

DOI 10.5817/MUJLT2025-2-3

OVERSIGHT BOARD – AN EFFECTIVE TOOL TO PROTECT FREE SPEECH? *

by

TAMÁS PONGÓ [†] RENÁTA KÁLMÁN [‡] RÉKA BRIGITTA SZANISZLÓ [§] NORBERT TRIBL [¶]

Meta has more than two billion daily active users, two and a half billion posts a day, and more than a million reports of violations of policies and guidelines every day. These large numbers justify the establishment of a body that goes beyond the usual content moderation and reporting system. Recognizing that its role in online speech cannot be managed solely through algorithmic decision-making and reporting, Meta has introduced a non-judicial grievance mechanism, the Oversight Board, which delivers its decisions solely in light of Meta's policies, values, and guidelines. The OB is committed to the Business and Human Rights regime, and in line with this commitment, the UNGPs are the governing international human rights law standard. This article cross-referenced the OB goals defined by its Charter, the UNGPs standards relevant to the goals of the OB, the statistics of the operation of the OB, and the doctrinal background of these goals, in particular, the accessibility, predictability, and transparency prongs as elements of an effective remedy. Through this in-depth analysis, we aim to explore whether the OB fulfills its predefined goals,

* The research was supported by the Digital Society Competence Centre of the Humanities and Social Sciences Cluster of the Centre of Excellence for Interdisciplinary Research, Development and Innovation of the University of Szeged. The authors are members of the Digitalization and Democracy research group.

[†] Tamás Pongó, Ph.D., senior lecturer, University of Szeged, Faculty of Law and Political Sciences, Institute of Public Law. E-mail: pongo.tamas@szte.hu

[‡] Renáta Kálmán, Ph.D., senior lecturer, University of Szeged, Faculty of Law and Political Sciences, Institute of Public Law. E-mail: kalman.renata@juris.u-szeged.hu

[§] Réka Brigitta Szaniszló, Ph.D., senior lecturer, University of Szeged, Faculty of Law and Political Sciences, International and Regional Studies Institute. E-mail: szaniszló.reka.brigitta@szte.hu

[¶] Norbert Tribl, Ph.D., senior lecturer, University of Szeged, Faculty of Law and Political Sciences, International and Regional Studies Institute. E-mail: tribl.norbert@szte.hu

and thus provides an effective non-judicial grievance mechanism regarding freedom of speech in the online environment.

KEY WORDS

Freedom of Speech, UNGPs, Oversight Board, Effectiveness, Accessibility, Transparency, Predictability

1. INTRODUCTION

“The purpose of the board is to protect free expression by making principled independent decisions about important pieces of content and by issuing policy advisory opinions on Meta’s content policies. The board will operate transparently and its reasoning will be explained clearly to the public, while respecting the privacy and confidentiality of the people who use Meta Platforms, Inc.’s services, including Facebook and Instagram (collectively referred to as “Meta”). It will provide an accessible opportunity for people to request its review and be heard.”¹

The Oversight Board’s Charter defines the framework for establishing the body and outlines its goals.² This unique institution, referred to as the Oversight Board (OB or Board), began operating in October 2020 to reform content moderation. We have found that the OB defined three specific elements to achieve the protection of free speech in the online environment, namely (1) transparency; (2) plain language; (3) access to remedy, i.e., users have the right to be heard.

The scope of this article is deliberately limited: we assess only three criteria of effective non-judicial grievance mechanisms under the UNGPs: (1) accessibility, (2) predictability, and (3) transparency. Other elements, such as plain language communication, while relevant, fall outside the scope of this study and would merit a separate analysis.

The OB is committed to the Business and Human Rights regime (BHR regime) and international human rights law (IHRL) as its Corporate Human Rights Policy (HRP) states. In line with this commitment, the UNGPs are the governing IHRL standard. According to the UNGPs, the OB categorizes itself as an independent non-judicial grievance mechanism. As such, the UNGPs’ effectiveness criteria for non-judicial grievance mechanisms should be used as a starting point to evaluate the Board’s decision-making mechanism.³

¹ Oversight Board. (2023) *Oversight Board Charter*. Available from: <https://www.oversightboard.com/wp-content/uploads/2023/11/3427086457563794.pdf>, p. 2.

² Oversight Board. (2023) *op. cit.* p. 13.

³ United Nations Human Rights Council. (2011) *UN Guiding Principles on Business and Human Rights*. [online] Available from: https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinessshr_en.pdf [Accessed 31 October 2024].

Therefore, this study cross-referenced the goals of OB defined by its Charter, the UNGPs standards relevant to the goals of the OB, the statistics of the operation of the OB, and the doctrinal background of these goals, in particular, the accessibility, predictability, and transparency prongs as elements of an effective remedy.

Through this in-depth analysis, we aimed to explore whether the OB fulfills its predefined goals and thus provides an effective grievance mechanism regarding freedom of speech in the online environment. This article addresses the following research questions:

RQ1: Is there a statistically significant difference in the likelihood of case selection between cases submitted by Meta and those submitted by users?

RQ2: Does the source of a case (Meta vs. user) influence the likelihood of the Board overturning Meta's original decision?

RQ3: Is there a disproportionate concentration of overturned or restored content in specific content categories (e.g., "dangerous individuals and organizations", "hate speech")?

The following sections proceed as follows: Part 2 outlines the BHR regime and its relevance to the OB; Part 3 situates the OB within IHRL; Part 4 presents our empirical statistical findings; Part 5 and 6 discuss the theoretical framework of effective redress; Part 7 assesses OB performance against these criteria; and Part 8 offers recommendations.

2. OVERVIEW OF THE BUSINESS AND HUMAN RIGHTS REGIME

In order to understand the importance and role of the OB, we shall briefly explore the framework in which it operates. As a private actor that delivers decisions on public law matters (free speech limitation), the existence of this body represents another step in the development of a global regime, of which the UN and the OECD are the driving forces.

The United Nations established the BHR regime through a several-year-long evolution. First, the United Nations Secretary-General established a position to study the human rights responsibility of businesses, which led to the appointment of John Ruggie as the UN Secretary-General's Special Representative on business and human rights in 2005.⁴ In 2008, the UN adopted the "Protect, Respect and Remedy" Framework, later operationalized through the UN Guiding Principles on Business and Human Rights (UNGPs) developed under John Ruggie's mandate. These Principles,

⁴ Wettstein, F. (2012) CSR and the Debate on Business and Human Rights: Bridging the Great Divide. *Business Ethics Quarterly*, 22 (4), pp. 739-770, p. 739.

while non-binding, have become a global reference point for corporate human rights responsibilities.⁵

The UNGPs system has three pillars: “the state duty to protect against business-related human rights infringements, the corporate responsibility to respect human rights, and increased access to remedy for victims of business-related human rights infringements.”⁶ Regarding the ICT sector and social media platforms, pillars two and three are the most relevant in our study since they should apply a corporate responsibility focusing on human rights protection and provide an accessible forum for redress. Basically, pillar two introduced the human rights due diligence method (HRDD) that requires businesses to identify, manage, and monitor adverse human rights impacts caused by their activities or any other activity connected to their operations.⁷ More precisely, HRDD covers the human rights violations caused by the company, or to which the company contributed, or which are directly linked to the company and its business relationships (UNGPS Princ. 17.(a)). When such an adverse human rights impact occurs, the business enterprise should use its leverage to prevent or manage these impacts. Leverage in this context refers to “the influence a business enterprise has with an entity or entities involved in a human rights abuse...”⁸ However, HRDD seems to be an umbrella term that involves human rights risk assessment and human rights impact assessment. Risk assessment focuses on avoiding or mitigating the negative outcomes a company may face, meanwhile, impact assessment requires strategic planning not just to avoid the adverse impacts, but to prevent them

⁵ Ruggie, J.G. (2014) Global Governance and “New Governance Theory”: Lessons from Business and Human Rights. *Global Governance*, 20 (1), pp. 5-18., p. 11.; Buhmann, K. and Olivera, R (2020) Human rights and social media platforms: the corporate responsibility to respect human rights in regard to privacy infringement involving photo posting. *Australian Journal of Human Rights*, 26 (1), pp. 121-141., p. 127.; Wettstein, F. (2012) *op. cit.*, p. 741.; Kun, A. (2017) Kellő gondosság alapú belső átvilágítás a munkajogban, mint a transznacionális vállalatok magatartásának új zsinórmértéke. *Miskolci Jogi Szemle*, 12 (2), pp. 285-297., p. 288.; Ruggie, J. G. (2007) Business and Human Rights: The Evolving International Agenda. *American Journal of International Law*, 101 (4), pp. 819-840.; Chronowski, N. (2023) Az alkotmány/bíraskodás potenciálja az üzleti szféra okozta emberi jogsérelmekkel szemben. *Alkotmánybírósági Szemle*, (1), pp. 2-14., p. 4.; Pap, L. A. Chronowski, N. and Nemessányi, Z. (2024) Corporate Human Rights Responsibility in Illiberal Regimes: The Example of the Ukrainian Refugee Crisis in Hungary. *Business and Human Rights Journal*, (9), pp. 32-53., p. 50 cites that France, Germany and Norway adopted laws and Switzerland put it on the agenda, unsuccessfully.

