4 The Oversight Board

4.1 Introduction

Announced in 2018 and operational since 2020, the Oversight Board immediately created quite a buzz. Since then, the Board's practices left a lasting impression on debates about how to control social media platforms' immense rulemaking and rule enforcement capacities. Some criticized the Board as a presumptuous Silicon Valley legitimacy-grab.[1] Others highlighted the ambiguous and potentially performative effects of the "Supreme Court" label that Meta had managed to pin on the Board.[2] Others again consider the Board an imperfect but innovative and potentially positive step for more platform accountability.[3] Scholars and practitioners have analogized the Board

[1] See esp. Evelyn Douek, 'Facebook's "Oversight Board:" Move Fast with Stable Infrastructure and Humility' (2019) 21 North Carolina Journal of Law & Technology 1; Julie Cohen, 'From Lex Informatica to the Control Revolution' (2021) 36 Berkeley Technology Law Journal 1017, 1038 et seq; Riku Neuvonen and Esa Sirkkunen, 'Outsourced Justice: The Case of the Facebook Oversight Board' (2022) 12 Journal of Digital Media & Policy; David Wong and Luciano Floridi, 'Meta's Oversight Board: A Review and Critical Assessment' (2023) 33 Minds and Machines 261; Evelyn Douek, 'The Meta Oversight Board and the Empty Promise of Legitimacy' (2024) 37 Harvard Journal of Law & Technology 373.

[2] Josh Cowls, Philipp Darius, Dominiquo Santistevan and Moritz Schramm, 'Constitutional Metaphors: Facebook's "Supreme Court" and the Legitimation of Platform Governance' (2024) 26 New Media & Society 2448. See further from a more normative perspective Thomas Kadri, 'Juridical Discourse for Platforms' (2022) 136 Harvard Law Review (Forum) 163; highlighting the general relevance of performative institutions in internet governance: Matthias C Kettemann, *The Normative Order of the Internet* (Oxford University Press 2020) 242; implicitly also: Rotem Medzini, 'Enhanced Self-Regulation: The Case of Facebook's Content Governance' 24 (2022) New Media & Society 2227.

[3] See especially: Kate Klonick, 'The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression' (2020) 129 Yale Law Journal 2232; Kate Klonick, 'Inside the Making of Facebook's Supreme Court' *New Yorker* (12 February

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

to the US Supreme Court,[4] US courts,[5] international courts,[6] administrative tribunals,[7] human rights tribunals,[8] "courts" in authoritarian regimes designed only to deflect responsibility,[9] human rights grievance mechanisms,[10] customer- service,[11] "fourth-branch" institutions, or transnational hybrid adjudicators.[12] The Board's founders have described it as a "non-judicial grievance mechanism" as envisaged by the terms of the UN's *Guiding Principles on Business and Human Rights*.[13] Meta also compares the Oversight Board to human rights grievance mechanisms, presenting it as "a new mechanism for independent review of certain content decisions."[14] Meta figures have also mentioned the possibility of future emulation of the Board by other actors as the company "hope[s] [the Board] serves as a useful model for other companies and initiatives."[15]

However, this self-description is merely stating the obvious. Clearly, the Board is not a traditional judicial institution but an organizational unit established and financed by a private entity, in this case a social media company. Arguably, it will manage to remedy a few particular grievances.

2021) <https://perma.cc/44Y7-S8VH>; see further: Laurence R Helfer and Molly K Land, 'The Facebook Oversight Board's Human Rights Future' (2023) 44 Cardozo Law Review 2223; Nele Meier and Angelo Jr Golia, 'The Emerging Normative System of Meta's Oversight Board: An Introduction' MPIL Research Paper Series No. 2022–09; Rishi Gulati, 'Meta's Oversight Board and Transnational Hybrid Adjudication: What Consequences for International Law' KFG Working Paper Series No. 53 March 2022; Paul Gowder, *The Networked Leviathan: For Democratic Platforms* (Cambridge University Press 2023) 140 et seq; Cowls, Darius, Santistevan and Schramm (n 2) 17 et seq.

4 Klonick, 'Inside the Making of Facebook's Supreme Court' (n 3).

5 Klonick, 'The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression' (n 3) 2476–2487. But see n 323 at ibid 2476.

6 Marlene Brosch, 'Alles neu macht das . . . Facebook Oversight Board?' [2021] MultiMedia und Recht (MMR) 26, 26.

7 Klonick, 'The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression' (n 3) 2477. Highlighting the Board's function as "ex post review one would expect from a judicial body," Evelyn Douek, 'Content Moderation as Systems Thinking' (2022) 136 Harvard Law Review 526, 568.

8 Laurence R Helfer and Molly K Land, 'Is the Facebook Oversight Board an International Human Rights Tribunal?' (*Lawfare*, 13 May 2021) <https://perma.cc/AZ4W-PSBR>.

9 Evelyn Douek, 'Facebook's "Oversight Board:" Move Fast with Stable Infrastructure and Humility' (n 1) 15 et seq.

10 Interview with Oversight Board Staffer 2 (13 April 2022).

11 Interview with Oversight Board Staffer 3 (6 May 2022).

12 Gulati (n 3).

13 The full quote reads that the UN Guiding Principles on Business and Human Rights "specify that non-judicial grievance mechanisms (such as the Oversight Board) should deliver outcomes that accord with internationally recognized human rights."

14 Meta, Corporate Human Rights Policy at 03 – accessible via: <https://perma.cc/EK3G-FC7E>.

15 Ibid.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

But the need to remedy individual grievances does not fully account for the Board's characteristics, origins, context, and function. Instead, and in contrast to the EU's out-of-court dispute settlement bodies (ODSs) discussed in the previous chapter, the Oversight Board firmly aims to elicit *structural* change at Meta beyond the resolution of (very few) disputes between the company and its users. The disputes and the Board's dispute resolution capacity must be understood mainly as a means to that end. This should come as no surprise, since the Board bends existing categories and draws bits and pieces from various role models that draw from both public- and private-law roots. However, viewing the Board solely from a legal perspective, especially through the lens of human rights, risks neglecting the legitimacy issues pertinent to it from an institutional perspective. In turn, dismissing the Oversight Board as a PR gimmick, a superficial take on content moderation, or even as a solely *private* body with little or no public relevance risks underestimating the possibly novel paradigm the Board (and bodies like it) may represent. In our increasingly globalized and interwoven world, private organizations wield ever more power.[16] This power manifests especially through rulemaking and rule enforcement.[17] Controlling this power requires new mechanisms that warrant at least experimentation beyond traditional, private law-inspired control instruments like liability or competition law. Notwithstanding Meta's self-serving interests, we can approach the Board as a pioneering experiment in bringing into being a new kind of control mechanism – even though, or precisely because, many of its features and underlying ideas replicate

[16] From the vast literature on private domination in the digital constellation see esp. Linnet Taylor, 'Public Actors Without Public Values: Legitimacy, Domination and the Regulation of the Technology Sector' (2021) 34 Philosophy & Technology 897; see in general: Alfred C Aman, 'The Globalizing State: A Future-Oriented Perspective on the Public/Private Distinction, Federalism, and Democracy' (1998) 31 Vanderbilt Journal of Transnational Law 769, 816 et seq; Jody Freeman, 'The Private Role in Public Governance' (2000) 75 New York University Law Review 543; Gunther Teubner, *Verfassungsfragmente: Gesellschaftlicher Konstitutionalismus in der Globalisierung* (Suhrkamp 2012) 18 et seq; illuminating on the reflexive relationship of legal and commercial globalization: Doreen Lustig, *Veiled Power: International Law and the Private Corporation 1886–1981* (Oxford University Press 2020); for free expression in today's mediated society: Jack M Balkin, 'Free Speech Is a Triangle' (2018) 118 Columbia Law Review 2011. See further already Matthias Ruffert, *Die Globalisierung als Herausforderung an das Öffentliche Recht* (Boorberg 2004).

[17] Douek highlighted that content moderation is more than billions of individual decisions based on preset rules. However, thinking of adjudicators' historical role in controlling administrative(-like) structures, rulemaking and rule enforcement are perhaps the practices most receptive to adjudicatory control (in contrast to, e.g., platform design). cf. Douek, 'Content Moderation as Systems Thinking' (n 7) 530 et seq.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

preexisting ones. Therefore, reduced to its essentials, the Board can best be characterized as an emulation (albeit imperfect) of *individual rights adjudication*. The Board adjudicates claims about individuals' rights to curtail Meta's power, particularly regarding rule enforcement and, increasingly, with respect to rulemaking. Board members exercise impartial judgment, and the body presents itself as operationally independent in many (if not all)[18] respects, decides disputes between individuals and a dominant power structure, and bases its decisions on norms, especially those enshrined in international human rights law. And yet it is a private body with a circumscribed adjudicatory authority. In other words, the Oversight Board is also an *Emulated Guardian*.

The key takeaway from this case study is as dialectic as it is cautiously optimistic. On the one hand, the Oversight Board, comparable to the Digital Services Act (DSA)'s ODSs, reflects the idea of using individual rights adjudication to control platforms' private bureaucracies, in this case Meta's content moderation machinery. But although Meta promoted the Oversight Board as "a Supreme Court for Facebook," the corporation never designed it to be a *potent* adjudicator. One could therefore dismiss the Board as a merely ceremonial appropriation of symbolism that is meant to publicly legitimate Meta's scolded content-moderation behemoth. However, and this is the dialectic aspect, as the case study shows, the Board's design, staffing, and the public's perception of it as an adjudicator enabled the Board to incrementally expand its guardianship over Meta. In other words, the Oversight Board is enough of an adjudicator to (somewhat successfully) play with the perception of being an adjudicatory guardian and thereby performatively construct its authority. To be clear, the Board is far from being a fully fledged adjudicative controller of Meta's administration-like rulemaking and rule enforcement. Yet the same mechanisms – juridification, judicialization, and individual rights adjudication – that provided the conceptual soil for the Board's creation in 2018, later enabled it to go beyond its in-built restrictions. The Board is a more "activist" adjudicator than initially thought – with all the positive and negative consequences this entails. These issues have become even more pronounced since Meta CEO Mark Zuckerberg announced in early 2025 that the company would scale back certain policies, including its collaboration with external fact-checkers and its rules on hate speech. How this will affect the Oversight Board remains to be seen, but a deep understanding of the Board's operation up to this point is essential for

[18] See at 4.3.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

assessing future developments in what appears to be a more conservative and libertarian era of online governance emerging in the mid-2020s.

The case study progresses in six steps. It recapitulates the Board's origins, analyzes the details of its institutional organization, the rules on which it bases its decisions, some of its decisions themselves, and whether Meta in fact implements them.

Before we begin, one last word on terminology. In 2021, Facebook, the parent company of Facebook, Instagram and Threads (social networks), WhatsApp (a chat app), and several other companies renamed itself Meta. Many of the Oversight Board's governing documents were crafted before said rebranding and therefore speak of Facebook instead of Meta. To represent existing textual material as directly as possible, the case study, therefore, uses sometimes also the term Facebook (especially, when directly quoting or referencing governing documents with the old terminology).

4.2 Origins

Others have already told the story of the Oversight Board's genesis at length.[19] This section therefore builds on existing accounts and reconstructs three aspects of the Board's origins which in turn highlight three factors driving its inception. Firstly, the Board was based from the start on the idea of *emulating* individual rights adjudication. Irrespective of the company's presumably self-serving (because profit-driven) motives, the organization its executives created utilizes personnel, institutional designs, normative material, and practices known from the traditional context of adjudicative control over (public) power. Second, this section sheds light on the decisive role played by lawyers and legal academics in the Board's origin (and practice). The Board is the offspring of these juridical arguments – that Meta executives decided to establish an adjudicator and not, say, an ombudsperson is no coincidence. In other words, we will examine the performative effects of juridical concepts presented in internal debates about how to make the platform's content moderation more accountable. Third, the section draws attention to excluded

[19] Extensively: Klonick, 'The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression' (n 3). See further: Douek, 'Facebook's "Oversight Board:" Move Fast with Stable Infrastructure and Humility' (n 1); Klonick, 'Inside the Making of Facebook's Supreme Court' (n 3); Kate Klonick, 'The New Governors: The People, Rules, and Processes Governing Online Speech' (2018) 131 Harvard Law Review 1598.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

perspectives and initial reactions from outside actors, some of whom were challenging the presumed 'rationality' of lawyers working for Meta who were defending the company's right to emulate an adjudicatory body. Contrasting these two perspectives highlights one major side effect of the Board's adjudicative design: a usurping of the public legitimacy attached to juridical arguments, personnel, and institutions. Fourth, we take a step back and look at the discursive effects of Meta's use of juridical concepts and constitutional metaphors in general. Discussions surrounding the Board's justification and purpose can be seen as a discursive intervention that shifted debates away from then prevalent discussions about election interference and privacy (remember Cambridge Analytica) and toward content moderation and accountability mechanisms. This external performative effect of the Board yields considerable path dependencies as debates about privacy and content moderation are separated – even though they share the same underlying problem – that is, "business models based on data collection, behavioral advertising, and other aspects of surveillance capitalism."[20]

4.2.1 *Emulation as the Board's Leitmotif*

Since the mid-2010s, social media platforms and especially Facebook, the world's largest social media network, have faced growing criticism. Politicians, journalists, and scholars started questioning platforms' role in a series of (arguably interrelated) political events, like Brexit, and the election of Trump as US President, or, later, the Rohingya's violent persecution in Myanmar.[21] Further, a string of scandals revealed social media companies' troubled relationship with privacy, content moderation, and

[20] Jack M Balkin, 'How to Regulate (and Not Regulate) Social Media' (2021) 1 Journal of Free Speech Law 71, 88.

[21] See esp.: Balkin (n 16); Klonick, 'The New Governors: The People, Rules, and Processes Governing Online Speech' (n 19); Tarleton Gillespie, *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media* (Yale University Press 2019); Julie Cohen, 'Law of the Platform Economy' (2017) 51 University of California Davis Law Review 133; Julie Cohen, *Between Truth and Power: The Legal Constructions of Informational Capitalism* (Oxford University Press 2019); for journalistic takes (in chronological order) see e.g.: Ben Mezrich, *The Accidental Billionaires* (Doubleday 2009); David Kirkpatrick, *The Facebook Effect: The Real Inside Story of Mark Zuckerberg and the World's Fastest Growing Company* (Virgin 2010); Steven Levy, *Facebook: The Inside Story* (Blue Rider 2020); Sheera Frenkel and Cecilia Kang, *An Ugly Truth: Inside Facebook's Battle for Domination* (Harper 2021); for more activist takes see e.g.: Jillian York, *Silicon Values: The Future of Free Speech under Surveillance Capitalism* (Verso 2021); Christopher Wylie, *Mindf*ck: Inside Cambridge Analytica's Plot to Break the World* (Profile Books 2019).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

potential external effects arguably exacerbated by their products.[22] In particular, the way platforms regulated the communication of their users, a process often euphemistically called content moderation, was characterized as unaccountable, arbitrary, and opaque.[23] These political events and platform-related scandals have led to an (ongoing) public relations crisis for Meta and have increased public pressure on regulators to ramp up their regulatory regimes for tackling possible abuses on (and by) platforms, which eventually resulted in the drafting and passing of the DSA (a complex process discussed in Chapter 3). These events also resulted in numerous hearings where company executives were brought before the US Congress and the European Parliament.

The story of the Oversight Board begins a few days before Mark Zuckerberg, Meta's founder and chief executive, was summoned to Capitol Hill in mid-April 2018. Anticipating angry lawmakers questioning the company's fumbled response to allegations of interference by the Russian regime in the 2016 Presidential election, its monopoly-like market power and, importantly, privacy concerns, Zuckerberg outlined his vision of "building a fair system" granting individual users "a right to appeal [Facebook's] decisions on individual posts."[24] Zuckerberg proposed emulating one of the oldest ideas in public law: individual rights adjudication. In other words, Meta was taking it upon itself to put in place *its own* system of adjudicative control over its content moderation bureaucracies. In the phrasing of this book, Meta vowed to establish its own Emulated Guardian. In an interview just a few days before his hearing on Capitol Hill, Zuckerberg said:

> [My goal is to create some] sort of independent appeal process. Right now, if you post something on Facebook and someone reports it and our community operations and review team looks at it and decides that it needs to get taken down, there's not really a way to appeal that. I think in any kind of good-functioning democratic system, there needs to be a way to appeal. And I think we can build that internally as a first step.

[22] See especially Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (PublicAffairs 2019).

[23] See e.g. Klonick, 'The New Governors: The People, Rules, and Processes Governing Online Speech' (n 19) 1630 et seq; Gillespie (n 21) 45 et seq; Douek, 'Content Moderation as Systems Thinking' (n 7) 535–564.

[24] Facebook, Social Media Privacy, And the Use and Abuse of Data – Joint Hearing before the Committee on Commerce, Science, and Transportation in the United States Senate and the Committee on the Judiciary in the United States Senate, Serial No. J-115–40 (n 121) 200 at Question 59(c) by Senator Roy Blunt. See also ibid 176, 179, 181, 186, 306, 320.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

> But over the long term, what I'd really like to get to is an independent appeal. So maybe folks at Facebook make the first decision based on the community standards that are outlined, and then people can get a second opinion. You can imagine some sort of structure, almost like a Supreme Court, that is made up of independent folks who don't work for Facebook, who ultimately make the final judgment call on what should be acceptable speech in a community that reflects the social norms and values of people all around the world.[25]

These statements are striking for several reasons. First, the company's founder is outlining the idea of his own company devising an adjudicatory entity that is somehow to be "almost like a Supreme Court." This "supreme court" metaphor is what the interview is famous for. However, there are at least two other notable remarks. That is, on the one hand, how Zuckerberg portrays *democratic* government as the role model for his company's governance of user communication. He draws inspiration from "any kind of good-functioning democratic system." Further, not only did he make reference to a "democratic system" as an inspiration and announced a "Supreme Court" as a governance solution, but he also mainly talked about introducing "a way to appeal" content moderation decisions – a thinly veiled reference to *administrative* governance.

Such references to legitimizing concepts like "supreme court" or democracy are by no means random. Speaking about his private company in public-law terms, his co-opting of constitutional vocabulary served to throw a veil of publicly legitimizing imagery over the at the time toothless accountability structures in place. Many media outlets picked up on the "supreme court" metaphor and started talking of the Oversight Board as Facebook's "Supreme Court."[26] However, despite nods to the notion of constitutional governance, the idea behind the Oversight Board's genesis and resulting practices indeed emulates policy choices and organizational designs that have historically been diffused as individual rights adjudication.[27] To be clear, the Board's actual powers are far more limited than initially foreseen, and allusions to a Supreme Court-like body were merely meant to be taken as metaphorical, a fact admitted by Zuckerberg himself.[28] However, even though initial plans were scaled down

[25] Ezra Klein, 'Mark Zuckerberg on Facebook's Hardest Year, and What Comes Next' *Vox.com* (2 April 2018) <https://perma.cc/5W3N-Y37S>.

[26] Cowls, Darius, Santistevan and Schramm (n 2) 5 et seq.

[27] For the theory and historical background on individual rights adjudication see at 2.2.

[28] In an interview conducted by a legal academic Meta's CEO said: "This is certainly a big experiment. It's certainly not as broad as everyone would like it to be, upfront, but I think there's a path for getting there," see Klonick, 'Inside the Making of Facebook's Supreme Court' (n 3).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

significantly, the Board is a tentative step to establishing a privately run adjudicatory body whose own goal is to incrementally increase its guardianship over one of the world's largest private governance operations: the way Meta "moderates" content on Facebook, Instagram, and Threads.

This distinction between, on the one hand, the Board's formal mandate and, on the other hand, the Board's practice is important. As we will see, due to performative self-legitimation by savvy lawyers, the Board managed to relatively quickly sewer its own institutional interest – that is, essentially, to do its own thing – from that of the company, whose executives arguably liked the idea of experimenting with the Board without giving it too much power.

The normative and institutional structures required to install such a novel guardian are reminiscent of public-law tenets and, especially, the principles underlying administrative law.[29] To assume this essentially public function, the Oversight Board combines, by way of emulation, public- and private-law elements.

4.2.2 Lawyers, the Supreme Court Label, and Juridification

The second context-related aspect of the Board's inception is the dominance of lawyers, legal scholars, and "constitutional metaphors" (in my coauthors' and my phrasing) or "juridical discourse" (in Thomas Kadri's phrase) in its creation.[30] From the start, the Board was conceived as a juridical, rights-focused, public law-inspired adjudicative body.[31] Even though many ambitious initial plans were scrapped, the underlying rationale seemingly always was to build a body that would be (perceived as) independent, norm-based, and with (some) authority to decide on disputes. In other words, creating the Board was always directed toward *emulating* an adjudicator.[32] Meta is a big and multilayered company. Presumably, different people working for it had differing visions for the Board at various times. What is remarkable though is that internal

[29] cf. Douek, 'Facebook's "Oversight Board:" Move Fast with Stable Infrastructure and Humility' (n 1) 28 et seq.

[30] See already Cowls, Darius, Santistevan and Schramm (n 2); Kadri (n 2).

[31] Interview with Academic, Consultant to Facebook in 2018 (1 November 2021). See further, Klonick, 'The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression' (n 3) 2476; Douek, 'Facebook's "Oversight Board:" Move Fast with Stable Infrastructure and Humility' (n 1) 28.

[32] Douek, 'Facebook's "Oversight Board:" Move Fast with Stable Infrastructure and Humility' (n 1) 28.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

documents show that the *perception* of whatever entity would emerge from this experiment seems to have been a crucial factor for Meta.[33]

4.2.2.1 From Supreme Court to Amorphous Adjudicator

The first to say that Facebook should subject its content moderation to a private *adjudicative* body was Noah Feldman. Feldman is a Harvard-based law professor specializing "in constitutional studies, with particular emphasis on power and ethics, design of innovative governance solutions, law and religion, and the history of legal ideas," to quote from his profile on the university's website.[34] As he later recalled, it all began on "a long bike ride up in the hills behind Palo Alto."[35] After talking to legal scholars at Stanford University as well as people working for Facebook, Feldman had the idea of reproducing institutional designs from liberal democracies for online platforms.[36] He says that

> platforms could benefit by the introduction of a governance mechanism that has traditionally been used, and intensively, in the last 50 or 60 years in liberal democracies, to manage the social conflict around what speech should be allowed and what should not be allowed, namely the constitutional court or the supreme court model. That is, in its essence, a model where there is an independent body that is not directly answerable to the primary, first-order decision-maker that has a set of principles that are clearly articulated on which it relies to make decisions.[37]

In late-January 2018, roughly two months before Zuckerberg announced his "Supreme Court" plan, Feldman sent a memorandum to his old college friend Sheryl Sandberg, Facebook's number two at the time.[38] Sandberg forwarded the memo to Zuckerberg.[39] The two-page document was titled

[33] An internal memo by Noah Feldman to Mark Zuckerberg from March 2018 states that one of two aims behind the whole operation "is to create a durable institution to deliver principled, reasoned decision-making *that would be widely understood as legitimate*," see Noah Feldman, 'Facebook Supreme Court: A Governance Solution (March 2018), published in: Global Feedback and Input on the Facebook Oversight Board for Content Decisions Appendix, Appendix D' 104 (emphasis added).

[34] See Harvard Law School, Faculty Directory, accessible via <https://perma.cc/95A8-BMXH>.

[35] Noah Feldman, at the GovAI Webinar – Noah Feldman on the Design of Facebook's Oversight Board, University of Oxford, Faculty of Philosophy, 16 September 2020, transcript accessible via: <https://perma.cc/NVH7-QH35>.

[36] Ibid.

[37] Feldman (n 35).

[38] For more details see the preeminent text on the Oversight Board's creation, which is Klonick, 'The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression' (n 3) 2448 et seq.; see further Kadri (n 2) 170 et seq.

