mediaLAWS

Anticipazioni

Framing the Facebook Oversight Board: Rough Justice in the Wild Web?*

Andrea Buratti

Abstract

Starting from the increasing awareness of the necessity to regulate social media's activity in content moderation, in this article I try to frame the peculiar model embodied by the Facebook Oversight Board.
The conflicts that arise from social media are all too relevant in the light of the protection of fundamental rights. Therefore, assessing actors and procedures involved in content moderation is crucial for shaping the founding principles of digital constitutionalism – a process aimed at protecting fundamental rights through the limitation of emerging private powers on the web.
In this perspective, FOB represents an excellent case study: it entails a brand-new model of content moderation which raises several questions, both related to its contextualization within legal categories as well as its consistency with legal and democratic standards. After having described FOB's design – in terms of rules, functions and procedures – I classify FOB in the alternative between contractual means of dispute resolution and a company's internal complaint-handling process. Subsequently, I try to contextualize FOB within the legal framework provided for by the Digital Service Act proposal of the European Union. Lastly, I answer the question this paper moves from, i.e. the compatibility of the institutionalization of private adjudication with digital constitutionalism's goals.

Summary

Keywords

Facebook Oversight Board - content moderation – social media platforms – free speech – delegation of judicial functions to private powers

* This paper is the outcome of a long-lasting dialogue with Marco Bassini and Oreste Pollicino, who introduced me to this area of research. I also want to thank Ludovica Durst, Juncal Montero Regules, Berdien van der Donk, Niranjan Nair Reghuvaran, Marco Fioravanti and Alberto Clerici for their advice.

1. Online Content Moderation in the Web's Risk Society

In contemporary societies, the web represents an extraordinary opportunity for the expansion of certain liberties, knowledge, human relations, economic entrepreneurship, collaboration and general improvement in the quality of our lives. At the same time, however, the "wild web" carries with itself a remarkable amount of threats to our safety, privacy, intellectual property and right to human dignity and equality.

The task of balancing opportunities and risks is a typical function performed by the law[1]. Nevertheless, in dealing with the digital ecosystem, the legal patterns that have developed throughout the "Twentieth-century society of communication" are demonstrably inadequate. Firstly, because the web is an inherently borderless realm, this makes a state-law-based reaction narrow-minded and, in the end, ineffective. Secondly, because legal patterns that have been developed for tackling dangers and imposing responsibilities with regard to the media protagonists of the "Twentieth-century society of communication" are nowadays obsolete. Finally, because private players on the web, with their legitimate interests and expectations, occupy a central role, which deserves to be contextualized in regard to its relationship with communities' values[2].

These challenges are strikingly evident in the assessment of conflicts arising from user-generated contents on social media platforms: here, indeed, (i) fragmentation of national regulations on illicit content and diversity of jurisdictions raise difficulties and often hinder reaching homogeneous and consistent solutions; furthermore, (ii) online platforms' activity of content moderation (i.e. detection, filtering, removal of content, suspension and deplatforming of users) requires the involvement of artificial intelligence, due to the huge amount of data to handle: this poses a number of problems related to algorithms' design, transparency and liability; lastly, (iii) content moderation lies at the intersection of two processes, i.e. the delegation of public functions to private actors, and the ongoing effort of the web's private players toward self-regulation.

Truly, the conflicts that arise from social media are all too relevant in the light of the protection of fundamental rights and therefore the management of such conflicts should not be carried out by mere private mechanisms of content moderation, or on the basis of rules and guidelines established autonomously and unilaterally by online platforms imposed to users through terms of contract[3].

Framing and assessing actors and procedures in online content moderation is, hence, crucial for shaping «due data process» as a founding principle of digital constitutionalism – a process aimed at protecting fundamental rights through the limitation of emerging private powers on the web[4].

In this perspective, the Facebook Oversight Board (FOB) represents an excellent

---

1 U. Beck, *Risk Society: Towards a New Modernity*, London, 1992.

2 J. Balkin, *Free Speech is a Triangle*, in *Columbia Law Rev.*, 7, 2018, 2011.

3 G. De Gregorio, *Democratising Online Content Moderation: A Constitutional Framework*, in *Computer and Security Law Rev.*, 2020.

4 O. Pollicino, *Judicial Protection of Fundamental Rights on the Internet: A Road Towards Digital Constitutionalism?*, London, 2021, 184 ss.; G. De Gregorio, *The Rise of Digital Constitutionalism in the European Union*, in *Intern. Journ. of Const. Law*, 1, 2021, 41; G. De Gregorio, *Digital Constitutionalism in Europe. Reframing Rights and Powers in the Algorithmic Society*, Cambridge, 2022.

case study: it entails a brand-new model of content moderation which raises several questions, both related to its contextualization within legal categories as well as its consistency with legal and democratic standards. FOB represents, indeed, an attempt toward the «institutionalization of private adjudication»[5]: because it entails substantive delegations of public functions to private actors, it must be carefully assessed.

In this work, I will take into account, first, FOB's design – in terms of rules, functions and procedures (§ 2), and then, the origins of FOB as a response to increased responsibilities of social media platforms in content moderation and legal diversity (§ 3). In order to reach a framing of FOB under legal categories, I will examine a preliminary issue, related to the legal identity of online platforms in the current digital ecosystem (§ 4), and I'll try to classify FOB in the alternative between contractual means of dispute resolution and a company's internal complaint-handling processes. Subsequently, it will be possible to contextualize FOB within the legal framework provided for by the Digital Service Act proposal of the European Union (§ 6). Lastly, I'll try to answer the question this paper moves from, i.e. the compatibility of the institutionalization of private adjudication with digital constitutionalism's goals (§ 7).