⁶ Buhmann, K. and Olivera, R (2020) *op. cit.*, p.125.

⁷ Pap, L. A. Chronowski, N. and Nemessányi, Z. (2024) *op. cit.*, p.45.

⁸ Buhmann, K. and Olivera, R (2020) *op. cit.*, p. 129.

and explore the opportunities in the course of establishing a strategy.⁹ Some scholars think that only risk assessment was included in the UNGPs.¹⁰

Nonetheless, the UNGPs influenced the OECD and the European Union to adopt the human rights due diligence approach. The OECD Guidelines for Multinational Enterprises on Responsible Business Conduct (updated 2023 version) highlights in the General Policies part that enterprises should “respect the internationally recognized human rights of those affected by their activities.”¹¹ In this part (II.A.11-13.), risk-based due diligence to identify, prevent, and mitigate adverse impacts is recommended, and such an approach aligns with the UNGPs standards, since it covers impacts that are caused by, or to which the company contributed, or directly linked to its operations, products or services. According to the Commentary, “due diligence is understood as the process through which enterprises can identify, prevent, mitigate and account for how they address their actual and potential adverse impacts as an integral part of business decision-making and risk management systems.”¹² The OECD Guidelines thereby strengthen pillars two and three of the UNGPs and thus intend to establish a global standard for human rights due diligence among business enterprises. Also, it acknowledges that the nature and extent of due diligence are affected by the specific circumstances of the given case and company, such as the size of the company, the severity of the adverse impact, and the involvement of the given enterprise.¹³ These statements are being reinforced in Chapter IV. Human Rights part, which was a result of Ruggie’s mandate and the adoption of the UN Framework 2008.¹⁴ In addition to the updated MNE Guidelines, the OECD issued the Due Diligence Guidance for Responsible Business Conduct, which seeks to help enterprises implement the UNGPs.¹⁵ “This Guidance was overseen by the OECD Working Party on Responsible Business Conduct (WPRBC) and involved a multi-stakeholder process with OECD and non-OECD countries and representatives from business, trade unions and civil society.”¹⁶ Even though the OECD adopted sector-specific guidelines for several fields, such as minerals, agriculture, garment, or footwear, no specific rules are available for social

⁹ Salcito, K. and Wielga, M. (2018) What does Human Rights Due Diligence for Business Relationships Really Look Like on the Ground?. *Business and Human Rights Journal*, (3), pp. 113-121., p. 114-115.

¹⁰ Salcito, K. and Wielga, M (2018) *op. cit.*, p. 115.

¹¹ OECD. (2008) *OECD Guidelines for Multinational Enterprises on Responsible Business Conduct*, OECD Publishing, II. General Policies A.2.

¹² OECD. (2008) *op. cit.* Commentary 15.

¹³ OECD. (2008) *op. cit.* Commentary 19.

¹⁴ Ruggie, J. G. (2014) *op. cit.*, p. 11.

¹⁵ OECD. (2018) *Due Diligence Guidance for Responsible Business Conduct*, Foreword.

¹⁶ *Ibid.*

media companies. Therefore, we need to look for a sector-specific toolkit or guidelines. Since our research focuses on Meta, we explored its Corporate Human Rights Policy, which referred to the Global Network Initiative (GNI) in addition to the UNGPs.¹⁷ The GNI is a multistakeholder framework with a diverse membership that issued its General Principles on Freedom of Expression and Privacy. The preamble of the GNI Principles declares that the Principles are based on the core international human rights conventions and are informed by the UN Protect, Respect, and Remedy Framework, the UNGPs, and the OECD Guidelines.¹⁸ Therefore, we can discover a BHR regime established by the UN, strengthened by the OECD, and reinforced by a technology-specific set of rules under the GNI principles with particular attention to freedom of expression and privacy. In addition to the Principles, the GNI adopted Implementation Guidelines, which seek to provide help for companies to put the Principles into practice.¹⁹ The Guidelines encourage companies to carry out HRDD in compliance with the UNGPs.²⁰ Also, it provides examples when a human rights impact assessment is needed as part of the HRDD.²¹

As discussed previously, the UNGPs were influential both at the global and regional level, and as a result, the EU has adopted two directives to strengthen the BHR regime in Europe, and we can see that through these directives the EU is trying to move the HRDD rules from the soft law to the hard law category.

The *Corporate Sustainability Reporting Directive* (CSRD)²² was adopted after five years of debate. The Directive was formally signed on 14 December 2022 and entered into force on 5 January 2023.²³ CSRD pivots on three design choices that connect it directly to the UNGPs logic: (i) double materiality and value-chain perimeter: companies must disclose

¹⁷ Meta (2021). Meta Corporate Human Rights Policy. [online] Available from: <https://about.fb.com/wp-content/uploads/2021/03/Facebooks-Corporate-Human-Rights-Policy.pdf> [Accessed 31 October 2024].

¹⁸ Global Network Initiative Principles. Available from: <https://globalnetworkinitiative.org/gni-principles/> [Accessed 30 April 2025].

¹⁹ Global Network Initiative. (2025) *Global Network Initiative Implementation Guidelines*. [online] Available from: <https://globalnetworkinitiative.org/implementation-guidelines/> [Accessed 30 April 2025].

²⁰ Global Network Initiative. (2025) *op. cit.* 2.4.

²¹ Global Network Initiative. (2025) *op. cit.* 2.6.

²² Cf. Hummel, K. and Jobst, D. (2024) An Overview of Corporate Sustainability Reporting Legislation in the European Union. *Accounting in Europe*, 21 (3), pp. 320–355.

²³ Directive of the European Parliament and of the Council of 14 December 2022 Amending Regulation (EU) No 537/2014, Directive 2004/109/EC, Directive 2006/43/EC and Directive 2013/34/EU, as regards corporate sustainability reporting. *Official Journal of the European Union* (2022/2464) 16 December. Available from: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32022L2464> [Accessed 30 April 2025].

both outward human-rights/environmental impacts and financial risks and opportunities;²⁴ (ii) European Sustainability Reporting Standards (ESRS):²⁵ ESRS S1–S4 oblige undertakings to describe their HRDD process, echoing UNGP Principle 17; and finally (iii) governance guarantees – mandatory external assurance and a director’s duty to oversee the sustainability statement strengthen accountability mechanisms.²⁶

Another pillar of EU regulation is the so-called *Corporate Sustainability Due Diligence Directive* (CSDDD)²⁷, which could be considered as a hard-law counterpart to the UNGPs in Europe.²⁸ The CSDDD was adopted on 13 June 2024 and entered into force on 25 July 2024.²⁹ The CSDDD – in Articles 5–11 – codifies the core UNGP HRDD cycle: identify, prevent, mitigate, and bring to an end adverse human-rights and environmental impacts in own operations, subsidiaries, and established business relationships. It should be noted that, unlike the UNGPs’ “should” language, CSDDD uses the legally binding “shall”.

While the UNGPs and OECD Guidelines articulate expectations, the CSRD and CSDDD embed those expectations in EU company law. HRDD thus shifts from reputational best practice to a compliance imperative with audit trails, enforcement by national supervisory authorities, and, under CSDDD, civil liability exposure. Besides, CSRD creates a transparency pull:³⁰ companies must publicly describe HRDD policies, processes, and outcomes.³¹ Nevertheless, CSDDD creates a substantive push: companies must perform HRDD to an adequate standard.³² We can say that together the two directives operationalise Pillars II and III of the UNGPs within the EU single market.

