for power struggles within the company. Whether this transpires remains to be seen – and will be verifiable only after some time and only with the help of sufficient data and transparency from the company. As yet, we only have anecdotal evidence pointing to this possibility, like the story about a powerful Meta executive (who helped design the Board), who now heads the company's content policy unit, increasingly considering the Board's (potential) view of the company's decisions.[55]

Yet we do not know how representative these indications for Meta's responsiveness to Board recommendations are. Judging from publicly disclosed material and accounts of the company's internal decision-making by Frances Haugen, we might put a (big) question mark on whether all Meta's departments are indeed that welcoming of the Board.[56] These observations are not mutually exclusive since the above-mentioned Meta executive is arguably one of the people at Meta "supporting" the Board; these are perhaps not representative but could, for example, be presented as such to a journalist writing a piece about the Board for a news story.[57] Admittedly, this scenario seems humbler than

[52] See the statement by an interviewed activist at 4.2.3.

[53] The concept of "legitimacy assets" refers to communicative strategies and performative constructions of expertise, tradition, etc. employed by actors in complex, multi-stakeholder governance process to ascertain what Venzke and Mendes call "relative authority." See Ingo Venzke and Joana Mendes, 'The Idea of Relative Authority in European and International Law' (2018) 16 International Journal of Constitutional Law 75.

[54] Eleanor Westney, *Imitation and Innovation The Transfer of Western Organizational Patterns to Meiji Japan* (Harvard University Press 1987) 27.

[55] Steven Levy, 'Inside Meta's Oversight Board: 2 Years of Pushing Limits' (*Wired*, 8 November 2022) <https://perma.cc/SUP5-RQCV>.

[56] In her written testimony to congress, former Meta executive Frances Haugen argued that the company would choose "profit over safety" every single time, cf. 'Statement of Frances Haugen' (n 43).

[57] That executive is Monika Bickert, a Harvard-educated lawyer who served in various leading policy-related positions at Meta and Facebook. Bickert also appears recurrently in Klonick's earlier accounts on the development of content moderation at Meta and the Oversight Board.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

commonly held perceptions of Herculean judicial guardians striving to control power cleaning out the Augean stables of corporate over-reach.[58] But from a sociolegal perspective, the mechanisms that put transnational adjudication to work are often multifaceted, ranging from formal and idealistic aspects to self-interested horse-trading.[59] Hence, such slow, incremental expansion of authority is normal for novel adjudicators.

5.3.2 *Performative Authority*

Authority does not necessarily derive from mandate. It may derive from perception. This leads us to perhaps the ultimate question regarding the Board's authority: does the Board manage to performatively build its authority vis-à-vis the company (and public regulators) through appropriating public law symbolism, its legalese, its references to rights, international law and courts, and overall appearance?

If the Board fails to further build its authority through such discursive, rather than formal legal means it will likely remain largely ceremonial. Skeptics as well as optimists could both point to developments to back up their views.

Crucially, some of Meta's highest echelons (claim to) support the Board. In a company whose controlling shareholder and CEO is a single individual, that person's support is key. However, it is unclear how widespread the support for the Oversight Board is at Meta. Here, we reach a clear methodological limitation of my empirical work, as I did not have the opportunity to investigate empirically these matters within Meta.[60] However, one Oversight Board staffer confided that, in their personal opinion, some people at Meta might "not care for the Board."[61] How representative such skepticism toward the Board was for Meta executives remains unclear. Yet Meta seems to operate in an untransparent way in its relationship with the

[58] For the latter see famously Ronald Dworkin, 'Hard Cases' in *Taking Rights Seriously* (Harvard University Press 1977) 105.

[59] For example, one key factor widely considered contributing to the success of EU law's preliminary reference procedure was lower courts' valorization through direct dealings with the ECJ – bypassing their superior domestic courts, Joseph HH Weiler, 'A Quiet Revolution: The European Court of Justice and Its Interlocutors' (1994) 26 Comparative Political Studies 397, 523.

[60] In 2021, I reached out to several current and former Meta employees who deal or dealt with the Oversight Board but to no avail. In 2022, I decided not to broaden the scope of potential interviewees to Meta as this would have been beyond time and resource capacities. Empirical research into how Meta implements those decisions would be a fruitful avenue for further research.