Let me anticipate my conclusion on this critical issue. Some scholars see FOB as a contribution to the development of digital constitutionalism[6], an opportunity for the expansion of individuals' rights and guarantees *vis-a-vis* online platforms' powers[7]. Especially in looking at the decisions FOB has adopted during its first years of activity, many may value the quasi-judicial shape FOB displays, emphasizing its reliance on a human rights-based discourse. Conversely, I'll try to demonstrate that FOB does not own any feature of a judicial body, nor does it reproduce the structure of an alternative dispute resolution mechanism: consequently, it sets an environment where substantive individuals' rights and interests are determined in a flawed framework lacking any rule of law. I will propose that FOB should be classified within the set of a company's internal tools of content-moderation. Even in this perspective, however, I'll show how the pattern of content moderation carried out by FOB is not consistent with basic constitutional standards, and in particular with the set of principles provided for by the Digital Service Act proposal of the EU Commission with the aim to regulate transparency of social media platforms. In my perspective, FOB is not consistent with, and is actually contradictory to, digital constitutionalism's goals, and should not be encouraged by public institutions, nor reproduced by other online platforms.

---

[5] O. Pollicino – G. De Gregorio *Shedding Light on the Darkness of Content Moderation: The First Decisions of the Facebook Oversight Board*, in *VerfBlog*, 5 February 2021.

[6] See G. De Gregorio, *Digital Constitutionalism in Europe*, cit., 6, who sees FOB as «a path toward constitutionalisation beyond the traditional boundaries of modern constitutionalism».

[7] K. Klonick, *The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression*, in *Yale Law Journ.*, 2020, 2418.

2. The Design of Facebook Oversight Board

FOB is a pillar of a more complex architecture, meant to moderate user-generated content on Facebook/Instagram platforms: this architecture – the outcome of a process of consolidation of an internal, autonomous, system of content moderation – is made of (i) normative standards (i.e., Facebook community standards), (ii) enforcement mechanisms (i.e., mechanisms of content detection, filtering and removal as well as temporary or permanent deplatforming), and (iii) a process of review of its internal system of standards' definition and enforcement (FOB)[8].

Focusing now on the latter, FOB's governance is a Chinese-box system aimed at granting its independence from Facebook/Meta. It includes the Oversight Board Trust, FOB's LLC, FOB's co-chairs and FOB's members. The Trust has been created with the aim of safeguarding FOB's independence from the Company. Its main function is related to granting the financial resources FOB needs. However, the Trust's economic resources are entirely granted by Facebook/Meta itself, who also appoints trustees. Although financially dependent on the Trust, FOB is incorporated as an LLC under U.S. law, and as such it owns a broad degree of autonomy in its administrative organization. As for FOB's co-chairs, they are members of FOB initially appointed by Facebook, who perform specific tasks in the management of FOB's activities.

According to the Charter, which is FOB's basic rules, the number of Board members varies from 11 to 40. They are selected for a three-year term, for a maximum of three terms, by a hiring committee of the Board. Both Facebook and the community can propose candidates to the Board: selection is based on expertise, but must be guided by the goal of enriching diversity, too. After being selected by the internal hiring committee, FOB's members are formally appointed by trustees, who also possess the power to remove a member in case of misconduct. Initially, appointment is a power exercised jointly by co-chairs and Facebook itself. Though, since the length of such an initial stage is not exactly defined, the appointment process mainly relies on the role of co-chairs, whose relations with the Company are evident.

Therefore, bearing in mind the inconsistencies in ensuring separation between the Trust and the Company and the fragility of the appointment process, it is possible to conclude that the goal of establishing FOB as an independent body has not been reached.

This conclusion offers the opportunity to strike a clear distinction between bodies, such as FOB, created within a company and directly dependent on it, and Social Media Councils, proposed in recent years by civil society actors such as private institutions, which are meant to overview social media content moderation activities. Whatever opinion one may have on the role Social Media Councils could play in the digital ecosystem, the basic tenet of these institutions should be the openness to civil society, their spreading over a broad community of social media, well beyond the borders of a single company, and the independence from stakeholders. The distance between the two models could not be wider.

[8] For a comprehensive presentation, see K. Klonick, *The Facebook Oversight Board*, cit.

FOB's jurisdiction is defined by the Charter and Bylaws. There are two main avenues for lodging a complaint: (i) when users disagree with Facebook's action on content, and if Facebook's internal appeal process has been exhausted, they can raise complaints; (ii) Facebook's power to raise appeals is defined in a broader way than users', for it can submit a request of review «including additional questions related to the treatment of content beyond whether the content should be allowed or removed completely». In both cases, complaints are submitted to a rigid procedure of case-selection, which is guided by the priority «to consider cases that have the greatest potential to guide future decisions and policies». So far, the rate of selected cases has been very low, with FOB providing no explanation on the methodology of case-selection.
Additionally, the Charter provides two special procedures of review, which circumvent case-selection: (i) in exceptional circumstances «when content could result in urgent real world consequences», Facebook has the power to ask for an «automatic and expedited review», that the Board must carry out «as quickly as possible» (FOB's Charter, art. 3.7.2); (ii) Facebook may request advisory opinions of «policy guidance» from the Board, regarding «the clarification of a previous decision by the board or guidance on possible changes to Facebook's content policies». (FOB's Charter, 3.7.3).
An imbalance is evident in the procedure for actual access to the Board, with users having very limited chances to see their complaints selected for review.
Concerns have been raised regarding the effectiveness of FOB's decisions. According to the Charter, indeed, «the board's resolution of each case will be binding and Facebook will implement it promptly», and the Bylaws specify that Facebook must implement decisions within seven days. However, FOB does not own means to impose the enforcement of its decisions, nor formalized tools to oversee the actual implementation of its decisions by the Company.