²⁴ Cf. Recital 27, and Article 19a of Directive 2022/2464

²⁵ Cf. Preamble, and Chapter 6a of Directive 2022/2464

²⁶ Cf. Recital 62, Article 19a, and Article 34 of Directive 2022/2464

²⁷ Directive of the European Parliament and of the Council of 13 June 2024 On corporate sustainability due diligence and amending Directive (EU) 2019/1937 and Regulation (EU) 2023/2859. *Official Journal of the European Union* (2024/1760) 5 July. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202401760 [Accessed 30 April 2025].

²⁸ Cf. Buen, N. et al. (2024) The EU Directive on Corporate Sustainability Due Diligence (CSDDD): The Final Political Compromise. *Business and Human Rights Journal*, 9 (2), pp. 294–300, p.296.

²⁹ The detailed analysis see: Ciacchi, S. (2024) The newly-adopted Corporate Sustainability Due Diligence Directive: an overview of the lawmaking process and analysis of the final text. *ERA Forum*, 25, pp. 29–48.

³⁰ Cf. Operato, L. et al. (2025) Navigating CSRD reporting: Turning compliance into sustainable development with science-based metrics. *Environmental Development*, 54.

³¹ Cf. Recital 30, Article 19a and Article 29a of Directive 2022/2464

³² Cf. Article 1, Article 5, Articles 7 to 16 and Recital 19 of Directive 2024/1760

Both directives adopt the UNGPs' concept of leverage over business relationships, but CSDDD gives it legal force by requiring termination or suspension (as a last resort) and contractual cascading of due diligence clauses.³³ By the beginning of 2027, for the largest EU and non-EU companies (and expanding in 2029), undertakings must actively carry out human rights due diligence: they must identify, prevent, mitigate, and, where necessary, remedy adverse impacts in their operations, their subsidiaries, and their established business relationships.³⁴ National supervisors can impose significant penalties for non-compliance, and victims gain a civil cause of action where harm was reasonably preventable and a causal link exists.

The UNGPs and the OECD Guidelines are soft-law instruments: they articulate what companies should do, but compliance ultimately rests on reputational incentives and peer pressure. By contrast, the European Union has converted the same expectations into binding law through two directives of CSRD and CSDDD. It should be noted that the CSDDD introduces a substantive duty of conduct as well. In short, the CSRD pulls companies toward better practice by exposing gaps to public scrutiny, while the CSDDD pushes them by making adequate due diligence a legal obligation backed by sanctions and liability. Together, they transform the voluntary standards of the UNGPs and OECD Guidelines into an enforceable European framework that strengthens access to remedy and explicitly links corporate human-rights management to climate-transition objectives.

This analysis indicates that Meta is committed to the BHR regime according to its Corporate Human Rights Policy, which encompasses the framework introduced above. Accordingly, our research on the OB as an effective grievance mechanism focuses on the UNGPs' effectiveness criteria since the UNGPs are the cornerstone of the current BHR regime, and Meta is committed to them as we introduced in Part 3. Although EU hard law rules do not yet prevail globally, it is likely that Meta will need to comply with them in the coming years to maintain its presence in the European market. Having outlined the BHR regime as the global framework, the next section situates the Oversight Board within international human rights law to clarify its institutional position.

³³ Cf. Article 10 and Article 11 of Directive 2024/1760

³⁴ Cf. Article 1(1)(a), Article 5(1), and Article 37(1)(a)-(c) of Directive 2024/1760

3. ESTABLISHMENT OF THE OVERSIGHT BOARD AND ITS CONNECTION TO IHRL

Meta has more than two billion daily active users,³⁵ two and a half billion posts a day, and more than a million reports of violations of policies and guidelines every day.³⁶ These large numbers justify the establishment of a body that goes beyond the usual content moderation and reporting system. Recognizing that its role in online speech cannot be managed solely through algorithmic decision-making and reporting, Facebook has introduced the OB, which delivers its decisions solely in light of Facebook's policies and values.

The establishment of the OB dates back to January 2018, when Noah Feldman wrote a brief note saying that Facebook needed a body to address the platform's content moderation issues.³⁷ *"Feldman, however, had something others didn't: he was close friends with Facebook COO Sheryl Sandberg. Feldman gave the memo to Sandberg, who forwarded it to Zuckerberg."*³⁸ Zuckerberg supported the idea and began the preparation for the establishment.³⁹ Highlighting its importance in the early stages, the Facebook founder compared the OB with the Supreme Court of the United States.⁴⁰ As a next step, the draft Charter was published in January 2019,⁴¹ which was followed by a global public consultation in June 2019. As a result of this process, the OB became fully operational in October 2020. Klonick has described this institutional creation as an unprecedented experiment in platform self-governance⁴², while Douek cautioned that procedural design choices would ultimately determine whether independence was more than symbolic.⁴³ Our findings on case-selection disparities, discussed in Part 4, provide empirical context for this scholarly debate.

³⁵ In December 2023, an average of 2.11 billion daily active users, and 3.98 billion monthly active users. Meta. (2023) *Meta Annual Report*. [online] Available from: https://www.annualreports.com/HostedData/AnnualReports/PDF/NASDAQ_FB_2023.pdf [Accessed 31 October 2024].

³⁶ Douek, E. (2019) Facebook's „Oversight Board:“ Move Fast With Stable Infrastructure and Humility. *North Carolina Journal of Law & Technology*, 21 (1). pp. 1-77.

³⁷ Klonick, K. (2020) The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression, *Yale Law Journal*, 129 (8). p. 2449.

³⁸ Ibid.

³⁹ Zuckerberg, M. (2021). *A Blueprint for Content Governance and Enforcement*. [Facebook] 5 May. Available from: https://www.facebook.com/notes/751449002072082/?hc_location=ufi [Accessed 31 October 2024].

⁴⁰ Klein, E. (2018) *Mark Zuckerberg on Facebook's hardest year, and what comes next*. [online] Vox. Available from: <https://www.vox.com/2018/4/2/17185052/mark-zuckerberg-facebook-interview-fake-news-bots-cambridge> [Accessed 31 October 2024].

⁴¹ Clegg, N. (2019) *Charting a Course for an Oversight Board for Content Decisions*. [Facebook] 28 January. Available from: <https://about.fb.com/news/2019/01/oversight-board/> [Accessed 31 October 2024].

⁴² Klonick, K. (2020) *op. cit.* p. 2449.

⁴³ Douek, E. (2019) *op. cit.* pp. 1-77.

This initiative was in line with the increased commitment to human rights of Facebook, which strives to respect and apply IHRL in its operations, in particular the UNGPs.⁴⁴ In Meta's Human Rights Policy (HRP), it is explicitly stated that "[we] implement our commitment to human rights using approaches set out in the United Nations Guiding Principles on Business and Human Rights (UNGPs). These approaches include... (3) providing access to remedy."⁴⁵ In (3) of the HRP, it is clearly expressed that „the Oversight Board is intended to increase access to remedy for Facebook and Instagram users, and other relevant rights holders. When designing the Oversight Board, we reviewed its proposed charter, bylaws, and operations against the effectiveness criteria for operational grievance mechanisms and access to remedy set out in the UN Guiding Principles on Business and Human Rights.”⁴⁶

The OB considers the UNGPs to be its guidelines. In February 2022, in the Submission to the Office of the High Commissioner for Human Rights, United Nations⁴⁷ on the practical application of the UNGPs to the activities of technology companies, the OB emphasized that “in its corporate human rights policy, Meta affirmed that the Board was conceptualized in line with the effectiveness criteria of operational grievance mechanisms and access to remedy under the UN Guiding Principles on Business and Human Rights (UNGPs). In its submission, the OB notes as well that “the Oversight Board has described itself in several decisions as an independent non-judicial grievance mechanism,⁴⁸ in addition to citing to the UNGPs as part of its framework for assessing Meta's compliance with its human rights responsibilities.”⁴⁹ Mazúr and Gramblichkova similarly note that the OB's self-characterization sits uneasily with its dependency on Meta for case referrals and implementation.⁵⁰ By situating our statistical results within this institutional context, we extend their critique with quantitative evidence.