[61] Interview with Oversight Board Staffer 6 (n 45).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

Board. In one instance, and to the "disappoint[ment]" of the Board, Meta withdrew a request for the Board to draw up a policy advisory statement regarding the company's handling of content in the context of Russia's invasion of Ukraine.[62] Meta executives seem not to grant the Board access to internal data to verify whether it implements the Board's recommendations, as the company claims to do. Further, Meta only grudgingly offers the Board relevant technical assistance to assess the technical context of specific cases.[63] At least in one instance, the company flagrantly lied to the Board – its denial that it had done so was later debunked by a series of leaks to an American newspaper.[64] Therefore, it can be validly claimed that the Board remains largely ceremonial – despite its sophisticated naming and shaming efforts, its well-equipped and capable team, and its crucial turn to human rights law. If the controlled organization simply does not acknowledge a guardian's oversight, the guardian cannot be a meaningful control mechanism. Lacking a rule of law-style enforcement system, then normative output controls the company as much as the company voluntarily subjects to this control. This situation epitomizes what Meyer and Rowan described as a decoupling of ceremonial façades conforming to popularized "myths" from an organization's actual work requirements.[65]

Yet, despite well-warranted caution, not all seems lost. I believe entirely dismissing phenomena like the Oversight Board on the ground that they are not yet as effective as we would like them to be underestimates two aspects, which, in the long run might give reason to hope. On the one hand, one can hardly overestimate the challenge of the task Emulated Guardians face. Their work is not about improving content moderation (this alone would be phenomenally difficult), but the challenge is to fundamentally reform the organizational structure, culture, and, by

[62] Ibid; Interview with Oversight Board Staffer 4 (27 May 2022). See further Oversight Board, 'Protecting Freedom of Expression and Human Rights in Ukraine and Russia (Press Release)' (May 2022) <https://perma.cc/59Z2-WLKY>.

[63] An example for the latter is the company's reluctance to grant the Board access to a tool called "CrowdTangle" which would have allowed the Board to analyze how many people see and interact with Facebook and Instagram posts. Yet, as one interviewee claimed, the Board had to go wrestle the company at painstaking lengths to eventually get access to the tool. Interview with Oversight Board Staffer 6 (n 107); see also Levy (n 55).

[64] This concerned the prevalence of the so-called "Cross Check" system, which allowed the accounts of famous people, such as then US President Donald Trump, to communicate in ways otherwise prohibited. A prominent example was his statement "when the looting starts, the shooting starts," which was widely considered dangerous in a situation of civil unrest and could be read as thinly veiled encouragement to vigilantism, and given the specific context, even racialized lynching. See further Levy (n 55).

[65] See Meyer Rowan (n 3) 356 et seq.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

extension, economic rationale of humongous digital corporations. This is incredibly complex. Unsurprisingly, no single actor could do that. Fundamentally reforming the power dynamics in the digital economy requires a multipronged approach.

Second and this reflects the bit of public law idealism left in me, one should not underestimate the performative power that societal imaginaries still bestow on adjudicatory bodies and public law symbolism. Remember, public law is relevant here not because these companies are "like states." Public law is relevant because it is a well-tested and widely supported set of tools to tie organized power structures to majoritarian will, egalitarian principles, and the public good. Public law, that is constitutional, administrative, and civil rights law, is not about the state per se – it is about controlling organized power. If that organized power manifests beyond the state, we might as well think about how to make the ideas of public law work in these private, profit-oriented contexts.

The more the Board is seen as a credible, serious, and (at least partly) legitimate adjudicator, the more likely it would be to become one eventually. A neo-institutionalist approach would suggest that Meta set up the Board because it wanted to conform to certain societal institutions, namely that of individual rights adjudication.[66] Yet public expectations do not necessarily cease once a new organization has been set up to somehow remedy these rights – institutionalized expectations toward Meta might eventually demand results. If the Board manages to capitalize on these institutionalized images and expectations and become ever more like an adjudicator, effectuating the Board's decisions and recommendations will become a source of legitimation for the company. The Board can, through performance, "de-ceremonize," or, in Evelyn Douek's framing, "de-theater" itself. Eventually, accepting the Board's authority might, eventually, be in the company's own interests and this development would thereby, reflexively, performatively construct and reaffirm the Board's authority. This would indeed be a revolutionary – and perhaps unlikely – development. But who could have foreseen many of the developments we consider natural today?

Similarly, the general perception of the Board appears to be shifting from an early lack of appreciation and toward a grudging respect by human rights organizations, academics, and public bodies like the UN

[66] See in general at 2.4 and 4.2.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

High Commissioner for Human Rights.[67] Evelyn Douek, a vocal critic who (arguably accurately) declared the company's announcement of a "supreme court" an authoritarian legitimacy-grab, later said that "[p]eople thought it would be a total fiasco. But in some real ways, it has brought some accountability to Facebook."[68] The reasons for these minor successes, lie in many of the discourses of our era – namely juridification, judicialization, and the sociopolitical importance of courts, rights, and adjudication, as well as narratives about individual agency. The following section will look at the practical problems arising from some of these aspects, most notably the DSA's novel adjudicators that promise only circumscribed enforcement prospects, as well as EU policymakers' apparently continuing belief that individuals are easy to mobilize for the greater good, even bereft of immediate rewards. The Board's dwindling appeals docket may be a warning sign in this regard. It also remains to be seen how the second presidency of Donald Trump, along with the internal policy shifts at Meta announced in January 2025, will affect the Board. However, these developments should be assessed with nuance, rather than through quick or fatalistic conclusions.