[39] Klonick, 'The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression' (n 3) 2448 et seq.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

"A Supreme Court for Facebook."[40] In this memo, Feldman framed Facebook's problem with hate speech as a battle between committed advocacy groups who want Facebook to ban certain types of speech and the general, societal interest in free speech.[41] In a binary interpretation, he argues that the "problem is the asymmetry between those who seek censorship and those who favor freedom."[42] Implicitly, Feldman seems to view the company as on the side of societal interests in safeguarding free speech. In Feldman's eyes, Facebook lacked a committed voice and a specific institution to help it defend free speech.[43] Without such an institution, well-organized lobbying groups demanding "censorship" could easily advance their agenda as "nobody is pushing hard on the other side of the door."[44] However, Feldman argues, "there is a solution, one borrowed from the free-speech lessons experienced by governments over the last hundred years."[45] This solution, Feldman says, is for Facebook to adopt free expression as its most fundamental normative basis, as a kind of "constitutional rule," and to establish an adjudicative body to defend it.[46] Courts, he clarifies, unlike majoritarian institutions such as legislators, "can say 'no' . . . [to] demands of advocates seek[ing] to limit speech."[47]

Feldman portrays adjudication as *the* model for Facebook to counter demands for "censorship" and to upgrade how it exercises power over its users.[48] His memo advocated that platforms should set up "their own quasi-legal systems" and that "we need a Supreme Court for Facebook" to mitigate the "dangers social media platforms pose to free speech."[49] In March 2018, Feldman sent the company a second memo, titled "Facebook Supreme Court: A Governance Solution."[50] This outlined how "a Facebook Supreme Court . . . [a]long with lower appeals courts" should adjudicate free speech on

[40] Noah Feldman, 'A Supreme Court for Facebook (January 2018), Published in: Global Feedback and Input on the Facebook Oversight Board for Content Decisions Appendix, Appendix D' (2018) 101 et seq.
[41] Ibid 101.
[42] Ibid.
[43] Ibid 102.
[44] Ibid 101–102.
[45] Ibid 101.
[46] Ibid.
[47] Ibid.
[48] The overly broad use of the word "censorship" should be seen rather critical in this context as it levels the difference between legitimate demands to take down content, e.g., because it is defamatory, and illegitimate censorship of oppositional speech in authoritarian regimes.
[49] Feldman (n 40).
[50] Feldman (n 33).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

Facebook and "create a durable institution to deliver principled, reasoned decision-making that would be widely understood as legitimate."[51] One interviewee, who worked at Facebook at the time, said that the "supreme court framework" was there "because of Noah Feldman and because we lawyers tend to think in adjudicative terms."[52] But apart from his attempt to pass off his metaphorically named entity as something "constitutional" and his argument that individual rights adjudication could be emulated, Feldman's influence seems to have gradually waned once more people got involved. One interviewee, who had an advisory role in the Board's creation, pointed out that various ideas were discussed and that

> [t]here are other models of self-regulation. One could be a human rights tribunal, for example, instead of a supreme court. One might be a self-regulatory body. There are many ways to describe tribunals that work along principles other than those of a supreme court.[53]

Nowadays, many interviewed staffers reject the portrait of the Oversight Board as a "supreme court."[54] Further, internal memos from Facebook from late 2018 and early 2019 indicate a wider search for the ideal type of organization. As one interviewee involved in the Board's creation at Facebook remembers, they looked at "the universe of possible solutions."[55] Correspondingly, an internal memo identified different "families" of control regimes for large, administrative power structures, namely "investigative institutions; supervisory institutions; arbitral adjudication processes, administrative adjudication bodies, national judicial systems, including both European continental-style appellate courts and American appeals courts, and international judicial systems."[56] This same memo listed the benefits and downsides of "judicial and quasi-judicial" control of power.[57] Examples range from the European Court of Justice, to Human Rights Courts, French appeals courts, the US Supreme Court, and, notably, to "the [US] system of administrative appeals in federal executive agencies."[58]

[51] Ibid 104.
[52] Interview with Oversight Board Staffer 4 (27 May 2022).
[53] Interview with Oversight Board Trustee (21 September 2021).
[54] Interview with Oversight Board Staffer 2 (n 10); Interview with Oversight Board Staffer 1 (25 March 2022); Interview with Oversight Board Staffer 4 (n 52); Interview with Oversight Board Staffer 5 (2 June 2022).
[55] Interview with Academic, Consultant to Facebook in 2018 (n 31).
[56] Facebook, 'Global Feedback and Input on the Facebook Oversight Board for Content Decisions Appendix' 138 <https://perma.cc/SP5M-Q7G7>.
[57] Ibid 140.
[58] Ibid 157. Later, the memo presumes that such judges in US administrative appeals are "unlikely to be independent" (ibid). Yet, generally dismissing the independence of

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

This reference to "administrative appeals" in US administrative law is worth commenting on as it plays into the point about the increasing "administrification" of platform governance in Chapter 2.[59] To recapitulate, Meta's rulemaking and rule enforcement, just like administrative law, concerns a powerful institution (Meta), a group of individuals (users), and rights that are traditionally understood as coming under the purview of public law (freedom of speech). Further, the main function of all normative material governing Meta's governance of communication is supposed to ultimately delineate how Meta exercises its power vis-à-vis a growing and anonymous (indeed, anonymized) group of users. This picture structurally resembles the function of administrative law.[60] Controlling such a complex "system of administration" (following Douek's description) would seem to require a panoply of guardians.[61] As outlined in Chapter 2, various guardians of administrative or administration-like power exist, ranging from traditional, generalist courts to other adjudicatory institutions to various notionally independent operators like ombudspersons or auditors.

4.2.2.2 Juridification and the "Appropriate" Policy Choice

However, since Meta tasked mainly lawyers with building the edifice that would eventually become the Oversight Board, the "appropriate" policy choice to emulate, in the legal team's eyes, was *individual rights adjudication*. As one interviewed legal scholar, who advised Meta on creating the Board, put it:

> Personnel matters in these kinds of things too. When you have a bunch of lawyers ... work[ing] on this thing, inevitably, what's going to come out is going to be something familiar to lawyers, and just the way in the world that lawyers work.[62]

This is a telling explanation. Because lawyers dominated the Board's creation process (and, so far, its later practice), the project to design

adjudicatory bodies attached to US agencies could be considered simplistic. Many authors consider judges beyond Article III courts in the US in general sufficiently independent to fairly adjudicate in their jurisdiction. cf. James E Pfander, 'Article I Tribunals, Article III Courts, and the Judicial Power of the United States' (2004) 118 Harvard Law Review 643.

[59] See also Moritz Schramm, 'Platform Administrative Law: A Research Agenda' (SSRN, 24 July 2024) <https://perma.cc/JH2A-8ZW2>.

[60] cf. at 2.3.

[61] Douek, 'Content Moderation as Systems Thinking' (n 7).

[62] Interview with Academic, Consultant to Facebook in 2018 (n 31).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

a guardian to oversee Meta's massive power quickly became funneled into emulating bits and pieces from different models of adjudicatory oversight over public power. This is why the Oversight Board, despite its generic and inoffensive-sounding name, is more than a mere panel of experts, an ombudsperson, or some other type of "guardian." From the start, its creators were perhaps subconsciously united by their respective preconceptions of what adjudicative control over power might look like. In turn, the structure that eventually emerged reflects the Board's limitations as a private law-inspired entity.[63]

From the perspective of organizational sociology, one cannot overstate the preeminence of this mélange of lawyer-dominated staff, uncertainty, and the woolly aim to "create an independent appeals mechanism." Uncertainty is the seed of emulation.[64] As DiMaggio and Powell said, "[w]hen goals are ambiguous, or when the environment creates symbolic uncertainty, organizations may model themselves on other organizations."[65] Emulating preexisting institutions, like individual rights adjudication, to increase an organization's public legitimacy is a common driver of new formal structures. As explained in Chapter 2, emulating institutions and formal structures is relatively straightforward and an "easy sell" to the public. In contrast, it is much harder to implement more deeply rooted changes in policy, strategy, or practice. These are often harder to comprehend for laypeople and thus cannot be legitimized in people's perceptions in the same way as the idea of emulating existing institutions by putting in place formal structures.[66] As Evelyn Douek noted, the "fundamental purpose [of the Oversight Board] is to bring greater legitimacy to Facebook's content moderation ecosystem."[67] According to Douek and many other critics, the Board is only a means to an end for Meta to increase its own public legitimacy in societal discourse.[68] Thomas Kadri points out that platforms "exploit" juridical arguments and structures like the Oversight Board to buttress their public legitimacy.[69] Similarly, DiMaggio and Powell argued already in the 1980s that private, for-profit enterprises "tend to model themselves after

[63] See in detail at 4.3.
[64] Speaking of "imitation" and "mimetic" acts Paul J DiMaggio and Walter W Powell, 'The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields' (1983) 48 American Sociological Review 147, 151.
[65] Ibid.
[66] Ibid 52.
[67] Douek, 'Facebook's "Oversight Board:" Move Fast with Stable Infrastructure and Humility' (n 1) 46.
[68] Ibid.
[69] Kadri (n 2) 167.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

similar organizations in their field that they perceive to be more legitimate or successful."[70] As discussed in Chapter 2, individual rights adjudication is one of the most legitimate and successful institutions controlling wide-reaching rulemaking and rule enforcement enterprises.[71] Hence, corroborating the theorization about emulation outlined in Chapter 2, Meta's "Governance and Strategic Initiatives" team followed a common pattern in modeling the company's Board after existing institutions. This team was the main driver behind the Board's development.[72] It was run by Brent Harris, a lawyer who joined Facebook that year.[73] Apart from Harris, it consisted of ten people.[74] Further, "dozens" of "cross-functional" people would contribute to the governance team's work.[75] The number of people involved in creating the Board, either as employees or outside consultants, was described by interviewees as in the low double-digits.[76] All of the lawyers involved with Facebook were highly educated and most had received law degrees in the US.[77] One interviewee, who was an in-house adviser to the company, recollected the internal debates at Facebook as follows:

> I think that understandings of the way that courts relate to the people over whom they exercise authority definitely were in the room. Whether they were in the room because anybody wanted to consciously think about them, or whether they were in the room just because of the personnel and communicative context of the thing is not as clear.[78]

How far this emulation went – and whether it can reproduce at least some of the success of adjudicatory control over power in the context of national states or nation-spanning conglomerations like the EU – follows later in this chapter.

[70] DiMaggio and Powell (n 64) 152. Their argument, however, did not directly concern the emulation of state structures in private corporations but rather isomorphism among private actors. The here described trend, therefore, seems to be an evolutionary step beyond those observed in the past by DiMaggio, Powell, and many other sociologists.
[71] See esp. at 2.2.
[72] Klonick, 'The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression' (n 3) 2452–2453.
[73] Ibid 2452.
[74] See Klonick, based on an email by Caroline Glanville, Policy Communications (Governance Board), Facebook, to Klonick on (May 25, 2020), see in Klonick's text at note 154 – ibid 2453.
[75] Ibid.
[76] Interview with Academic, Consultant to Facebook in 2018 (n 31).
[77] Ibid.
[78] Ibid.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

4.2.3 *Excluded Perspectives*

Building an edifice with the help of (mostly) US-educated lawyers will necessarily exclude a range of options. To include more diverse perspectives but also to buy time, in 2019 Meta organized a series of workshops focusing on ideas for the Board's design. This global consultation included a series of consultations with stakeholders such as nongovernmental organizations (NGOs) and civil-society advocacy groups.[79] The company also held several "workshops and roundtables" in, to name just a few places, Singapore, Berlin, Oklahoma City, New York, Paris, Nairobi, Mexico City, and Taipei.[80] Presumably these roundtables were the first occasions when the ideas advanced by Meta's lawyers (presumably with Zuckerberg's sanction) were presented to activists, stakeholders, and practitioners from outside the company.[81] The first roundtable, held in Singapore, featured a moot court-like debate about specific cases.[82] Participants were split into groups and received a hypothetical case upon which they had to decide on the basis of predefined norms.[83] As one interviewed activist, who participated in the Berlin workshop in the summer of 2019, recalls:

> [T]hey gave us test case studies ... These were blunt cases, and the people invited to this thing were, again, Westerners. There was one done in Latin America and there was maybe one in Africa, but I know people who were at both of them, and they said that there were still a lot of white people at both. ... I would say it was more like a college debate team than a court. It felt nothing like a court to me. It was me sitting around a table with a bunch of other overprivileged white people, discussing a case where actual harm hadn't been done. ... I don't think it set them up well for understanding the cases that would come to them or what they would mean for people on the ground. I think it already came out, so to be clear, they focused solely on European, American right-wing propaganda kind of cases.[84]

Although it is impossible to identify who participated in these workshops, we can reasonably assume that Facebook wanted participants to discuss

[79] Facebook (n 56) 11 et seq.
[80] Ibid 39 at endnotes 40, 41.
[81] Apparently, Zuckerberg remained deeply involved in the whole creation process, e.g. editing various materials. See Klonick, 'Inside the Making of Facebook's Supreme Court' (n 3).
[82] Klonick, 'The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression' (n 3) 2451–2457.
[83] Ibid 2451–2452.
[84] Interview with Activist (1 December 2021).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

how they would like an adjudicative institution to look like. In contrast, *if* an adjudicative institution was even the right idea in the first place was not under discussion.

Further, the critical remarks by the participant indicate that Meta's governance team's preconceptions diverged from the needs of affected users. Especially in non-"Western" parts of the world, the same interviewee recalls that

> [T]here's just a lot of eye-rolling at the Oversight Board, to be honest . . . I think that there's a lot of "What's the purpose of putting all of this money behind the Oversight Board when we've been saying these things all along?" which is a fair critique. . . . I think the Oversight Board's most useful function . . . is, in fact, that it formalizes and makes explicit a lot of those recommendations, synthesizes them, really. To be honest, from my colleagues in the Middle East I haven't heard that many positive things about the Oversight Board. It's not so much negative as just "What's the point?"[85]

What the activist is alluding to here is important. The Oversight Board was constructed by those with a specific mindset governed by ideas about individual rights, adjudicative control over power, and shot through with the largely metaphorical language of "supreme" courts and "constitutions." As we will see, the Board somewhat falls short of these expectations. Most importantly, its authority is severely restricted, and it decides too few cases to offer any meaningful judicial protection to the users of its services. This reasonably brings us to question the institution's overall validity. As the activist put it: "What's the point?"

4.2.4 Foreclosed Debates

The dominance of lawyers in the discourse, as well as the predominance of constitutional metaphors and juridical framings, construed the Board as an adjudicative solution to Meta's unaccountable methods of content moderation, which was itself portrayed as an issue of how to police billions of free-speech acts. In other words, the question the Oversight Board was designed to handle was that of free speech.[86]

[85] Ibid.

[86] See e.g. Klonick, 'The New Governors: The People, Rules, and Processes Governing Online Speech' (n 19) 1667–1678. For a critique of Facebook's "exploit[ation]" of juridical discourse see Kadri (n 2) 167 et seq. For the relevance of constitutional metaphors in debates on platform governance see already Cowls, Darius, Santistevan and Schramm (n 2). For a critique of the rights-focused "standard picture" of content moderation see Douek, 'Content Moderation as Systems Thinking' (n 7).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

However, when we jump back to the beginning of the Oversight Board's story in early 2018, we see an unusual discursive shift. Back then, content moderation was arguably not Meta's biggest, let alone its only, problem. Instead, Meta was drawing strong criticism from scholars, journalists, and politicians for its handling of personal data and its interface design.[87] To begin with, Meta was using algorithms to "personalize" user experiences, leading to the potential creation of "echo chambers" and an entrenching of attitudes and opinions.[88] Then, Meta's flawed privacy governance schemes were allowing outside parties to access and retrieve huge amounts of personal data. The most infamous of these data breaches was orchestrated by a company called Cambridge Analytica.[89] Cambridge Analytica managed to gain access to the personal data of eighty-seven million US citizens and later allegedly used at least some of that data to tamper with the 2016 Presidential election.[90]

Zuckerberg announced the setting up of the Board a few days before his hearing on Capitol Hill, and around the time those data breaches became public, in April 2018. Although the hearing was also about content moderation, fundamental rights, and accountability, its main concerns were election interference and privacy in the context

[87] Insightful: Asunción Esteve, 'The Business of Personal Data: Google, Facebook, and Privacy Issues in the EU and the USA' (2017) 7 International Data Privacy Law 36; see further Zuboff (n 22) defining Google's "discovery of behavioral surplus" 74–87, for Facebook e.g. at 483–485, 497 et seq.

[88] Seth Flaxman, Sharad Goel and Justin M Rao, 'Filter Bubbles, Echo Chambers, and Online News Consumption' (2016) 80 Public Opinion Quarterly 298; Rune Karlsen and others, 'Echo Chamber and Trench Warfare Dynamics in Online Debates' (2017) 32 European Journal of Communication 257; regarding Facebook's "like button" Zuboff (n 22) 455–461; in general regarding the psychological underpinning of designing online interfaces Chris Nodder, *Evil by Design: Interaction Design to Lead Us Into Temptation* (Wiley 2013). See further Niklas Eder, 'Beyond Automation: Machine Learning-Based Systems and Human Behavior in the Personalization Economy' (2021) 25 Stanford Technology Law Review 1, 19 et seq.

[89] The scandal broke after a data analyst at Cambridge Analytica leaked several documents to the Guardian. The whistleblower later published a memoir describing how Cambridge Analytica, apparently with the help of far-right political strategist Steve Bannon and Russian state affiliates retrieved and utilized data about millions of American voters, see Wylie (n 21) 95–155.

[90] Issie Lapowsky, 'Facebook Exposed 87 Million Users to Cambridge Analytica' (*Wired*, 4 April 2018) <https://perma.cc/D3UV-KH3T>; Remarkably, social media users seem to meet privacy and data breaches with stoic neglect: 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.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

of the Cambridge Analytica scandal.[91] The Congress hearing took place under the title "Facebook, Social Media Privacy, and the Use and Abuse of Data."[92] Senators pressed Zuckerberg mostly on his company's "disturbing" privacy-governance systems that had ended up leaving the door open for Cambridge Analytica to steal personal data.[93]

As Ari Ezra Waldman noted, there are "numerous discourses at play in informational capitalism," some focusing on innovation and possible technological solutions to real-world problems, others highlighting social networks' effects on free speech.[94] Therefore, one can view Meta's project of the Oversight Board – and its metaphorical narration in constitutional-law terminology – as a performative shift in the discourse surrounding the wider role of social media companies in society.[95] Discussing how the Oversight Board might bring accountability to Meta's moderation of speech shifted debates away from Meta's still flawed handling of personal data and the morality of its business model. As revelations by whistleblower Frances Haugen in late 2021 indicated, Meta did not substantively improve its products, its handling of private data, and its overall business model.[96] Instead, the intensifying debates about content moderation approach platforms through the prism of individual rights adjudication but are increasingly detached from discussions regarding social media platforms' business models, ways of handling data, and privacy.[97] Debates

[91] Chairman Chuck Grassley, Senator from Iowa, stated that "the issues [the hearing] will consider range from data privacy and security to consumer protection and the Federal Trade Commission enforcement," 'Facebook, Social Media Privacy, And the Use and Abuse of Data – Joint Hearing before the Committee on Commerce, Science, and Transportation in the United States Senate and the Committee on the Judiciary in the United States Senate, Serial No. J-115-40' (n 24) 2.

[92] 'Facebook, Social Media Privacy, And the Use and Abuse of Data – Joint Hearing before the Committee on Commerce, Science, and Transportation in the United States Senate and the Committee on the Judiciary in the United States Senate, Serial No. J-115-40' (n 24).

[93] Statement of Senator John Thune, US Senator South Dakota, ibid 2. See further the statements by Senators Dianne Feinstein, Bill Nelson, Orin Hatch at ibid 3–8, 21–22, as well as Mark Zuckerberg's statement at 8–13.

[94] *Ari Ezra Waldman, Industry Unbound: The Inside Story of Privacy, Data, and Corporate Power* (Cambridge University Press 2021) 46.

[95] See Cowls, Darius, Santistevan, and Schramm (n 2); criticizing that platforms swayed discourse on content moderation toward "juridical discourses": Kadri (n 2); regarding companies' performative impact on our conception of privacy see Waldman (n 94) 224 et seq.

[96] 'Statement of Frances Haugen' 2–3 <https://perma.cc/3QN6-JMY4>.

[97] Remarkably, Evelyn Douek's comprehensive critique of narrow understandings of content moderation largely skips the aspect of privacy, which Douek considers to be a "related" context, cf. Douek, 'Content Moderation as Systems Thinking' (n 7) 530–533, 571.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

surrounding intrusive regulatory intervention when it comes to platforms' business models – or even arguments for breaking up social media companies altogether – are still rife today, although the academic and political winds seem to have shifted toward more reform-oriented approaches.[98] In other words, platforms' wide-ranging regulatory powers seem to be so entrenched that minimizing them, let alone eliminating them, appears increasingly impracticable. This tacit discursive premise in turn paved the way for various reformist ideas to come to the fore, the Oversight Board being just one of the most prominent.

To conclude, we shall recapitulate the two – slightly stylized of course – centripetal forces shaping the Board's creation and its later practice. These are, on the one hand, Meta's striving for public legitimacy and, on the other hand, as a consequence thereof, the nascent adjudicators' unique position and ability to incrementally build on the public's perceptions of what adjudicative authority entails. In other words, the fact that Meta decided to create the Board as an adjudicator and not as, say, an ombudsperson, potentially enables the Board to performatively play with the perception of being an adjudicator and thereby provide some accountability for Meta. Even though the next sections show the serious limitations built into the Board's institutional framework, the story of the Board's genesis and practice are thus evidence of the shortcomings, and dangers, but also of the potentially paradigmatic potential of Emulated Guardians.

[98] For example, the US Federal Trade Commission's (FTC) sued Meta for its acquisition of Instagram and WhatsApp in 2020. If successful, the FTC would require Meta to break up parts of its business. See Federal Trade Commission, 'FTC Sues Facebook for Illegal Monopolization (Press Release)' (9 December 2020) <https://perma.cc/N6RX-AX8Q>. Lina Khan, the FTC's chairwoman under the Biden administration, previously wrote a widely influential article in the Yale Law Review arguing for ramping up antitrust toward the platform economy. However, preparations for the case against Meta because of its acquisition of Instagram and WhatsApp began already under the first Trump administration. For the scholarly groundwork of Ms. Khan see Lina M Khan, 'Amazon's Antitrust Paradox' (2017) 126 The Yale Law Journal 710. Khan's title references (and problematizes) the foundational text of the dominant "consumer welfare" paradigm in US antitrust: Robert H Bork, *The Antitrust Paradox* (Basic Books 1978).

See further already the influential Stigler Report, which demanded, among other things, more "aggressive antitrust": Stigler Center for the Study of the Economy and the State (University of Chicago), 'Stigler Committee on Digital Platforms Final Report' (2019) 7–9, 16–17, 92–121.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

4.3 Institution

This section investigates the Board's institutional characteristics. To emulate individual rights adjudication, Meta needed to construct the Oversight Board as a private adjudicator that would adjudicate the public rights of users or, in other words, provide accountability for Meta's private governance by using public-law mechanisms. Now we look at the Board's organizational structure and personnel, as well as its jurisdiction, before reflecting on its impartiality and independence.

4.3.1 Organization and Personnel

Owing to its entirely private nature, the Board's institutional structure differs significantly from many other adjudicative bodies controlling traditional administrative power.

The Oversight Board does not exist as a uniform entity.[99] Instead, it consists of four actors working together, governed by a set of founding documents. These are the "Oversight Board Trust," an irrevocable Delaware noncharitable trust with several trustees ("the trust").[100] The trust funds the Board's operations and makes management decisions concerning its operation. The trust created a limited liability corporation (LLC) under Delaware law that rents office space and hires staffers.[101] Further, the trust filed the individual contracts of the Board members who bear ultimate responsibility for the Board's adjudication. At the time of writing, the Board had between seventy and ninety staffers, who are employed by the LLC and who run the Board's operations.

4.3.1.1 The Founding Documents

Four documents govern the Oversight Board. These are the *charter*, a vague and lofty text setting out the Board's purpose and function; the *trust agreement*, a detailed contract establishing a Delaware noncharitable trust; the *LLC agreement*, a contract establishing a limited liability

[99] Even formally uniform adjudicatory entities like "the" Bundesverfassungsgericht collapse into various complementary sub-units if one looks closely enough. See Matthias Jestaedt, 'Autorität und Zitat – Anmerkungen zur Zitierpraxis des Bundesverfassungsgerichts' in Mathias Mayer and Joachim Jacob (eds.), *Im Namen des Anderen – Die Ethik des Zitierens* (Wilhelm Fink Verlag 2010) 147–148.

[100] Trust Agreement Section 1. For the provisions concerning trustees see Section 6.

[101] Trust Agreement Section 2.1.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

corporation; and a set of detailed *bylaws*.[102] The charter, bylaws, and trust agreement were drawn up exclusively by Meta's lawyers. Meta then drafted the bylaws; the Board approved them and may amend parts of them subject to approval by Meta executives.[103] Further, several later crafted documents emerged as important throughout the case study and will be discussed in greater detail below. These are the code of conduct for Board members and staff as well as the Board's self-given *Rulebook for Case Review and Policy Guidance* of November 2020.[104]

In organizational terms, the trust agreement in particular has been described as perhaps the most consequential because the trust agreement is a formal contract, whereas the charter was just "declared" by the company.[105] The trust agreement, LLC agreement, and bylaws are much more detailed – and restrictive – than the charter. Yet, the exact relationship and order of precedence between those documents is complex. Especially in terms of organizational or operational hierarchy, it is not intuitively clear how some of the documents relate to one another. For example, the bylaws restrict the Board's jurisdiction over subject matter to removal of particular types of content, whereas the charter uses general language, allowing "people who disagree with the outcome of Facebook's decisions ... a request for review."[106] Intuitively, the use of terms like 'charter' and 'bylaws' would imply that rules established by the charter rank more highly than those set out in the bylaws. Similarly, the bylaws describe the charter as "the Board's primary governing document" whereas the bylaws only "detail the operational procedures."[107] Further, the bylaws explicate that "where there is a potential conflict between

[102] For a coeval overview, albeit omitting commentary on the crucial trust agreement, see the series of blogposts on different versions of the charter, Facebook's 'values,' and the Board's bylaws: Evelyn Douek, 'How Much Power Did Facebook Give Its Oversight Board?' (*Lawfare*, 25 September 2019) <https://perma.cc/54D7-X7YW>; Evelyn Douek, 'Why Facebook's "Values" Update Matters' (*Lawfare*, 16 September 2019) <https://perma.cc/527F-7GBZ>; Evelyn Douek, 'Facebook's Oversight Board Bylaws: For Once, Moving Slowly' (*Lawfare*, 28 January 2020) <https://perma.cc/YM8P-T5S2>; ibid; Evelyn Douek, 'What Kind of Oversight Board Have You Given Us?' (*The University of Chicago Law Review Online*, 11 May 2020) <https://perma.cc/6VPV-EZLR>.

[103] Article 5 Section 1 Bylaws.

[104] As the rulebook's introduction states: "The procedures outlined in the Rulebook were established by the Oversight Board Co-Chairs and Members," 'Rulebook for Case Review and Policy Guidance' of November 2020, which can be downloaded here: <https://perma.cc/4BV3-ZA9C>; The Code of Conduct is an appendix to the bylaws and can be downloaded on the Oversight Board website, accessible via <https://perma.cc/CPL4-VPRZ>.

[105] Interview with Oversight Board Trustee (n 53).

[106] Charter Article 2(1).

[107] Bylaws Introduction (p 5).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

interpretations of the charter and bylaws, the charter shall prevail."[108] The trust agreement reaffirms that the charter sets forth the "framework for creating the Oversight Board" and that only the charter "specifies the Oversight Board's authority and jurisdiction."[109] But at the same time, the charter states, in Article 6(2), that "the Board's operational procedures will be outlined in its bylaws. The charter and the bylaws will act as companion documents." Whether the charter's vague allusion to the bylaws as a 'companion document' indicates normative equality between charter and bylaws, primacy of the bylaws as *lex specialis* and/or *lex posterior*, or primacy of the charter as *lex superior*, remain an entirely unsettled matter.

The best way to understand the relationship between the founding documents is to keep in mind the documents' different drafting histories.[110] The charter was pinned down first and reflects the initial, high-flying goals from the early "Supreme Court" times.[111] It speaks of "people" instead of users, delineates the Board's "authority to review" and establishes that the "Board's resolution of each case will be binding."[112] The trust document and the LLC agreement, in turn, were crafted as waterproof legal documents by a probably larger und more specialized group of lawyers scrambling to find a private-law form operationalizing the lofty public announcements of Zuckerberg and the terms of the charter.