⁴⁴ Sissons, M. (2021) *Our Commitment to Human Rights*. [Facebook] 16 March. Available from: <https://about.fb.com/news/2021/03/our-commitment-to-human-rights/> [Accessed 31 October 2024].; Meta. (2021) *Corporate Human Rights Policy*. [online] Available from: <https://about.fb.com/wp-content/uploads/2021/03/Facebooks-Corporate-Human-Rights-Policy.pdf> [Accessed 31 October 2024].

⁴⁵ Meta. (2021) *op. cit.*

⁴⁶ Ibid.

⁴⁷ Oversight Board. (2022) *Facebook Oversight Board Submission to the Office of the High Commissioner for Human Rights, United Nations on the practical application of the UNGPs to the activities of technology companies, February 2022*. [online] Available from: <https://www.ohchr.org/sites/default/files/2022-03/Oversight-Board.pdf> [Accessed 31 October 2024].

⁴⁸ Facebook Oversight Board Case Decisions: 2020-006-FB-FBR (Claimed COVID Cure); 2021-004-FB-UA (Pro-Navalny Protests in Russia); 2021-005-FB-UA (“Two Buttons” Meme); 2021-006-IG-UA (Ocalan's Isolation); 2021-007-FB-UA (Myanmar Bot); 2021-001-FB-UA (South Africa Slur); See also Oversight Board. (2022) *op. cit.*

⁴⁹ Oversight Board. (2022) *op. cit.*

⁵⁰ Mazúr, J. and Gramblichkova, B. (2023) New Regulatory Force of Cyberspace: The Case of Meta's Oversight Board. *Masaryk University Journal of Law and Technology*, 17 (1), pp. 3-32.

When assessing the effectiveness of the OB decision-making mechanism, the first challenge concerns the Board's positioning within the IHRL protection system. Opinions praising the work of the OB emphasize⁵¹ its functioning as a judicial forum, comparable to the US Supreme Court, with its own precedent system, but its decisions in reality are difficult to interpret in the context of genuine judicial protection. Nevertheless, the OB represents an important institutional innovation and a notable attempt to constrain the near-unlimited power of Meta as a platform provider over online discourse. The OB aims to enhance the transparency and legitimacy of Meta's decision-making⁵² by facilitating dialogue about how and why the platform acts in this sphere. Although this objective is more modest than establishing an independent (external) source of law with general powers and thus a normative definition of freedom of expression, it remains ambitious for an institution operating in uncharted territory, given the absence of a preexisting model for classifying such a body.⁵³

Therefore, when examining the activity of OB and the norms and expectations that govern it, it is worth starting from the nature of the category in which it categorizes itself: an independent⁵⁴ non-judicial grievance mechanism. As such, the UNGPs' effectiveness criteria for non-judicial grievance mechanisms, which the OB considers to be its own, should be used as a starting point when assessing its decision-making mechanism.⁵⁵ Based on Principle 31, to ensure their effectiveness, non-judicial grievance mechanisms should be: (1) legitimate, (2) accessible, (3) predictable, (4) equitable, (5) transparent, (6) rights-compatible, and (7) a source of continuous learning. This framing aligns with Ruggie's view that operational mechanisms must be judged on both design and outcomes.⁵⁶ A point we operationalize in the statistical tests presented below.

This doctrinal background provides the foundation for our empirical investigation. The following section turns to statistical analysis of the OB's

⁵¹ Gradoni, L. (2021) *How platform governance had its Marbury v Madison. Constitutional Review via Facebook's Oversight Board*. [blog entry] 10 February 2021. Verfassungsblog. Available from: <https://verfassungsblog.de/fob-marbury-v-madison/> [Accessed 31 October 2024].

⁵² Mazúr, J. and Gramblichová, B. (2023) *op. cit.* p. 7.

⁵³ See also Douek, E. (2019) Facebook's „Oversight Board.“ Move Fast With Stable Infrastructure and Humility. *North Carolina Journal of Law & Technology*, 21 (1). pp. 1-78, pp. 76-77.

⁵⁴ The authors of this paper question the appropriateness of independence as an attribute used by the OB, but the criteria of independence are not the subject of this research.

⁵⁵ United Nations Human Rights Council. (2011) *UN Guiding Principles on Business and Human Rights*. [online] Available from: https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr_en.pdf [Accessed 31 October 2024].

⁵⁶ Ruggie, J. G. (2014) *op. cit.*

operation to test whether the theoretical commitments described are borne out in practice. In particular, we focus on whether the OB's grievance mechanism provides genuine accessibility, transparency, and predictability. Three essential criteria of effectiveness under the UNGPs. To assess this, we examine whether measurable patterns in the OB's operation reveal shortcomings in these areas. The result of our statistical analysis indicates that such concerns are well-founded.

4. OPERATION OF THE OVERSIGHT BOARD IN FIGURES

The Charter of the Oversight Board declares that *"people using Meta's services and Meta itself may bring forward content for board review."*⁵⁷ When users exhaust all appeal options within Meta, they may request the Oversight Board to review the content decision. Meta may refer cases it considers 'significant or difficult' to the Board, with 'significant' referring to real-world impact, and 'difficult' involving controversial or policy-sensitive issues: *[s]ignificant means that the content in question involves real-world impact and issues that are severe, large-scale, and/or important for public discourse [...], while difficult means the content raises questions about current policies or their enforcement, with strong argument on both sides for either removing or leaving up the content under review."*⁵⁸

Having clarified submission eligibility, we now examine the types of decisions the Board can issue and their governing criteria (see Table 1). The most common type is the standard case review. The case is reviewed by a panel of five board members and the aim is to issue a decision within 90 days from the referral by Meta or submission of an appeal by a user.⁵⁹ Although in standard case review the panel deliberates and decides cases, the full board approves the decisions.

In addition to the standard case reviews, there are special procedures that can only be initiated by Meta. These are the following: i) expedited review; ii) request for policy advisory opinions. *"In exceptional circumstances, including when content could result in urgent real-world consequences, Meta may send cases to the board with requests for the board to review Meta's decision under expedited review."*⁶⁰ The expedited review will be conducted within 30 days and will be voted on only by the panel.

⁵⁷ Oversight Board. (2023) *op. cit.* p. 6

⁵⁸ Oversight Board. (2023) *op. cit.* p. 22.

⁵⁹ Oversight Board. (2023) *op. cit.* p. 15.

⁶⁰ Oversight Board. (2023) *op. cit.* p. 22.

	Standard case review	Expedited review
Procedural time	90 days, which may be extended in exceptional circumstances or in the event of technical or operational incidents ⁶¹	maximum 30 days
Submitted by	users and Meta itself	only Meta
Decision maker	the panel makes the decision, but the full board approves	the panel decides without the approval of the full board

Table 1: Comparison of Standard and Expedited Oversight Board Case Review Processes

Meta may request policy advisory opinions from the Board, which is separate from individual case reviews. *“The policy advisory opinion may concern the clarification of a previous decision by the board or guidance on Meta’s content policies.”*⁶² Except for the policy advisory opinion, all decisions of the Board are binding on Meta, unless the implementation of the decision could violate the law.⁶³ The third type of case is summary decisions. *“Where Meta determines that the content in a particular case was incorrectly actioned by Meta and reverse its original decision, the case selection committee may choose to select the case for summary decision.”*⁶⁴ Similar to expedited review, in summary decisions, the panel votes without the approval of the full board. Since the operation of the OB, the board has made 37 summary decisions out of 98 cases.

Case selection is at the Board’s discretion, guided by the potential for precedent-setting or policy impact. The Board *“considers cases that have the greatest potential to guide future decisions and policies.”* The OB was not established to resolve all submitted cases, but to adjudicate those of greatest significance. It aims to make principled, independent decisions about important pieces of content.⁶⁵ The Board established its own set of procedures, which help its staff select a pool of cases from which it can choose those that are worth deciding. Every decision is based on Meta’s content policies and values. Any prior decision, where the facts, applicable policies, or other factors are substantially similar, has precedential value for each decision.⁶⁶

⁶¹ Oversight Board. (2023) *op. cit.* p. 15.