5.3.3 *Are Europeans Altruist Enough to Make Article 21 DSA Work?*

The previous section suggested that the Oversight Board has the potential to gradually develop into a more-than-ceremonial adjudicator. This section now examines whether ODSs also have access to paths that could lead to meaningful accountability. We learned about plans of some people to equally performatively expand the authority of ODSs through qualitative reasoning and close cooperation with legal academia, a strategy also not foreign to the Oversight Board. However, lawyers in general and legal academics in particular perhaps overestimate the real-world efficacy of their profession. As argued above, even if some ODSs become widely

[67] For example, in February 2022, the Oversight Board submitted a report to the OHCHR outlining its practice, Oversight Board, 'The Oversight Board: Operationalizing the UN Guiding Principles on Business and Human Rights: Submission to the Office of the High Commissioner for Human Rights, United Nations on the Practical Application of the UNGPs to the Activities in Technology Companies' (2022) <https://perma.cc/PS65-ENNM>.

For positive, human-rights focused evaluations of the Board in the literature see e.g. Meier and Golia (n 27) 40–41; Helfer and Land (n 27) 51–63; Levy (n 55); Kate Klonick, 'Inside the Making of Facebook's Supreme Court' *New Yorker* (12 February 2021) <https://perma.cc/44Y7-S8VH>; Rishi Gulati, 'Meta's Oversight Board and Transnational Hybrid Adjudication: What Consequences for International Law' KFG Working Paper Series No. 53 March 2022 19 et seq.

[68] Levy (n 55).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

accepted, they remain persuasive rather than binding accountability forums. Their normative force will hinge on the depth of their public support. Therefore, the decisive stakeholders for accountability in this scenario are EU citizens. Only if people across Europe become "active" – perhaps altruistic would be a better word – might the DSA's enforcement mechanisms work.[69] In short, one contextual factor may build ODSs' authority, while two contextual aspects may quickly undermine it.

First to the more optimistic aspect. Compared to Meta's Oversight Board, ODSs have one great authority-conferring and legitimizing advantage. That is that ODSs are the product of a democratic process. Out-of-court dispute settlement bodies are based on an EU regulation and are certified by the relevant authorities at the Member State level.[70] The European Commission arguably cares a great deal about making the DSA a success as its own output legitimacy depends on it. In interviews, representatives of three already certified ODSs – ADROIT, User Rights, and the Appeals Centre Europe – highlighted that they had the impression that the Commission will throw its weight behind them, even though their decisions are formally nonbinding.[71] The fact that ODSs spring from democratic, supranational lawmaking instead of from untransparent management efforts situates ODSs in a different sphere than the Oversight Board.[72] They face fewer legitimacy challenges as they are based on a public mandate as well as being certified by states and closely monitored by regulators. Moreover, EU lawmakers never intended to formally replace traditional judicial redress with ODSs, which are intended to be complementary to judicial redress in the Member States outside the DSA.[73] Ideally, these bodies will take some of the burden off Member State courts and provide a newer and more easily accessible path to justice. Hence, from a norm-oriented public law perspective, ODSs that base their decisions on EU and Member State law and whose decisions are (voluntarily) implemented would come close to the ideal of an Emulated Guardian. They could, for example, seek to apply a legitimated (though only in the EU), external normative framework to digital corporations' content moderation, remedy millions of individual rights violations and push companies toward improving their governance practices. If this

[69] cf. Martin Husovec, 'Will the DSA Work?' (*Verfassungsblog*, 9 November 2022) <https://perma.cc/FV8G-UPHZ>.

[70] See at Chapter 3, regarding certification especially at 3.4.2.

[71] Interview with ADROIT Founder (30 October 2024); Interview with User Rights Co-Founder (25 October 2024); Interview with Appeals Centre Director (1 November 2024).

[72] Regarding the importance of a democratic footing of international adjudicators see Bogdandy and Venzke (n 50) 215 et seq.

[73] See at 3.2 and 3.3.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

were indeed to transpire, it would provide meaningful accountability for digital corporations in Europe. And yet, because of the expected limitations of this model, especially the fact that EU and Member State law will only play a minuscule role when compared to the platforms' own terms of service, this idealized version of out-of-court dispute settlement bodies might not transpire.

However, despite these presumed setbacks, ODSs could offer what, from an individual's perspective, would be a key benefit compared either to the Oversight Board or to traditional courts. Out-of-court dispute settlement bodies will be much more accessible than traditional courts. They will also, unlike the Oversight Board, decide on (almost) every appeal brought to them.[74] Therefore, ODSs might in theory broaden access to justice in online-speech related disputes in the EU.[75] Given the sheer size of content moderation machineries and the structural, rather than individual root causes of recurring fundamental rights violations, the EU's focus on individual justice rather than structural reform seems overall questionable.