3. FOB as a Response to Strengthened Responsibility of Online Platforms and Legal Diversity in the Digital Ecosystem

According to Facebook, «the oversight Board was created to help Facebook answer some of the most difficult questions around freedom of expression online». On a social media platform that gathers more than two billion users, and which is now the main channel of dissemination of information, social and political campaigns, propaganda, etc., dilemmatic issues emerge every day, requiring an assessment with a cross-disciplinary expertise, embracing law, humanities and technology.
Nevertheless, many authors have adopted a skeptical approach, explaining FOB as a customers' satisfaction device: they underline how Facebook decided to create FOB in the aftermath of scandals, such as the "Cambridge Analytica" case, which undermined the reputation of the Company. In this perspective, FOB develops the role of preserving the reputation of the Company, as a tool of a peculiar corporate social responsibility strategy whose main commitment is the respect of individual rights.
Both these explanations do not completely satisfy our understanding of the very rea-

sons why a social media platform felt the need to create a body of experts aimed at supervising its content moderation process through a court-of-justice-fashioned structure and a legal-style discourse.

I propose to consider FOB as a response to two challenges social media platforms need to tackle: (i) strengthened responsibilities of online platforms in content moderation, and (ii) legal diversity in the digital ecosystem.

To begin with, the emergence of strengthened responsibilities of social media platforms in content moderation is the unavoidable outcome of a process triggered by the strategy of public actors, and especially of the European Union together with European countries, towards online platforms – a strategy aimed at emphasizing their responsibilities in content moderation, and more generally in managing conflicts arising on the web[9].

The affirmation and consolidation of a «notice and take down» liability of Internet service providers (ISPs) – that offers a safe harbour from liability only to ISPs who promptly remove illicit content as soon as they acquire knowledge of them – forced online platforms to actively engage in the moderation of all the content published on the platforms. In their support, states and supranational institutions have strongly encouraged online platforms to adopt codes of conduct and community guidelines for content moderation.

Consistent with this framework, EU soft law has always stimulated member states and economic actors to adopt out-of-court dispute settlement mechanisms, seen as a more efficient device with respect to traditional legal remedies, that seem inherently unfit to accommodate the exigencies of online interactions: recently, a 2018 EU Recommendation has encouraged ISPs to spontaneously adopt methods of out-of-court dispute resolution for issues related to content moderation[10]. Turning the attention to national legislation, it is worth mentioning the German Network Enforcement Act (2017), which provided for the possibility to establish self-regulated institutions, created by online service providers, and recognised by state authorities based on principles of independence and professional expertise, to be involved in the process of complaint handling[11].

In conclusion, the privatization of justice – at least at the level of user-generated content and the violation of individuals' rights and duties – appears as a natural trend in the digital ecosystem, that both online economic players and public institutions have striven for securing, because of its consistency with the specific dimensions and necessities of disputes related to online content[12].

The second complication that pushed a globally widespread social media platform, such as Facebook, to give birth to the Oversight Board is legal diversity.

---

[9] G. De Gregorio, *Digital Constitutionalism in Europe*, cit., 102.

[10] Commission Recommendation 2018/334 (on measures to effectively tackle illegal content online) § 14. For an analysis of the American approach see S.F. Kreimer, *Censorship by Proxy: The First Amendment, Internet Intermediaries, and the Problem of the Weakest Link*, in *University of Pennsylvania Law Review*, 1, 2006, 11 ss.

[11] Sect. 3 § 6.

[12] M. Bassini, *Fundamental Rights and Private Enforcement in the Digital Age*, in *European Law Jour.*, 2, 2019, 182.

In user-generated content moderation, social media companies apply their policies and guidelines, which are codified with the sake, among others, to comply with the law. But, what *law* are we speaking about? Overlaps and conflicts emerge whenever the legislation of a country diverges from another's, as well as whenever social media guidelines broaden the scope of illicit content to activities that (some) national laws do not qualify as such. Take the case of birth-control (or abortion, or euthanasia) propaganda, which is illicit in many countries and lawful in others; or take the case of fake news and disinformation, that Facebook community standards tackle through (apparently) severe policies of content moderation, whereas national legislation react in a different and heterogeneous way[13].

On one side, the unavoidable consequence of such a process is the consolidation of content moderation norms that often broaden the area of illicit content defined by national measures: since online platforms' main goal is avoiding liability for user-generated content, they incline to comply with standards of "the most rigorous country", thus raising serious concerns on the impact of community standards on users' right to free speech. On the other side, however, social medias' main business goal is expanding the sphere of online communication, avoiding any form of collateral censorship that may derive from an excessive codification of content-moderation's standards.

All of these difficulties push social media platforms to promote a process of spontaneous codification of a harmonized set of rules, a global *lex informatica* of content moderation, able to exclude ISPs' liabilities while, at the same time, not exceedingly restricting users' voices[14].

From Facebook/Meta's perspective, FOB, with its court-of-justice-fashioned structure and its legal-style discourse, should pave the way to such a process.

4. Background Issues about Online Platforms' Identity: Players in a Digital Free Marketplace or Public Utilities in a Shared Infrastructure?

After having analysed the normative architecture designed by Facebook/Meta, and the reasons that led to the creation of this body as a mechanism of content moderation, we can now try to define FOB's identity within legal categories, especially in the attempt to understand whether this pattern is consistent with the legal framework and, if so, whether it should be reproduced by other online platforms.

A satisfactory legal definition of FOB requires dealing with a preliminary background issue: questions on FOB's identity, as well as those on its possible generalization as a pattern of self-regulation of online adjudication, depend on the approach one assumes toward the identity and the role that ISPs, and especially social media platforms,

[13] S. Sassi, *Disinformazione contro costituzionalismo*, Napoli, 2021. C. Pinelli, *Disinformazione, comunità virtuali e democrazia: un inquadramento costituzionale*, in *Diritto Pubblico*, 1, 2022, 173 ss.

[14] E. Lee, *Moderating Content Moderation: a Framework for Nonpartisanship in Online Governance*, in *Am. Univ. Law Rev.*, 2021, 913 ss.

play, and should play, on the web[15]. In general, most of the obstacles that burden the effort toward regulation of the web are due to the divergent understandings of online platforms, which, in turn, is strictly related to cultures and traditions of liberties and rights in the world[16].