4.3.1.2 The Trust

The trust was set up in October 2019.[113] According to Section 2.1 of the trust agreement, the "purpose of the Trust . . . is to facilitate the creation, funding, management, and oversight structure that will permit and protect the operation of an Oversight Board." To achieve this, "the

[108] Bylaws Introduction (p 5) (at the very end).
[109] Trust Agreement Section 2.1.
[110] See also Evelyn Douek, 'Facebook's "Draft Charter" for Content Moderation: Vague, But Promising' (*Lawfare*, 31 January 2019) <https://perma.cc/52RW-7E7A>; Douek, 'Facebook's Oversight Board Bylaws: For Once, Moving Slowly' (n 102).
[111] See below at 4.3.2.
[112] cf. Charter Introduction, Article 2, Article 4. Apparently, said change was initiated by Meta CEO Mark Zuckerberg himself or his personal team. In an interview with Kate Klonick, who was embedded as an observer at Facebook, the head of the company's governance team said that "Feedback is coming from very high up in the company." Further, the CEO edited the charter to make it "more approachable." See Klonick, 'Inside the Making of Facebook's Supreme Court' (n 3).
[113] cf. Trust Agreement Section 1.1.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

Trustees," of which there are six, "shall form and fund a limited liability company ('the LLC') of which the Trust will be the sole member through its Trustees," according to Section 2.1 paragraph 2 of the trust Agreement. The purpose of the LLC is to "establish, administer, and attend to the ongoing operation of the group of individuals who make up the Board Members," the paragraph continues. In 2019, Meta announced a contribution of 130 million dollars to the trust.[114] In 2022, Meta allocated another 150 million dollars to the trust, bringing the total amount to 280 million dollars.[115]

4.3.1.3 The Board Members

The most public and perhaps the most important component of the Board are its members. Members are individual contractees of the trust, meaning they are not employed by the LLC, to maintain their independence. Their compensation, described as "six-figure salaries," is not disclosed and might differ from member to member as "each Board member contract shall be based upon provisions suggested by ... [the trustees] and the Oversight Board."[116] Members do not work full-time for the Oversight Board but are contractually expected to work fifteen hours a week for it.[117] However, several interviewees indicated that members often spend considerably more time preparing and deliberating.[118] Members need not be physically present at the Board's offices around the world. Many of their deliberations happen online, spanning several time zones and continents. The members bear ultimate responsibility for the Board's decisions, recommendations, and policy advice.[119] In November 2024, the Board had twenty-one members, though it can have up to forty.[120] According to the charter, Board members are appointed for a fixed term of three years, renewable twice.[121]

[114] Oversight Board, 'Securing Ongoing Funding for the Oversight Board (Press Release)' (22 June 2022) <https://perma.cc/Y859-KUEF>.
[115] Ibid.
[116] For the quote on the salaries see Klonick, 'Inside the Making of Facebook's Supreme Court' (n 3) 10. For the latter quote see LLC Agreement Section 5.3(a)(i).
[117] Ibid.
[118] Interview with Oversight Board Staffer 1 (n 54); Interview with Oversight Board Staffer 5 (n 54).
[119] For the details of the different types of decisions see below at 4.5.
[120] Charter Article 1 Section 1.
[121] Charter Article 1 Section 3.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

Powers

The members are vested with several "collective powers," according to Article 1(4) of the charter. They can request Meta to provide information, instruct Meta to allow or remove content and to explain its own decisions. According to Article 1(4)(2) of the charter, members may collectively "interpret Facebook's community Standards and other relevant policies ... in light of Facebook's articulated values." The Board can issue "policy guidance" but "will have no authority or powers beyond those expressly defined by this charter," according to Article 1(4) of the charter. Later sections – in Article 2 of the charter – refer to the Board's "authority to review."

Qualifications and Selection of Members

Qualifications to become a member are vaguely expressed. According to the charter, members shall "possess and exhibit a broad range of knowledge, competencies, diversity and expertise."[122] They should be able to work in a team and have "familiarity with matters relating to digital content and governance."[123] The Board appears to be relatively diverse when compared to traditional adjudicative bodies but not necessarily when compared to international adjudicators, which often have strict rules safeguarding proportional representation on their bench. According to its bylaws the Board must always include a "globally diverse set of members" representing all world regions.[124] Naturally, these regions are broadly defined – mainly by continent – and members might tend not to be familiar with the particular local contexts of appeals or the jurisdictional realities applying in the regions from where the appeals emanate. Nevertheless, the panels tasked with drafting decisions must include at least one member from the specific region.[125]

More than half of the Board's twenty-one members as of November 2024 have a law background.[126] Among the Board's members are several journalists, a former politician, a Nobel Peace Prize laureate, and a media and communications scholar.[127] As of November 2024, the membership's composition appears almost gender-balanced (eleven male, ten female) and features members from all world regions, however with a US

[122] Charter Article 1 Section 2.
[123] Ibid.
[124] Bylaws 1.4.1.
[125] Bylaws 3.1.3.
[126] Meet the Board section at the Board's website, accessible via <https://perma.cc/PYP5-L9M9>.
[127] Ibid.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

dominance (six members), followed by Asia (five members), Europe (four members), South and Central America, the Middle East (two members each), and Africa and Oceania (one member each).[128] The initial four members of the Board were selected by Meta.[129] The later members were selected by the Board's cochairs, who are also members, and Meta.[130] Since the Board became operational, its members autonomously decide who joins their ranks (but seemingly only based on a shortlist reviewed by Meta).[131] From a comparative perspective, this is a rather unusual structure, as appointment and nomination procedures tend to tie adjudicators to the entities they are tasked with controlling. Given the spotlight on the Board's independence and Meta's seeming inability to offer a more legitimate appointment or nomination process, this model appears to be a sensible one. We should not forget that many such arrangements owe their existence to the fact that Meta remains a private company – with all the structural reluctance toward external oversight that comes with it. If ever, external, meaningful guardianship over the world's largest content moderation machinery will build incrementally, not overnight.

Juridified Practice

Although the Board needs to function as an adjudicative body, its members do not have to be lawyers. Yet, as mentioned above, more than half of its twenty-one members active in November 2024 either studied law and/or work in law-related fields like human rights activism or are legal scholars. The dominance of lawyers among the Board members has several consequences.[132] For instance, the company's executives presumably viewed lawyers as particularly capable of building the internal procedures, substantive reasoning, and external legitimacy of a body that was supposed to *function* and *be perceived as* adjudicative. Intuitively, this makes sense. Who else but lawyers could (or would) build an adjudicative institution? It also mirrors the dominance of lawyers in the Board's creation at Facebook.[133] US lawyers dominated the internal debates at Facebook's "Governance Team" that rolled out the Board's institutional design between 2018 and 2020. Arguably, this lawyer-dominated group

[128] Ibid.
[129] LLC Agreement Section 5.3(a)(i).
[130] Charter Article 1 Section 8.
[131] Bylaws 5.3 (a) (i); cf. Steven Levy, 'Inside Meta's Oversight Board: 2 Years of Pushing Limits' (*Wired*, 8 November 2022) <https://perma.cc/SUP5-RQCV>.
[132] For an illuminating analysis of the impact of 'juridical discourse' on the Board see Kadri (n 2) 171–176.
[133] See above at 4.2.2.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

contributed to the Oversight Board being "an adjudicative body" (and not, say, one with a more parliamentary flavor). As one interviewee, who advised Meta's governance team when creating the Board, recalls:

> The fact of the matter is that the Oversight Board clearly started down one path. They could have started totally different paths. They could've issued three-page decisions, like the court of cassation does. It could've done that. It didn't. There's nothing in the charter or whatever that will tell you that it didn't do. It didn't do because of sociological factors, because they put a bunch of US lawyers in charge. What the fuck happens when you put a bunch of US lawyers in charge? Well, they started doing things that they're familiar with.[134]

Similarly, most people staffing the institution are lawyers as well. Whereas no member can have held a prior position at Meta, some high-level staffers at the Oversight Board were part of Facebook's effort in creating the Board in the first place.[135]

The dominance of lawyers among members pushed two normative aspects higher up the Board's agenda than initially foreseen. These are, first, human rights and, second, public- and administrative-law principles. The expertise in public law and human rights law of many Board members explains why the Board took a different turn than the company, or at least some people in the company, had presumably intended at its inception. The charter posits that the "Board will review content enforcement decisions and determine whether they were consistent with Facebook's content policies and values."[136] In contrast, the Board should only "pay particular attention to the *impact* of removing content in light of human rights norms protecting free expression."[137] Therefore, based on a strict reading of the charter, it seems as though Meta initially envisaged normative material unilaterally produced by Meta itself (its policies and values) and not human rights law as the key normative yardstick for the Board's decision-making.[138] Of course, many people were involved in creating the Board, and competing visions may have circulated at Meta and within the governance team.

[134] Interview with Academic, Consultant to Facebook in 2018 (n 31).
[135] See above at 4.2.2.
[136] Charter Article 2(2).
[137] Ibid (emphasis added).
[138] In that sense also Klonick, 'The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression' (n 3) 2477–2479 (not even mentioning international human rights law and pointing to other cautious, and rather limiting formulations in the charter).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

That said, the Board seems to focus particularly on international human rights law and to a much lesser extent on Meta's own devised rules. As we will explore in detail in Section 4.5 below, much of the Board's normative reasoning revolves around, and is framed in, the language of international human rights law, most notably Article 19 of the International Covenant on Civil and Political Rights (ICCPR).[139] For example, the Board allows for restrictions of freedom of expression only if "provided by law" and "necessary" to protect the rights of others, national security, public order, public health or morals, in the sense of Article 19(3) of the ICCPR. This external and arguably stricter normative framework can be understood as going beyond the initial design in the charter. The human rights background of many Board members (and staffers, as we will see in the following section) likely contributed to the Board's tendency to focus on international human rights law instead of Meta's vague and self-given "values," "charter," or "community standards." The policies and values that Meta executives had devised are still referenced in all decisions, but most of the Board's ink is dedicated to human rights-related reasoning.[140] This focus on human rights also constitutes, as we will soon see, one of the key differences from the DSA's ODSs, the other Emulated Guardian examined in this work.

Further, the Board used this human rights framework to infuse administrative-law principles into their recommendations based on legality and proportionality assessments, as specified in Article 19 of the ICCPR. It often deliberates on the basis of arguments like normative clarity and accessibility, users' right to remediation, the duty of the company to give reasons, and to translate its decisions and rules, for example. In this way, the Oversight Board aims to incentivize Meta to improve the normative framework for its content moderation. As one interviewee put it:

> The Board is attempting to infuse content moderation with juridical norms and procedures. We might call this a process of juridification . . . In partnership with Meta, and in light of existing human rights norms, the Board articulates norms by which Meta's content moderation decisions should be settled. It seeks to ensure that these norms will be enforced according to public law procedures. The upshot is to take a private for-profit corporation like Meta and to juridify its content-moderation operations.[141]

The rules that are supposed to "juridify" Meta's content moderation are public-law rules. Legality, proportionality, right to a remedy, and the duty

[139] See below at 4.4 and 4.5.
[140] See below at 4.5.2 et seq.
[141] Interview with Oversight Board Trustee (n 53).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

to give reasons for a decision, for example, are basic, almost universal administrative-law principles. The interviewee continues:

> To me, the issue is the infusion into management of public law principles. Now, I don't know whether that's the precise idea of juridification, but that seems to me the precise challenge of the Oversight Board.[142]

Similarly, in its decision regarding whether Facebook was justified in banning former US President Trump from its platform after his role in the attack on the Capitol on January 6, 2021, the Board assessed whether Facebook's rules were "clear and accessible."[143] In classic administrative-law fashion, the Board "reiterates that the patchwork of applicable rules makes it difficult for users to understand why and when Facebooks restricts accounts, . . . [this] raises legality concerns."[144] In the same vein, the Board argued that "[a]ppropriate limits on discretionary powers are crucial to distinguish the legitimate use of discretion from possible scenarios around the world in which Facebook may unduly silence speech not linked to harm or delay action critical to protecting people."[145] The discursive might of such public law arguments is remarkable. While the Board had used general public law arguments to contain (and justify) Meta's power in its 2021 *Former President Trump's suspension* decision, the Board further expanded this approach in 2023. Then, when faced with incendiary posts by another head of state, the Board used more substantive human rights arguments and weighed, in quite some detail, the conflicting arguments for and against shutting down the profile of the Cambodian Prime Minister, who had used his Facebook account to incite violence in the run up to an election. Both cases dealt with Facebook's infamous "newsworthiness allowance," a rule that essentially excludes important people from otherwise enforced rules against hate speech and the incitement of violence. In 2023, the Board demanded that the company would not only delete content posted by the Cambodian Prime Minister but recommended to Meta to temporarily suspend his profile altogether.[146] That head of state, Hun Sen, had disseminated incendiary rhetoric through his Meta accounts, which arguably led to physical violence.[147]

[142] Ibid.
[143] Former President Trump's suspension [2021] Oversight Board 2021–001-FB-FBR.
[144] Ibid at 25, 26.
[145] Ibid at 26.
[146] Cambodian Prime Minister [2023] Oversight Board 2023–003-FB-MR.
[147] Ibid.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

Such clear borrowings from public law would be impossible if not for the preponderance of lawyers among the Board's members (and staff). Hence, the dominance of lawyers among the Board's creators, members, and staffers aligned its design and practice closely with what lawyers are familiar with – that is, adjudicatory institutions. This again led to direct borrowings from public- and especially administrative-law principles to "juridify" Meta's content management operations.

4.3.1.4 The Staff

Apart from its members, the Board relies on the efforts of a considerable number of staff. Between seventy and ninety people work in offices around the world. In 2024, the London office was the largest, with roughly forty staffers working there.[148] As Board members only work part-time for the Oversight Board and are scattered around the globe, the staff have assumed a key – and largely overlooked – role in shaping the Board's strategic vision and practice. Staffers fulfill a dual function. On one hand, staffers are sherpas, helping the Board members navigate intricate legal and technical terrain. On the other, staffers are strategists themselves. By dint of their full-time and long-term positions working for the Board and their expertise in the field, they know many cases by heart and infuse their own normative and strategic considerations into the Board's deliberative remit. Staff members have thereby played a crucial part in fashioning the Board as a formal structure independent of Meta.

Permanent Position and Expertise

In light of many social theorists' focus on practice and cultural perspectives,[149] recent sociological and legal scholarship has highlighted the often overlooked but prime importance of staffers, secretariats, and "back-room personnel" in the business of international adjudication.[150]

[148] Interview with Oversight Board Staffer 1 (n 54) 1.

[149] The foundational work is arguably: Pierre Bourdieu, *Outline of a Theory of Practice* (Cambridge University Press 1977); for an overview see: Andreas Reckwitz, 'Toward a Theory of Social Practices: A Development in Culturalist Theorizing' (2002) 5 European Journal of Social Theory 243; for a theory of social norms through a praxeological lens see: Christoph Möllers, *The Possibility of Norms: Social Practice beyond Morals and Causes* (Oxford University Press 2020).

[150] For EU law Tommaso Pavone, *The Ghostwriters: Lawyers and the Politics behind the Judicial Construction of Europe* (Cambridge University Press 2022) 129; already Antoine Vauchez, *Brokering Europe: Euro-Lawyers and the Making of a Transnational Polity* (Cambridge University Press 2015) 24–35. Antonin Cohen and Antoine Vauchez, 'The Social Construction of Law: The European Court of Justice and Its Legal Revolution Revisited' (2011) 7 Annual Review of Law and Social Science 417. For a similar perspective on

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

Especially in diversely staffed international adjudicative institutions, advisers, lawyers, and staffers are an indispensable pillar for organizational efficacy. Typically, their role is particularly prominent wherever adjudicating personnel work in globally dispersed, ad hoc, or part-time conditions. Yet most courts, tribunals, and other adjudicators remain coy about the precise roles their staff play. Instead, adjudicators often seek to "maintain the 'fiction' of rulings being delivered from on high …, whereas in fact, a large team of permanent staff is devoted to handling increasingly complex disputes."[151] Other scholars have suggested that this applies for other international and transnational adjudicators like the European Court of Justice and the World Trade Organization (WTO)'s appellate bodies, and this is equally true for the Oversight Board.[152]

The main reason for the staff's prominent role in the Board's practices is that, in contrast to Board members, staff are full-time employees. Plus, there are a lot more staffers than Board members. To repeat the figures, while the staff totals between seventy and ninety people, the Board has only roughly twenty members. Around thirty-five staffers, deployed to the so-called "content team," work directly on all cases. The content team has several subteams. These are: a case selection team; a project management team; a case and policy team; and the data and implementation team. All staffers work under the supervision and at the direction of Board members. Yet, the content team's permanent positions, expertise, and involvement with all aspects of each case seem to allow staffers to dig deep into issues they consider relevant and bring questions to Board members' attention.[153] Staffers contribute – and thereby shape – the Board's practices and strategies in various ways. They create longlists for consideration by a subteam

international adjudication see Tommaso Soave, *The Everyday Makers of International Law: From Great Halls to Back Rooms* (Cambridge University Press 2022); Jeffrey L Dunoff and Mark A Pollack, 'International Judicial Practices: Opening the "Black Box" of International Courts' (2018) 40 Michigan Journal of International Law 47. For a sociological perspective on the influence of specialized staffers see RR Ritti and Fred H Goldner, 'Professional Pluralism in an Industrial Organization' (1969) 16 Management Science 233, 233–246; DiMaggio and Powell (n 64) 150, 155. See further for so-called "fourth branch" institutions: Mark V Tushnet, *The New Fourth Branch: Institutions for Protecting Constitutional Democracy* (Cambridge University Press 2021) 45–46.

151 Pauwelyn and Pelc (n 15) 538 (referencing a quote by Thomas Cottier who himself was a WTO Appellate Body member).

152 See for an empirical account of the WTOs' Appellate Body's secretariat: ibid. See for reminiscent phenomena in national judiciaries already Owen M Fiss, 'The Bureaucratization of the Judiciary' (1983) 92 Yale Law Journal 1442.

153 Interview with Oversight Board Staffer 1 (n 54); Interview with Oversight Board Staffer 2 (n 10); Interview with Oversight Board Staffer 3 (n 11); Interview with Oversight Board Staffer 4 (n 52).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

of members, the Case Selection Committee, which then draws up a shortlist of, in their view, the most serious cases.[154] Staff also provide evidentiary material to inform deliberations at the direction of the Board members and help with fact-finding in relation to complex technical and normative questions. During the deliberations of Board members, staffers are also allowed to respond to questions about the briefing materials when asked.[155]

Further, Oversight Board staffers are highly qualified. Many staffers are lawyers, and many have a human rights background.[156] This enables staffers to contribute to the Board's adjudicative practice while reflexively reinforcing the Board's adjudicative nature and perceptions thereof. Appropriately enough, staffers and members are subject to the same code of conduct.[157] Requirements to disclose conflicts of interest or the obligation to be independent, for example, amount to virtually the same thing as the organization's code of conduct always refers to "members and staff."[158] This indicates the staff's importance for the decision-making process and the Board's overall strategy. Another factor reinforcing the staff's position within the institution is staff members' identification with the Oversight Board's mission. A strong sense of commitment and the perceived societal relevance of their work was quite prevalent in every interview and supplementary discussion. Many of the interviewed staffers working in content teams perceive their work as helping to resolve problems for society.[159] Many have a background in NGO work, courts, or other judiciary-related jobs.[160] One staffer describes the staff as having a "very strong sense of identification" with the Board and its mission,[161] while another mentioned their work's main objectives to "protect individual rights" and to "mitigate the negative systemic effects on the basis of individual cases."[162] Even though these qualitative insights are

[154] cf. Rulebook (November 2021) (n 104) 8; Email Oversight Board Staffer 7 (on file with the author) (3 February 2023).
[155] Interview with Oversight Board Staffer 1 (n 54).
[156] Interview with Oversight Board Staffer 1 (n 54); Interview with Oversight Board Staffer 2 (n 10); Interview with Oversight Board Staffer 3 (n 11); Interview with Oversight Board Staffer 4 (n 52).
[157] See Code of Conduct attached to the Bylaws.
[158] See e.g., Code of Conduct Sections "Introduction," "Independence and Impartiality."
[159] Interview with Oversight Board Staffer 1 (n 54); Interview with Oversight Board Staffer 2 (n 10).
[160] Interview with Oversight Board Staffer 1 (n 54); Interview with Oversight Board Staffer 2 (n 10).
[161] Interview with Oversight Board Staffer 1 (n 54).
[162] Ibid.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

inherently subjective and reflect, at best, the perspective of the interviewed staffers and not the entire group of employees, we should consider them tentative indicators of many staffers' attitudes to their work.

However, staff expertise also raises one problematic issue. Several staffers are former Facebook employees, some even from the Governance and Strategic Initiatives team that created the Oversight Board.[163] On the one hand, intimate knowledge of how Meta works is crucial for the Board's success. On the other, staffing a guardian institution with former employees of that same "guarded" organization brings with it connotations of personal connections between controller and controlled. In that sense, the questions surrounding the Oversight Board's institutional design are not at all new or unprecedented. In fact, they remind us of Iuvenal's classic "*quis custodiet ipsos custodes*?" (who guards the guardians?) question.[164] Correspondingly, the Oversight Board's code of conduct considers "being a former ... employee of Facebook, Inc." a conflict of interest for "members *and* staffers."[165] Further, Board members and staffers must disclose several potentially disqualifying facts, one of them being previous employment at Facebook.[166] However, former employment does "not necessarily automatically disqualify" someone from working for the Board, as these cases would "be reviewed by the trustees to determine the appropriate course of action."[167] A reading of the applicable governance documents indicates that such conditions (and the accompanying potential conflicts of interest) were reviewed and validated by the trustees in several cases. From a pragmatic perspective this makes sense, because outside hires would presumably be less familiar with the inner workings of Facebook's content-moderation activities. However, this episode adds to the impression that the normative material applicable to the Board is met by considerable centripetal

[163] Interview with Oversight Board Staffer 4 (n 52).

[164] Iuvenal, *Satires* VII 347–348.

[165] Code of Conduct, Article 1 bullet point 1 No 1 (emphasis added). Verbatim that paragraph reads: "[M]embers, staff, and their immediate family will not create or maintain conflicts of interest, which include but are not limited to the following: ... Personal or professional conflicts of interest that may impact (or appear to impact) the ability of Board members or staff to perform their duties, such as: (1) being a former or current employee of Facebook, Inc. or having an immediate family member employed by Facebook, Inc," Code of Conduct Article 1 bullet point 2.

[166] Article 3 Code of Conduct in conjunction with Appendix A (to the code of conduct) 'Disqualifications and Required Disclosures' (Bylaws January 2021, p 45).

[167] Appendix A (to the code of conduct) 'Disqualifications and Required Disclosures' (Bylaws January 2021, p 45).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

forces of seeking legitimacy (by the use of constitutional metaphors and strict codes of conduct, for example) while maintaining practical operability (in other words, needing people who actually know what Meta is doing).[168]

"Fishing Mode"

Driven by what appeared like a high degree of personal motivation, the staffers working on the Board's adjudicative processes operate and help fine-tuning the Board's focus, "jurisprudence," and procedures. Each of the content subteams consists of several staffers, and there are usually fewer than ten on each subteam.[169] These content subteams fulfill specific tasks concerning the Board's adjudication and are present at every step while a case makes its way through the Oversight Board.[170] Staffers in the Case Selection Team, which is made up of several lawyers and political scientists, aid in selecting cases at the direction and supervision of the Case Selection Committee, which comprises a rotating membership of Board members.[171] Staffers also prep Board members in substantive, procedural, and factual matters.[172] Even though staffers do not write formal drafts of the Oversight Board's decisions or advisory opinions, it is plausible that the extensive reports written by the Case and Policy Team lay the groundwork for the later draft produced by Board members.[173] At least some of the "content team" staffers are present when Board members deliberate but do not contribute to deliberations and are not encouraged to share their opinions.[174]

And so, even though they are not vested with formal decision-making authority, staffers are an indispensable part of the decision-making process in a broader sense. One striking example of the staff's value and strategic vision was what one interviewed staffer called "fishing mode."[175] This describes how a group of fewer than ten staffers draws up a longlist of a few dozen cases from the many thousands submitted to the Board by

[168] Whereas it might be even questionable whether former members of the Governance and Strategic Initiatives team would be experts in content moderation as the team's job was precisely not to look at what Facebook is doing but to invent an institution that would control Facebook.
[169] Interview with Oversight Board Staffer 1 (n 54).
[170] Ibid.
[171] Ibid.
[172] Ibid.
[173] Ibid.
[174] Email Oversight Board Staffer 7 (on file with the author) (n 154).
[175] Interview with Oversight Board Staffer 3 (n 11).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

users of Meta's services. Board members provide staffers with criteria, which staffers then apply when filtering appeals.[176] A high-ranking staffer described these case-selection criteria as a blend of specific content, like "breasts of transitioning people that are deleted as female nudity," with general aspects like "posts from conflict regions."[177] Hence, filtering appeals works in an aim-oriented fashion. "Fishing mode" depicts the case selection team's angling for "good cases" after first filtering the cases based on criteria like metadata, region, language, views, and violated community standard.[178] Like Captain Ahab and the crew of the Pequod, the case selection team stands in the crow's nest looking for tell-tale ripples in the water, hinting at a whale under the surface. But how to spot a metaphorical whale in this ocean of cases? The same interviewee mentioned that a key consideration in longlisting is the user statement.[179] It increases the chances of selection where the user shows serious engagement with the matter and an understanding of what "the problem" is.[180] This has a practical background. It makes it more likely that the user will keep their post up and does not delete it once informed that the Oversight Board is investigating it.[181] The staffer described finding good cases as follows:

> It's not like finding the needle in the haystack. It's more like finding a particularly nice-looking piece of hay in the haystack.[182]

Clearly, one key question is what makes a "good" case. It is not only the criteria but also the structural issues underlying Meta's decision-making processes that the case is potentially addressing.[183] Yet the staffer points out that it is hard to interpret these structural aspects from the limited amount of information they are able to glean from the submissions.[184] Often, the Board recommends Meta to change its rules to be more concise and transparent and align its procedures more with principles of public law, administrative law, and good governance. To buttress its argument, the Board refers to public-law standards as they appear in the wording of legislation like Article 19 of the ICCPR, for example – standards like

[176] Bylaws Article 1 Section 3.1.2.
[177] Interview with Oversight Board Staffer 3 (n 11).
[178] Ibid.
[179] Ibid.
[180] Ibid.
[181] Ibid.
[182] Ibid.
[183] Ibid.
[184] Ibid.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

legality and proportionality.[185] Further, the Board repeatedly recommended that Meta should notify users about content-moderation decisions, that it should state reasons and, where necessary, provide timely translations of its rules and decisions.[186]

4.3.1.5 The LLC

Lastly, there is the limited liability company, which is the legal person at the center of much of the Oversight Board's contracting. The purpose of the LLC is solely organization related. The sole member of the LLC is the Oversight Board Trust, according to Article 2.2 of the LLC agreement. The trustees serve as the LLC's managers. The corporation is obliged to rent office space, draw up contracts with staffers (as employees), and draw up individual contracts with Board members.[187]

4.3.2 Jurisdiction

Because Emulated Guardians are adjudicators, they must have jurisdiction. To be effective controllers of organizational power, said adjudication must also be compulsorily applied to the organization these guardians are tasked with controlling.