The question then is whether platforms will implement (hundreds, thousands, millions of) ODS decisions even though these will not be binding. As explained in Chapter 3, I am deeply skeptical as to whether Article 21 of the DSA will inflict heavy enough costs on platforms to incentivize the latter to make substantive and positive changes to their operations.[76] In addition and given the expected diversity of ODSs across Member States, we should not assume that orchestrated naming and shaming campaigns will exert pressure on platforms to enforce ODS decisions. Whether dismissing individual decisions may generate enough public outcry to nudge social media platforms into obeying ODS decisions is also doubtful. In a similar vein, interviewed staffers from the European Parliament, who were involved in legislating the ODSs into existence, voiced skepticism as to whether these economic incentives would work – especially for very large online platforms.[77] Even the academic who helped the Commission draft its proposal for ODSs has recently called the version that survived the legislative process mere "expert panels."[78]

[74] They can, for example, dismiss cases that were already decided, cf. Article 21(2) subpara 2 DSA.

[75] For an introduction, coming, however, more from a contractual, consumer protection perspective see Richard E Susskind, *Online Courts and the Future of Justice* (Oxford University Press 2019) 95 et seq.

[76] See at 3.7.4.

[77] Interview with European Parliament, Staffer 2 (n 11); European Parliament, Staffer 1, Email to the Author (on file with the author) (17 November 2021).

[78] For the latter see Husovec (n 69).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

Further, ODSs focus on individual justice paired with the absence of authority to bindingly remedy rights violations risks *demotivating* appellants. It is unclear whether there will be a steady stream of people applying for dispute resolution to fuel the DSA's cost-inflicting machinery. As Husovec argued, the DSA depends on "active individuals" as well as societal structures that would help push platforms toward improved governance.[79] Yet such structures "need to be built bottom-up and sometimes locally in each Member State. If their creation fails, the regulatory promises might turn out to be a glorious aspiration."[80] As of now we cannot reasonably assess whether such structures will emerge and whether people will indeed become active in demanding their rights are enforced. What we do know, however, is that given ODSs' current authority, rights enforcement remains a euphemism. The DSA's ODSs cannot *enforce* users' rights in any immediate way. Interviewed Parliament staffers openly questioned whether individuals would be motivated in sufficient numbers – especially once they learn that ODSs cannot guarantee *binding* remedies to successful appellants.[81]

Alternative strategies for how to enable ODSs bodies to help generate naming and shaming campaigns were not discussed while they were being legislated for.[82] After all, until pretty much the end of the legislative procedure ODS decisions should have been binding. Yet, as analyzed in Chapter 3, a concerted effort of some Member States, activist groups, and industry actors advocated gutting the authority of ODSs. This is, by the way, not necessarily a bad thing. Vesting those bodies with the authority to bindingly adjudicate the online free expression of hundreds of millions of Europeans without having devised any procedural law or a mechanism to maintain jurisprudential coherence between those bodies would have likely ended in chaos.

Now, it seems that the DSA may itself risk partially succumbing to ceremony. Simply put, the DSA – or especially the political discourse surrounding it – might have overpromised compared to what it actually delivers. At first sight, some DSA provisions apparently promise to protect fundamental rights online. However, if one takes a closer look,

[79] Ibid.
[80] Ibid.
[81] Interview with European Parliament, Staffer 2 (n 11); European Parliament, Staffer 1 (n 77).
[82] A possible alternative route might be using the Digital Services Coordinators' reports, see in detail 3.7.3. Source of the descriptive statement: Interview with European Parliament, Staffer 2 (n 70).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

they fail to set forth an effective remedy for violations of those rights. Once people begin to understand this, the whole idea of a mass of appeals to ODSs in the hope that they will inflict costs that would motivate platforms to change their rules might burst like a bubble. Here, the continuously declining appeal numbers to the Oversight Board are a warning sign. If even the well-staffed and loaded Oversight Board attracts less than a quarter of the cases it did just two years ago, how are we expecting a diverse and arguably quite chaotic ecosystem of ODSs fare better? Especially so as the practical relevance of these ODSs will arguably hinge on how easy it is to access them. The easiest way to do so will be buttons and automatic redirections on social media websites themselves. How this potentially concentrates markets around several particular ODSs – especially the Appeals Centre – and if this constitutes a problem from an independence perspective remains to be seen.

To be fair, at the time of this writing, the DSA just became applicable. As with the Oversight Board, many things might change once more ODSs become operational and the few that already are have a few years of practice under their belt. Depending on the staff they employ, the publicity they seek, and the quality of the decisions they issue, ODSs may fare better than their current institutional design would suggest. As we saw in the third chapter, the creativity unleashed by the business opportunity to start an ODS and make it profitable by designing a large language model might yield positive results – if not for European internet users at least as general experiments of algorithmic justice systems.[83] We will likely see great variety among ODSs, with some managing to eke out incremental progress toward functional accountability while others remain largely on the ceremonial side.