One approach deems online platforms as mere private business entities, who have autonomously developed a "tech-know-how" and have engaged in a free economic undertaking, in an environment, the internet, inherently open to competition, as characterized by an indefinite availability of space. Although theoretically distant, the same assumption that online platforms must be regarded as self-regulated community of users[17] converge in the same conclusion having regard to the viability of a regulation of the web. In this scenario, indeed, the web must be regarded as a space impermeable to state actors' interferences: consequently, ISPs' duties and liabilities are only those based on contractual terms of service.

Those assumptions shape the whole legal framework of the web in the U.S.: sect. 230 of the Communication Decency Act (1996) shields ISPs from liability for content published by third parties, on the assumption that they are neutral economic actors, offering a service whose content is out of their control[18]. When dealing with social media platforms, this approach leads to their definition as «public fora», where a free marketplace of ideas takes place[19]; as such, and due to a rigid application of the state action doctrine[20], ISPs are not burdened with a duty to protect users' fundamental rights[21]. As the Supreme Court affirmed in *Reno v. Aclu* (1997), «as a matter of constitutional tradition, in the absence of evidence to the contrary, we presume that governmental regulation of the content of speech is more likely to interfere with the free exchange of ideas than to encourage it. The interest in encouraging freedom of expression in a democratic society outweighs any theoretical but unproven benefit of censorship»[22]. Following the deplatforming of former President Donald Trump from Instagram and Facebook, while Democrats have proposed to enhance the liability of social media platforms in content moderation, Conservatives are sponsoring laws

---

15 B. van der Donk, *Circumventing ambiguous qualifications and national discrepancies: a European roadmap to define social media platforms*, in *ssrn.com*.

16 M. Bassini, *Libertà di espressione e social network, tra nuovi "spazi pubblici" e "poteri privati". Spunti di comparazione*, in *Rivista di diritto dei media*, 2, 2021, 67 ss.

17 D.R. Johnson – R. Post, *The New "Civic Virtue" of the Internet: A Complex System Model for the Governance of Cyberspace*, in C. Firestone (ed.), *The Emerging Internet*, *Ann. Rev. of the Institute fo Information Studies*, Denver, 1998, 81 ss.; R. Post, *Anarchy, State and the Internet: An Essay on Law-Making in Cyberspace*, in *Journ. of Online Law*, 1995.

18 M. Bassini, *Internet e libertà di espressione*, Roma, 2019, 109 ss.

19 U.S. Supreme Court, *Reno v. American Civil Liberties Union*, (1997). See also U.S. Supreme Court, *Packingham v. North Carolina*, (2016). For critical remarks, A. Morelli – O. Pollicino, *Metaphors, Judicial Frames, and Fundamental Rights in Cyberspace*, in *Am. Journ. of Comp. Law*, 2020, § 7.

20 According to which obligations of fundamental rights protection deriving from the U.S. Bill of Rights are applicable only to the U.S. Government and, through the due process clause of the 14th Amendment, the States. See D.A. Strauss, *State Action After the Civil Rights Era*, in *Univ. of Minn. Law School Scholarship Repository*, 1993.

21 E. Volokh, *Treating Social Media Platforms like Common Carriers?*, in *Journ. of Free Speech Law*, 2021, 377.

22 U.S. Supreme Court, *Reno v. American Civil Liberties Union* (1997).

aimed at prohibiting social media companies from adopting any kind of content-based censorship. Both these approaches, though, seem incompatible with current constructions of the First Amendment's protection of the liberty of social media platforms to shape their own environment as private property.

There is a solid cultural background in the American tradition of fundamental rights that helps to explain such an outcome: ideological roots of this approach go back to the very origins of the web, where a cyberlitarians movement took shape, radicalizing values inherent to the American political identity[23].

Nonetheless, the backwardness of such an approach is evident[24]. It ignores that the evolution of the web in the last decades has created an oligopoly – a "platformization" –, thus giving birth to new private powers, the "lords of the web", who prevent to define the web as a free marketplace[25]: as it happened to many economic markets throughout the history of liberal societies, concentration of power in the hands of few actors forced states to abandon a *laissez-faire* approach for engaging in regulatory intervention, aimed at preserving competition among economic players and protecting consumer rights, whose knowledge, position and contractual force is incomparable to big economic players. This is particularly relevant in the field of online content moderation: in a television talk-show broadcasting studio, the editor retains the power to invite and exclude a guest speaker at any moment of the programme, thus affecting guest's constitutional right to free speech: private property prevails, here, over free speech. However, interactions on social media can hardly be compared to such a context: here, the fragmentation of users results in varying relations with platforms, thus requiring a more stringent regulation of terms of service, in the direction of affording a fair and equal service in a safe environment.

A second, more mature and modern approach assumes that large online platforms own assets and occupy hubs belonging to the digital infrastructure, with an impact on general interests and fundamental rights, thus requiring the development of a comprehensive public law regulation. The understanding of online platforms as «privately owned public spaces» paves the way to the imposition of duties and liabilities, including the duty of protection of fundamental rights of users – whose nonuniform relationship with platforms deserves an enhanced attention[26].

This assumption drives EU law of the web: the same E-Commerce Directive, although aimed at expanding the growth of online business, introduced duties for active ISPs in content removal[27]: an approach upheld and even strengthened by the case-law of the

---

[23] M. Bassini, *Internet e libertà di espressione*, cit., 40 ss.

[24] R. Tushnet, *Power Without Responsibility: Intermediaries and the First Amendment*, in *Geo. Wash. L. Rev.*, 2008, 101.

[25] M. Moore – D. Tambini, *Digital Dominance: The Power of Google, Amazon, Facebook and Apple*, New York, 2018.

[26] S. Poier, *Privately Owned Public Spaces: The Internet and the Shaping of a New Breed of Consumers. From Participants to Users*, in *Oñati Socio-Legal Series*, 1, 2015; B. van der Donk, *Circumventing Ambiguous Qualifications and National Discrepancies*, cit., 30.

[27] Art. 15 of Directive 2000/31/EC (E-Commerce Directive) established a shield from liability for active ISPs (i.e., an ISP who owns technology that allows to know and control user-generated content) who promptly remove illicit content after acquiring knowledge of them, even following notices by users.