4.3.2.1 Subject Matter Jurisdiction

The Board's purpose is to control Meta's content-moderation activities. To achieve this aim, the Board's founding documents outline the Board's "jurisdiction." The charter states that Meta users "may bring forward content for Board review" whenever they "disagree with the outcome of Facebook's decision and have exhausted appeals."[188] In contrast, the bylaws are formulated more narrowly. They had initially restricted the Board's jurisdiction to review only "individual pieces of content, such as specific posts, photos, videos, and comments" that had been "removed for violations of . . . 'Community Standards' or 'Community Guidelines.'"[189] Because of this, the Board appeared to be barred from reviewing a multitude of other actions and design choices impacting how Meta

[185] See e.g. Pro-Navalny Protests in Russia [2021] Oversight Board 2021–004-FB-UA at p 13 to 16 (of the PDF version) for the "legality" and proportionality assessment as well as recommendation No 6 (notification).
[186] See e.g. COVID lockdowns in Brazil [2021] Oversight Board 2021–008-FB-FBR at 10.1.
[187] LLC Agreement Section 5.3.(a)(i) and (vii).
[188] Charter Article 2 and 2(1).
[189] Bylaws Article 3 Section 1.1.1.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

exercises power over its users. Scholars have repeatedly criticized these reduced terms of reference as the "biggest disappointment" of the Board.[190] Allowing users *other than* the original uploader to appeal moderation decisions is essential, as it will typically be the former whose rights may be infringed. For example, to protect the rights of victims in cases of defamation, libel, or revenge porn, victims must be able to challenge moderation decisions. Correspondingly, respective provisions were included in the list of possible appellants to ODSs in the EU's DSA.[191] Since April 2021, and pursuant to a clause in the bylaws,[192] which opened the door for future expansions of jurisdiction, the Board now accepts appeals *for* the removal of content from Facebook and Instagram (that is, appeals *against* Meta's decision to leave content up).[193] In 2024, the Board expanded its jurisdiction to Meta's Threads, a social network the company set up in 2023 to compete with X after Elon Musk's takeover.

Interestingly, the bylaws already indicate that the Board might move toward a more comprehensive interpretation of its powers in the future.[194] Apparently, this issue emerged during the global consultation period, prompting the company to concede, in an attachment to the global consultation report, that "[o]ver time, the Board may look to decide upon other actions (e.g. downranking or applying interstitial warnings)."[195] Correspondingly, Article 3 Section 1.1.2 of the bylaws states that "in the future, people will have the opportunity to request the Board's review of other enforcement actions." As mentioned previously, the Oversight Board has already made use of that provision and expanded its jurisdiction. Further, in October 2022, the Board ventured further down the path of these more graded forms of moderation as it demanded Meta to put up warning screens like "sensitive or disturbing content."[196] In several interviews, staffers and Board members indicated their willingness to focus on

[190] Douek, 'What Kind of Oversight Board Have You Given Us?' (n 102). See further Helfer and Land (n 3) 12–15.

[191] See at 3.3.4.1.

[192] Bylaws Article 3 Section 1.1.2.

[193] Oversight Board, 'The Oversight Board Is Accepting User Appeals to Remove Content from Facebook and Instagram (Press Release)' (April 2021) <https://perma.cc/FTD9-8SE2>.

[194] Bylaws Article 3 Section 1.1.2.

[195] See at tenth question, topic: resolution under "implementing feedback" in Facebook, Capturing Feedback: Charter for the Oversight Board, accessible via: <https://perma.cc/XZ37-MBQG>. See also Douek, 'How Much Power Did Facebook Give Its Oversight Board?' (n 102).

[196] Oversight Board, 'Oversight Board Publishes Transparency Report for Second Quarter of 2022 and Gains Ability to Apply Warning Screens (Press Release)' (October 2022) <https://perma.cc/6C3F-WHPT>.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

"other enforcement actions" like the amplifying or demoting of content, which may highlight or restrict the visibility (and therefore harmfulness) of content without deleting it.[197]

Therefore, prominent criticism that the Oversight Board would exemplify a simplistic, "stylized" understanding of content moderation as an endless string of customized decisions seems not entirely fair.[198] Although such a viewpoint may have colored the earliest visions for the Board, its practice indicates a nuanced and comprehensive understanding of the structural challenges of content moderation.[199] As we will shortly see, the Board's key normative output is *not* made up of individual decisions but general, structure-focused recommendations, which aim at improving Meta's governance procedures and structures instead of deciding only on a few paradigmatic cases.[200] Thus, the Oversight Board is not a solely binary adjudicator allowing "individual appeal to a human in every case."[201] Quite the opposite, because the Board decides only very few individual appeals. Instead, the Board uses the few cases it adjudicates to formulate general normative sentences that its members believe should apply to Meta. Based on these general normative sentences, the company might then improve its governance and decision-making practices.

In conclusion, the Board's jurisdictional focus rests on one important, and publicly visible, aspect of Meta's content moderation: the removal or retention of individual pieces of content for violation of community standards. However, the company exercises power over its users in a more wide-ranging way than this. Meta retains content even though other users may want it removed. The company shows content according to algorithmic metrics and has, after all, designed the technical infrastructure that allows users to communicate in the first place. All these decisions, one way or another, "moderate" the communication of users on the platform. For the time being, the Board's "jurisdiction" seems not to have plunged further into those deeper layers of moderation. Ultimately, those deeper and darker waters is where the

[197] See for an informative introduction to this governance practice on the network TikTok by Jing Zeng and Bondy Valdovinos Kaye, 'From Content Moderation to Visibility Moderation: A Case Study of Platform Governance on TikTok' (2022) 14 Policy & Internet 79.

[198] For the criticism see Douek, 'Content Moderation as Systems Thinking' (n 7) 528, 567 et seq.

[199] For the Board's origins see above at 4.2.

[200] See at 4.5 and 4.6.

[201] Douek, 'Content Moderation as Systems Thinking' (n 7) 568.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

whale is. Yet, on the evidence of decisions made up until now and the tentative expansion of the Board's jurisdiction following the remit allowed by the bylaws' provisions, it appears that the Board might incrementally venture toward a more comprehensive jurisdiction over the various aspects of content moderation which go beyond individual decisions that enforce rules relating to individual pieces of content.

4.3.2.2 Compulsory Jurisdiction

In principle, the Board's jurisdiction is *compulsory* for Meta for the appealable content outlined above. According to Article 2(1) of its charter, the company "will commit to the Board's independent oversight." However, appeals may be ineligible for Board review if they concern content that is "criminally unlawful" or if a Board decision was likely to trigger "criminal liability" allegations or "adverse government reactions" for Facebook, its employees, or the Board.[202] For such pieces of content, Meta can contest the Board's "jurisdiction." The Board and the company found a somewhat peculiar way to ensure that the Board stays within the bylaws' and charter's limits. The Board typically shows the company the shortlisted cases prior to selecting the appeals for review. At this prior stage, the company might then object to the Board choosing, or technically: shortlisting, a specific case for the above-mentioned reasons.[203] Neither the charter nor the bylaws had stipulated any requirement for this prior "legal review" by the company.[204] But it appears to be now standard practice that the Board shares shortlisted cases for prior "legal review" with Meta. The company's legal department then may "exclude" certain appeals from the list due to legal problems in the cases' respective jurisdictions. As stated in the Board's *Rulebook for Case Review and Policy Guidance*, which Board members adopted in November 2020: "shortlisted cases will be sent to Facebook for legal review. Facebook may exclude from the shortlist cases that are not eligible for review by the Oversight Board, in accordance with the Bylaws, such as cases that could result in criminal liability or adverse government action if reviewed."[205]

This is remarkable. In principle, Meta's interest in avoiding negative legal consequences by subjecting itself to the Board appears legitimate.

[202] Bylaws Article 2 Section 1.2.2.
[203] Interview with Oversight Board Staffer 1 (n 54) 1.
[204] See Charter Article 3 and Bylaws Section 3.
[205] Rulebook (n 104) Step 2 No 1 (p 8)

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

Despite all the metaphorical bravado, the Board remains an Emulated Guardian that resides, in contrast to the DSA's ODSs, outside the realms of democratically set law, and only time will tell whether the Board will incrementally build adjudicatory authority.[206] However, the fact that the company has retained at least the possibility of weighing in on the Board's case selection is a blot on the latter's authority. From the perspective of individual rights adjudication, Meta should be able to *contest* the eligibility of cases but not *decide* whether they are ineligible. Thus, at the end of the day, we see how Meta's interest in sustaining autonomy collides with the Board's public image and function as (adjudicatory) guardian. Yet, historically, adjudicatory control over bureaucratic power structures grows gradually and incrementally.[207] Thus, possibly existing communicatory channels leading from the company to the Board should warn us not to enthusiastically drape the Board with the ceremonial garb of "supreme court," as the Board is still in the process of building up its organizational muscles.

4.3.2.3 Appeals, Referrals, and Policy Advice

Users whose content has been deleted by Meta or who filed an unsuccessful notice for other users' content to be deleted have "standing" to file an appeal to the Board.[208]

Further, Meta may "directly refer cases to the Board that are significant and difficult."[209] For example, in January 2021, the company referred the decision to the Board as to whether it was justified in permanently deleting the Facebook and Instagram accounts of then US President Trump after his role in the January 6, 2021 storming of the Capitol. Using Article 19 of the ICCPR's criteria for legality, the Board harshly criticized the company's incomplete and imprecise sanctioning policies. Ultimately, the Board argued that the company had no right to permanently ban Trump's accounts because such a sanction was not stipulated in its terms of service at the time. However, the Board allowed that a temporary ban on Trump's accounts would have been a proportionate response given the threat of real-world harm emanating from many of the posts on his accounts.

[206] See at 3.7, 4.5, 4.6, and 5.3.
[207] See at 2.2.
[208] Bylaws Article 3 Section 1.1.1 and see above at 3.2.1.
[209] Bylaws Article 2 Section 2.1.1.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

Lastly, separate from individual case review, the company may request "policy guidance" and a "policy advisory statement from the Board."[210] These are general normative sentences on structural issues. Even though the various documents and the Board's decisions themselves do not use consistent vocabulary, we can distinguish between policy advisory statements and specific recommendations following *general* requests by the company, and policy advisory statements and "recommendations" concerning the underlying governance issues at the company, but which are retrospectively issued as a byproduct of individual decisions.[211]

4.3.3 *Impartiality and Independence*

To emulate adjudicative control over administrative power, the Oversight Board must be *impartial* and *independent*. As outlined in Chapter 2, independence refers to the adjudicating institution's structural design, whereas impartiality signifies the fair-mindedness and objectivity of the person adjudicating.[212] Adjudicators must be externally autonomous – that is, not bound by hierarchical command or factual constraint – and internally objective and fair-minded in terms of the procedure's outcome and the rules applied.[213] When applying these standards, we should remind ourselves of the words of Martin Shapiro who, based on extensive historical work, cautioned: "persons and institutions to which we normally award the titles judges and courts, . . . in reality . . . are simply at one end of a spectrum rather than constituting an absolute distinct entity."[214] For example, specialized adjudicators that control administrative power often have a legal, organizational, and practical nexus to the organization they control. Often, sub-court administrators like UK administrative tribunals, American Article-I courts, or EU Boards of Appeal may even operate attached to specific administrative agencies. Yet they all have structural safeguards in place to shield their adjudicative decision-making processes and people from undue influence by the controlled entity. Therefore, certain organizational overlaps are often necessary in

[210] Bylaws Article 1 Section 2.1.3.
[211] See at 4.5.
[212] For a US perspective see James E Moliterno, 'The Administrative Judiciary's Independence Myth' (2006) 41 Wake Forest Law Review 1191, 1199–1208.
[213] cf. the recent case-law by the ECJ on Poland's judiciary, see Independence of the Polish Supreme Court [2019] European Court of Justice (Grand Chamber) C-619/18, ECLI:EU:C:2019:531 paras 71–77.
[214] Martin M Shapiro, *Courts: A Comparative and Political Analysis* (University of Chicago Press 1981) 8.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

practice or, for the purposes of gaining greater expertise, even desirable. Adjudicatory impartiality and independence hence call for an appropriate institutional, personal, and procedural setup that helps to safeguard externally autonomous decision-makers and internally objective decision-making while maintaining operability.

Viewed against this yardstick, I argue that the Board is – albeit organizationally and financially intertwined with Meta – in practice sufficiently autonomous and its members fair-minded to be considered impartial. However, when it comes to the Board's independence we see a more blurred picture. On the one hand, the (irrevocable) trust installed a seemingly effective organizational buffer between Board and company. On the other, Meta retained avenues of influence over who gets nominated[215] and may "exclude" cases from review if "they are not eligible for review by the Oversight Board."[216] To what extent Meta treads these paths of influence is unclear and it should be noted that during interviews nobody indicated any undue influence being applied.[217] However, these two organizational and procedural aspects are clear "no-go's" from a classic perspective on adjudicatory independence. The crux is: the Board is only of value to Meta if it is publicly perceived as sufficiently independent. This is arguably why, for example, Meta's preliminary "legal review" of the cases entering the Board's docket is not explicitly stipulated in the provisions of the charter or in the bylaws but only in a rather ominous sounding "rulebook," which "nonbindingly" specifies the bylaws.[218]

The following section analyzes the decisive aspects of independence and impartiality for adjudicators controlling administration-like operations. The Board is not entrenched in Meta's hierarchy, though the company nevertheless retains some influence over its docket. Further, we reflect on the trust's role as mediator between the Board and Meta, and on how "dangerous" the Board effectively is for Meta. We then look more deeply into the Board's nomination and removal procedures and investigate its operational and financial autonomy.

[215] See also especially at 4.3.3.1.

[216] For the Stephen Levy article see Levy (n 131) 14–16; for the quote see Rulebook (n 104) Step 2 No 1 (November 2020).

[217] The journalist Stephen Levy reported in a leading outlet for tech journalism that "Meta employees remain deeply involved in hiring and are a factor in why the Board is still far short of the total number set out in its charter." Levy also reported that Meta directly prevented one critical researcher from becoming a Board member, cf. ibid 14.

[218] See Rulebook (n 104) Step 2 called "Facebook Legal Review" at 8; in contrast Bylaws Section 3.1.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

4.3.3.1 No Hierarchy, but Docket Influence?

For the Board to be independent, its members must not be subject to hierarchical command governing their decision-making. The Board's independence was a key concern during its creation.[219] Article 1 Section 1 of the bylaws, as well as Article 1 of the charter, stipulate that members "will exercise neutral, independent judgment and render decisions impartially." Even though the founding documents retain some veto powers with the company when it comes to certain management decisions by the trust (for example, changing one or more bylaws), Meta cannot issue any "commands" to Board members.[220] Correspondingly, no interviewee voiced concerns about company overreach. Even though their responses do not guarantee that there was no overreach by the company,[221] the way the Board's decisions developed suggests that it indeed operates largely autonomously from the company's influence. In many decisions, the Board has thrown light on and condemned Meta's governance practices. For example, in the Trump decision, the Board criticized the company's request to *formulate* rules. It stated:

> [t]he Board rejects Facebook's request for it to endorse indefinite restrictions, imposed and lifted without clear criteria. Appropriate limits on discretionary powers are crucial to distinguish the legitimate use of discretion from possible scenarios around the world in which Facebook may unduly silence speech not linked to harm or delay action critical to protecting people.[222]

The Board's forceful repudiation of the company's formally referred request, the founding documents' focus on the Board's independence, as well as the Board members' status as independent contractees, makes hierarchy-based subordination of the Board to the company unlikely. Similarly, several interviewees said that they had seen a drop in public criticism of the Board's purported independence after the Trump decision.[223]

[219] See for the role of the trust in safeguarding independence at 4.3.3.2; see for the Board's genesis in general at 4.2.

[220] cf. Bylaws Article 5. See further Klonick, 'The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression' (n 3) 2481 et seq.

[221] We should not forget that, after all, the interviewees working for the Board have a self-interest to be perceived as independent as possible.

[222] When Facebook referred the case to the Board, it specifically requested "observations or recommendations from the Board about suspensions when the user is a political leader." cf. Former President Trump's suspension (n 143) 5.

[223] Interview with Oversight Board Staffer 1 (n 54); Interview with Oversight Board Staffer 2 (n 10); Interview with Oversight Board Staffer 4 (n 52).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

Yet even though the company does not appear to be exercising undue influence on the Board's decision-making and especially its deliberations, Meta retains at least the possibility of exerting influence over the Board's docket. Board members who are part of the Case Selection Committee are in theory free to compile a shortlist of possible cases to take forward. However, before the Case Selection Committee decides which cases to adopt for decision, that shortlist, according to the Board's internal "rulebook" *must* be sent to Meta. It is unclear what would happen if longlisted cases were not sent to legal review by Meta. As mentioned previously, the company "may exclude from the shortlist cases that are not eligible for review by the Oversight Board, in accordance with the Bylaws, such as cases that could result in criminal liability or adverse government action if reviewed."[224]

Naturally, if Meta ends up removing cases from the shortlist, the Case Selection Committee cannot select them, and the Board cannot decide whether to take them forward. Neither the bylaws nor the charter regulate this "Facebook Legal Review," although the company's intervention is supposed to be just part of the "non-binding ... practical guidance ... detailing case review and policy guidance procedures."[225] Judging from transparency reports, Meta indeed conducts such legal reviews for each case and, in early 2021, Meta did exclude at least one case, and possibly more, from the shortlist through its use of this "legal review."[226]

An adjudicator whose docket is subject to any caveats emanating from the entity it is supposed to be controlling does not match up to the traditional standards of independence familiar in other adjudicative bodies. Even nonadjudicative guardians like ombudspersons or other "fourth branch" actors often enjoy more autonomy in selecting and rejecting cases. As Klonick observed, Meta's "delegation of authority to the Trust [and the Board in general] is far from absolute."[227] In sum, Meta granted the Board independence in the most literal sense, as the Board is

[224] Rulebook (n 104) step 2 No 1.

[225] Rulebook (n 104) Introduction.

[226] The transparency report only states "Of the 39 cases shortlisted in this period. 34 were confirmed as eligible by Facebook's legal team and five cases were removed from the Board's review either by the person who appealed the content or through Facebook's legal review." Whether it was just one of those five, two, three, or four, which Meta excluded we do not know. If it would have been four (it cannot be five as that would not leave one for the user removal), Meta would have excluded more than 10 percent of the longlisted cases in that quarter. See Oversight Board Transparency Report Q 4 2020, Q1 & Q2 2021, 30 – accessible via: <https://perma.cc/JTJ8-DTHD>.

[227] Klonick, 'The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression' (n 3) 2481. See also again Levy (n 216) 14–17.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

not formally entrenched into Meta's hierarchical structure. However, depicting the Board as independent in the traditional, adjudicatory sense of the word would, at the moment, be inaccurate for the simple reason that Meta retains some influence over the Board's docket.

4.3.3.2 The Role of the Trust

Even though the Board does not receive directives from the company, organizationally speaking, the two bodies are intertwined. Central to these ties – but also, simultaneously, to the Board's far-reaching practical autonomy vis-à-vis the company – is the trust. The irrevocable, noncharitable "Oversight Board Trust" (established in Wilmington, Delaware) serves as a buffer between Meta (as a private company) and the Board, which, although privately organized, performs the public law-like function of controlling Meta's administration-like content moderation. Six trustees are obliged to isolate the Board's adjudicatory decision-making from the company as much as possible. Board members and trustees communicate mostly in regular (though none too frequent) meetings, where they discuss organizational requirements and administrative issues, for instance.[228]

According to Article 2.2 of the trust agreement, the trust and the individual trustees have to "protect the independent judgment of the Board Members and their ability to fulfil their stated purpose." To do so, many organizational aspects are delegated to the Board itself. For example, the Board can propose new members, establish new procedural rules (within the limits of the bylaws), or engage in new practices like issuing quarterly reports to measure the impact of its decisions and policies. In turn, the trustees are obliged to allocate resources to the LLC for office space, travel, accommodation, workshops, and work retreats for teams scattered around the globe. The trust also handles contracts for Board members and employees. Trustees do not participate in case selection, neither do they take part in deliberations, voice opinions, or anything else that could directly influence the Board's adjudicatory work. One trustee describes their involvement with the Board's day-to-day work as follows:

> We [the trustees] stay completely independent of the Board when it comes to selecting and deciding cases. But with respect to the actual administration of the OSB entity, there are channels of communication. There must be, because we need to set up an organization that works for the Board members, but we must do so in ways that are consistent with our own fiduciary responsibilities.[229]

[228] Interview with Oversight Board Trustee (n 53).
[229] Ibid.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

Hence, beyond formal channels of communication, informal avenues between trustees and Board members arguably exist. After all, the group of Meta executives and employees creating the Board, its members, and staffers is fairly tight knit. Most of them went to (and sometimes teach at) elite US law schools. The effects of these personal and habitual connections among practitioners, lawyers, and legal scholars have been identified as a common factor for fostering the emergence of normative orders.[230] Such "field effects" (borrowing from Pierre Bourdieu) should not be underestimated when it comes to the makeup and operations of the Oversight Board.

Correspondingly, even though most people working for the Board, or the trust seem to have, at some time or other, shared their strategic ideas for how the Board should operate, the Board's practice isolates its decision-making from the company and even the trustees. Trustees seem largely isolated from personal or strategic considerations relating to the people working for the Board. One trustee said:

> Of course, I stay completely away from the decision-making processes. That has nothing to do with me. But the larger strategic positioning and functioning of the Board is an entirely different question. We need to establish a model that will provide criteria of success with respect to these larger strategic issues. I myself have many ideas, but in the end the Trustees and Board members will need to reach a common agreement.[231]

This trustee alludes to a thread emerging during the conversations and interviews conducted for this present work: the Oversight Board's ongoing soul-searching. The Board's complicated institutional and internal structure, the public demands on it (some of which Meta and the Board itself drummed up), and the challenge of using private-law mechanisms to accommodate its own emulation of individual rights adjudication – at heart an intrinsically public function – complicates the Board's soul-searching. In short, controlling one of the largest experiments in (private) global governance while being an entirely private and relatively small institution is a conflicting, ambiguous, and perhaps impossible task. Again, the same trustee explained that:

> We're using the structures of private law to create an entity that's essentially responsible for establishing public law norms. We have to work out exactly how

[230] See for early community law Pavone (n 127). For EU law in general: Cohen and Vauchez (n 127); Vauchez (n 127). See further Yves Dezalay and Bryant G Garth, *Dealing with Virtue: International Commercial Arbitration and the Construction of a Transnational Legal Order* (University of Chicago Press 1998); Pauwelyn and Pelc (n 151).

[231] Interview with Oversight Board Trustee (n 53).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

these two distinct legal paradigms intersect. The role of the trustees is to guarantee the independence of the Board. That obligation is clear enough when it comes to actually selecting and deciding cases. But what about if the Board decides that to be independent it requires certain staffing levels or administrative assistance? How do we reconcile *that* kind of independence with our own fiduciary responsibility to the *res* of the trust? These kinds of issues raise complicated and unresolved questions. We're trying to feel our way towards solutions.[232]

Reflecting on the trust's role in setting down a procedural framework for the Board's external autonomy and internal fair-mindedness, we arrive at some (partly conflicting) conclusions. For a start, the trust constitutes a formal as well as informal link between the Oversight Board as an organization, its members as individuals, and Meta as a company. Communication exists and, all symbolic distance aside, it would be naïve to assume that the Board operates in isolation from the company. The issue is then whether such communication impairs the Board's external autonomy and internal fair-mindedness. Reasonably assessing the fair-mindedness of Board members (and staffers), if it were possible at all, would require an array of psychological surveys that is far beyond the scope of this book. However, if this author's subjective impressions based on the conducted interviews are of any value, a "pro-Meta bias" was not detected among the interviewed staffers or Board members. Some interviewees even described the Board's internal culture as expressly "adversarial" toward Meta,[233] while others compared it to a "human rights NGO."[234] At the same time, interviewees characterized the Board's impact as relatively limited and rejected the idea that it had any "supreme court"-like influence over Meta.[235] But they seemed confident that their adjudicatory practice, based largely on human rights and administrative law, could be valuable for the interests of users – and that they were not in their jobs just to do Meta a favor.[236]

This leaves us with the question of the Board's external autonomy. This autonomy would be impaired if the trust or the company could somehow pressure the Board to decide in a certain way. The trustees' fiduciary duty is to enable the Board to exercise "independent judgment."[237] Trustees

[232] Ibid.
[233] Interview with Oversight Board Staffer 5 (n 54).
[234] Interview with Oversight Board Staffer 6 (23 November 2022).
[235] Interview with Oversight Board Staffer 1 (n 54); Interview with Oversight Board Staffer 2 (n 10); Interview with Oversight Board Staffer 3 (n 11).
[236] Interview with Oversight Board Staffer 1 (n 54); Interview with Oversight Board Staffer 2 (n 10); Interview with Oversight Board Staffer 3 (n 11); Interview with Oversight Board Staffer 6 (n 234).
[237] Trust Agreement, Section 2.2.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

have no obligations vis-à-vis the company, and none of the interviewees suggested that the company exerts influence via the trustees. Further, baseline communication between adjudicator and controlled entities seems normal and does not necessarily impair external autonomy or internal fair-mindedness. Based on the limited impressions from the conducted interviews, therefore, the trust seems to work reasonably well as a buffer between company and Board.