5.4 Culture

We have seen that both Emulated Guardians show mixed results regarding their normative framework, putting them on the more ceremonial side of the scale, although the Oversight Board, since its turn to human rights, is slightly nearer to the scale's "control" endpoint. As the previous two sections showed, Emulated Guardians' success largely depends on factors beyond their immediate control. Another key factor, and the third

[83] See in detail at 3.8.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

constitutive element of any functional accountability regime, is a *culture* of accountability. As Jack Balkin argued, "[s]ocial media companies have to become key institutions for fostering a healthy public sphere. They can't just serve economic incentives. They have to adopt public-regarding professional norms related to the important public function that they serve in the digital public sphere."[84]

Balkin is right. But infusing such "public" norms into private corporate cultures will prove fiendishly difficult. Establishing an accountability culture would require habitualized bona fide application of the normative framework by the guardian, public demand for said control, and internal adaptation on the part of companies. Given Emulated Guardians' recent emergence, evaluating such an emerging accountability culture necessarily remains cursory. Pointing out, however, that accountability culture is decisive for the success of any Emulated Guardian – any guardian institution frankly – appears mandatory. I conclude this chapter by briefly reflecting on the internal and external elements of a culture of accountability and examining the first signs of accountability that have emerged from the case studies.

5.4.1 *Emerging Accountability Cultures*

The notion of accountability cultures refers to a minimum of public support for, and the institutional pattern of, bona fide practices, expectations, customs, and traditions regarding the accountability of administration-like power.[85] Accountability cultures develop over time,

[84] Jack M Balkin, 'How to Regulate (and Not Regulate) Social Media' (2021) 1 Journal of Free Speech Law 71, 82.

[85] See for example the work of Tarunabh Khaitan who highlights the role of "political cultures" in enabling and restricting fourth branch institutions (which could be understood as closely related to Emulated Guardians as they also shall be independent and, in one way or the other, safeguard the public interest). Khaitan highlights the performative effects of accountability cultures: "Political cultures matter a great deal. Providing institutional guarantees in a relatively stable, peaceful, egalitarian, affluent society with a high degree of political trust and low levels of corruption is likely to be a lot easier than in a post-conflict transitional context marred by poverty, inequality, and mistrust. If anything, effective design matters more rather than less in the latter context. In either case, it is not enough that the initial design of the institution is credible – it must remain credible over time. A guarantor institution can, therefore, become an ordinary regulatory agency by losing credibility, and a formerly ruling-party-controlled regulator could, over time, acquire enough effectiveness and credibility to qualify as a guarantor institution," Tarunabh Khaitan, 'Guarantor Institutions' (2021) 16 Asian Journal of Comparative Law 40, 45, see also at 53.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

closely link to societal institutions, and transcend legal obligations.[86] They require structured and habitualized practices paired with the expectations these practices produce.[87] Culture, in this broader sense, does not refer to traditional "cultural" phenomena like literature or music, but rather to the comprehensive institutionalization and symbolism of a guardian's accountability practice.[88] It includes an organization's practice of rules as much as the relevant public's expectations regarding this practice. In this conception, culture includes *internal* and *external* perspectives on adjudication. Adjudication as an accountability mechanism is therefore both individual – because it deals with individual rights – as well as public-facing in that it lets people assess whether a given power structure is acts accountable.[89]

This internal–external aspect should not, however, make majority opinion an absolute. It just points out the salience of public opinion

[86] In some contexts, the term culture comes with baggage. It can be used exclusionary. As Adorno noted, culture may still convey "brutal claims to societal power." I do not refer to such an exclusionary meaning here but use the term "culture" to highlight a holistic, sociocultural perspective on accountability in practice. For the quote see Theodor W Adorno, *Gesammelte Schriften 9.2 Soziologische Schriften II (Zweite Hälfte)* (Suhrkamp 1975) 276 (translation by the author).

[87] I use the term "(accountability) culture" instead of "practice" as the latter already closely relates to my understanding of rules and authority. To understand a rule, one must look at the practice of it, not only its letter. To build authority, adjudicatory practice is crucial. Peter Cane, on whose three-pronged distinction of control systems I structured this chapter, instead speaks of "practices" and not of culture. Cane argues that said control practices refer to "non-normative patterns of behaviour or, in other words, the way control institutions in fact perform their controlling functions, and the ways they interact with each other and with the . . . agencies they can control." Culture also puts a bigger emphasis on societal expectations and, to speak with Bourdieu, its structure-making, eventually habitualizing effect.

For Cane's account see (n 1) 2 et seq. For a reconstruction of what could be described as transnational law's specific culture see e.g. Martti Koskenniemi, 'Performing Legal Expertise: Reflections on the Construction of Transnational Authority' in Emilia Korkea-aho and Päivi Leino-Sandberg (eds.), *Law, Legal Expertise and EU Policy-Making* (Cambridge University Press 2022). For the importance of 'culture' in guiding actors' behavior see e.g. R Jepperson and J Meyer, 'Reflections on Part II: Institutional Theory' in John W Meyer and Ronald L Jepperson (eds.), *Institutional Theory: The Cultural Construction of Organizations, States, and Identities* (Cambridge University Press 2021) 130 et seq.