European Court of Justice[28]. The spreading of the "notice and take down" liability is just the most evident symptom of this approach, which also entails duties of enhanced cooperation in fields of particular risk, such as counter-terrorism, protection of minors, and duties of compliance with higher standards of privacy and data protection[29]. In this framework, the recent Copyright Directive represents the most enhanced attempt to establish a general responsibility of online platforms in preventing violations of users' copyright[30].

Quoting the European Court of Justice, «the internet has now become one of the principal means by which individuals exercise their right to freedom of expression and information. In the light of their accessibility and their capacity to store and communicate vast amounts of information, internet sites, and in particular online content-sharing platforms, play an important role in enhancing the public's access to news and facilitating the dissemination of information in general, with user-generated expressive activity on the internet providing an unprecedented platform for the exercise of freedom of expression»[31].

This approach is shared by the Council of Europe. The Court of Strasbourg has envisaged the necessity of fundamental rights balancing on the web: «… user-generated expressive activity on the Internet provides an unprecedented platform for the exercise of freedom of expression. […]. However, alongside these benefits, certain dangers may also arise. […] Bearing in mind the need to protect the values underlying the Convention, and considering that the rights under Articles 10 and 8 of the Convention deserve equal respect, a balance must be struck that retains the essence of both rights»[32].

Even more precise is the view of the Venice Commission, which recently affirmed: «The small number of very powerful private actors that literally own the information highways have commercial interests and rights that tend to collide with both civil and political rights and electoral principles. These internet providers have taken up the

ECJ's case-law has blurred the rigid distinction between active and passive service providers, with the aim of broadening the area of responsibility (starting with ECJ, C-324/09, *L'Oréal SA e a. c. eBay* (2012)). See G. Morgese, *Moderazione e rimozione dei contenuti illegali online nel diritto dell'Unione Europea*, in *Federalismi.it*, 1, 2022, 87.

[28] Consider, in particular, the seminal ECJ, "*Google Spain*" (ECJ, Grand Chamber, C-131/12, *Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González* (2014)), where the Court affirmed the duty of a web search engine to comply with the request to remove content that affects individual's right to be forgotten, independent from the behaviour of the author of the content. See O. Pollicino, *Judicial Protection of Fundamental Rights on the Internet*, cit.

[29] M. Manetti, *Regolare Internet*, in *Rivista di diritto dei media*, 2, 2020, 35 ss.

[30] Art. 17 of the EU Directive 2019/790, exempting for-profit online content sharing service providers from liability who actively engage in preventing and tackling breaches of copyright, by submitting agreements to right holders who share their content or, in case no agreement is adopted, by making their best efforts for preventing and tackling breaches. See C. Geiger – B.J. Jütte, *Platform liability under Article 17 of the Copyright in the Digital Single Market Directive, Automated Filtering and Fundamental Rights: An Impossible Match*, in *GRUR International*, 6, 2021, 517.

[31] ECJ, Grand Chamber, C-401/19, *Republic of Poland v. Parliament and Council* (2022), § 46.

[32] ECtHR, Grand Chamber, app. no. 64569/09, *Delfi AS v. Estonia* (2015), § 110. For a comprehensive overview of ECtHR's case law on online platforms, see B. Sander, *Democratic Disruption in the Age of Social Media: Between Marketized and Structural Conceptions of Human Rights Law*, in *European Journ. of Int. Law*, 1, 2021, 180 ss.

gatekeeping role which originally belonged to the traditional media, without however having adopted the ethical obligations of the media»[33].

Even in this second viewpoint, however, there is no general agreement on what role private actors should play, and what their level of involvement should be in the process of regulation and adjudication on the web. Emphasizing the private dimension of online platforms in a regulated digital market, while at the same time preserving the value of pluralism on the web, one could reach the conclusion that a fair regulation should limit itself to the creation of a framework of rules aimed at both granting fair competition among economic actors and protecting the rights of users, leaving platforms with a high margin of discretion in shaping their business model and the terms of their service[34]. By contrast, stressing the overwhelming importance of fundamental rights in the digital ecosystem and the enabling function online platforms have assumed throughout the years, it would be possible to conclude that, although remaining private entities, platforms represent nowadays a public service, a gateway to the digital infrastructure instrumental to the general interest of society, assuming the identity of public utilities in a universal public service: a thesis that would lead to a more detailed and stringent regulation of online platforms duties and responsibilities, expanding the area of public control over the internal processes of business activity, and limiting the contractual liberties of the parties more strictly[35].

5. FOB Between Alternative Dispute Resolution and Internal Complaint-Handling

Following the assumption that online platforms are mere private actors, whose relations with users depend on the service contract they submit to, one could view FOB as a contractual means of alternative dispute resolution.

In general, the (voluntary or mandatory) devolution of consumer issues to alternative dispute resolution mechanisms, preliminary and alternative to traditional channels of judgement, is a strategy consistently encouraged by the business sector, and welcomed even by consumer associations and institutions. With regard to the web, this strategy has been emphasized by European law – at both a supranational and national level.

Whatever conclusion one may reach on the nature of online platforms, the hypothesis of framing FOB as a contractual means of alternative dispute resolution must be rejected, in both broad and specific reasoning.

Even if one looks at online platforms as mere private actors unfettered by any legal

[33] *Joint report of the Venice Commission and of the Directorate of information society and action against crime of the Directorate General of Human Rights and Rule of Law (DGI) on the Use of digital technologies and elections* (2019), § 145.

[34] B. van der Donk, *Circumventing Ambiguous Qualifications and National Discrepancies*, cit., 30. According to van der Donk, the main difficulty to the identification of social media platforms as universal services is related to the extensive competition which characterizes online platforms' market, thus excluding an essential condition of such category (spec. 12).