In that context, it is important to keep a realistic perspective when it comes to how big a "threat" the Board really is for the company.[238] Apart from individual decisions on specific pieces of content, Board decisions are not binding. Arguably, and given the hundreds of millions of dollars Meta already invested, it would not make sense for Meta to weigh in the Board's decision-making process because the Board's main normative feedback to Meta, the nonbinding recommendations (see at Section 4.6.2), require Meta's involvement for them to be implemented anyway. If Meta policymakers want to circumvent the Board, it would be much easier to delay and tacitly thwart the implementation of the Board's decisions, recommendations, and policy advisory statements.[239] Direct meddling with the Board tends to result in disastrous publicity and thwarts the (assumed) aim of Meta to (also) use the Board as a publicly legitimizing formal structure.

Take as example an episode about Meta's withdrawal of a policy advisory opinion on the company's handling of hate speech in the context of Russia's war in Ukraine.[240] The company requested a policy advisory opinion from the Board, whose importance several interviewees described as equal to that of the Trump case. Yet, very late in the process, Meta reversed and withdrew its request. In an unusual press release, the Board subsequently criticized the company stating that it "believe[s] that the request raises important issues and ... [the Board is] disappointed by the company's decision to withdraw it." Press reactions to Meta's withdrawal were equally negative, some depicting the move as "undercutting" the Board's authority whose "only role is to handle the easy questions," prompting the question "why bother with it at all?" Neither charter nor bylaws explicate that Meta had the right to withdraw an already referred request for a policy advisory opinion. Article 2

[238] See below at 4.6 and Chapter 5.

[239] See at 4.6.

[240] cf. for the whole paragraph: Oversight Board, 'Protecting Freedom of Expression and Human Rights in Ukraine and Russia (Press Release)' (May 2022) <https://perma.cc/59Z2-WLKY>; Interviewees: Interview with Oversight Board Staffer 2 (n 10) 2; Interview with Oversight Board Staffer 3 (n 11); Interview with Oversight Board Staffer 4 (n 52) 4.

For the press coverage see: Casey Newton, 'How Facebook Undercut the Oversight Board' (*The Verge*, 5 December 2022) <https://perma.cc/H6RE-S57F>.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

Section 2.1.3 of the bylaws states that the company "may request" such statements and that the Board "may accept or reject" such requests.

4.3.3.3 Nominations and Removal of Members

Another factor relevant for independence are the nomination and removal procedures for Board members. A body consisting only of people hand-picked or waved through by the controlled entity raises concerns about independence. Moreover, adjudicators' personal sense of fair-mindedness requires that they should not worry about being sacked for any decisions they might make. The removal aspect does not appear particularly problematic. According to the charter and the bylaws, Board members can be prematurely removed from the Board only for just cause but not "due to the content decisions they have made."[241] Instead, removing Board members requires a "two-thirds vote of the Board (not counting the member(s) in question), subject to approval of the trustees . . .; and may be considered only for a violation of the code of conduct."[242]

The code of conduct regulates in detail how members and staff "must conduct themselves, both in their professional and personal capacities" which means "avoiding even the appearance of impropriety and flagging all potential conflicts through proper channels."[243] Therefore, members, staff, and their immediate family should not, for example, receive gifts, loans, or financial assistance on the basis of their position on or relationship with the Board.[244] Further, the idea of removal just because Meta wants a Board member to be removed appears unlikely. The trustees' fiduciary duty is to protect the Board's independence, not Meta's interests, and Board members themselves would hardly vote against their own institutional interests. Hence, the Board's removal regime protects Board members from removal by the company.

However, nomination procedures for members are not fully beyond Meta's grasp. Neither the charter nor the bylaws detail the process of nominating new Board members. Article 4 Section 2 of the bylaws stipulates that "trustees will formally appoint" Board members. None of the governing documents further specifies the nomination procedure except for allusions to the preferred professional qualities of potential nominees. According to an article by Steven Levy, Meta has wielded its influence over

[241] Charter Article 1 Section 8.
[242] Bylaws Article 1 Section 1.2.2.
[243] Code of Conduct, Introduction.
[244] Code of Conduct, Articles 1, 4.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

the Trust to prevent people its executives perceived as too critical from becoming Board members.[245] According to Levy, a candidate's previous and repeated criticisms of Meta resulted in some of its management (internally of course) voicing "reservations," which ultimately "put the kibosh on the nomination."[246] Such reporting is perhaps even more damaging for Meta than it is for the Board, as it emphasizes how reluctant the company is to cede authority to the Board and subject itself to truly independent external feedback. That is, however, neither surprising nor particularly worrying – authority of external guardians emerges incrementally, through continuous wrestling and discourse, rather than immediate submission by the controlled power structure.

4.3.3.4 Operational Autonomy and Budget

When focusing on its day-to-day work, the Board is simultaneously autonomous – as the company does not meddle with the Board's decision-making – and dependent on operational support by the company. In particular the Board's IT relies heavily on Meta's support. From a pragmatic perspective, such organizational semi-dependence is inevitable. Yet at times, it may be delicate as highlighted by severe disruptions of the Board's web infrastructure in 2021. For example, the bylaws state that "Facebook will maintain responsibility for the development and maintenance of tools that the Board uses to review and decide upon cases, including the Board's website and case management tool."[247] This web infrastructure is crucial for the Board to be able to receive user requests and handle its massive caseload. This "case management tool" refers to "the platform used by the Oversight Board to receive and review case submissions and collect and store case files."[248] What at first sounds mundane in fact generally shapes case selection. One staffer remembered that the case management tool malfunctioned when the Board started to accept the first appeals in early 2020.[249] Another staffer, in May 2022, mentioned that the case management tool, albeit working well overall, was still having occasional technical issues.[250] Therefore, organizationally, the Oversight Board, with its workforce of only seventy to ninety people,

[245] Levy (n 216) 14 et seq.
[246] Ibid 14.
[247] Bylaws Article 2 Section 1.3.2 (at the very end).
[248] Bylaws, Definitions "case management tool," at 34.
[249] Interview with Oversight Board Staffer 2 (n 10).
[250] Interview with Oversight Board Staffer 5 (n 54).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

remains a relatively small entity for controlling the rulemaking and rule enforcement of the world's largest social network, an organization with tens of thousands of employees. This necessarily leads to a considerable degree of dependence when it comes to day-to-day operations because the Board often requires Meta's engineers to provide, for example, IT support.

Lastly, the Board depends on Meta's contribution to the irrevocable trust.[251] The Board itself cannot raise any money.[252] As mentioned before, in 2019 Meta contributed 130 million dollars to the trust. Meta initially estimated that the first tranche would last six years, but in July 2022 the company announced putting another 150 million dollars into it.[253] Presumably, the Board spent more money than anticipated but this supposition could not be confirmed by any of the interviewees.

The trustees retain prerogative over all major spending decisions.[254] The trustees are independent of Meta and have a fiduciary duty to "facilitate the creation, funding, management, and oversight" of the Board.[255] Therefore, trustees have the power to invest in and sell property, establish the LLC, open accounts and exercise transactions, enter into contracts, and anything else necessary to keep the Board running.[256] To date, the biggest investment seem to have been the fifteen million seed funding for the newly founded Appeals Centre Europe – an Oversight Board spin off that was certified as an ODS under Article 21 DSA in late summer 2024 and might become an independent funding stream for the Board.[257] However, forming additional "companies for the purpose of effectuating the purpose" of the trust requires Meta's consent.[258] Even though Meta initially conceived of and initiated the Board as a six-year experiment, the early recapitalization of the trust in July 2022 is a sign of the Board's financial independence for the foreseeable future. However, in line with 2023 cost-cutting measures at Meta, the company appeared poised to gradually limit the Board's financial basis, prompting the Board to moderately consolidate in 2024.[259]

[251] Trust Agreement Section 1.4.
[252] Although the Trust an LLC agreement allows the LLC to be commercially active.
[253] Oversight Board, 'Securing Ongoing Funding for the Oversight Board (Press Release)' (n 114).
[254] Trust Agreement Sections 2.1, 2.3, and 4.
[255] Trust agreement Sections 2.1 and 6.8.4.
[256] Trust Agreement Section 4.
[257] See at 3.8.1.3.
[258] Trust Agreement Section 4.7.
[259] Naomi Nix, 'Meta's Oversight Body Prepares to Lay off Workers' *Washington Post* (29 April 2024) <https://perma.cc/Z8SS-MKSW>.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

4.4 Rules

This section introduces the different normative layers governing content moderation and the Board's adjudication. We can differentiate between internal rules devised by Meta and external international human rights law. By way of context, the Board has so far not substantively interacted with EU law. The key takeaway of this section has two parts. We find out that the practically relevant internal rules of content moderation are, in contrast to what is often argued, not the terms of service but a variety of subsidiary rules that detail the normative regime. Then, we also discuss Article 19 of the ICCPR, which provides the Board with a normative scaffolding to evaluate what Meta is doing.

4.4.1 Internal Rules: Values and Content Policies

The initially intended normative focus of the Board were the rules that Meta (or Facebook) had unilaterally and internally devised. These were a set of "values," and Meta's "content policies" which are terms of service and, crucially, the huge list of rules governing user communication known as "community standards."

4.4.1.1 Meta's "Values"

In early 2018, when the specific shape of the Oversight Board was still unclear, Meta executives considered setting up the company's "own quasi-legal system where clearly stated policy rules can be applied to the plethora of real-world expression challenges."[260] The normative core (a privatized Kelsenian *Grundnorm* so to speak), should have been an "iconic, one-sentence values commitment Facebook would adopt."[261] To a degree, the Board's foundational documents still reflect this idea, stating that the Board "will review content enforcement decisions and determine whether they were consistent with Facebook's . . . values."[262] However, identifying these values is not that simple. Neither the charter nor the bylaws explain what these values are. Instead, both charter and bylaws premise that Meta "articulate" them.[263] There is no public process or document of perceptible legal quality establishing what these values ought to be. The only source naming and detailing the values is a blog post by Meta's then Vice President for Global Policy Management

[260] Feldman (n 40) 101.
[261] Feldman (n 33) 104.
[262] Charter Article 2 Section 2.
[263] Charter Article 1 Section 4 No 3.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

on the company's own public relations blog.[264] The executive writes that the company was "expanding the values that serve as the basis for [its] decision-making."[265] According to this post, the "paramount" value is "voice," whereas other values – "authenticity," "safety," "privacy," and "dignity" – might limit the free expression of that "voice."[266] The text goes on to briefly explain what these values mean. "Voice" relates to a far-reaching, US-informed interpretation of free expression.[267] Of course, voice and expression on Meta are exercised by way of users posting content, which is the basis of Meta's business. "Authenticity" means ensuring that users do not "misrepresent who they are or what they're doing."[268] "Safety" prohibits intimidating other users, as well as the dangers of excluding, or silencing content.[269] "Privacy" aims to "give people the freedom to be themselves," while "dignity" refers to the executive's assertion that Meta "believe[s] that all people are equal in dignity and rights."[270]

In short, various Meta employees' own (unexamined) definitions of "values" should serve as the primary normative framework for the Board's assessment of Meta's rule enforcement and arguably also its rulemaking.[271] Given the vagueness of their definitions, as well as the limited public visibility and legitimacy of those expounding the ideals upon which these "values" are purportedly based, "values" cited in a blog hardly provide appropriate normative material to control what Meta is doing.

4.4.1.2 Content Policies

As argued above, the Board is supposed to assess whether Meta's enforcement actions comply with its "content policies." The charter mentions that "Facebook's Community Standards and other relevant policies" are "collectively referred to as content 'policies.'"[272] These are unilaterally devised rules that manifest in at least three layers.

[264] Monika Bickert, 'Updating the Values That Inform Our Community Standards' (*Meta Newsroom*, 12 September 2019) <https://perma.cc/6TPQ-MW35>.
[265] Ibid.
[266] Ibid.
[267] Ibid.
[268] Ibid.
[269] Ibid.
[270] Ibid.
[271] Douek, 'Why Facebook's "Values" Update Matters' (n 102).
[272] Charter Article 1 Section 4 No 2.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

Terms of Service

The top layer of social media companies' rules are their terms of service. These represent nothing less than the contractual basis governing the relationship between Meta and the users of its services. The terms of service are relatively brief and do not regulate user communication substantively. When it comes to content moderation, the terms of service only establish that users must not violate the "community standards" and "other terms and policies." These community standards and other terms and policies make up a second layer of rules. These are a variety of normative texts, appear under different titles, and are not immediately obvious as a part of the terms of service. Instead, the terms of service only reference the community standards and other rules. Therefore, in general, the actual terms of service are only germane because they signpost the much more decisive second normative layer, which are the community standards (and the other, unspecified, rules).

Community Standards and Other Policies

The decisive normative layer governing online communication in Europe and beyond consists of unilateral rules devised by platforms detailing what communication is allowed on a given platform. These community standards are binding with respect to users because the platforms' terms of service (which the user has to agree to) reference them. In the case of Meta, community standards claim to "apply to everyone, around the world, and to all types of content."[273] They are split by topic into more than twenty different "standards," dealing with issues such as "hate speech" to "bullying and harassment" to "cybersecurity."[274] Each standard has a "policy rationale" which outlines the aim of the specific community standard. The community standards – with varying concision and specificity – then provide definitions and examples of prohibited communication acts. To illustrate this, we can look at Facebook's community standard covering hate speech. As of November 2024, almost thirty updated versions of the standard were published on Meta's website. The first hate-speech community standard published by Facebook dates to May 25, 2018.[275] Since then, Meta has amended the hate-speech standard roughly every few months.[276] The hate-speech standard's "policy rationale" states that people should be able to "use their voice and connect more freely when they don't

[273] See Meta, Community Standards accessible via <https://perma.cc/6PBT-KJSD>.
[274] Ibid.
[275] Ibid.
[276] See how this might complicate the EU's rationale behind Article 21 DSA at 3.5 and 3.7.4.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

feel attacked on the basis of who they are." There then follows a "definition" of hate speech. Facebook defines this as

> a direct attack against people – rather than concepts or institutions—on the basis of what we call protected characteristics: race, ethnicity, national origin, disability, religious affiliation, caste, sexual orientation, sex, gender identity and serious disease. We define attacks as violent or dehumanizing speech, harmful stereotypes, statements of inferiority, expressions of contempt, disgust or dismissal, cursing and calls for exclusion or segregation.[277]

The practical complexity of such a definition is obvious. How to draw a line between "direct" and "indirect" attacks against people? How to handle necessarily subjective, value-laden, and context-dependent adjectives like "violent," "harmful," or "dehumanizing"? Which diseases are "serious"? Does only the content of a communication act make it "hate speech" or also the attitudinal stance(s) of the person who expresses it? What about the temporal and territorial space in which an utterance was made?[278] Are we even capable of identifying the past, current, and future social, political, and cultural contexts of an utterance with sufficient certainty? What about contexts that are culturally foreign to Meta's predominantly North American rule makers? Interviewed practitioners described content moderation as "phenomenally difficult,"[279] and scholars[280] and tech journalists[281]

[277] See Community Standard Hate Speech, accessible via <https://perma.cc/GGB3-XST8>. Loosening this standard was one of Zuckerberg's most significant public gestures of acquiescence to the incoming Trump administration in January 2025. However, the first sentence's original wording remained in place until the time of this manuscript's revisions in late February 2025, with the arguably largely symbolic change of renaming "hate speech" to "hateful content." In practice, the specifics of the standard were scaled back, and the already somewhat convoluted definition of what constitutes an attack appears to have been pared down. Ultimately, the key issue is not the exact wording of these standards but how Facebook enforces and implements them. Changing the wording, in turn, is a process that, like many other aspects examined in this book, may remain largely ceremonial.

[278] In essence, these questions point to one of the central aporias of the philosophy of language. See Judith Butler's critique of J. L. Austin's "total situation" approach. In his foundational work 'How to do things with words' Austin argued that context – meaning mostly spatial and temporal aspects – would, if "totally comprehended" arguably allow for an absolute interpretation of certain speech acts. Butler disagrees. cf. Judith Butler, *Excitable Speech: A Politics of the Performative* (Routledge 1997) 2–13. See further John L Austin, *How to Do Things with Words* (Oxford University Press 1962); John R Searle, *Speech Acts: An Essay in the Philosophy of Language* (Cambridge University Press 1969). For a US law perspective see especially Robert Post, 'Participatory Democracy and Free Speech' (2011) 97 Virginia Law Review 477.

[279] Interview with Oversight Board Staffer 3 (n 11).

[280] Douek, 'Content Moderation as Systems Thinking' (n 7) 533, 568.

[281] Mike Masnick, 'Masnick's Impossibility Theorem: Content Moderation At Scale Is Impossible To Do Well' (*TechDirt*, 20 November 2019) <https://perma.cc/Q97F-CFHB>.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

have even pronounced content moderation an impossible endeavor. Platforms' increasing reliance on automated moderation (despite the well-known shortcomings of natural-language processing)[282] only exacerbates these difficulties.

To drain the seemingly endless swamp of possible contextual layers, each community standard therefore provides further, generalized explanations. These general "explanations" often consist of normative statements relating to those whom the company wishes to protect. For example:

> We also prohibit the use of harmful stereotypes, which we define as dehumanizing comparisons that have historically been used to attack, intimidate, or exclude specific groups, and that are often linked with offline violence.[283]

The sections then elaborate – in a section entitled "do not post" – on specific examples of the types of content that meet the above-mentioned definitions. The "do-not-post" content subdivides into several "tiers." A "tier 1" attack includes, for example,

> (d)ehumanizing speech or imagery in the form of comparisons, generalizations, or unqualified behavioral statements . . . to or about . . . insects, animals that are culturally perceived as intellectually or physically inferior, . . . sub-humanity.

In sum, Meta's community standards reflect the reactive and incremental nature of their creation. They are lengthy, convoluted, sometimes contradictory, and highly subjective. As one interviewee recalls, "[t]he process for making content rules is basically a room full of US lawyers sits down and makes the rules after the latest crisis."[284] Accordingly, "ad-hockery," imprecise terminology, and structural deviations between (and within) standards characterize the resulting "community standards." Simply put, if you want to take down a communication act (usually a post), you are likely to find a community standard that prohibits this post. Yet most standards and their cascading "explanations" are crammed with adjectives and imprecise terminology that can be interpreted in many ways. To an extent, such normative uncertainty is inevitable – especially if you are aiming to regulate the communications of two or more billion people. However, like any

[282] Robert Gorwa, Reuben Binns and Christian Katzenbach, 'Algorithmic Content Moderation: Technical and Political Challenges in the Automation of Platform Governance' (2020) 7 Big Data & Society 1. See also Paul Friedl, 'Dis/Similarities in the Design and Development of Legal and Algorithmic Normative Systems: The Case of Perspective API' (2023) 15 Law, Innovation and Technology 25.

[283] See <https://perma.cc/2YRS-K46C>.

[284] Interview with Academic, Consultant to Facebook in 2018 (n 31); see further Klonick, 'The New Governors: The People, Rules, and Processes Governing Online Speech' (n 19).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

sanctioning approach, restricting "speech" on social media requires a minimum level of normative foreseeability and consistency if it is not to be totally arbitrary. We will re-encounter these deeper levels of platform rulemaking in Chapter 6 because they point to one practical shortcoming of many scholarly or regulatory takes on reforming content moderation.[285] Just infusing fundamental rights, which are broad normative categories, without specifying how this affects these sublayers of specific rules, will not solve content moderations' normative shortcomings.

4.4.2 External Rules: International Human Rights Law

The third set of norms the Board uses to control Meta is international human rights law. Meta is a private company and therefore cannot, of course, be a signatory of international human rights treaties like the ICCPR. Furthermore, despite scholarly criticism, international law does not provide for the immediate application of human rights law to private actors even if they are powerful transnational companies.[286] This section briefly outlines how the charter largely omits human rights, introduces the UN Guiding Principles as the tool that makes international human rights law relevant for corporations like Meta, and introduces the ICCPR, and especially Article 19, as the normative fulcrum that underpins the Board's reasoning.

4.4.2.1 The Charter's (Superficial) Reference to Human Rights

The charter only superficially references human rights. In Article 2 Section 2 – entitled "basis for decision-making" – the charter explains that "the Board will review content enforcement decisions and determine whether they were consistent with Facebook's content policies and values." In other words, the charter implicitly excludes human rights as a normative standard for the Board to "review" actions by the company.[287] Instead, it merely stipulates that "[w]hen reviewing decisions, the Board will pay

[285] See at 6.1.2.

[286] For recent criticism see from a general perspective the late Antonio Cassese (ed.), 'Introduction' in *Realizing Utopia* (Oxford University Press 2012) xix. See specifically Andrew Clapham, *Human Rights Obligations of Non-State Actors* (Oxford University Press 2006) 195; Marco Fasciglione, 'A Binding Instrument on Business and Human Rights as a Source of International Obligations for Private Companies: Utopia or Reality?' in Martina Buscemi and others (eds.), *Legal Sources in Business and Human Rights* (Brill Nijhoff 2020). For the term "transnational corporation" see Clapham 199–201.

[287] In that sense also Klonick, who considers this "far from a broad or robust adoption of human rights," Klonick, 'The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression' (n 3) 2478.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

particular attention to the *impact* of removing content in light of human rights norms protecting free expression."[288] The other relevant documents, especially the bylaws, only partially reproduce the charter's wording regarding the community standards while entirely omitting human rights law in the context of case review.[289] Therefore, Meta's initial vision for the Board did not include a prominent role for anything but rules made by Meta executives themselves, especially not for international human rights law.[290] Instead, at least one strong idea in the creation process was to base the Oversight Board's control entirely, or at least predominantly, on rules and "values" devised by the company itself. Hence, just as with Article 21 of the DSA's review of the criteria governing "incompatibility," Meta's idea for the Oversight Board was to create an adjudicator that would, however, only apply Meta's internally devised rules to control the company.[291]

4.4.2.2 UN Guiding Principles on Business and Human Rights

Although international human rights law is not binding for private companies and even though the Board's foundational documents do not grant it a prominent place, normative bridges between international human rights law and bodies like the Oversight Board do exist. The main bridge connecting corporations and international human rights are the *United Nations Guiding Principles for Business and Human Rights* (UNGP).[292] According to its author, UN Special Representative John Ruggie, this UN normative framework rests, apart from states' (and the EU's) duty to protect individuals from human rights abuses by corporations, largely on a voluntary "responsibility" of corporations "to respect human rights" as well as on "remediation through legitimate processes."[293] Also, the UNGP demand "judicial and non-judicial remedies" by states.[294] Further, they stipulate that "[i]ndustry . . . should ensure

[288] Charter Article 2 (2) (emphasis added).
[289] cf. Bylaws Article 1 Section 3.
[290] See at 4.2 and especially the Feldman memo referenced in n 33.
[291] ODSs review content moderation on grounds of "illegality," which is a normative framework external to platforms, and "incompatibility" with platforms' terms of service, which remain an internal normative framework. See in detail at 3.5.
[292] For an introduction see Radu Mares, 'Business and Human Rights After Ruggie: Foundations, the Art of Simplification and the Imperative of Cumulative Progress' in Radu Mares (ed.), *The UN Guiding Principles on Business and Human Rights* (Brill Nijhoff 2012) 1–49.
[293] UN Guiding Principles para 6 and Principle 22.
[294] UN Guiding Principles para 6.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

that effective grievance mechanisms are available."[295] However, as betrayed by its title, the UN's "principles" in this context are "guiding," not legally binding. They also tend often to be vaguely formulated, asserting for example that "[b]usiness enterprises should respect human rights. . . . [which] refers to internationally recognized human rights."[296] How stringently corporations are expected to honor their "responsibility to respect" in this UN framework is not a question of legal accountability but of managerial competence and operational feasibility against the backdrop of the profit motive.

4.4.2.3 The ICCPR, Especially Article 19

Among these internationally recognized human rights mentioned in the UNGP is the ICCPR. The ICCPR contains several human right provisions. Crucial for our context is Article 19, which establishes an individual right to freedom of expression. A close reading paired with a word-count analysis of the Board's decisions taken up to November 2024 shows that Article 19 of the ICCPR – and not Meta's values or content policies – is the prime norm-oriented focus of the Board. The justification requirements of Article 19 of the ICCPR inform the Board's core normative considerations. The decisive parts of Article 19 of the ICCPR read:

> 2. Everyone shall have the right to freedom of expression; . . .
> 3. The exercise of the rights provided for in paragraph 2 of this article carries with it special duties and responsibilities. It may therefore be subject to certain restrictions, but these shall only be such as are provided by law and are necessary:
>
> (a) For respect of the rights or reputations of others;
> (b) For the protection of national security or of public order (ordre public), or of public health or morals.

Based on these paragraphs, the Oversight Board came up with a three-pronged test for Meta's compliance with its "human rights responsibilities," as the Board refers to them. The subject of this test can be either Meta's rule enforcement (for example, because it deleted a post), or Meta's rulemaking (because it prohibited a post). Often, both aspects are tested together, lacking a formal distinction.[297] The three prongs are: legality; legitimate aim; and necessity and proportionality. While the "legality"

[295] UN Guiding Principles 30.
[296] UN Guiding Principles 11 and 12.
[297] At least a formal distinction that would show in a specific title, header, or marking.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

section mainly tests the quality of Meta's rulemaking, the "necessity and proportionality" section often mixes rulemaking and rule enforcement.

4.4.3 *Detached from EU Law*

Despite its engagement with international human rights law, the Board does not engage with regional human or fundamental rights regimes. Notably, neither does it reference or apply the EU's Charter of Fundamental Rights (CFR) to cases dealing with content posted within the bloc or by users residing in the EU.