[88] Albeit I borrow from Bourdieu, e.g., the concepts of practice, structure, and habitualization (not habitus in a strict sense as I do not study the personal behavior of people), my understanding of "a culture of accountability" is not synonymous with culture as used in most of Bourdieu's work dealing with societal distinction based on "cultural capital." For the latter see e.g. Pierre Bourdieu, *Distinction* (Harvard University Press 1984) 260 et seq.

[89] cf. for actors' embeddedness to individual and collective cultural frames: Jepperson and Meyer (n 87) 129 et seq.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

when it comes to motivating organizations to act in certain ways. Hence, no particular type of culture (like a rule of law-inspired culture of judicial review) can be a one-size-fits-all solution.[90] Varying sets of rules and guardians work differently depending on the context. Ombudspersons, for example, are often found in centralized systems of government as in the UK or Australia and are embedded in cultures of transparency and public accountability common in Scandinavian countries. The establishment of a European Ombudsman for a highly diffused European governance system lacking such an "ombudspersons culture" has hence been seen by some commentators as puzzling.[91] And indeed, as we had to painfully learn from, not all EU institutions indeed "care" all too much about the European Ombudsman. Comparable experiments transplanting the Ombudsperson model into the US administrative state have likewise produced mixed results.

In other words, institutions, rules, and cultures necessarily reflect context. Cultural studies-inspired perspectives on law have argued that looking at the cultural underpinning reproduced by legal institutions helps us gain insights into these institutions.[92] To turn this around, this means that guardians will not be able to function "properly" if they operate in, and with, a culture opposed to their mission. Rules and institutions require bona fide practices that truly aim to control administration-like power.[93]

The effectiveness of an accountability system, correspondingly, hinges not just on formal mechanisms but also on cultural, informal, and societal dynamics. A guardian's cultural alignment with, and influence on, a governance constellation's institutional framework often outweighs doctrinal analyses of formal structures or the predictive allure of principal-agent and rational-actor theories. As Martin Shapiro demonstrated in his comparative study of courts – from US courts to Islamic tribunals to arbitral systems in imperial China – functional guardians adapt reflexively to their cultural environments. Imposing guardians that

[90] See for a reflection on positionality at 1.5.

[91] Peter Cane, 'A Framework for Historical Comparison of Control of National, Supranational and Transnational Public Power' in Susan Rose-Ackerman, Blake Emerson and Peter L Lindseth (eds.), *Comparative Administrative Law* (2nd ed., Edward Elgar Publishing 2017) 611.

[92] Paul W Kahn, *The Cultural Study of Law: Reconstructing Legal Scholarship* (University of Chicago Press 1999).

[93] Similar arguments have been recently made about the robustness of constitutional regimes for preventing democratic backsliding, see Steven Levitsky and Daniel Ziblatt, *How Democracies Die* (Penguin Books 2019). For a historical perspective see Benjamin Carter Hett, *The Death of Democracy: Hitler's Rise to Power and the Downfall of the Weimar Republic* (Windmill Books 2019).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

lack this cultural resonance onto governance systems, however well intentioned, tends to end in failure.[94] In much, the same vein, Karen Alter, Laurence Helfer, and Mikael Madsen explained international adjudicators' sometimes-sluggish path toward authority in terms of struggles to adapt to contexts and cultures. International adjudicators "are for the most part new institutions. As such, their rulings may conflict with, and seek to displace, well-established or assumed interpretations of legal rules or social norms. Displacing entrenched ideas and practices is always difficult."[95] Indeed, despite the internet's breakneck speed, governance reforms often move slowly, grinding through overlapping spheres of authority, endless contestations, and persistent uncertainties.

5.4.2 Globalizing Accountability Culture

The previous section pointed to the relevance of cultural factors for Emulated Guardians to function as meaningful controllers of power structures. Given Emulated Guardians' relatively recent origins, expecting them to have already developed "cultural" aspects in their practices may be premature. But Emulated Guardians' very existence points to an underlying cultural shift in societal expectations that private governance should be held to account. Indeed, the speed of events is striking. In 2018, Kate Klonick criticized platforms' lack of accountability in the Harvard Law Review.[96] That same year, Meta announced its plans for the Oversight Board and, while many countries were entering a second Covid lockdown in late 2020, the European Commission published its proposal for a comprehensive regulation advancing accountability for these platforms. Related regulatory plans have also been presented in the US and the UK.[97] India, the largest democracy on the planet and with a history of ethnically and religiously driven hate speech set up three so-called "Grievance Appellate Committees" to adjudicate on disputes over content

[94] cf. Martin M Shapiro, *Courts: A Comparative and Political Analysis* (University of Chicago Press 1981).