⁶² Oversight Board. (2023) *op. cit.* p. 9.

⁶³ Oversight Board. (2023) *op. cit.* p. 25.

⁶⁴ Oversight Board. (2023) *op. cit.* p. 23.

⁶⁵ Oversight Board. (2023) *op. cit.* p. 3.

⁶⁶ Oversight Board. (2023) *op. cit.* p. 7.

One of the main goals of the OB is „to provide an accessible opportunity for people to request its review and be heard.”⁶⁷ Since its establishment in October 2020, the Oversight Board has received millions of submissions – predominantly from users, though some have originated from Meta itself. To date, the Board has issued three annual reports; however, no annual report has been made available for 2024, despite it now being May 2025.⁶⁸ This prolonged reporting gap undermines institutional transparency and hampers meaningful evaluation of the Board’s performance and impact over time. This undermines ongoing public oversight and academic inquiry.

To evaluate whether the OB’s functions as an accessible and effective forum for remedy, as intended by its Charter, we conducted a statistical analysis of all publicly available decisions issued between October 2020 and 1 June 2024. This section addresses RQ1-RQ3 introduced in Part 1, presenting statistical tests that examine selection likelihood, decision outcomes, and content category concentration.

A chi-square test of independence was conducted to examine the relationship between the source of case submissions (Meta vs. users) and whether the cases were selected for review by the Oversight Board (see Figure 1). As previously indicated, since the 2024 Annual Report is still unavailable, here, we have used data from 2020 to 2023 for all cases submitted and for cases accepted for decision. The results revealed a statistically significant association between the source and selection outcome, $\chi^2(1, N = 2,841,680) = 81,886.56$, $p < .001$. While Meta submitted only 107 cases during the observation period, 16 (14.95%) were selected for review. In contrast, of the over 2.8 million user-submitted cases, just 67 (0.00235%) were accepted – revealing a stark disparity in selection likelihood.

⁶⁷ Oversight Board. (2023) *op. cit.* p. 3.

⁶⁸ Oversight Board. (2024) *Annual Transparency Reports*. [online] Available from: <https://www.oversightboard.com/transparency-reports/> [Accessed 30 April 2025]. Even though Meta has published bi-annual reports on the Oversight Board, its structure and focus are different than the Board’s Annual Report. The only relevant information from Meta’s report is that 2 out of 7 Meta case referrals were selected for review by the Board.

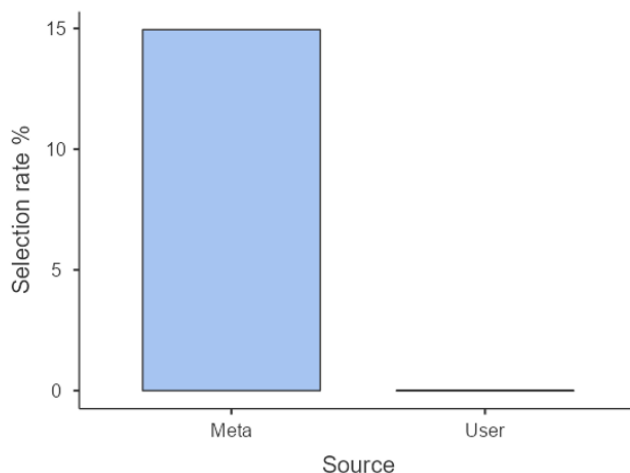


Figure 1: Selection Rates by Source of Submission

To further assess whether the outcome of a case – namely, whether the Board upheld or overturned Meta’s original decision – was associated with the source of submission (Meta vs. user), we conducted a chi-square test of independence (see Figure 2). The results revealed a statistically significant association between the source of the case and the Board’s decision, $\chi^2(1, N = 97) = 14.56, p < .001$.

Notably, Meta-submitted cases were upheld in 52.9% of instances, while user-submitted cases were upheld in only 12.5% of cases. Conversely, 87.5% of user-submitted cases resulted in the Board overturning Meta’s original decision, compared to just 47.1% of Meta-submitted cases. This disparity raises concerns that the Oversight Board may be more inclined to uphold Meta’s decisions when the referral originates from Meta itself, calling into question the consistency of its independent decision-making function.

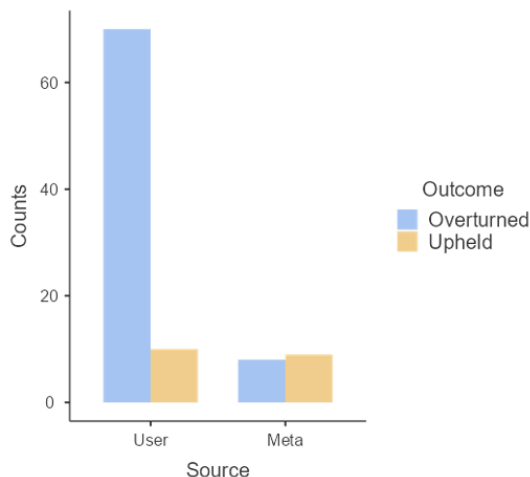


Figure 2: Proportion of Upheld and Overturned Decisions by Source of Submission

To further investigate the thematic focus of the Oversight Board's decisions, we conducted a chi-square goodness-of-fit test to determine whether the distribution of case categories significantly deviates from a uniform distribution (see Figure 3). This test assessed whether each content category – Nudity, Violence and incitement, Hate speech, Dangerous individuals and organizations, and Other – was represented equally among the publicly available decisions. The analysis yielded a statistically significant result, $\chi^2(4, N = 63) = 17.24$, $p = .002$, indicating that certain categories are disproportionately represented. Specifically, cases concerning “Dangerous individuals and organizations” constituted nearly 40% of all decisions, whereas categories such as “Nudity” and “Other” were considerably less prevalent. This uneven distribution may reflect either underlying differences in the frequency or severity of content moderation issues across categories or institutional priorities guiding the Board's selection process.

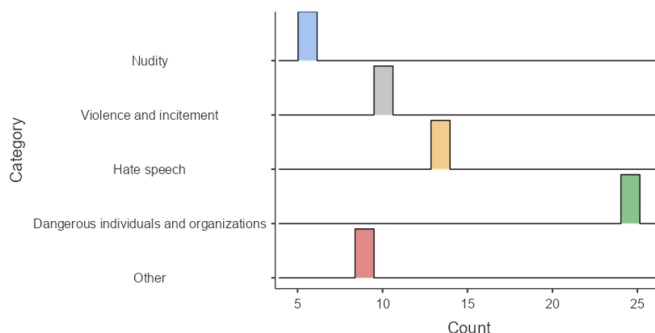


Figure 3: Distribution of Case Categories

These empirical results reveal disparities in accessibility and outcomes. We now relate these findings to the UNGPs’ effectiveness criteria to evaluate whether the OB’s decision achieves its stated goals.

5. CRITERIA OF EFFECTIVENESS IN THE UNGPS

Before proceeding with the analysis, it is necessary to provide an overview of the OB’s decision selection mechanism, set out in Article 2 of its Charter: “[t]he board has the discretion to choose which requests it will review and decide upon. In its selection, the board will seek to consider cases that have the greatest potential to guide future decisions and policies. [...] The board will establish its own set of procedures that its staff will use to select a pool of cases from which the board can choose. Once a case is selected, the submitting person, the person who originally posted the content, and Meta will be notified of the case selection.”

OB identifies potential cases from those referred to it by Meta or appealed to it by users and selects cases based on the following criteria: (1) the content’s critical importance to public discourse, (2) if the impugned content directly or indirectly affects a substantial number of individuals, (3) and/or the content raises significant questions about Meta’s policies.⁶⁹ Under its description, the OB selects cases that reflect the user base of Facebook and Instagram and ensures regional and linguistic diversity of cases.⁷⁰ According to the OB, “the cases the Board has selected reflect these considerations, with the majority of the cases affecting countries outside the US & Canada (notwithstanding that most appeals originate from this region), and several cases arising from situations of conflict or civil unrest occurring in Ethiopia, Israel/Palestine, and Myanmar. This focus reflects

⁶⁹ Oversight Board. (2022) *op. cit.*

⁷⁰ *Ibid.*

Principles 12, 23, and 24 of the UNGPs, which highlight the need for increased efforts to minimize human rights risks where the consequences are most severe."⁷¹

Therefore, based on Article 2 of the Charter, the OB develops its procedures to select the hundreds of thousands of applications⁷² that reach the decision-making mechanism of the Board. However, neither the exact selection mechanism nor the applications that did not pass the set eligibility filter are known.⁷³ A review of the reports published by the OB⁷⁴ provides detailed statistics on the decisions taken concerning the received cases and the principles according to which the decisions were made. However, it is not possible to examine the screening mechanism used to select the reports, nor is it possible to find out which reports were excluded from the admissible category, which therefore raises the question of the exhaustion of the requirements of effective legal protection.