In principle, the CFR binds the EU and Member States when applying EU laws.[298] In principle, it does not bind corporations. However, the European Court of Justice (ECJ) applies several CFR provisions and other guarantees that could be understood as fundamental rights – such as nondiscrimination and equal pay – to private actors.[299] Further, various member-state domestic courts have regularly infused the principle of fundamental rights into their assessment of relevant contractual relations, subjecting the corporation then known as Facebook to due-process requirements.[300] As we saw in the previous chapter, at least one newly certified ODS in the EU outrightly applies the charter between social media platforms and individuals.[301] This descriptive part is not the place to discuss whether the Board might be under a legal obligation to apply EU law, and particularly the CFR, in cases legally located in the EU. For the purposes of this section, it is sufficient to note that the Oversight Board operates entirely detached from the normative layers of supranational or domestic law. For example, cases "where the underlying content has already been blocked, following the receipt of a valid report of illegality, and not removed for Community Standards violation" are inadmissible. In other words, the Board does not aim to interfere with regimes like Germany's now defunct Network Enforcement Act (*NetzDG*) or the DSA. Rather it aims to operate free from such legal constraints, using only Meta's own standards as well as

[298] Article 51 CFR.

[299] See e.g. Defrenne v Belgian State [1971] European Court of Justice C-80/70; Defrenne v SABENA [1976] European Court of Justice C-43/75; Egenberger [2018] European Court of Justice C-414/16; Stadionverbot [2018] Federal Constitutional Court (Bundeserfassungsgericht) 1 BvR 3080/09, BVerfGE 148, 267–290.

[300] See e.g. Hassrede [2021] Federal Court of Justice (Bundesgerichtshof) III ZR 179/20 and III ZR 192/20.

[301] See at 3.8.1.2.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

international law. From the global perspective in which Meta operates, such an approach makes sense, especially for a private company. From an economic standpoint, diversifying standards for different regions across the world would make Meta's governance structure more costly.[302] But it also means that the Board's norm-based reasoning does not reflect regional (specifically EU) legal frameworks.

4.5 Decisions

After the previous section introduced the normative layers guiding the Board's decision-making – Meta's values and content policies as well as international human rights law – this section investigates how the Oversight Board applies these rules in its decisions. To do so, the section presents qualitative insights as well as a quantitative analysis of the word count in different parts of Board decisions. To anticipate: almost all Board decisions follow the same overall structure for their different normative layers as well as, where applicable, within the human rights assessment it produces. This allows us to compare how much the Board writes and argues on certain aspects by counting and comparing the word distribution among these sections. Word count does not necessarily equate to the quality or "importance" of a text's normative reasoning. However, insights from interviews and a close reading of Board decisions indicate that international human rights law, and especially Article 19 of the ICCPR, is *the* normative focus of the Board. Thus, a thorough understanding of how many words the Board allots to analyzing Article 19 of the ICCPR, in comparison to other normative phenomena like values or community standards, might hint at the relevance of human rights law for the Board's practice.

After a brief introduction to the general layout of Board decisions, this section then engages with some of the Board's reasoning to gain an impression of how the Board uses human rights thinking in its deliberations. After outlining how using word count as the basis for a quantitative analysis may corroborate these qualitative insights and statements made by interviewees, the section proceeds to examine how

[302] Such "nondivisibility" of business operations is one of the key enabling factors for the so-called Brussels Effect, see Anu Bradford, *The Brussels Effect: How the European Union Rules the World* (Oxford University Press 2020) 25 et seq. For a critical review whether the DSA may really unfold such a global effect see Moritz Schramm, 'Curb Your Enthusiasm: Why the Digital Services Act Might Not Have a "Brussels Effect"' (2025) 16 European Journal of Legal Studies (Special Issue) 61.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

much the Board writes in terms of the three normative layers – content policies, values, and human rights – as well as in the three subsections within the human rights assessment. This provides us with a quantitative basis upon which to draw several conclusions.

4.5.1 What Board Decisions Look Like

The Board published its first decisions on January 28, 2021. Since then, it has issued more than 100 decisions (as of November 2024). Initially, all decisions followed the same procedure and textual structure. However, after some criticism that the Board was too hesitant and slow, the Board began issuing three different types of decisions. These three types are "standard," "summary," and "expedited." Standard decisions are in-depth reviews of specific and often systemic issues, whereas the function of summary decisions largely seems to be to check on the implementation of prior standard decisions. Expedited decisions are relatively rare and offer fast-tracked procedures in situations of imminent real-world harm, for example regarding videos of abducted Israelis in the context of the October 7 attacks or the aftermath of an Israeli strike on Al-Shifa hospital in Gaza. Whereas summary decisions offer only cursory analysis and reasoning, expedited decisions go into considerable detail. Both do not include a public comment section. The following largely focuses on standard decisions. Until the time of writing in November 2024, the Board issued sixty-eight standard decisions (one of which was essentially empty as the user withdrew their appeal, so it is effectively sixty-seven decisions).

All decisions are published on the Board's website.[303] They are published in English and, at least in standard and expedited decisions, the language(s) of the original dispute. The texts have a serial number and a title.[304] Titles are crisp and reference the subject matter of the case, for example "Former President Trump's suspension" when the Board decided whether Meta was right to indefinitely suspend Trump's accounts from its platforms after his role in the 6 January 2021 riots.[305] Other titles point to the wide range of topics and contexts the Board deals with, for instance "Alleged crimes in Raya Kobo" (about communication related to ethnic tensions in Ethiopia), "Colombian protests" (about criticisms of the Colombian President),

[303] Accessible via: <https://perma.cc/BK4R-ZYUD>.

[304] All cases following the year-number-platform-type of referral pattern, e.g., 2021-001-FB-FBR, which stands for first case of 2021 on Facebook (FB) which was referred by Facebook (FBR). Further, all decisions receive a name, e.g. Öcalan's Isolation [2021] Oversight Board 2021-006-IG-UA.

[305] Former President Trump's suspension (n 143).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

"Mention of the Taliban in news reporting" (covering reporting on terrorist regimes), "Claimed COVID cure" (relating to misinformation about the Covid-19 pandemic in France), or "Cambodian Prime Minister" (when the Board recommended to Meta to suspend the account of Hun Sen).

All standard decisions contain roughly the same sections, typically in the same sequence, with slight variations over time. Usually, decisions begin with a case summary and a summary of the Board's key findings. The decision's text is separated into several parts, typically nine or ten. These are, a decision summary, a description of the case, a few remarks on the authority and scope of the Board's review, and a description of the relevant normative standards. In this part, the Board typically cites the relevant paragraphs of the community standard and sets out the human rights-related norms it sees as pertinent to the case. Here, the Board does not engage in substantive reasoning. Further, the decisions summarize the statements made by the user in their appeal, as well as Meta's explanation for its decision, and submissions by third parties. Then, the Board engages in a normative assessment called "Oversight Board Analysis." In earlier decisions this was always part eight and since mid-2024 part five. This is the only part of the decision where the Board assesses facts, weighs different interpretations, and engages in normative reasoning. Standard decisions usually conclude with one or two brief parts consisting of the Board's nonbinding recommendations and its binding decision. These last two parts are sometimes combined.

Two things are noteworthy about this structure. To begin with, it is quite striking that the Board rather strictly adheres to it. All sixty-seven standard decisions published at the time of this writing follow that structure, with only minor deviations along the way. That a transnational body – one that deliberates via video call and whose decision-makers are spread across five continents – can commit to such a uniform working method already tells us something about the Board's strategic vision and discipline. Further, the sequencing of the parts appears to be deliberate. Outlining facts and normative standards before engaging in normative reasoning to arrive at a conclusion consisting of a decision and several recommendations implies that the Board members see themselves – or at least want to present themselves – as appliers rather than creators of the law. This viewpoint is reminiscent of Montesquieu's famous observation that "*Le juge est le bouche de la loi,*" and mirrors the practice of many international adjudicators.[306] Setting out a decision only at the very end of a twenty- to thirty-page

[306] Nußberger implicitly made a similar argument about the European Court of Human Rights (ECtHR)'s "pedantic" practice to spell out the entire text of each disputed charter

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

document resembles the style of decision-framing known from practitioners at the European Court of Justice, the International Court of Justice, and the European Court of Human Rights but also EU Boards of Appeals. In contrast, domestic courts like those in Germany or the United States begin their judgments by presenting the decision; they only explain it later.

In its decisions, mostly within the "Oversight Board Analysis" part, the Board engages with various normative materials. These are, international covenants and authoritative commentary, most notably the ICCPR and the UN Human Rights Committee's General Comments.[307] Further, the Board quotes decisions of international courts (especially the European Court of Human Rights),[308] the US Supreme Court,[309] and legal scholars.[310] It engages with decisions and other normative material produced by a wide range of adjudicators, including the UN Human Rights Committee[311] and the Committee on the Elimination of Racial Discrimination (CERD) Committee.[312] In particular, the Board's legality and proportionality assessments engage in some detail with the acquis of international human rights law, most notably Article 19 of the ICCPR.[313] In several decisions the Board reiterated criticism of the company's rule-making and rule enforcement.

provision, see Angelika Nußberger, 'Wer Zitiert Wen? Zur Funktion von Zitaten Bei Der Herausbildung Gemeineuropäischen Verfassungsrechts' (2006) 61 JuristenZeitung 763, 763.

[307] See Former President Trump's suspension (n 89) at 8.3, Depiction of Zwarte Piet [2021] Oversight Board 2021-002-FB-UA at 8.3, Pro-Navalny Protests in Russia (n 81) at 8.3 to general comment, Two Buttons Meme [2021] Oversight Board 2021-005-FB-UA at 8.3 to general comment, Öcalan's Isolation [2021] Oversight Board 2021-006-IG-UA at 8.3, COVID lockdowns in Brazil [2021] Oversight Board 2021-008-FB-FBR at 8.3, Colombia Protests [2021] Oversight Board 2021-010-FB-UA at 8.3, South Africa slurs [2021] Oversight Board 2021-011-FB-UA at 8.3, Wampum Belt [2021] Oversight Board 2021-012-FB-UA at 8.3, Ayahuasca brew [2021] Oversight Board 2021-013-IG-UA, Knin Cartoon [2022] Oversight Board 2022-001-FB-UA at 8.3, Colombian police cartoon [2022] Oversight Board 2022-004-FB-UA at 8.3.

[308] See e.g. Depiction of Zwarte Piet (n 95) at 8.3, Two Buttons Meme (n 95), Ayahuasca brew (n 95) at 8.2. In Alleged Crimes in Raya Kobo [2021] Oversight Board 2021-014-FB-UA at 8.3 III the Board quoted the International Criminal Tribunal for Rwanda.

[309] See Ayahuasca brew (n 95) at 8.1 and 8.2.

[310] See Pro-Navalny Protests in Russia (n 81) at 8.1.

[311] See in Claimed COVID cure [2021] Oversight Board 2020-006-FB-FBR at 8.3, Colombia Protests (n 95) at 8.3. II, Knin Cartoon (n 95) at 8.3. II.

[312] Depiction of Zwarte Piet (n 307).at 8.3 (to decisions), South Africa slurs (n 95) at 8.3 to recommendation and reports, Wampum Belt (n 95) at 8.3 to general recommendations.

[313] See in detail below at 4.5.5.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

4.5.2 *Tracing the Board's Turn to Human Rights*

Several interviewees pointed to human rights as determinative normative category in their day-to-day work.[314] Many staffers and several Board members have a background in human rights, and the Board's first Director, Thomas Hughes, previously worked for an NGO literally called "Article 19."[315] Hughes would later run the Oversight Board spin off, the Appeals Centre, in Dublin.[316] Because Meta's community standards and values are so convoluted and vaguely worded, they are of little help in deciding cases. Therefore, what began as a mission to control the enforcement of rules quickly faced the reality that, not only was enforcement problematic, its underlying rules were not fit for purpose as well. To enhance the quality of Meta's rules, especially its community standards, the Board would have to resort to *external* standards. Especially from a legitimation perspective, citing and thereby tying oneself to international law autonomizes and legitimizes the Board vis-à-vis the company as well as in relation to national regulators. From the Board's perspective, this hits two birds with one stone. One interviewed Board member described the system in this way:

> [A] purely formalist review system is meaningless. Yes, there are important things to do in making sure that Facebook is applying its own rules correctly and consistently, but the big issues are substantive. . . . We need some sort of external standards here. . . . Assuming we're not dealing with the legal and we don't have any democratic footing, it seems that human rights is the best bet . . . for . . . a scope of externally justifiable values that have received a lot of consensus-based legitimacy and been drawn out over 50 years.[317]

A close reading of the Board's decisions corroborates important aspects of this member's claim. The Board confines most of its argumentation, reasoning, and citing of other adjudicatory material – for example, by referring to material from the UN Human Rights Committee or the European Court of Human Rights, other texts, such as reports by UN Special Rapporteurs, or the "Rabat Plan of Action," and the arguments of legal scholars – to the "Oversight Board Analysis." In one subsection of that "Analysis" section, the Board analyzes whether Meta is complying

[314] Interview with Oversight Board Member (15 November 2022); Interview with Oversight Board Staffer 1 (n 54); Interview with Oversight Board Staffer 2 (n 10); Interview with Oversight Board Staffer 4 (n 52).
[315] Interview with Oversight Board Staffer 2 (n 10).
[316] See at 3.8, esp. 3.8.1.3.
[317] Interview with Oversight Board Member (n 314).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

with its human rights responsibilities in the case at hand.[318] Virtually every standard decision features this human rights-focused subsection. Citing other adjudicative decisions, primary and secondary legal material as well as scholarship are key ingredients of adjudicative reasoning.[319] This material situates and thereby legitimizes normative utterances by an adjudicator in the more general system comprising the norms, applications, institutions, and practices of legality.[320] The Board also sometimes discloses internal disagreement in this human rights section. Several decisions feature passages contrasting majority and minority positions among Board members in the Board's human rights assessment.[321] For example, in a decision concerning communication in the context of the Azerbaijani-Armenian conflict, the Board stated:

> The *majority* noted that the post, when read as a whole, made clear the user's choice of slur was not incidental but central to the user's argument that the target group was inferior. Moreover, the post in question was widely disseminated at the height of an armed conflict between the user's State and the State whose nationals the post attacked. . . . A *minority* found Facebook's deletion of the post was not proportionate, on the basis that the risks cited by the majority were too remote, and were not foreseeable.[322]

[318] For quotes on ECtHR see Two Buttons Meme (n 307); Öcalan's Isolation (n 304); Ayahuasca brew (n 307). For quotes on scholars see Pro-Navalny Protests in Russia (n 185).

However, this trend is not absolute. For example, in one decision the Board distinguished the case at hand from court cases cited by Meta, including from the US Supreme Court in the section pertaining to Meta's values. Yet, this deviation may even corroborate the argument that the Board focuses more on human rights, whereas the company looks primarily to its own rules. Arguably, the Board only engages with the US Supreme Court in that particular section because Meta referred to it. In contrast, when the Board argues on its own terms, argumentation and court-citation takes place in the human rights section.

[319] David J Walsh, 'On the Meaning and Pattern of Legal Citations: Evidence from State Wrongful Discharge Precedent Cases' (1997) 31 Law & Society Review 337, 398. See for a transnational, qualitative perspective differentiating between "domestic," "transnational European," and entirely "foreign" citations: Nußberger (n 306).

[320] For reflections on the normativity of applying norms see Möllers (n 149) 107–114.

[321] See vividly COVID lockdowns in Brazil (n 178) 8.3.III presenting alternating arguments by the majority and minority, ultimately asserting that "[t]he majority understands the minority's view but respectfully disagrees with it." See further Colombia Protests [2021] Oversight Board 2021-010-FB-UA, were, in a case concerning protests in Colombia, the minority argued that apart from Article 19 ICCPR the right to freedom of assembly, Article 21 ICCPR, should have been assessed in detail. See further Protest in India against France [2021] Oversight Board 2020-007-FB-FBR; Former President Trump's suspension (n 143); Depiction of Zwarte Piet (n 307); Two Buttons Meme (n 307); Alleged Crimes in Raya Kobo (n 308).

[322] Armenians in Azerbaijan [2021] Oversight Board 2020-003-FB-UA at section 8.3(c) (emphasis added).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

Contrasting minority and majority positions within a multiperson adjudicator has several effects.[323] It might delegitimize the adjudicator as its decisions could then be perceived as the outcome of bargaining or as a subjective interpretation and not as a clear evidence-based consequence of the interpretation of a particular legal text. On the upside, it could be argued that admitting to norm-related disagreements among adjudicators paints a more realistic picture of their actual deliberations and offers a broader variety of perspectives on the decision. This might bolster an adjudicator's legitimacy as it increases transparency and presents adjudication as what it is: a process of reasoning that will usually allow for more than one conclusion.

In contrast, the Board only occasionally engages with Meta's values. For example, in its first closely reasoned decision, the Board wrote 320 words on Meta's compliance with its own community standards, 691 words on compliance with international human rights law, and a meager forty-six words regarding compliance with Facebook values.[324]

4.5.3 *Introducing Word Count as a Quantifiable Unit*

But how does the Board reason, and on what issues? And how can this be studied objectively? This section draws on several quantitative analyses of Board decisions to show that the Board increasingly grounds its reasoning in international human rights law—rather than Meta's self-made rules—and subjects Meta's rulemaking itself to growing scrutiny.

Based on the interviews, a close reading of the decisions, and academic works looking at the value of argumentation and citation in adjudicative texts, we can assume that the Board writes most about what it considers normatively important or otherwise relevant.[325] Writing more will increase word count. Hence, within the Board's normative reasoning (which is always confined to the "Oversight Board Analysis" part), each section's word count, that is, that of the section on Meta's community standards, the company's "values," or crucially, international human rights law, may serve as a proxy indicator for that section's importance.[326] Hence, analyzing the word count

[323] Disclosing minority opinions in judgments is not the same than dissenting opinions, which are much more detailed. See further Ruth Bader Ginsburg, 'The Role of Dissenting Opinions' (2010) 95 Minnesota Law Review 1; Steven Peterson, 'Dissent in American Courts' (1981) 43 Journal of Politics 312.

[324] Myanmar post about Muslims [2021] Oversight Board 2020-002-FB-UA. The very first decision, 2020-001-FB-UA, was moot as the user deleted the post before the Board decided the case.

[325] For more details underlying the methodology of the word-count analysis, see at 1.4.3.

[326] Note: A similar claim could not be made regarding the comparison of the argumentative part ("Oversight Board Analysis") and the more descriptive parts (rest). In these other

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

aims to answer two questions. First, how much does the Board write about Meta's self-given community standards, "values," and international human rights law respectively? Second, how much does the Board write that relates to the different subsections (i.e., legality, aim, necessity, and proportionality) within its human rights assessment? This again may tell us something about which issues and norms the Board considers particularly important – and which arguments it uses to lend weight to its largely nonbinding normative input vis-à-vis the company.

4.5.4 *Quantifying the Board's Focus on Human Rights*

Let us begin with the first question: How much does the Board write when it comes to each of the three sets of normative material? That is, how many words does the Board allot to community standards, values, and human rights in its decisions? To repeat, all standard decisions follow the same structure. In every decision, the Board's normative reasoning is contained to one specific part labeled "Oversight Board Analysis." Initially, this Oversight Board Analysis part cascaded into three sections, one for each normative level, that is, Meta's community standards, Meta's values, and international human rights law. These "values" are presented by way of five vaguely worded normative statements unilaterally made by the company,[327] whose purpose is to "guide" Meta's enforcement and rulemaking.

However, before digging further, we may have a look at the overall development of the word count of the "Oversight Board Analysis" section. They key takeaway is: on average, it grew considerably. Figure 4.1 shows the total word count of each decision in columns and the trend (exponential regression) as a dotted line. Thus, the Board wrote more and more over time in its normative analysis.

Now, we unpack the different subsections of the "Oversight Board Analysis." As previously noted, the Oversight Board Analysis contains the Board's normative reasoning and initially divided into three, since late 2022 into two parts. That missing third element are Meta's values, which were the shortest section from the very beginning. In all but one of its decisions, fewer than 500 words of text were devoted to discussing Meta's values. Often it was considerably lower, averaging out at 181 words. Later, as those values were

parts, the Board merely describes existing material. That it does so at great length indicates, on the one hand, the complexity of the issues at stake, and on the other hand, the presumed aim of the Board to present its reasoning as flowing from the normative material and the facts of a case.

[327] See at 4.4.1.1.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

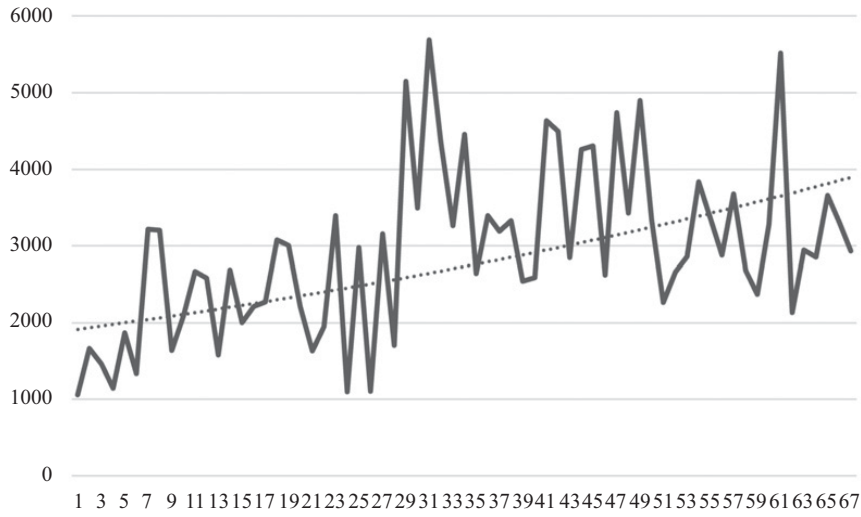


Figure 4.1 Word count Oversight Board Analysis total (2021–11/2024)

ostensibly useless to properly review Meta's exercise of power, the Board largely ceased referencing them. Since late 2022, decisions do not contain a specific section dealing with Meta's "values" anymore.

In turn, the Board's engagement with Meta's community standards and international human rights law seems to have grown consistently over the years. In the first two years, the Board usually allotted between 500 and 1,000 words to its engagement with Meta's community standards. Since 2023, that number climbed considerably to often more than 1,500 words. The overall average as of November 2024 being 855 words. While the growth in words allotted to the community standards appears, on average, steady but not exponential, the human rights section ballooned over the years. Here, the average word count was much higher from the beginning, being roughly 1,500 words in the first two years, reaching almost 2,000 words overall in November 2024 and 2,350 if one looks at only the decisions since 2023. In many decisions, the section on human rights comprised more than 2,000, 3,000 and in one case even more than 5,000 words. In all decisions, human rights standards "received" more words than the sections regarding Meta's values; and almost always more than in the section relating to Meta's community standards.

Figure 4.2 shows the different word counts distinguished by applied normative standard, namely Meta's values and policies, as well as

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

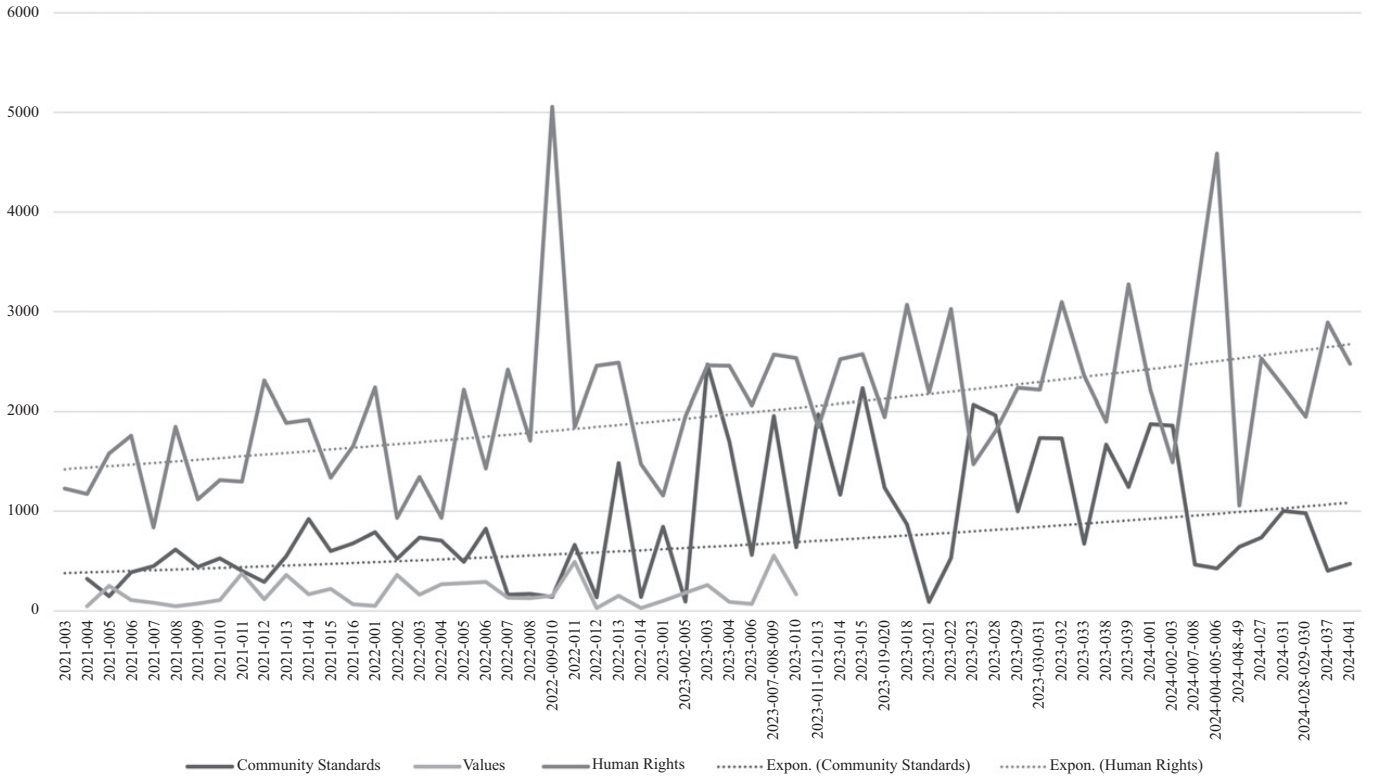


Figure 4.2 Word count of Meta values, Meta policies, and human rights law in the normative reasoning part of Oversight Board decisions (2021–11/2024)

international human rights law. Further, comparing the trends of Figures 4.1 and 4.2 we see that the Oversight Board Analysis and especially its section on human rights became longer over time. In contrast, the section on Meta's values was largely stable and then disappeared while the section where the Board analyzes only Meta's community standards, grew but by far not as strongly as the human rights section. Thus, the expanded word count in the human rights section appears to be the main driver for the (over time) greater amount of text devoted to normative reasoning in the Board's decisions.