[95] Karen J Alter, Laurence R Helfer and Mikael Rask Madsen, 'How Context Shapes the Authority of International Courts' in Karen J Alter, Laurence R Helfer and Mikael Rask Madsen (eds.), *International Court Authority* (Oxford University Press 2018) 26–27.

[96] Kate Klonick, 'The New Governors: The People, Rules, and Processes Governing Online Speech' (n 15) 1668.

[97] In the US and in the UK, legislative proposals were tabled that partly mirror the DSA (yet none has a direct emulation of Article 21 DSA). See for the US, Lori Trahan, Adam Schiff and Sean Casten, Digital Services Oversight and Safety Bill [introduced], H.R. 6796 – 117th Congress (2021–2022) 2022; UK Online Safety Bill [HL Bill 87-EN].

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

moderation.[98] These government-run committees decide appeals by Indian users against content moderation measures within thirty days.[99] Establishing these appeal bodies was necessary, as the relevant Indian ministry has said: "due to large numbers of grievances being left unaddressed or unsatisfactorily addressed by internet Intermediaries. GAC is expected to create a *culture* of responsiveness amongst all internet Platforms and Intermediaries toward their consumers."[100]

For the DSA, these developments in India might indicate that administrifying platforms and fostering, as the Indian bill put it, a "culture of responsiveness" is a pressing though difficult task.

However, whenever the DSA's attempts to make platforms better administrators lack clear definitions and are at odds with social media platforms' business interests, the platforms' further administrification risks remaining ceremonial rather than substantive – even within the DSA framework. Many of the DSA's procedural impositions are vague, qualified, and adjective-heavy. For example, platforms are expected to "handle complaints submitted through their internal complaint-handling system in a timely, diligent and objective manner," according to Article 15(3) of the DSA. The DSA does not detail precisely what "diligent" and "objective" procedures would look like. This might improve through later practice and litigation. However, although the administrification of platform governance, in part, has already been evidenced in empirical studies, we know relatively little about platforms' actual internal operations, the habitualized practices of their employees, and the norms and values underpinning their choices. Most of what we do know – often from leaks – points to a corporate culture favoring growth, user engagement (the euphemism for the time users spend on a program), and profit, over sound, balanced practices regarding rulemaking and rule enforcement (not to mention firms' datamining activities).[101] So, even

[98] Ministry of Electronics and Information Technology (India), 'Three Grievance Appellate Committees (GACs) Notified on the Recently Amended "IT Rules 2021" (Press Release)' (28 January 2023) <https://perma.cc/98M3-5DQG>.

[99] Ibid.

[100] Ibid (emphasis added).

[101] cf. Jillian York, *Silicon Values: The Future of Free Speech under Surveillance Capitalism* (Verso 2021); Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (PublicAffairs 2019); and especially for Facebook (although similar things presumably apply to all tech giants): Jeremy B Merril and Will Oremus, 'Five Points for Anger, One for a "Like": How Facebook's Formula Fostered Rage and Misinformation' *Washington Post* (26 October 2021) <https://perma.cc/7E55-RPZL>; Sheera Frenkel and Cecilia Kang, *An Ugly Truth: Inside Facebook's Battle for Domination* (Harper 2021).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

though the sheer size and functional logic of platforms partially mirrors (or even supersedes) traditional administrative institutions, we should assume that their normative underpinning and cultural embedding continues to reflect their private and profit-oriented nature.

Implementing individualized, in-depth assessments of individual cases in the context of Articles 14, (moderation), 17 (reason giving), or 20 (internal complaints) of the DSA could give rise to potentially huge costs for platforms. It seems therefore likely that platforms will instead adopt "ceremonial" procedures that invoke the perception of legitimacy without the need for extensive modifications to governance. In turn, in sociological parlance: platforms might design "a formal structure that adheres to the prescriptions of myths in the institutional environment" and thereby demonstrate that they "act ... on collectively valued purposes in a proper and adequate manner."[102] Yet these formal structures often remain ceremonial and decoupled from organizations' actual work activities.[103] Hence, the DSA's goal of holding platforms accountable and protect fundamental rights may advance administrative ceremonies but not necessarily trigger a fundamental shift toward more "responsible and diligent behavior" by platforms, which the EU rightly considers "[e]ssential for a safe, predictable and trusted online environment and for allowing Union citizens and other persons to exercise their fundamental rights."[104] In short, installing a meaningful accountability culture within digital corporations might be indeed the thorniest of all three elements required to establish an accountability regime. The struggle over these cultural and regulatory frameworks has only just begun, with allegiances becoming ever more complex. Key corporations in the tech sector began aligning themselves with the second Trump administration, while the European Union and United States governments finds themselves in a challenging geopolitical position as matters of digital, trade, and security policy appear increasingly intertwined.