*"Based on the UNGPs, a grievance mechanism can only serve its purpose if the people it is intended to serve know about it, trust it, and are able to use it. These criteria provide a benchmark for designing, revising, or assessing a non-judicial grievance mechanism to help ensure that it is effective in practice and – based on the commentary of UNGPs – poorly designed or implemented grievance mechanisms can risk compounding a sense of grievance amongst affected stakeholders by heightening their sense of disempowerment and disrespect by the process."*⁷⁵

In general, it is very difficult to define what is considered effective in the field of legal protection, given that effectiveness can only be determined in the light of the practice of the body implementing the legal protection and the examination of the rules that apply to it.⁷⁶ As a general principle, effectiveness is that *"law in general, and rights and obligations thereunder, should*

⁷¹ Ibid.

⁷² The Board has received over 800,000 appeals since it began accepting cases in October 2020 up to the end of September 2021. In the whole year the number of appeals were around 1 million in 2021, 1.3. million in 2022, and 400,000 in 2023. See also Oversight Board. (2024) *op. cit.*

⁷³ It should be noted - using the analogy with the constitution or supreme courts - that in the centralized Constitutional Court model, the procedure of the OB is most similar to that of a constitutional complaint, where there is a strict set of eligibility requirements and rejected applications are made public, together with the reasons for rejection. See also Váradi, Á. (2022) Access to constitutional complaint procedures: A real chance? Remarks on the availability of legal aid in constitutional court procedures in Austria, Germany and Hungary. *Hungarian Journal of Legal Studies*, 61 (4), pp. 372–385.

⁷⁴ See also Oversight Board. (2024) *op. cit.*

⁷⁵ United Nations Human Rights Council. (2011) *op. cit.*

⁷⁶ Arnall, A. (2011) The Principle of Effective Judicial Protection in EU law: An Unruly Horse? *European Law Review* (1). p. 51.

be effective rather than not, expressed sometimes in the maxim ut res magis valeat quam pereat."⁷⁷

We face even greater difficulties when it comes to defining what is effective in the field of human rights and business, specifically when it comes to defining the criteria for compliance with the UNGPs. Theories and concepts from norm scholarship can shed light on one of the most significant and persistent gaps in business and human rights research, which is the lack of understanding regarding the effectiveness of corporate human rights accountability norms and instruments. In general, effectiveness could be defined as comprising three dimensions: (1) acceptance and uptake by relevant actors in the field; (2) the extent of the behavioral change induced; and (3) intended as well as the unintended consequences of that behavioral change and interventions.⁷⁸

In our view, these dimensions presuppose the accessibility, predictability, and transparency concerned by the UNGPs' effectiveness criteria. It is as much as saying that accessibility, predictability, and transparency are prerequisites to achieving the triad of effectiveness. Having set out the UNGPs' effectiveness criteria, the next section examines how legal theory on access to justice and effective redress can sharpen our assessment of the OB.

6. ELEMENTS OF ACCESSIBLE AND EFFECTIVE REDRESS IN THEORY

Currently, the HRDD rules (Part 2) can be interpreted as a stand-alone field of law; however, there is a definite shift toward their recognition as hard law, particularly through EU directives. Furthermore, Meta acknowledged in its HRP both the diversity of legal frameworks across the jurisdictions in which it operates and its commitment to respect domestic laws.⁷⁹ Accordingly, we shall briefly highlight the fundamental requirements for effective judicial protection, in particular, the essential elements of an effective and accessible remedy applicable to states.

The limitations of the OB in terms of accessibility, effectiveness, and transparency – identified through Meta's internal data and assessed against the UNGPs – necessitate a broader theoretical examination. Legal scholarship provides established criteria for assessing redress mechanisms, particularly regarding the barriers that hinder access to justice. Garth and Cappelletti's theory of "Effective Access" conceptualizes accessibility and effectiveness as

⁷⁷ Grant, J. P. and Barker, J. C. (2009) *Encyclopædic dictionary of international law*. 3rd ed. New York: Oxford University Press. p. 177.

⁷⁸ Wolfstetter, R. and Li, Y. (2022). Business and Human Rights Regulation After the UN Guiding Principles: Accountability, Governance, Effectiveness. *Human Rights Review* 23 (1). p. 8.

⁷⁹ Meta. (2021) *op. cit.*

interdependent, impeded by factors such as the cost of litigation, disparities in party capabilities, and the structural difficulties associated with diffuse interests.⁸⁰ These factors reveal systemic inequalities that compromise procedural fairness and obstruct meaningful redress.

Central to any normative account of justice is the work of John Rawls, whose theory emphasizes fairness, equality of opportunity, and the moral duty of institutions to be structured in a way that benefits the least advantaged. From this perspective, accessibility is not merely about procedural entry points but about ensuring substantive inclusion. A redress system that fails to account for socioeconomic, linguistic, or cultural disparities cannot be considered just.⁸¹ Building on Rawls, scholars such as Deborah Rhode⁸² and Fernand de Varennes⁸³ argue for the necessity of inclusive legal systems – ones that incorporate multilingual access, protections for marginalized communities, and mechanisms to ensure equal standing before the law.

Effectiveness in redress extends beyond resolution; it requires enforceability, deterrence, and procedural timeliness. Without practical enforcement, redress remains theoretical, as emphasized by Owen Fiss⁸⁴ and Richard Posner⁸⁵. The timely administration of justice, a point underscored by E. Allan Lind, Tom R. Tyler⁸⁶, is critical both for remedying harm and for preserving the legitimacy of the redress mechanism. Delays not only exacerbate the experience of injustice but also risk diminishing trust in the legal system itself.

Trust, while difficult to quantify, constitutes a foundational element of an effective redress process. It is cultivated through transparency, public awareness, confidentiality, and responsiveness to feedback. Transparency in decision-making, clear articulation of procedures, and adaptive mechanisms – particularly in the context of digital grievance systems like the OB – are essential. In alignment with Rawlsian principles, trust must be seen as a

⁸⁰ Garth, B. G. and Cappelletti, M. (1978) Access to Justice: The Newest Wave in the Worldwide Movement to Make Rights Effective. *Buffalo Law Review* 27. pp. 181-292.

⁸¹ Rawls, J. (1971) *A Theory of Justice*. Cambridge: Harvard University Press.

⁸² Rhode, D. L. (2004) *Access to Justice*. Oxford: Oxford University Press.

⁸³ de Varennes, F. (2001) Language Rights as an Integral Part of Human Rights. *International Journal on Multicultural Societies* 3 (1). pp. 15-25; de Varennes, F. (2012) Language as Right in International Law: Limits and Potentials. In: Dagmar Richter et al. (eds.). *Language Rights Revisited – The Challenge of Global Migration and Communication*. Berlin: Wolf Legal Publishers. pp. 43-52.

⁸⁴ Fiss, O. M. (1979) The Civil Rights Injunction. *The University of Chicago Law Review* 47 (1). pp. 199-211.

⁸⁵ Posner, R. A. (2011) *Economic Analysis of Law*. New York: Aspen Publishers.

⁸⁶ Lind, E. A. and Tyler, T. R. (1988) *The Social Psychology of Procedural Justice*. New York: Springer New York.

reflection of institutional legitimacy and moral accountability, both of which are indispensable to ensuring that redress systems are genuinely accessible, inclusive, and just. Building on this understanding, the next step is to examine whether the OB, as a non-judicial grievance mechanism, satisfies the prerequisites of accessibility, predictability, and transparency, which form the basis of the triad of effectiveness set out by the UNGPs.