In total, the values received roughly 6,000 words, the content policies roughly 60,000 words, and human rights a bit more than 130,000 words. Even though such total values mix together lengthy and more briefly worded decisions, the disparity indicates that the Board's normative reasoning mainly takes place in that part of the decision that focuses on human rights law.

What can we conclude from Figure 4.3? The increase in words may not necessarily indicate that the Board considers human rights more "important" and that therefore the decision's authors take it upon themselves to write more. The text may simply be more imprecise, lengthy, or repetitive.

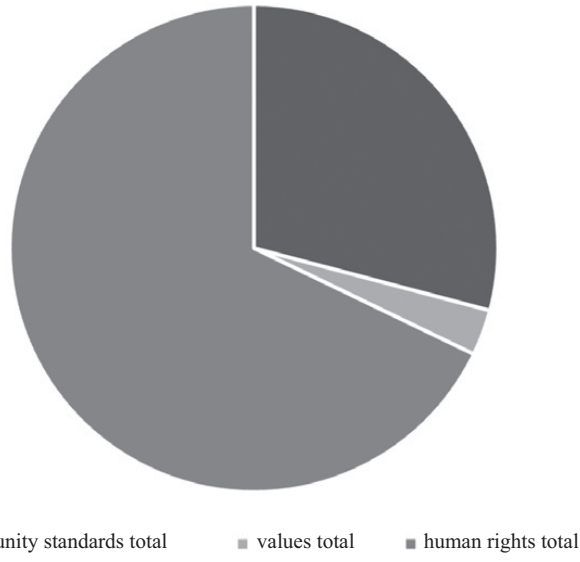


Figure 4.3 Total word count "Oversight Board Analysis" per normative category (2021–11/2024)

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

However, as a close reading of the decision shows, most of the text allotted to the human rights subsection does indeed read like legal argumentation. For example, in its human rights analysis in the decision relating to the banning of former US President Trump, the Board cautiously constructed a normative standard against which Meta's actions could be measured:

> In international law on freedom of expression, the principle of legality requires that any rule used to limit expression is clear and accessible. People must be able to understand what is allowed and what is not allowed. Equally important, rules must be sufficiently clear to provide guidance to those who make decisions on limiting expression, so that these rules do not confer unfettered discretion, which can result in selective application of the rules. In this case, these rules are Facebook's Community Standards.[328]

After outlining this normative standard by recourse to Article 19 of the ICCPR, the Board applies the facts, cites experts (like a UN Special Rapporteur, for instance), and refers to previous decisions. The same decision argues:

> The clarity [of Meta's community standard] leaves much to be desired, as the Board noted in a prior decision (case 2020–005-FB-UA). The UN Special Rapporteur on Freedom of Expression has also raised concerns about the vagueness of the Dangerous Individuals and Organizations Standard (A/HRC/38/35, para 26, footnote 67). As the Board has noted previously in case 2020–003-FB-UA, there may be times in which certain wording may raise legality concerns, but as applied to a particular case those concerns are not warranted. Any vagueness under the terms of the Standard does not render its application to the circumstances of this case doubtful. The January 6 riot at the Capitol fell squarely within the types of harmful events set out in Facebook's policy, and Mr. Trump's posts praised and supported those involved at the very time the violence was going on, and while Members of Congress were calling on him for help.[329]

In no decision does the Board employ comparable argumentative resources to assess Meta's values and only in a few decisions (and to a lesser extent within those decisions) to Meta's community policies. Although the Board's engagement with Meta's community standards grew over time, the main forum seems to remain human rights law. Therefore, the qualitative insights from the close reading of paragraphs like the above and the insights from the interviews substantiate the claim that we can plausibly assume that the Board has over time shifted its attention away from Meta's own rules and toward human rights. Whether this is because human rights

[328] Former President Trump's suspension (n 143) at 8.3.I.
[329] Ibid.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

are perceived as more legitimate, more pertinent, or simply because many employees of the Oversight Board previously worked in a human rights context, is uncertain.

So, to answer the first question, we can conclude that the Board allots by far the most words to the human rights section of its decisions. It writes considerably less in both sections dealing with either Meta's community standards of the later defunct section on Meta's values.

4.5.5 *The Two Forums: Legality and Proportionality*

Now let us move to the second question: how many words does the Oversight Board allot to the three subsections *within* its assessment of Meta's human rights responsibilities? Does the Board write more in the sub-section dealing with the normative basis of a measure, which the Board calls based on the framing of Article 19 ICCPR "legality?" Or does it focus more on a measure's "legitimate aim," or the "necessity and proportionality" of a measure? And what can this tell us about the Board's normative focus?

In Figure 4.4 we see that the Board writes the least in the subsection dealing with the "legitimate aim" of a measure. Here, it typically allocates less than 200 words, averaging out at only roughly 140 words. More usually, but not as a consistent trend across all the decisions, the subsection on "legality" has a greater amount of text. This averages at roughly 440 words but peaks at 1,457 words and dips as low as thirty-five words. In several cases the legality subsection had the highest number of words devoted to it and, when this occurred, this was, at least in the first two years of practice, often in inverse proportion to the word count of the subsection covering a measure's "necessity and proportionality." In other words, when the part of a decision dealing with legality is longer, then the part focusing on issues of necessity and proportionality tends to be shorter, and vice versa. However, this correlation seems to have largely ceased. Decisions of 2023 and 2024 appear to have a more equidistant relationship of the sections on legality and necessity and proportionality. This might be plausibly explained through incrementally manifesting practices within the Board as to its approach and style of reasoning.

Furthermore, we see a trend occurring over time in changes to the word count in the subsections covering legality, aim, and necessity and proportionality.[330] Whereas the (legitimate) aim's word count remains largely stable, the word count increases much stronger over time for the legality and, particularly, the necessity and proportionality sections.

[330] Identified using the exponential regression trend function in Excel.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

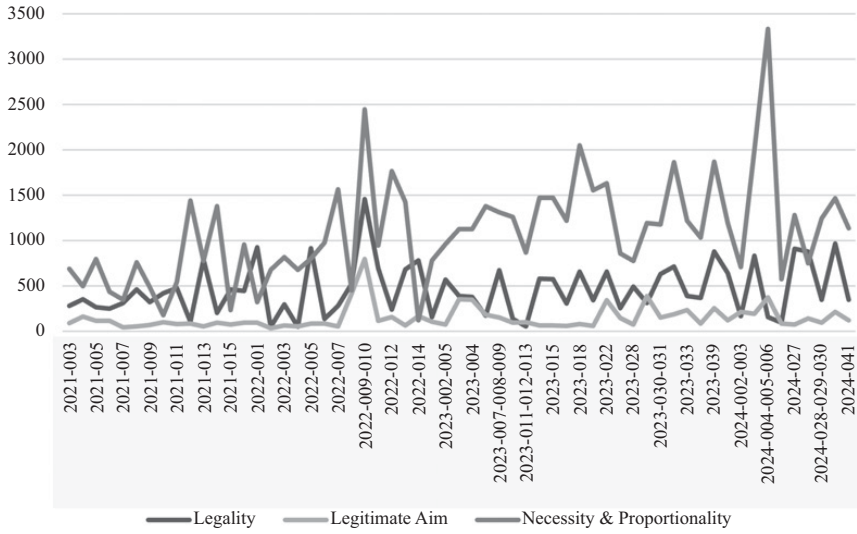


Figure 4.4 Word count of legality, legitimate aim, necessity, and proportionality subsections in the "Oversight Board Analysis" part (2020–11/2024)

Simply put, the Board writes increasingly more about legality and necessity and proportionality than about the legitimate aim.

This could mean that the legality *and* the necessity/proportionality parts are the two main forums for normative argumentation *within* the "Oversight Board Analysis". If that were the case, this would be remarkable. Whereas the necessity and proportionality section typically concentrates on an actor's *enforcement action*, the legality section focuses solely on this enforcement action's legal basis. Traditionally, Art 19 ICCPR requires any restrictions speech to be based on a *legal rule* such as an anti-libel or anti-hate speech law passed by parliament. The Board, however, resorts to Meta's self-made rules as "legal" basis for speech restrictions. In other words, when the Board devotes a lot of words to the legality of Meta's actions and not only their proportionality, this might suggest it is engaging in greater scrutiny of Meta's rulemaking. Simultaneously, the increasing word count in the necessity and proportionality section might also indicate the Board's increased scrutiny of Meta's enforcement procedures.

We now proceed with a close reading of the Board's conclusions about the "legality" of Meta's moderation efforts and their necessity and proportionality.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

4.5.5.1 The Board's "Legality" Test

Let us now have a look at the legality section. Legality, a concept rooted in Article 19 of the ICCPR, feels somewhat out of place in the entirely private context we find ourselves navigating here. As envisaged by the United Nations' Guiding Principles, the Board requires Meta's restrictions of free expression to be "provided by law."[331] But what is Meta's "law" in the context of Article 19(3) of the ICCPR? Meta has huge problems producing normative material that works on the worldwide scale on which the company operates. Normative material regulating human behavior must be concise and clear. As we have seen above, most of the rules established by Meta are neither concise nor clear. Further, the notion of "law" in Article 19(3) of the ICCPR is traditionally understood in a rather narrow sense. Based on the Human Rights Committee's "General Comments," this "may include laws of parliamentary privilege ... [but not] traditional, religious or other such customary law."[332] The Board, however, never discussed whether notions like "law" or "legality" are applicable for entities like Meta. It just applied them. Meta does not feature a rule-based let alone participatory process of rulemaking.[333] How can its privately devised and ad hoc rules be "law" within the meaning of Article 19 of the ICCPR?

Frictions will always occur when transplanting intrinsically state-centered rules into a private-law context, but this awkward grafting is still striking. On the one hand, the Oversight Board, for good reasons, generally avoids discussions of "sovereignty" or whether an entity like a transnational company or a social media firm is "state-like."[334] On the

[331] Article 19(3) ICCPR.

[332] Human Rights Committee (United Nations), 'General Comment No. 34, Article 19 Freedoms of Opinion and Expression, CCPR/C/GC/34' para 24.

[333] In 2009, Facebook experimented with democratic elements but failed to mobilize enough users to participate in voting. See Nicolas P Suzor, *Lawless: The Secret Rules That Govern Our Digital Lives* (Cambridge University Press 2019) 10 et seq.

[334] A minor digression, but I remain skeptical of analogizing corporations to sovereigns. Sovereignty is an elusive, normatively charged concept with a troubled history, making it ill-suited to scrutinize or describe private power. In fact, sovereignty might not even be particularly helpful for capturing the reality of statehood itself, given the vast disparities in states' power and their ability to exercise the rights tied to idealized notions of sovereignty. Sovereignty, rather, should be understood primarily as a normative, self-legitimizing construct, rooted in Enlightenment-era European conceptions of statehood. Platforms, by contrast, are capitalist corporations. Their internal structures are shaped by the demands of corporate and securities law, not by political projects. They lack participatory – let alone democratic – elements, a shared political community, or the historical continuity that defines states. While platforms wield undeniable power – arguably surpassing that of states in some respects – this does not render them state-like

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

other side, Article 19 of the ICCPR's legality requirement offers the Board an opening to review not only Meta's "enforcement decisions" but also its rulemaking. In several decisions, the Board's evaluation in the legality section assessed the content and quality of Meta's community standards. Often such assessments result in comprehensive recommendations to Meta outlining what aspects of the community standards the company should change. For example, in a 2021 decision regarding support for the – later deceased – Russian opposition politician Alexei Navalny, the Board demanded a complete overhaul of one of the then twenty-four community standards.[335] Its criticism was that the different subsections and rules of the standard were unclear, that it included many words with unclear definitions (what is the difference between bullying and harassment?) and obscured rather than elucidated the true meaning of the standard. Part of its decision read:

> Facebook should amend and redraft its Bullying and Harassment Community standard to … explain the relationship between policy rationale and the "Do Nots" as well as the other rules restricting content that follow it. … Distinguish between bullying and harassment and provide definitions that distinguish the two acts. … Clearly define its approach to [different groups of users] … Include illustrative examples of violating and non-violating content.[336]

Such criticism is often renewed in different, subsequent cases, especially in following "summary" decisions. The Board regularly criticizes the company for offering too few examples, for its slow progress in translating standards and decisions, and its unclear and incoherent definitions. For example, in a 2022 decision regarding communication about the Taliban, the Board reiterated that it considers the relationship between two definitions in one particular standard to "remain unclear to users."[337]

In sum, the Board extensively uses its decisions' legality subsection to control – or at least comment on – Meta's rulemaking. The increase in word count for this part of its decisions, coupled with statements by

or sovereign. Their power is of a fundamentally different nature, grounded in markets and profit motives rather than in the intricate, historically contingent construct of sovereignty.

[335] Pro-Navalny Protests in Russia (n 187) Recommendation N 10.

[336] Ibid at 10.1–4. However, Meta declined implementation of that recommendation, cf. Oversight Board Recommendation Tracker, accessible via <https://perma.cc/76MD-NN32>.

[337] Colombian police cartoon [2022] Oversight Board 2022-004-FB-UA at 8.3.1 speaking about the "reporting" allowance in the context of the dangerous individuals and organizations standard.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

interviewees, therefore indicates that a prime focus of the Board's deliberation is Meta's *rulemaking*.

4.5.5.2 Necessity and Proportionality

Whereas the legality section focuses on the mostly formal requirements (clarity and consistency) of Meta's rulemaking, the necessity and proportionality section are where the Board scrutinizes Meta's policies of rule enforcement. Typically, the Board's necessity and proportionality assessment has a double focus. On the one hand, it concentrates on the specific enforcement action as it was conducted for the case it is reviewing. For example, in a case dealing with the wrongful deletion of a post promoting breast cancer awareness, the Board established that: "The Board is concerned that the content was wrongfully removed by an automated enforcement system and potentially without human review or appeal. . . . In this case, these technologies failed to recognize the words 'Breast Cancer' that appear at the top left of the image in Portuguese."[338]

However, the Board normally embeds assessments of a specific enforcement act in more general statements pertaining to the issue at stake. Therefore, assessing an action's necessity and proportionality allows the Board to say something about the company's procedures in general. For example, in the breast cancer awareness case, the Board inductively drew several general conclusions from its individual assessment of the particular case. Drawing attention to a specific failure of Meta's automated moderation system, the Board criticized Meta's over-reliance on automated tools and pointed out the "limitations" of the technology, especially its "limited ability to understand context." This part of the decision reads:

> This reflects the limitations of automated technologies to understand context and grasp the complexity of human communication for content moderation. . . . The Board accepts automated technologies are essential to the detection of potentially violating content. However, enforcement which relies solely on automation, in particular when using technologies that have a limited ability to understand context, leads to over-enforcement that disproportionately interferes with user expression. . . . Regardless, automated removals should be subject to both an internal audit procedure explained under section 9.2 (I) and appeal to human review should be offered, allowing enforcement mistakes to be repaired. Automated content moderation without necessary safeguards is not a proportionate way for Facebook to address violating forms of adult nudity.[339]

[338] Breast Cancer symptoms and nudity [2021] Oversight Board 2020-004-IG-UA at 8.3.I. c.
[339] Ibid at 8.3.3 (in-text references redacted).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

What is striking is how the Board references its own recommendations in a later section of the same decision. It states that "automated removals should be subject to . . . an internal audit procedure explained [below]."[340] In the referenced section, the Board issued a series of recommendations regarding how Meta should improve its overall procedures. Among these recommendations we see again an increasing degree of generalization. The Board issued some recommendations with a specific connection to the case. For example, the company should "[i]mprove the automated detection of images with text-overlay to ensure that posts raising awareness of breast cancer symptoms are not wrongly flagged for review."[341] However, other recommendations are more general. The company should, for example, notify users about the specific rule of the community standard Meta considers to be violated, and it should put in place an "internal audit procedure to continuously analyze a statistically representative sample of automated content removal decisions to reverse and learn from enforcement mistakes."[342] Similar arguments and connections are made in other cases. As will be argued in Section 3.6.2 below, the Board's recommendations constitute a crucial part of its normative output.[343] Therefore, tying them to arguments made in the necessity and proportionality section highlights that section's importance.

In sum, a close reading of the decisions shows that the Board invests considerable argumentative resources on the necessity and proportionality section. Even though it is typically longer than the legality section, this is not always the case, which might indicate the beginnings of a trend toward stricter scrutiny of Meta's rulemaking by the Board. However, the heart of the Board's human rights assessment is making recommendations for Meta to tighten up its content moderation. In other words, the Board "uses" the individual case to inductively generate procedural advice for the company. This advice might be restricted to a particular subject or community standard in question. However, it can often be of a more general nature and based on principles of good governance and administration.

4.5.6 Interim Conclusion

This section detailed how the Board uses various normative categories – Meta's community standards, Meta's values, and international human

[340] Ibid (emphasis added).
[341] Ibid at 9.2.1 first bullet point.
[342] Ibid at 9.2.1 fifth bullet point.
[343] See this at 4.6.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

rights law – in its decisions and arrives at several conclusions. First, the Board does not engage substantively with Meta's "values." Second, the Board does however engage in some detail with the company's community standards as a way to control Meta's enforcement actions. Third, the Board's main normative focus – in every single decision and increasing over time – is *international human rights law*, especially Article 19 of the ICCPR. This is remarkable since the charter only mentions human rights law in passing and the bylaws do not mention it as a basis for review at all. Fourth, as part of its human rights assessment, the Board engages in an increasingly detailed analysis modeled on elements of Article 19(3) of the ICCPR. Typically, this includes a "legality" subsection in the Board's written decisions that tests the normative "clarity" and "accessibility" of Meta's rulemaking, as well as an extensive "proportionality" assessment split into "(legitimate) aim" and "necessity and proportionality" subsections. Sixth, whereas in its earliest decisions the Board tended to focus on the "necessity and proportionality" subsection, we have seen a relative increase over time in words allotted to the "legality" subsection. In the latter, the Board assesses whether Meta's own rules, especially the "community standards", comply with international human rights law. Even though it is too early to speak of a stable trend, we can tentatively conclude that the Board's attention focuses both on controlling Meta's *enforcement* (by questioning the proportionality of Meta's enforcement) and controlling Meta's *rulemaking* (by questioning the clarity and consistency of the rules Meta enforces). So far, the focus on controlling rulemaking has remained at the mainly formal level (relating to the clarity and consistency with which standards are formulated, for instance).

The Board's turn to human rights aligns with remarks by staff members in the interviews who said they were limited in their ability to control individual decisions but wanted to "change how Meta makes these decisions."[344] This turn to human rights might also be explained by the prevalence of people with a background in human rights among staffers (who draft a first legal analysis of each case) and among Board members.

4.6 Implementation

We now explores how (and whether) Board decisions are implemented. The empirical basis for assessing the degree to which Meta implements Board decisions is thin because the company remains coy about how exactly (and

[344] Interview with Oversight Board Staffer 2 (n 10).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

if at all) it implements the Board's normative output. However, based on transparency reports by Meta and the Board, several interviews, and a close reading of Board decisions, we might be able to pin down the contours of how Meta implements, or chooses not to implement, Board decisions.

The Board generates several distinct kinds of normative output. Formally, it "decides" individual cases and offers "policy advisory opinions" based on general questions referred to it by Meta.[345] Its decisions feature an operational part that is binding for Meta for that specific case. But the document – known as the "decision" – also includes "policy advisory statements" and/or "recommendations." These are nonbinding, generalized normative recommendations relating to how Meta should improve its rulemaking and rule enforcement. A close reading as well as remarks from several interviewees indicate that these policy advisory statements are the Board's primary normative contribution to Meta. Further, even though policy advisory statements are not binding, many interviewees felt them to be more effective than the (necessarily individual) decision due to the general terms in which these statements were formulated and because of their focus on the company's operations. The Board has begun to build up recurring normative patterns that could be interpreted as doctrines and call for "naming and shaming"-inspired measures to influence the company's behavior.

Due to Meta's largely untransparent implementation of Board decisions, the latter's efficacy cannot be assessed comprehensively. For an adjudicator, this lack of transparency is hard to accept. Even though the Board has quickly adapted – bringing out increasingly sophisticated quarterly transparency reports and, later, a "recommendation tracker" – questions remain regarding Meta's transparency and accountability when it comes to how, or even whether, it implements nonbinding policy recommendations and advice from the Board.

4.6.1 Set Board Decisions Binding Precedent?

Each Board decision features a brief "operational part" that assesses whether a specific piece of content should stay on Meta's services or not. This decision is "binding" for Meta, but only for that specific enforcement action relating to the case under review.[346] Decisions do not have precedential status beyond the individual case. Meta should only evaluate whether it is "technically and operationally feasible to apply the Board's

[345] The latter is not to be confused with individual referrals of cases by Meta according to Bylaws Article 2 Section 2.

[346] Charter Article 4; Bylaws Article 2(2).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

decision" in cases with "identical content and parallel context," according to Article 4 of the charter. What "parallel context" means is obscure, but "identical content" can perhaps be pinned down more successfully. Further, even if these criteria are met, the resulting obligation is nothing more than the requirement on Meta's part to exercise due diligence.[347] Thus, Meta is not bound by the Board's decision in cases other than the one decided – their executives need only "take action by analyzing" whether parallel implementation would be feasible. Hence, the "binding" power of the Board's decisions is negligible for the Board's practice. Especially so given the Board's leeway in case selection (subject to Meta's approval) and its relatively small caseload.

4.6.2 Infusing Public Law Principles Into Private Governance?

Almost all the Oversight Board's decisions feature nonbinding, more general normative statements called either "recommendations" or "policy advisory statements." These statements inductively generate principles for Meta's rulemaking and rule enforcement. Interviewees described these statements as the normative core of the Board's decisions.[348] One interviewee emphasized that "albeit the Board makes its decisions in a way that can be understood as adjudicative, our main way of exercising our power are the policy recommendations."[349] These are attached to each decision in their own section, typically the tenth. In these recommendations, the Board formulates general guidelines and principles recommending what the Board "should" do to improve its rulemaking and rule enforcement. As one trustee interviewed for this book put it:

> The Oversight Board is going to decide maybe 40 or 50 cases in a year, out of billions of posts. It follows that the Board itself is not the manager of disputes. The Board is not like a police force. Instead, it's infusing the content moderation operations of Meta with public law principles and procedures. It's articulating substantive norms that interpret Meta's community standards in light of human rights principles.[350]

What the interviewee describes as an infusion of "public-law principles" into Meta's managerial way of governing its users are the nonbinding recommendations and advisory statements that the Board attaches as

[347] cf. UK drill music [2022] Oversight Board 2022-007-IG-MR [8.4].
[348] Interview with Oversight Board Staffer 1 (n 54); Interview with Oversight Board Staffer 2 (n 10); Interview with Oversight Board Staffer 3 (n 11).
[349] Interview with Oversight Board Staffer 2 (n 10).
[350] Interview with Oversight Board Trustee (n 53).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

supplements to most of its decisions. In its third decision, the Board explicitly formulated such recommendations for the first time.[351] Since then, each decision has featured at least one such "policy advisory statement." These statements draw upon normative or factual assessments made by the Board in its analysis, often with a special focus on human rights-related implications.

For example, in a case concerning posts referring to the – then still alive – Russian opposition figure Alexei Navalny, the Board used the legality assessment of Meta's bullying and harassment community standard to assess this standard's normative coherence. It found that the whole standard was constructed in an overly complex way, making it hard for users to know "what is and what is not allowed."[352] The Board criticized several categories Meta introduced in the standard as it were not clear if they "serve[d] simply as a guidance for the specific rules that follow, or if they must be interpreted conjunctively with the rules."[353] To back up this criticism, the Board then formulated a 'policy advisory statement' outlining what Meta should do. The Board recommended that "[t]o comply with international human right standards, Facebook should amend and redraft its Bullying and Harassment Community Standard."[354] It then formulated six tangible reforms. For example, the Board should "explain the relationship between the policy rationale and the 'Do nots' as well as the other rules restricting content that follow it."[355] Further, the standard should differentiate between bullying and harassment and "should clearly explain to users how bullying and harassment differ from speech that only causes offense and may be protected under international human rights law."[356] Meta should "[c]learly define its approach to different target user categories and provide illustrative examples for each category (i.e. who qualifies as a public figure)."[357]

Most of the Board's recommendations in fact reflect rationales, principles, and notions familiar from public-law contexts. Regarding Meta's rule enforcement, the Board recommended, for example, that the company should provide reasons for its decisions.[358] It should ensure that users are

[351] Armenians in Azerbaijan (n 322).
[352] Breast Cancer symptoms and nudity (n 318) at section 8.3.I.
[353] Ibid at section 8.3.I.
[354] Ibid at section 10.
[355] Ibid.
[356] Ibid.
[357] Ibid.
[358] See e.g. in Armenians in Azerbaijan (n 299) at section 9.2; Breast Cancer symptoms and nudity (n 318) at section 9.2.I second bullet point; Nazi Quote [2021] Oversight Board 2020-005-FB-UA at section 9.2 first bulletpoint; Depiction of Zwarte Piet (n 298) at section 10 No 2.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

notified that their content has been removed and which particular violation had led to its removal.[359] It should also provide timely translations so that its rules are intelligible to users.[360] Users must have the right to appeal, sometimes directly linked to the ICCPR's right to an effective remedy. Several recommendations explicitly referred to the principle of proportionality.[361] Further, the Board recommended specific procedures to Meta for investigating whether its enforcement algorithms might be biased.[362]

In terms of Meta's rulemaking, the Board regularly recommends that the company includes more examples,[363] as well as rules that are worded in a clearer and more concise way.[364]

In sum, the Board's policy advisory statements synthesize principles from external normative material – human rights law and secondary sources (like General Comments from the UN Human Rights Committee), non-governmental or international organizations i documents (the Rabat Plan of Action), and scholarship – and the statements offer recommendations for Meta to implement in order to improve its rulemaking and rule-enforcement efforts.