[35] K.S. Rahman, *The New Utilities: Private Power, Social Infrastructure, and the Revival of the Public Utility Concept*, in *Cardozo Law Rev.*, 2018, 1621.

responsibility, their service contract is still subject to contract law, where general principles on the establishment of dispute settling methods agreed by the parties have been identified.

This is why the already mentioned EU Recommendation on measures to effectively tackle illegal content online explains that any mechanism of out-of-court dispute settlement should be «easily accessible, effective, transparent and impartial and should ensure that the settlements are fair and in compliance with the applicable law»[36]. The Recommendation reproduces here a set of principles that have been laid down in EU soft law instruments for more than twenty years, with the aim of facilitating the expansion of alternative dispute resolution means in consumer law, while at the same time providing effective guarantees for customers.

Against this normative background, it becomes evident that FOB does not comply with standards of accessibility, effectiveness and impartiality that must govern a contractual dispute resolution mechanism: FOB cannot be seen as an independent third party; there is no guarantee of equity between the Company and its users in many stages of FOB's procedures, starting with the right to access, which is varying; and the same enforcement of FOB's decision is left to discretional compliance by the Company.

Moreover, looking in detail at the service contract proposal that the Company submits its users to, it's impossible to identify any legal provision aimed at establishing a preliminary dispute resolution mechanism agreed by the parties: the contract does not entail any explicit reference to the Board and its authority, at least in its present wording; it only references community standards, which mention FOB within the content moderation complaint system. This set of references is not sufficient in structure as a contractual provision on dispute resolution, especially taking into consideration that community standards are open to modification without any obligation for Facebook/Meta to inform its users.

Under a legal point of view, FOB does not lie on any contractual provision: its legitimacy is, therefore, unilateral.

It is appropriate, rather, to contextualize FOB within the category of internal complaint-handling systems: the approach to shape business' internal processes and mechanisms of consumer disputes resolution as quasi-judicial remedies has already been identified by Rory van Loo as a means of privatization of justice[37]. This pattern is highly consistent with the peculiar structure of online platforms and their activity of content moderation, where online service providers seem to play an intermediary role in a free marketplace of ideas fuelled by users.

In this perspective, FOB can be described as a body autonomously established by Facebook/Meta within its system of complaint-handling, with the aim to advise and guide the internal dispute-settling process, in a stage preluding formalized legal remedies.

On one side, as a device of internal dispute settlement, FOB can contribute to offer

[36] Commission Recommendation (EU) 2018/334, 1 March 2018, on measures to effectively tackle illegal content online, § 14.

[37] R. van Loo, *The Corporation as a Courthouse*, in *Yale Journ. on Regulation*, 2016, 547.

fast, cost-free, substantive remedies to users' claims, avoiding the need for engagement in the more stressful set of solutions provided for by the law – which in the case of disputes arising from online content are often ineffective. On the other side, as an institution placed within – and subject to – a specific business model, FOB raises serious concerns related to transparency and consistency.

In a realm, such as online content moderation, where adjudication plays a crucial role in a codification process which is still open and malleable, privatization of justice must be evaluated with the utmost care: although paying attention to the scope of online platforms' liberty of organization, internal complaint-handling processes must be regulated under compulsory legal principles, especially when they involve forms and procedures of a quasi-judicial nature.

6. Assessing FOB in the Awaiting of the Digital Service Act

Since it represents a consistent step toward the constitutionalisation of the digital ecosystem[38], the Digital Services Act (DSA) proposal of the EU Commission allows us to assess FOB's model within the set of legal principles it provides for.

In general, EU's DSA Regulation proposal is an attempt to set a more organized regulatory framework on the responsibilities of ISPs, also identifying the scope of «very large online platforms», whose role and dimension entails strengthened obligations[39]. With this aim, it lays down duties and obligations ISPs must respect in the organization of their business and in their relations with users; in addition, it starts to build a more efficient system of public control over online service providers, through the establishment of a national digital service coordinator, entrusted with effective powers of supervision and enforcement, also in cooperation with equivalent authorities of member states (DSA proposal, art. 41 *et seq.*)[40]; having regard to «very large online platforms»[41], the system of supervision, investigation, enforcement and monitoring is more detailed, and involves the European Commission itself (art. 50 *et seq.*), which assumes relevant functions of investigation – including a structured dialogue with online platforms – and sanction, with the possibility to impose fines calculated in proportion to a company's turnover.

Focusing on online content moderation, article 12 of DSA sets the basic principles of transparency of content moderation activity and makes clear its reliance on the respect

[38] G. De Gregorio, *Digital Constitutionalism in Europe*, cit., 26.

[39] G. De Gregorio – O. Pollicino, *The European Constitutional Road to Address Platform Power,* in VerfBlog, 31 August 2021; P. Dunn, *Il Digital Markets Act: tra logiche concorrenziali e istanze costituzionali*, in *Diritticomparati.it*, 17 February 2022.

[40] B. Wagner – H. Janssen, *A first Impression of Regulatory Powers in the Digital Services Act,* in VerfBlog, 4 January 2021.

[41] According to art. 25, para. 1, of DSA, very large online platforms – whose more stringent regulation and responsibilities derive from the systemic risk they entail – are those that «provide their services to a number of average monthly active recipients of the service in the Union equal to or higher than 45 million».

of fundamental rights. When regulating litigation that could arise from content moderation activity, articles 17 and 18 of the DSA draw a clear distinction between «internal complaint-handling systems» and «out-of-court dispute settlement» bodies.

According to art. 18, a certified out-of-court dispute-settlement body must meet prerequisites of impartiality, independence, accessibility, transparency, and legal enforcement of its decisions. Art. 18, therefore, reproduces the principles already stated in the 2018 Recommendation, relying on the pattern of FSM (the German Association for Voluntary Self-Regulation of Digital Media service providers), which is the first independent body recognized under the German Network Enforcement Act to develop the function of an external self-regulated body involved in settling disputes arising from content moderation. This provision has been harshly criticized, which still confirms the excessive reliance of EU law strategies over the resources of private, self-regulated, mechanisms of dispute resolution, while the trend should be that of entrusting – and perhaps even expediting and simplifying – state courts to arbitrate this kind of litigation[42].