7. EFFECTIVENESS OF THE OB'S REDRESS MECHANISM

Considering the foregoing, we need to explore whether the OB, as a non-judicial grievance mechanism, met the requirements of accessibility, predictability, and transparency, which are the prerequisites of the triad of effectiveness under the UNGPs.

Our statistical findings suggest that OB is not accessible to users since 14,95% of cases submitted by Meta were selected to decide, meanwhile, only 0,00235% of user-submitted cases were decided by the OB. Consequently, a user has no real possibility of being heard by the OB. Also, the report form for users has a very limited space to elaborate on the issue (approximately 1000-2000 characters) and no attachment is allowed.⁸⁷

In addition to this, the OB has developed an internal case selection system with eligibility criteria, but such a system has not been published, thus, users cannot act accordingly. Due to the lack of publicly available eligibility requirements, users cannot predict the outcome of their case, which results in a non-transparent grievance procedure. Also, the rejected cases are not publicly available, thus, users cannot explore the grounds of rejection and adjust their conduct accordingly in the future.

Moreover, the names of panel members, who prepare the standard review process and decide on the expedited review, are also not published, and neither are the minority opinions. According to the Bylaws, this practice of anonymity is required to protect the safety and security of the Board members,⁸⁸ but it undermines the transparency of the decision-making. By contrast, some other non-judicial adjudicatory bodies balance safety concerns with transparency through partial disclosure. For example, the Internet Corporation for Assigned Names and Numbers (ICANN), the global body overseeing internet domain governance, publishes its final independent Review Process (IRP) declarations promptly and commits to redacting only

⁸⁷ Pánek, M. (2021) Appealing to Facebook's Oversight Board. Mmister [blog entry] 9 September. Available from: <https://mmister.com/appealing-to-facebook-oversight-board/> [Accessed 31 October 2024].

⁸⁸ Bylaws Article 1. Section 3. 3.1.3.

the minimal necessary details, demonstrating a high degree of procedural transparency.⁸⁹

When users send their report concerning content, they cannot adjust themselves to predefined eligibility criteria and predict the admissibility of their report. As a result, trust in the redress system and user confidence, which are elements of effectiveness as stated above in Part 6, are lacking. Such a tendency can be detected in the total number of reports, which significantly decreased from 1.290.942 (2022) to 398.597 (2023).

Furthermore, as the relative party capability suggested above by Garth & Cappelletti, the party with greater resources and capabilities will be favored. The fact that only 0,00235% of user-submitted cases were selected to decide, compared to the 14,95% of Meta-submitted cases, proves that the OB procedure does not meet the accessibility element of effectiveness. Moreover, only Meta can initiate an expedited review to achieve a faster decision, which further reduces the users' accessibility to the grievance mechanism.

In the practice of the Board, the intertwined nature of timeliness and enforceability, emphasized by Lind & Tyler above, is being realized both in the length of the procedures (90 days; 30 days) and the short implementation of Board decisions (7 days)⁹⁰. Even though these procedural deadlines are short, they are not always appropriate. For instance, in the Brazilian case of breast cancer, the content was removed in October and reinstated a few months later in compliance with the procedural deadlines, but the purpose of the content was to raise awareness of breast cancer in October, because that month was dedicated to this activity.⁹¹

In summary, analyzing the non-judicial grievance mechanism of the OB in light of both the UNGPs' effectiveness criteria and the academic literature, we found that even though under the Charter the OB will operate transparently and aims to provide an accessible opportunity for people to request its review and be heard, these goals are not being fully achieved. Taken together, these patterns suggest that the OB functions less as an equal-access grievance forum and more as a selective body whose decisions reinforce Meta's strategic interests. While this does not invalidate its value as an experiment in private governance, it stresses the gap between its stated Charter goals and its operational reality.

⁸⁹ ICANN. (2019) *ICANN Publication Practices for Board, Litigation, and Accountability Processes*. [online] Los Angeles: ICANN. Available from: <https://www.icann.org/en/board-activities-and-meetings/publication-practices> [Accessed 21 August 2025].

⁹⁰ Bylaws Article 3. 1.2.4.

⁹¹ Facebook Oversight Board Case Decision: 2020-004-IG-UA (Breast Cancer Symptoms & Nudity Case).

8. CONCLUSIONS AND SUGGESTIONS

Answers to research questions:

- RQ1: Our statistical analysis confirms a substantial disparity in case selection likelihood: 14.95% of Meta-submitted cases were selected compared to only 0.00235% of user-submitted cases;
- RQ2: Case source also correlates with decision outcomes: Meta-submitted cases were upheld in 52.9% of instances, whereas only 12.5% of user-submitted cases were upheld, with the remainder being overturned;
- RQ3: Overturned or restored content is disproportionately concentrated in three categories: dangerous individuals and organizations, hate speech and violence, and incitement. The first category alone represents nearly 40% of decisions.

In the course of the research, first, we explored the BHR regime, which serves as the global framework in this field, and the connection between the IHRL and the OB through Meta's HRP, and discovered the effectiveness criteria of a non-judicial grievance mechanism under the UNGPs. As a result, we found that accessibility, predictability, and transparency are prerequisites of the UNGPs' effectiveness criteria. Therefore, we formulated our research hypothesis that the grievance mechanism system of the OB is not accessible and has a significant negative impact on the transparency and predictability of the decision-making mechanism – and consequently the effectiveness – of the body. In order to prove or rebut our hypothesis we reviewed the case law of the OB in terms of numbers. Also, we examined the effectiveness criteria both under the UNGPs and the academic standpoints. Finally, we cross-referenced these three approaches to the Board and revealed that our hypothesis was proved since the current operation of the OB does not meet the accessibility, predictability and transparency requirements, which are prerequisites of an effective grievance mechanism. As a result of our research, we make a few recommendations to improve the operation of the OB and thereby help to achieve the goals defined in its Charter.

Our recommendations are:

- publish the eligibility criteria, thus users can act accordingly;
- publish the rejected cases with the reasoning of the rejection; establish a database for the rejected cases and make it possible to search within this database;
- abolish the anonymous decision-making process in all types of procedures;
- allow users to initiate the expedited review procedure.

Implementing these recommendations would improve the effectiveness of the OB, a unique institution within the field of business and human rights. While operational improvements remain necessary, the mere existence of the Board represents a significant step in the protection of free speech in the online environment. Moreover, it serves as a model that other companies could emulate to strengthen the human rights due diligence in an increasingly digitalized world.

LIST OF REFERENCES

- [1] Arnulf, A. (2011) The Principle of Effective Judicial Protection in EU law: An Unruly Horse? *European Law Review* (1). pp. 50-70.
- [2] Buen, N. et al. (2024) The EU Directive on Corporate Sustainability Due Diligence (CSDDD): The Final Political Compromise. *Business and Human Rights Journal*, 9 (2), pp. 294-300.
- [3] Buhmann, K. and Olivera, R (2020) Human rights and social media platforms: the corporate responsibility to respect human rights in regard to privacy infringement involving photo posting. *Australian Journal of Human Rights*, 26 (1), pp. 121-141.
- [4] Bylaws Article 1. Section 3. 3.1.3.
- [5] Bylaws Article 3. 1.2.4.
- [6] Chronowski, N. (2023) Az alkotmány/bírászkodás potenciálja az üzleti szféra okozta emberi jogsérelmekkel szemben. *Alkotmánybíróági Szemle*, (1). pp. 2-14.
- [7] Ciacchi, S. (2024) The newly-adopted Corporate Sustainability Due Diligence Directive: an overview of the lawmaking process and analysis of the final text. *ERA Forum*, 25, pp. 29-48.
- [8] Clegg, N. (2019) *Charting a Course for an Oversight Board for Content Decisions*. [Facebook] 28 January. Available from: <https://about.fb.com/news/2019/01/oversight-board/> [Accessed 31 October 2024].
- [9] de Varennes, F. (2001) Language Rights as an Integral Part of Human Rights. *International Journal on Multicultural Societies* 3 (1). pp. 15-25.