4.6.3 *Developing Doctrine*

The Board also began to develop normative sentences that could be described as *doctrines*. Although conceptions of what counts as doctrine differ greatly among legal cultures, several characteristics seem identifiable.[365] The function of doctrine is to *stabilize* a normative order

[359] Öcalan's Isolation (n 288) at section 10 No 9; Colombia Protests (n 298) at section 10 No 4; South Africa slurs [2021] Oversight Board 2021-011-FB-UA at section 10 No 1; Wampum Belt [2021] Oversight Board 2021-012-FB-UA at section 10 no 1; Pro-Navalny Protests in Russia (n 177) at section 10 No 6.

[360] Myanmar post about Muslims (n 301) at section 10 No 1; Punjabi Concern over the RSS in India [2021] Oversight Board 2021-003-FB-UA at 10 No 1; Shared Al Jazeera Post [2021] Oversight Board 2021-009-FB-UA at section 10 No 2.

[361] COVID lockdowns in Brazil (n 186); Former President Trump's suspension (n 143).

[362] Shared Al Jazeera Post (n 341) at section 10 No 2.

[363] Nazi Quote (n 358) at 10 No 2, 3; Pro-Navalny Protests in Russia (n 185) at 10 No 3, 4; Öcalan's Isolation (n 304) at 10 No 6; Shared Al Jazeera Post (n 360) at 10 No 1; Colombia Protests (n 307) at 10 No 1, 2.

[364] Former President Trump's suspension (n 143) section 10.; Öcalan's Isolation (n 304) at section 10 No 5; Colombia Protests (n 307) at section 10 No 3; Asking for Adderall® [2022] Oversight Board 2921-015-FB-UA at section 10 No 1.

[365] For an overview see e.g. René Smits, 'What Is Legal Doctrine? On the Aims and Methods of Legal-Dogmatic Research' in Rob van Gestel, Hans-Wolfgang Micklitz and Edward L Rubin (eds.), *Rethinking Legal Scholarship: A Transatlantic Dialogue* (Cambridge University Press 2017) 207; Rob van Gestel, Hans-W Micklitz and Miguel Poiares Maduro, 'Methodology in the New Legal World, EUI Working Papers, Law 2012/13' 3 et seq.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

by establishing hierarchy, clarity, and foreseeability. Doctrine creates general normative sentences guiding future behavior.[366] The Board uses citations and references to incrementally construct "doctrines."

On the one hand, the Oversight Board increasingly returns to issues it brought up in earlier decisions but considers insufficiently remedied. The Board clarifies or reiterates prior normative statements and criticizes Meta for insufficient implementation. Such repetitions often include citations of earlier Board decisions. For traditional courts, citing decisions' titles or excerpts is widely considered as a useful tool for maintaining normative stability and developing precedential law and bodies of legal doctrine.[367] Further, the Board references international courts and tribunals. For example, in a case concerning hate speech in the context of ethnic conflicts in Ethiopia, the Board referred to prior decisions as well as to judgments by the International Criminal Tribunal on Rwanda (ICTR). The Board said that it "came to a similar conclusion in decision 2020–003-FB-UA. [And that] . . . cumulative [hate speech] can amount to causation through a 'gradual build-up of effect,' as happened in Rwanda where calls to genocide were repeated [then quoting an ICTR case]."[368]

Citing older cases may not automatically mean that an adjudicator is accurately recreating earlier normative assessments.[369] Citations are also sometimes used "ritualistically."[370] However, for a young but ambitious organization like the Oversight Board, such "ritualistic" practices are important for gaining public legitimacy.

Some of these "doctrines" directly challenge traditional conceptions of international law. The Board increasingly formulates abstract

[366] Doctrine's stabilizing function only lasts as long as its proponents in judiciary, practice, and legal academia reaffirm it. See for a daunting example from German history: Bernd Rüthers, *Die unbegrenzte Auslegung: zum Wandel der Privatrechtsordnung im Nationalsozialismus* (9th ed., Mohr Siebeck 2022).

[367] The exact difference between the term precedent and doctrine is elusive. Precedents are necessarily court decisions. In principle, doctrines may also result from scholarly writing. However, in practice most doctrines seem to equally originate from judgments. The concepts hence differ mostly regarding their context.

[368] Alleged Crimes in Raya Kobo [2021] Oversight Board 2021-014-FB-UA at 8.3.III.

[369] Joost Pauwelyn and Manfred Elsig, 'The Politics of Treaty Interpretation: Variations and Explanations across International Tribunals' in Jeffrey L Dunoff and Mark A Pollack (eds.), *Interdisciplinary Perspectives on International Law and International Relations: The State of the Art* (Cambridge University Press 2012) 456.

[370] Wolfgang Alschner and Damien Charlotin, 'The Growing Complexity of the International Court of Justice's Self-Citation Network' (2018) 29 European Journal of International Law 83, 86.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

interpretations of Meta's normative material and human rights law. Perhaps the most striking example of such an abstract "doctrine" fashioned by the Board is its reference to a state of "public emergency" and Meta's resulting recourse to "emergency powers" to suspend its fundamental rights obligations. We will discuss said "doctrine's" ripple effects in greater detail in the next chapter.[371] For descriptive purposes, we will just briefly look at that decision's decisive paragraph. In a case dealing with potentially incendiary speech about Tigrayans in Ethiopia. The content in question was posted by an individual user, but the Ethiopian government was reported to post similar content. The Board asserted:

> 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).[372]

In other words, the Board emboldens a private, for-profit company to step in whenever it considers governmental institutions to be failing in their human rights obligations in situations of armed conflict. But not only that. The Board essentially emboldens Meta, again, a private company, to declare a state of emergency vis-à-vis a government to justify policing that government tighter than before. We will discuss this in more detail in the following chapter. Yet, I take the liberty of highlighting the audaciousness of such doctrines already here. Arguments like these directly challenge core tenets of international law. They are reminiscent of the earliest plans for the Board, outlined in an internal memo which argued that "by acting to a degree like a sovereign in establishing courts, Facebook presents itself as a responsible decision-maker, prepared to use a method of decision-making deemed valid and legitimate (at least in principle) by most governments on earth."[373] The Board follows this path – even though its actual authority falls far short of earlier ambitions expressed by some within the company.[374] Asked about the audacity of such state-of-emergency arguments, one Board member interviewed, who was not party to the deliberations relating to that specific decision, affirmed that "the Board members

[371] See at 5.2.
[372] Alleged Crimes in Raya Kobo (n 308) at section 8.3.III.
[373] Feldman (n 33) 112.
[374] See e.g., Feldman's memo of March 2018, which envisaged a multitier court system for Facebook and envisaged the company as a "quasi-sovereign," ibid 107, 112 et seq.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

are acutely aware of the potential effects of statements like that."[375] Time will tell what the effects will ultimately be, but the significance of the potential downstream effects is hard to overstate.

4.6.4 *The Implementation Conundrum*

Yet, despite intermittently emerging doctrines, the Board's recurring focus on Meta's governance, and the Board's lists of recommendations, the question remains: does it work? Does Meta actually implement the Board's nonbinding recommendations?

As argued in Chapter 2, nonbinding recommendations and doctrines can also have lasting effects. When an adjudicator manages to *convince* the controlled actor, this may lead to more deeply rooted policy changes than the exercise of adjudicatory force through issuing binding decisions.[376] Because adjudicators seek to "adjust future behavior, not compensate for past wrong," their involvement is prolonged as they construct an "ongoing regime of performance."[377] Accordingly, the Board regularly returns to issues it raised in earlier decisions to "guide" the company. Dialogue-based and nonbinding adjudicatory techniques have been extensively explored by legal scholars and, to a degree, might also work for the benefit of the Oversight Board.[378] However, many dialogue-based forms of adjudicatory control are exercised by actual courts, which still have the ultimate tool of comprehensive and binding decisions at their disposal. In contrast, as shown above, the Board's authority to issue binding decisions exists mainly on paper, so the Board – although seen by some as a potent adjudicator – cannot bindingly adjudicate at scale. To make its normative output more effective, the Board could draw on the perceived legitimacy attached to its partly adjudicative status but ultimately it must convince Meta executives to implement its recommendations. Such "good faith"-based implementation reflects the private nature of the Board and of Meta. After all the Board can revert to neither the institutions nor the norms of the rule of law.

[375] Interview with Oversight Board Member (n 314).

[376] See at 2.2.6.

[377] cf. Chayes' characterization of the judge in what he calls "public law litigation," Abram Chayes, 'The Role of the Judge in Public Law Litigation' (1976) 89 Harvard Law Review 1281, 1298.

[378] Chayes (n 377); Judith Resnik, 'Managerial Judges' (1982) 96 Harvard Law Review 374; Abram Chayes and Antonia Handler Chayes, 'On Compliance' (1993) 47 International Organization 175; critical in the context of transnational human rights adjudication Veronika Fikfak, 'Against Settlement before the European Court of Human Rights' (2022) 20 International Journal of Constitutional Law 942.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

As a consequence, Meta executives regularly choose not to put in place the Board's recommendations in a verifiable way. The story is of fumbled implementing efforts, and the operational complexity of the recommendations. Meta staffers argue that the

> size and scope of the Board's recommendations go beyond the policy guidance that we first anticipated when we set up the Board, and several require multimonth or multiyear investments.[379]

This aligns with statements from interviewed Board staffers who said their counterparts at the company were surprised and even unhappy with the depth and intrusiveness of Board recommendations – which arguably would be a good thing.[380]

To track implementation of its recommendations, the Board first began publishing quarterly transparency reports that assess Meta's implementation efforts. Especially since 2022, the Board scrutinizes how Meta implements its recommendations and uses its quarterly published transparency reports as a "mechanism to force better behavior from Meta."[381] In the first quarterly report of 2022, the Board published a comprehensive assessment of Meta's implementation of all recommendations until then. The report distinguishes six categories of implementation, ranging from "implementation demonstrated through published information" to "recommendation omitted, declined, or reframed."[382] This was later expanded into a section called "recommendation tracker" on the Board's website. The recommendation tracker contains two tables, one listing all published recommendations and one adding two more categories to the six categories just mentioned.

Neither the 2022 quarterly report, which introduced those six implementation categories, nor the later introduced recommendation tracker on the Oversight Board website publish their underlying methodology or the used data. So, the Board's assessment that implementation in these two categories has indeed been "demonstrated" remains at times rather untransparent to outside observers.

Of the 288 recommendations issued by the Board up until November 2024, only fifty-seven recommendations were demonstrably and completely implemented according to the Board. Another forty-one

[379] Jennifer Broxmeyer, 'Facebook's First Quarterly Update on the Oversight Board' (*Meta Newsroom*, 16 July 2021) <https://perma.cc/2YDU-E62C>.
[380] Interview with Oversight Board Staffer 2 (n 10).
[381] Ibid.
[382] Oversight Board Q1 2022 Transparency Report, p 21.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

recommendations assessed the Board to be verifiably partially implemented. For many recommendations, Meta reported progress (sixty-seven) or reported implementation without providing corroborating information (forty-six). Meta declined to put in place several recommendations after a feasibility assessment (fourteen), omitted or reframed recommendations (twenty), or declined to implement the recommendation out of principle (thirty-one).

According to Figure 4.5 all but the two top-right categories (implementation demonstrated; partial implementation demonstrated) show recommendations whose normative content was not verifiably adopted. According to the Board, Meta verifiably implemented 21 percent of its recommendations as of November 2024. Further, the Board considers another 15 percent of its implementations to be verifiably "partially" implemented. However, the exact efficacy and relevance of "partial" implementation remains unclear. Also, the Board does not disclose its verification of Meta's implementation. Rather, until this day, the Board and Meta remain in a constant struggle toward, hopefully, incremental improvements in transparency. For example, in a 2023 decision the Board still deemed it necessary to explicate that it "expects Meta to share more data to enable the Board to assess improvements in performance over time."[383] Therefore, this category may be treated with some skepticism. Nonetheless, when comparing the data available in November 2024 with that of the first comprehensive transparency report in 2022, we see that the full implementation indeed increased over time, climbing from 14 percent in 2022 to 21 percent in November 2024; "verifiable partial implementation" more than doubled from 7 percent in 2022 to 15 percent in November 2024.[384]

Between the Board's inception in 2021 and November 2024, Meta reported – without corroborating their self-reporting – either "progress" or even "implementation" on another 113 Board recommendations. However, "progress" only means that "Meta has made a commitment to implementing this recommendation but has not yet completed all necessary actions."[385] Whereas modifications to its publicly accessible normative material can be assessed by the Board, Meta's post-recommendation handling of automated enforcement procedures remains obscure. In particular, it is still unclear how significant an implementation must be

[383] United States posts discussing abortion [2023] Oversight Board 2023-01-IG-UA, 2023-012-FB-UA, 2023-013-FB-UA at 8.2.IV.
[384] Oversight Board, 'Annual Report 2022' 29 <https://perma.cc/MZ3J-86PE>.
[385] Oversight Board Q4 2021 Transparency Report, p 16.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

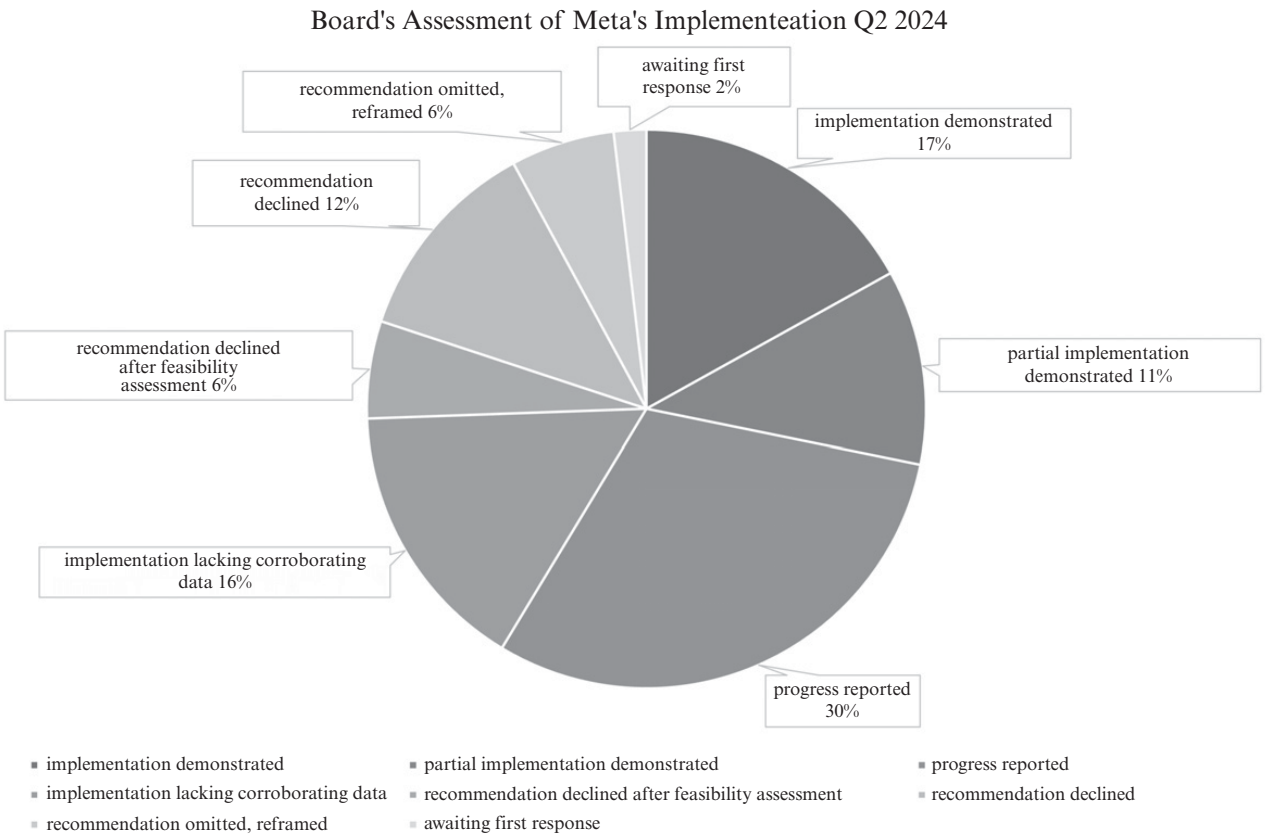


Figure 4.5 The Board's assessment of Meta's implementation of Board recommendations (as of 11/2024)

to count as "progress" and whether "implementation" always pertains to a comprehensive, "word-for-word" implementing of a recommendation.

The success or otherwise of transparency reports as enforcement mechanisms is beyond the scope of this study. However, judging from the anecdotal insights from the interviews as well as a close reading of Meta's updates to its community standards, the tentative answer must be that the Board's recommendations have had mixed success. To some extent, they have been verifiably implemented. Yet 178 out of 276 recommendations until November 2024 were either not fully implemented, their implementation has not been corroborated by any (publicly accessible) data or otherwise failed to materialize. Only roughly a fifth (21 percent) of the total are verifiably implemented recommendations. Conversely, for almost four-fifths (79 percent) of the recommendations made in that period Meta did not corroborate full implementation. Even if one deducts here the 15 percent verifiable partial implementations, one is left with roughly two-thirds of Oversight Board recommendations whose implementation cannot be seriously assessed or is even flat-out rejected by the company.

These figures present a paradox. On one hand, if only one out of five recommendations is fully implemented, it would be premature to declare success. In short, Meta remains obstinate and opaque, a fact the Board itself repeatedly criticized in its decisions. It is also important to look at *which* recommendations Meta implemented and which it rejected. It seems that at least some very salient recommendations were flat-out rejected. Remember the 2023 case on the Cambodian Prime Minister we discussed above?[386] The Board recommended that

> Meta should immediately suspend the official Facebook page and Instagram account of Cambodian Prime Minister Hun Sen for a period of at least six months under Meta's policy on restricting accounts of public figures during civil unrest. The Board will consider this recommendation implemented when Meta suspends the accounts and publicly announces that it has done so.[387]

And indeed Mr. Sen's Facebook page went dark shortly after the Board issued its recommendation to Meta. However, that was not because the company disabled the account. Quite the opposite. The Cambodian Prime Minister appears to have shut down his Facebook account himself, arguably in response to the Oversight Board's decision.[388] The company, in contrast,

[386] See at 3.3.
[387] Cambodian Prime Minister (n 146).
[388] Mike Ives and Sun Narin, 'Cambodian Ruler's Facebook Account Goes Dark after Spat With Meta' *New York Times* (30 June 2023) <https://perma.cc/PPJ9-SCKF>.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

expressly "declined" the Oversight Board's recommendation to suspend Mr. Sen's account.[389] Presumably, that is because the company does not wish to cede decision-making capacity in such a vital question to the Board – a remarkable rollback from the stance the company took in its 2021 Trump referral. Back then, the Meta asked the Board to come up with guidelines how the company should handle malevolent heads of state – and the Board tarnished the company for its lack of clear rules on that matter.[390] That Meta now expressly rejects a Board recommendation – based on rules Meta had devised in the meantime – to suspend an account of a government figure who was arguably inciting real-world violence seems deeply troubling. Especially so as all executives at the company must be aware that Facebook infrastructures played – according to the United Nations – a significant role in the displacement and murder of many Rohingya people in Myanmar in 2018. At the same time, the fact that the Cambodian Prime Minister self-suspended his account could be understood as testament to the normativity of Board decisions. Yet, alas, it did not last. Only a month after his exit, shortly before Cambodians went to the polls for a general election, Mr. Sen reactivated his account.[391] In the following election Mr Sen's party won almost all seats – 120 out of 125 – in the Cambodian Parliament.

On the other hand, there is some positive movement. The rate of verifiable implementation in 2024 has significantly improved compared to 2022, which is a crucial indicator of progress. Perhaps, the Board's focused and detailed efforts – such as the implementation tracker, or even external factors such as this very book – are exerting pressure on the company to improve its adherence to recommendations and overall improvement of its content moderation regime.

Looking ahead, two key points warrant attention. First, the Oversight Board was astute in transforming a vice into a virtue by publishing Meta's initially dismal implementation rates. Highlighting how uncooperative the company was during the Board's early years, and likely still is in some respects, appears to have prodded Meta into demonstrably adopting more recommendations. At present, the Board openly calls out Meta's shortcomings in fully implementing its guidance, a strategy that paradoxically risks undermining the Board's own credibility. However, Meta has a clear interest in sustaining the perception of

[389] As per the Oversight Board's Recommendation Tracker, accessible via <https://perma.cc/7RY8-CSLH>.

[390] Former President Trump's suspension (n 143).

[391] Martin Petty, 'Cambodia PM's Facebook Page Reactivated Days before Election' *Reuters* (20 July 2023) <https://perma.cc/KSD7-3NHV>.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

the Board as a (somewhat) effective oversight mechanism. As long as the company seeks to benefit from the legitimacy bestowed by external, adjudicative oversight, it remains incentivized to comply – at least enough to avoid entirely eroding the Board's authority. Therefore, it seems entirely plausible that we will see a further uptick in implementation. This would then warrant further, more qualitative research as the mere number of implemented recommendations tells us little about the actual quality of the reforms.

Second, some patience is warranted. The Oversight Board is still a young organization, and many well-established human rights bodies within the UN system arguably have similarly low implementation rates, despite decades of experience. The core challenge, as consistently highlighted in interviews with Board members, staff, and trustees, lies in the social norms and internal culture at Meta – factors that do not change overnight. Consequently, the gradual but steady implementation of Oversight Board decisions is a hopeful sign.

Looking to the future, however, it will be critical for the Board to be transparent about how it assesses "verifiable" implementation, so that independent third parties – particularly researchers – can effectively scrutinize the compliance process.

To summarize, having limited jurisdiction or relying on the good faith of digital corporations may not be immediately fatal for Emulated Guardians. Especially so in post-state governance where many power relations are regulated in vague, nonbinding arrangements. However, a "verifiable" success rate of only 21 percent cannot be seen as representing effective implementation.[392] Maybe Meta implements more than that – but the company's unwillingness to be transparent about measures it puts into practice constitutes a legitimacy problem for the Board, although the latter's transparency reports are bold and innovative in their approach. Eventually, these transparency reports may either gradually push Meta into compliance or, as a result of that transparency itself, their contents might end up chipping away at the Board's image as a credible guardian of power, and thereby undermine its authority.

[392] I am not sure whether many adjudicatory guardians of post-state and international governance could expound higher implementation quotas. See regarding the methodological difficulties assessing implementation, Erik Voeten, 'Domestic Implementation of European Court of Human Rights Judgments: Legal Infrastructure and Government Effectiveness Matter: A Reply to Dia Anagnostou and Alina Mungiu-Pippidi' (2014) 25 European Journal of International Law 229.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

4.6.5 *Prospects*

Assessing possible improvements in the Board's efficacy and Meta's implementation will be the work of further studies. To a large degree, this change will depend on a change of culture at Meta. We know how hard and unlikely such a cultural change may be.[393] Also, the whole tech sector shifted or appears poised to shift to the political right, a trend epitomized by but not restricted to Elon Musk's role in the 2024 US presidential election.[394] How lasting and impactful such shifts are remains to be seen. Nonetheless – or perhaps particularly since that election – the Oversight Board, despite its ceremonies, can be seen as a first step toward important yet contested cultural change. In addition, the fact that Meta increased its funding of the trust earlier than it needed to indicates a long-term and perhaps enhanced commitment to adjudicatory oversight. In contrast, other developments foreshadow the long journey still ahead for Emulated Guardians. Evolving such a culture, reflexively, requires institutions and rules to pave the way. The Board and the references to human rights in its decisions may be encouraging such cultural change from the outside. But they will not and cannot *compel* this change.

If, ten years from now, we consider the Oversight Board a failure (or do not recall it at all), it will likely not have been the Board's fault but the company's unwillingness to subject itself to adjudicative oversight. However, for the time being, impatience seems misguided. If the Board turns out to be the next chapter in the saga of adjudicative control over organizational power, assuming authoritative guardianship will take time. Nonetheless, it seems equally possible that the Board – even if it persists in its current form – will become yet one more quaint guardian-like experiment sidelined by more dynamic and powerful organizations.

In any event, I argue the Oversight Board should not be entirely forgotten, as emulation inevitably casts ripples. As we saw in the other case study, we already witness policy diffusion of aspects of the Oversight Board in Europe, most notably through the Appeals Centre Europe but arguably also in User Rights.[395] I suspect this trend will extend beyond content moderation, likely reappearing, *mutatis mutandis*, in future

[393] See esp. 'Statement of Frances Haugen' (n 96). See further for anecdotal examples of unsuccessful governance reforms at FIFA, Miguel Maduro and Joseph H H Weiler, '"Integrity," "Independence" and the Internal Reform of FIFA' in Arnout Geeraert and Frank van Eekeren, *Good Governance in Sport* (Routledge 2021).

[394] cf. Anja Bossow, Marie Müller-Elmau and Moritz Schramm, 'The Messiah and His Oligarchs' (*Verfassungsblog*, 1 November 2024) <https://perma.cc/55QN-HM2S>.

[395] See at 3.8.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004

governance frameworks. Anthropic, a leading artificial intelligence company, has established a so-called Long-Term Benefit Trust, which is supposed to infuse (some) public-interest considerations into both the company's management and the composition of its board. For future corporations and post-state power structures, someone might once again come up with the same, apparently simple, idea to emulate individual rights adjudication or other canonical ideas of public law.

4.7 Conclusion

This chapter examined the genesis and practice of the Oversight Board, a privately established Emulated Guardian that was set up ostensibly to control the world's largest social network's rulemaking and rule-enforcement efforts. It looked at the ways in which lawyers and juridical concepts shaped the Board's creation as an adjudicatory institution, and examined its institutional structure and personnel. We have seen that the dominance of lawyers among Board members and staffers enabled the Board to attempt to incrementally overcome some of its in-built handicaps. Among these are a complex relationship between, on the one side, an able, impartial group of members and staffers and, on the other, several backdoor channels used by Meta to influence the Board's adjudicative practices, which eventually undermined the Board's hoped-for independence. Crucially, the Board turned to international human rights law as an external normative yardstick against which to measure at least some of Meta's internally devised rules. Although the Board's attempts to infuse public-law principles into a private-law context represent sometimes bold challenges to the status quo, for example when developing a doctrine of "public emergency," its recommendations on the basis of these "infusions" often end up not being verifiably implemented by Meta. Nevertheless, the Board's skillful adaptation of juridical personnel, argumentation, practices, and reasoning might indicate that its incremental build-up of authority has, perhaps, only just begun.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 18:19:46, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009639750.004