From this perspective, the French Constitutional Council took a rigorous stance in its decision of 2020, striking down relevant provisions of a statute imposing severe economic sanctions on ISPs which failed to remove hateful content, following a notice by users[43]. According to the Council, although hate speech must be tackled with rigorous limitations, a mechanism based only on the assessment of online platforms, under the menace of a severe economic sanction, does not respect the basic guarantees of freedom of expression. The Council's main concern is not related to the abstract possibility to impose limitations on free speech – which actually represents a consistent orientation of its case-law – but rather is related to the lack of a judicial assessment on the very nature of the posted content, and to the delegation to private actors of an evaluation- an evaluation very relevant in the light of fundamental rights protection.

Whatever opinion one should have on the scope of private, certified, mechanisms of dispute resolution, as those envisaged in the DSA, in the more limited sake of this paper, it must be said that FOB does not comply with almost any of the requisites art. 18 imposes: in looking at this normative framework, it is evident that FOB cannot be framed in the category of a certified out-of-court dispute-settlement body.

If instead categorizing FOB as a means of the internal complaint-handling process, it is more appropriate to assess it in the light of art. 17.

Nonetheless, even in this perspective, striking concerns arise, related to the voluntary and unregulated processes of content moderation adopted outside a rule of law-based framework[44].

In DSA, the organization of internal complaint-handling systems is left to the will of online platforms in the respect of a set of basic principles: FOB's procedure, with its

[42] D. Holznagel, *The Digital Services Act Wants You to "sue" Facebook over Content Decisions in Private* de facto *Courts*, in *Verfblog*, 24 June 2021.

[43] Conseil Constitutionnel, déc. n. 2020-801 DC, du 18 juin 2020.

[44] See the *Joint Report of the Venice Commission and of the Directorate of information society and action against crime of the Directorate General of Human Rights and Rule of Law (DGI) on the Use of digital technologies and elections*, (2019), § 145.

limited and unequal right to access, does not comply with the principle of «easy access» (DSA proposal, 17.2), while the duty for the service provider to promptly implement the outcome of the assessment (art. 17.3) is here impaired from the discretion Facebook/Meta retains on the enforcement of FOB's decisions. A further issue refers to online platforms' duty to «inform complainants of the possibility of out-of-court dispute settlement» (art. 17.4): although formally nothing in the design of FOB can be interpreted as preventing users from access to an independent dispute resolution body, or to courts, the part of a justice that FOB plays can be misleading for the public, preventing users from possessing a clear understanding of the system of remedies provided for by the law. The true goal that business' internal complaint handling mechanisms pursue is diverting users from formalized avenues of litigation[45].

Even emphasizing the advisory role FOB plays, as a mere "supervisor" of the internal complaint-handling process, this pattern seems inconsistent with the exigencies of transparency that should govern the design of companies' internal process in a field, such as content moderation, where fundamental rights and public interests are at the forefront.

With the aim of facilitating companies' compliance to legal standards, DSA foresees a model based on independent auditing: according to art. 28, indeed, very large online platforms are compelled to adopt a risk assessment system, where audits realized by independent bodies take place. Personally, I believe that this pattern of review of internal complaint-handling process is preferable, since it grants independent evaluations, favours the circulation of best practices among private actors for an increasing compliance to generally agreed standards, and offers users grounds to evaluate the accuracy of complaint-handling by the companies, to be used in case of legal disputes.

At the same time, in my opinion even the DSA's model of platform's internal processes' assessment is not sufficient in consideration of how much internal complaint-handling systems have reached individuals' rights and society as a whole. I propose, rather, to rely on a supranational administrative agency entrusted with powers of standard setting, examination, and expedited review on content moderation and internal dispute resolution mechanisms, as those existing in many fields of consumer law[46]: such a model would channel online platforms' internal complaint-handling mechanisms toward consistent, rule of law-based standards, while providing a fast and complimentary remedy which ensures an independent – and transparent – assessment.

My conclusion is that FOB embodies a model of internal complaint-handling that should not be encouraged by public actors nor reproduced by other online platforms: although it could represent, in particular circumstances, a substantive remedy for individuals' claims, it enacts a staging, pretending to be a judge although lacking all of the basic features a justice owns, thus blurring the system of remedies provided for by the law. Whatever sources of human rights standards they take into account as a paramount for their review, and whatever methods of reasoning they develop for carrying

[45] R. van Loo, *The Corporation as a Courthouse*, cit., 562.

[46] R. van Loo, *The Corporation as a Courthouse*, cit., 554. See also, reaching the same conclusion having regard to Copyright-related disputes, C. Geiger – B.J. Jütte, *Platform Liability under Article 17 of the Copyright in the Digital Single Market Directive*, cit., 517.

out their judgement – bodies, such as FOB, created by social media platforms and under their direct control, do not fulfil the requirements of legitimacy, independence and equity that should govern dispute resolution on the web, and furthermore they offend democratic standards of political communities[47].
In the light of enabling «due data process» legal instruments, FOB appears inconsistent and antithetical.