- [10] de Varennes, F. (2012) Language as Right in International Law: Limits and Potentials. In: Dagmar Richter et al. (eds.). *Language Rights Revisited – The Challenge of Global Migration and Communication*. Berlin: Wolf Legal Publishers. pp. 43-52.
- [11] Directive of the European Parliament and of the Council of 14 December 2022 Amending Regulation (EU) No 537/2014, Directive 2004/109/EC, Directive 2006/43/EC and Directive 2013/34/EU, as regards corporate sustainability reporting. *Official Journal of the European Union* (2022/2464) 16 December. Available from: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32022L2464> [Accessed 30 April 2025].
- [12] Directive of the European Parliament and of the Council of 13 June 2024 On corporate sustainability due diligence and amending Directive (EU) 2019/1937 and Regulation (EU) 2023/2859. *Official Journal of the European Union* (2024/1760) 5 July. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202401760 [Accessed 30 April 2025].
- [13] Douek, E. (2019) Facebook’s „Oversight Board:” Move Fast With Stable Infrastructure and Humility. *North Carolina Journal of Law&Technology*, 21 (1). pp. 1-77.
- [14] Facebook Oversight Board Case Decision 2020-004-IG-UA (Breast Cancer Symptoms & Nudity Case).
- [15] Facebook Oversight Board Case Decision 2020-006-FB-FBR (Claimed COVID Cure);
- [16] Facebook Oversight Board Case Decision 2021-001-FB-UA (South Africa Slur);
- [17] Facebook Oversight Board Case Decision 2021-004-FB-UA (Pro-Navalny Protests in Russia);
- [18] Facebook Oversight Board Case Decision 2021-005-FB-UA (“Two Buttons” Meme);
- [19] Facebook Oversight Board Case Decision 2021-006-IG-UA (Ocalan’s Isolation); 2021-007-FB-UA (Myanmar Bot);
- [20] Fiss, O. M. (1979) The Civil Rights Injunction. *The University of Chicago Law Review* 47 (1). pp. 199-211.
- [21] Garth, B. G. and Cappelletti, M. (1978) Access to Justice: The Newest Wave in the Worldwide Movement to Make Rights Effective. *Buffalo Law Review* 27. pp. 181-292.
- [22] Global Network Initiative Principles. Available from: <https://globalnetworkinitiative.org/gni-principles/> [Accessed 30 April 2025].

- [23] Global Network Initiative. (2025) *Global Network Initiative Implementation Guidelines*. [online] Available from: <https://globalnetworkinitiative.org/implementation-guidelines/> [Accessed 30 April 2025].
- [24] Gradoni, L. (2021) *How platform governance had its Marbury v Madison. Constitutional Review via Facebook's Oversight Board* [blog entry] 10 February 2021. Verfassungsblog. Available from: <https://verfassungsblog.de/fob-marbury-v-madison/> [Accessed 31 October 2024].
- [25] Grant, J. P. and Barker, J. C. (2009) *Encyclopædic dictionary of international law (third edition)*. New York: Oxford University Press.
- [26] Hummel, K. and Jobst, D. (2024) An Overview of Corporate Sustainability Reporting Legislation in the European Union. *Accounting in Europe*, 21 (3), pp. 320-355.
- [27] ICANN. (2019) *ICANN Publication Practices for Board, Litigation, and Accountability Processes*. [online] Los Angeles: ICANN. Available from: <https://www.icann.org/en/board-activities-and-meetings/publication-practices> [Accessed 21 August 2025]
- [28] Klein, E. (2018) *Mark Zuckerberg on Facebook's hardest year, and what comes next*. [online] Vox. Available from: <https://www.vox.com/2018/4/2/17185052/mark-zuckerberg-facebook-interview-fake-news-bots-cambridge> [Accessed 31 October 2024].
- [29] Klonick, K. (2020) The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression, *Yale Law Journal*, 129 (8). pp. 2418-2499.
- [30] Kun, A. (2017) Kellő gondosság alapú belső átvilágítás a munkajogban, mint a transznacionális vállalatok magatartásának új zsinórmértéke. *Miskolci Jogi Szemle*, 12 (2), pp. 285-297., p. 288.
- [31] Lind, E. A. and Tyler, T. R. (1988) *The Social Psychology of Procedural Justice*. New York: Springer New York.
- [32] Mazúr, J. and Gramblickova, B. (2023) New Regulatory Force of Cyberspace: The Case of Meta's Oversight Board. *Masaryk University Journal of Law and Technology*, 17 (1), pp. 3-32.
- [33] Meta. (2021) *Corporate Human Rights Policy*. [online] Available from: <https://about.fb.com/wp-content/uploads/2021/03/Facebooks-Corporate-Human-Rights-Policy.pdf> [Accessed 31 October 2024].
- [34] Meta. (2023) *Meta Annual Report*. [online] Available from: https://www.annualreports.com/HostedData/AnnualReports/PDF/NASDAQ_FB_2023.pdf [Accessed 31 October 2024].

- [35] OECD. (2008) *OECD Guidelines for Multinational Enterprises on Responsible Business Conduct*
- [36] OECD. (2018) *Due Diligence Guidance for Responsible Business Conduct*
- [37] Operato, L. et al. (2025) Navigating CSRD reporting: Turning compliance into sustainable development with science-based metrics. *Environmental Development*, 54.
- [38] Oversight Board. (2022) *Facebook Oversight Board Submission to the Office of the High Commissioner for Human Rights, United Nations on the practical application of the UNGPs to the activities of technology companies, February 2022*. [online] Available from: <https://www.ohchr.org/sites/default/files/2022-03/Oversight-Board.pdf> [Accessed 31 October 2024].
- [39] Oversight Board. (2023) *Oversight Board Charter*. Available from: <https://www.oversightboard.com/wp-content/uploads/2023/11/3427086457563794.pdf>
- [40] Pánek, M. (2021) *Appealing to Facebook's Oversight Board*. Mmister [blog entry] 9 September. Available from: <https://mmister.com/appealing-to-facebooks-oversight-board/> [Accessed 31 October 2024].
- [41] Pap, L. A. Chronowski, N. and Nemessányi, Z. (2024) Corporate Human Rights Responsibility in Illiberal Regimes: The Example of the Ukrainian Refugee Crisis in Hungary. *Business and Human Rights Journal*, (9), pp. 32-53.
- [42] Posner, R. A. (2011) *Economic Analysis of Law*. New York: Aspen Publishers.
- [43] Rawls, J. (1971) *A Theory of Justice*. Cambridge: Harvard University Press.
- [44] Rhode, D. L. (2004) *Access to Justice*. Oxford: Oxford University Press.
- [45] Ruggie, J. G. (2007) Business and Human Rights: The Evolving International Agenda. *American Journal of International Law*, 101 (4), pp. 819-840.
- [46] Ruggie, J.G. (2014) Global Governance and “New Governance Theory”: Lessons from Business and Human Rights. *Global Governance*, 20 (1), pp. 5-18.
- [47] Salcito, K. and Wielga, M. (2018) What does Human Rights Due Diligence for Business Relationships Really Look Like on the Ground?. *Business and Human Rights Journal*, (3), pp. 113-121.
- [48] Sissons, M. (2021) *Our Commitment to Human Rights*. [Facebook] 16 March. Available from: <https://about.fb.com/news/2021/03/our-commitment-to-human-rights/> [Accessed 31 October 2024].
- [49] United Nations Human Rights Council. (2011) *UN Guiding Principles on Business and Human Rights*. [online] Available from: https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr_en.pdf [Accessed 31 October 2024].

- [50] Váradi, Á. (2022) Access to constitutional complaint procedures: A real chance? Remarks on the availability of legal aid in constitutional court procedures in Austria, Germany and Hungary. *Hungarian Journal of Legal Studies*, 61 (4), pp. 372–385.
- [51] Wettstein, F. (2012) CSR and the Debate on Business and Human Rights: Bridging the Great Divide. *Business Ethics Quarterly*, 22 (4), pp. 739-770.
- [52] Wolfsteller, R. and Li, Y. (2022). Business and Human Rights Regulation After the UN Guiding Principles: Accountability, Governance, Effectiveness. *Human Rights Review* 23 (1). pp.1–17.
- [53] Zuckerberg, M. (2021). *A Blueprint for Content Governance and Enforcement*. [Facebook] 5 May. Available from: https://www.facebook.com/notes/751449002072082/?hc_location=ufi [Accessed 31 October 2024].