5.5 Conclusion

Our evaluation of the DSA's ODSs and the Oversight Board offers an ambiguous picture. The case studies showed that some pioneering thinkers – whether academics, politicians, executives, or public servants – from both

[102] Meyer and Rowan (n 3) 349.
[103] Ibid 357 et seq.
[104] DSA Recital 3.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

sides of the Atlantic had a similar idea at the same time: subjecting private power to individual rights adjudication. Yet, the formal mandates of both Emulated Guardians are limited. The Oversight Board gives nonbinding feedback on structural governance problems within Meta. The EU created a class of actors, the ODSs, which shall remedy millions of individual grievances, however, without being able to issue binding decisions. Simultaneously, neither the Oversight Board nor the European Union seem to dare, yet, to take on the truly crucial issues with content moderation, such as platforms' business model or, at least, the undemocratic nature of platform rules.

Nonetheless, dismissing Emulated Guardians as useless ceremonies would be premature. Especially the first years of the Oversight Board also offer indications for some cautious optimism: with the people who work for the Board, the financial resources they have, and the comprehensive, structure-focused interpretation of their mission, one would be justified in viewing the Oversight Board a worthwhile experiment, shining a spotlight on the important normative and organizational issues of content moderation and the private governance structures of global companies. As the Oversight Board shows, they may performatively play with judicial symbolism and discursively construct rather than legally exercise their authority vis-à-vis the companies they shall control. Public law ideas still seem to carry significant discursive might – even if public law's enforcement through "classic" institutions and mechanisms appears ever more challenging.

Remarkably, the external elements of an accountability culture are emerging. This is a good sign. That platforms constitute problematic power structures today seems widely – even if not unequivocally – accepted.[105] This alone is a significant achievement compared to many governance discourses between the 1980s and the recent "techlash" that has been around since roughly 2016. This shift in overall narrative and political climate may very well be the soil in which Emulated Guardians

[105] The Supreme Court's jurisprudence on this issue – and the broader American political discourse underpinning it – could easily fill an entire comparative study. In Europe, government regulation of content moderation tends to be a rallying cry for the left, while in the United States, it became predominantly a cause championed by the right. Many US conservatives essentially view the entire content moderation apparatus as a liberal enterprise designed to marginalize conservative viewpoints. As a result, although the current majority on the US Supreme Court largely maintains a libertarian stance on the issue, the most pointed critiques of platform power have come from Justices Thomas and Alito, both on the right end of the Court's composition. In contrast, liberal justices like Elena Kagan and Sonia Sotomayor appear untroubled by this particular concentration of corporate influence, cf. Moody v Netchoice 603 US ____ (2024) (United States Supreme Court).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

grow to become "masters of their own destinies" and incrementally build their authority vis-à-vis social media platforms.[106] Further, we are already seeing other jurisdictions debating whether to pass regulations (loosely) modeled on the DSA. India has recently established adjudicatory mechanisms to control content moderation that could perhaps be described as Emulated Guardians (although we do not know whether India purposely emulated the Commission proposal in that regard).

However, the emergence of an accountability culture and successful narratives of "self"-regulation may bolster platforms' all too natural "instinct" to erect ceremonial façades instead of embarking on structural reform. Fostering an institutionalization of meaningful control above and beyond the mere appearance of accountability will be a pressing task for the coming years. Although both Emulated Guardians show promising features, they can only be the first arenas for a new set of actors entering the power struggle between public, value-based, and ideally democratic governance of digital communication and its current domination by private and profit-driven interests.

Because we are just setting out on this somewhat ethereal journey for more accountability in the digital economy (and beyond), the idea behind Emulated Guardian remains malleable. As they rest on social and discursive rather than formal legal authority anyway, they are potentially quick in adapting and improving. We may hence remedy the shortcomings of the two studied specimens in the future.

The following final chapter therefore makes some tentative generalizations as a result of insights from the two case studies and this analysis chapter. The key ideas I will discuss in this last chapter are the following. First, I believe that focusing on individual rights – as much of current political and scholarly debates do – is insufficient. We should focus less on individual rights and more on structural deficits in the governance architectures and infrastructures of the digital economy. Second, we should learn more from administrative law to improve private governance. Third, perhaps the most important intake from administrative law for private governance is that private rulemaking in content moderation should be much more participatory. Fourth, looking back at the whole research and writing process of this book, I argue that we observe the

[106] This, of course, references an optimistic perspective on international courts that remains prominent in much legal and political science scholarship. However, even the initial optimists of "classic" international courts employ many more caveats than in the early 2000s, see Karen J Alter, Laurence R Helfer and Mikael Rask Madsen (eds.), *International Court Authority* (Cambridge University Press 2018) 435 et seq.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005

emergence of a new type of public law-inspired yet privately enforced *accountability law* in globalized, private, and often profit-driven contexts. These *New Accountability Norms* transgress classic public-private distinctions. They respond to public demands for accountability of private, profit-driven actors. I studied those norms here for social media companies. But I would not be surprised to find similar phenomena in other digital or otherwise globalized contexts. The next chapter will sketch out the contours of these New Accountability Norms.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 12:18:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.005