7. Escape from the Metaverse: From the Wild Web to Digital Constitutionalism

The application of the normative categories adopted in the DSA proposal allowed us to frame FOB avoiding the risk of immerging ourselves into the metaverse.
We must disregard Facebook's superficial creation with the purpose to only increase customer satisfaction, consolidate the Company's reputation, and feed the fascinating illusion of the web as a self-regulated environment. Notwithstanding their purpose of creating an alternative legal environment, "avatar" of our legal orders, where they could perform the role of an enlightened and benevolent lord, a Solon who gifts basic laws to the community, and a Solomon who exercise his wise *iurisdictio* – internet service providers and online platforms remain economic actors who behave in a democratic society: in content moderation they are – and cannot be otherwise – parties involved in the dispute, with an economic interest that does not coincide with the interests of users nor with the interest of the political community as a whole.
More broadly, delegation of functions of adjudication to private entities is a process that should not be encouraged. Whether we look at the institutionalization of private adjudication as the outcome of a process of hyper-privatization of cyber law, or if we link it to the emergence and consolidation of a self-regulated digital ecosystem, many problems arise.
The delegation of sovereign functions to private actors and the establishment of private adjudication are not brand-new phenomena. Facing the sudden availability of new, unchartered, territories, where new society was taking shape and new economic actors were quickly sinking their roots, the states have reacted through the delegation of sovereign functions to private players who are now able to expand their control and power in those realms.
It's what happened in the colonization process, both in the Indian subcontinent as well as in the Atlantic Colonies, where chartered trading companies assumed functions of territorial government, including the administration of justice. Company-states – as they are now labelled – possess powers normally associated with sovereign states, «in-

[47] In a recent comment, D. Keller (*Lawful but Awful? Control over Legal Speech by Platforms, Governments, and Internet Users*, *The Univ. of Chicago Law Rev. Online*, 28 June 2022) has proposed to involve the community of users into the function of setting rules on content moderation, reproducing a bottom-up approach to the web that resembles the originary pattern of the web as a self-regulated community. These approaches ignore that what happens in the web is not relevant only to the people of the web, but rather affects lives, liberties and values of the community as a whole, including those who don't surf the web. It's the people, those of flesh and bone, which must be entrusted with the power to deliberate on those issues.

cluding the capacity to mint currency and administer civil and criminal justice within their forts and factories, as well as to raise military forces, wage war, and conduct diplomacy with non-European powers»[48]. The most blatant example is the British East India Company, a chartered joint-stock Company, authorized by the Crown to establish a militia and move war, with the aim to expand British control over trade in Indian Ocean; in about two Centuries, the Company was able to dismantle local dynasty, settle urbans aggregates, conquer territories and affirm itself as a political actor exerting sovereign functions in the name of the British Crown[49].

A similar, although democratically fashioned, process of privatization of public functions took place in American frontier lands in the 19th Century, where forms of spontaneous and self-regulated mechanisms of dispute resolution and criminal justice were explored[50]: in the wild frontier *Order* anticipated *Law*[51]. In pioneers' communities spontaneous methods of dispute resolution, based on «custom and contract», arose[52], as in the case of fur traders in Wisconsin[53], or in the case of cowboy factions in Wyoming[54]; communal administration of justice was typical of miners' communities in California, Montana and Colorado, where autonomous codes, courts and vigilantes committees took shape[55]. The practice of *Regulators* and *Vigilantes* was constant in the West[56], both as a sudden, spontaneous, need of the people to «take the law into their own hands», as well as under authorization by sheriffs[57] or even by states legislation, as in the case of Texas Rangers[58].

Perhaps, in the context of a wild and unorganized society, these patterns of private justice granted basic and primitive performances of the administration of justice. However, the price payed for delegation of sovereign functions to private actors was – almost everywhere – violence, corruption, inequality, violation of basic standards of procedure, and the lack of effectiveness of justice. As a result, after a transitional stage of territorial government, those raw experiences were substituted by the unavoidable

---

48 A. Phillips – J.C. Sharman, *Outsourcing Empire. How Company-States Made the Modern World*, Princeton - Oxford, 2020, 7. See also A. Weststeijn, *The VOC as a Company-State Debating Seventeenth-Century Dutch Colonial Expansion*, in *Itinerario*, 1, 2014, 13.

49 W. Dalrymple, *The Anarchy: The Relentless Rise of the East India Company*, London, 2020.

50 A. Buratti, *La frontiera americana. Un'interpretazione costituzionale*, Verona, 2016, 83 ss.

51 W. Gard, *Frontier Justice*, Norman, 1949, VI.

52 A.P. Morris, *Hayek and Cowboys: Customary Law in American West*, in *N.Y.U. Journ. of Law and Liberty*, 1, 2005, 35 ss., spec. 42.

53 D.P. Kommers, *The Emergence of Law and Justice in Pre-Territorial Wisconsin*, in *Am. Journ. of Leg. Hist.*, 1, 1964, 28.

54 A.P. Morris, *Hayek and Cowboys: Customary Law in American West*, cit., 43.

55 J.F. Burns, *Taming the Elephant. An Introduction to California's Statehood and Constitutional Era*, in J.F. Burns – R.J. Orsi (eds.), *Taming the Elephant. Politics, Government, and Law in Pioneer California*, Berkeley – Los Angeles – London, 2003, 6.

56 P.B. Nolan, *Vigilantes on the Middle Border. A Study of Self-Appointed Law Enforcement in the States of the Upper Mississippi from 1840 to 1880*, New York – London, 1987, 4-8, spec. 11; C.G. Fritz, *Popular Sovereignty, Vigilantism, and the Constitutional Right of Revolution*, in G.M. Bakken – B. Farrington (eds.), *The American West. Interactions, Intersections, and Injunctions*, New York – London, 2000, 93 ss.

57 P.B. Nolan, *Vigilantes on the Middle Border*, cit., spec. 35, 42-43.

58 W. Gard, *Frontier Justice*, cit., 217 ss.

expansion of public law and legitimate political institutions.

For many decades, the web has been a new uncharted frontier land, made available by digital technology. After an initial stage of self-regulation, we should now lead it back under the authority of public, democratic norms and institutions. The fate of digital constitutionalism passes through a process of re-imagination of online platforms: as private economic actors, their freedom must be valued; as gateways to the digital infrastructure, they are authentic global commons[59], instrumental to the exercise of liberties and to the protection of rights – deserving, therefore, a more stringent public regulation of their duties and responsibilities.

[59] The issue is broadly discussed in M. Kwet, *Fixing Social Media: Toward a Democratic Digital Commons*, *Markets,* in Globalization & Development Rev., 1, 2020; see also, although with a focus on data rather than on platforms, J. Shkabatur, *The Global Commons of Data*, in *Stanford Tech. Law Rev.*, 2019, 354.