



**FRIEDRICH NAUMANN
FOUNDATION** For Freedom.

PLATFORM- PROOFING DEMOCRACY

**Social Media Councils as Tools to Increase
the Public Accountability of Online Platforms**

Matthias C. Kettmann and Martin Fertmann

Imprint

Publisher

Friedrich-Naumann-Stiftung für die Freiheit
Truman-Haus
Karl-Marx-Straße 2
14482 Potsdam-Babelsberg

/freiheit.org

/FriedrichNaumannStiftungFreiheit

/FNFreiheit

/stiftungfuerdiefreiheit

Authors

PD Dr. Matthias C. Kettmann, LL.M. (Harvard) and Martin Fertmann

Editor

Referat Globale Themen im Fachbereich Internationales
Ann Cathrin Riedel, Themenmanagerin Digitalisierung & Innovation

Contact

Telefon +49 30 220126-34

Telefax +49 30 690881-02

E-Mail service@freiheit.org

As of

May 2021

Notes on using this publication

This publication is an information service of the Friedrich Naumann Foundation for Freedom. The publication is available free of charge and not for sale. It may not be used by parties or election workers for the purpose of campaigning (Bundestag, regional and local elections and elections to the European Parliament).

License

Creative Commons (CC BY-NC-ND 4.0)

Table of contents

KEY FINDINGS OF THE STUDY	4
1. DEFINITION OF OBJECTIVES	5
2. INTRODUCTION.....	5
3. HISTORY AND DESIGN OF SOCIAL MEDIA COUNCILS	7
3.1. Overview and definition.....	7
3.2. Origins of the concept.....	7
3.3. Requirements.....	8
3.4. Design decisions	9
3.4.1 Areas of activity.....	9
3.4.2 Regulatory implementation: self-regulation or co-regulation.....	9
3.4.3 Mission: advisory, quasi-legislative, quasi-judicial.....	10
3.4.4 Membership: councils of experts or citizens	10
3.4.5 Geographic jurisdiction: national, regional or global	10
3.4.6 “Material” jurisdiction: platform-specific or industry-wide	11
3.4.7 Sources of inspiration of existing institutions of self-regulation.....	11
4. THE FACEBOOK OVERSIGHT BOARD	12
4.1. Evolution	12
4.2. Operating principle.....	12
4.3. First decisions	13
5. FURTHER EXAMPLES OF SOCIAL MEDIA COUNCILS.....	13
5.1. Other company initiatives.....	13
5.2. Planned “Social Media Council” in Ireland	14
5.3. “Institutions of regulated self-regulation” under the German Network Enforcement Act (<i>Netzwerkdurchsetzungsgesetz, NetzDG</i>).....	14
6. SUMMARY AND CONCLUSION: THE POTENTIAL OF SOCIAL MEDIA COUNCILS	15

Key findings of the study

- Enabling political participation on the internet is an important challenge of the digital age.
- The core achievement of the democratic revolutions was to successfully fight for the right to participate in determining the rules that shape what can be said in any given jurisdiction.
- Platforms define rules for communication spaces which are essential for opinion aggregation and articulation – almost entirely without democratic oversight and subject to very limited judicial control.
- Our democracy is in need of an update aimed at democratising the increasingly private disposition of the communication order: Democracy has to be platform-proofed, while the platforms have to meet more stringent requirements regarding their democratic legitimacy.
- Social media councils are a promising concept for mitigating existing shortcomings of corporate norm setting and enforcement. Existing comparable media regulating institutions, such as press or broadcasting councils, can serve as a source of inspiration, but should not be copied one-to-one as the control requirements are significantly different.
- Currently, there is insufficient evidence regarding the optimal design of social media councils. According to the current state of knowledge, a combination of a complaints institution (quasi-judiciary) and participation in designing the rules (quasi-legislature) would appear optimal. In any case, participation in assessing and designing measures that affect the visibility of users' content to others without their knowledge is critical.
- Social media councils may be able to check possible violations of terms and conditions or community standards on a case-by-case basis. But their real benefit lies in the systematic improvement of the governance systems of companies beyond just individual cases, which is made more likely if their membership is representative.
- Quasi-judicial social media councils are not suited to controlling and correcting corporate speech governance decisions at scale. Instead, their potential lies in checking a small number of "leading cases" to help improve general systems.
- The core risk of social media councils is that companies may use them as fig leaves to hide abuses or only provide selective relief; to counter this, civil society control must be brought to bear.
- Social media councils should report on the level of actual implementation of the systemic improvements they propose. They should also provide researchers with the required data access to verify this.
- Despite remaining shortcomings, iterative improvement (even at a small scale) by means of greater social accountability for private corporate decisions can only be positive. Spreading out the platforms' decision-making powers even slightly should not be rejected as a matter of principle ("more could be done") if one follows separation of power arguments.
- The Facebook Oversight Board represents a first important example of a social media council. It provides valuable material for analysis – both regarding advantages and disadvantages – as the first decisions have already been published. But it should not be elevated to a role model or monopolise the debate around social media councils conceptually or terminologically.
- The current development of platform regulation at the European level contains new proposals regarding the instrumentalisation of private governance systems, but at this stage provides no corresponding creative attempts at increasing citizen involvement in platform norm setting.
- A promising route to pursue is to develop a model for social media councils based on the extensive literature and constitutional jurisprudence on broadcasting councils.
- The debate around the potential of social media councils to reimport democratic values into the private orders of public communication has only just begun. Despite current uncertainties around their exact designs, they represent a good opportunity to increase the legitimacy of these orders, strengthen the protection of individual rights, and promote social cohesion.

1. Definition of objectives

It is common cause that the evolution of the internet has had an impact on private and public communication behaviour. The internet has become one of the most important tools we use to exercise our rights, especially the right to information and the right to freedom of expression. As the European Court of Human Rights put it in 2015, the internet provides “essential tools for participation in activities and discussions concerning political issues and issues of general interest”.¹ The Committee of Ministers of the Council of Europe emphasised that “the internet plays a particularly important role with respect to the right to freedom of expression”.²

But where exactly does “communication on the internet” take place? Very often on and via platforms. We understand platforms to mean service providers offering internet-based Web 2.0 applications, linking user-generated content by means of application-specific user profiles. Platforms regulate access to the online communication space; indeed, they help constitute it. Consequently, the companies providing these services play an important role in the communication framework. This gives rise to certain questions: Given that this form of communication, with its high relevance for democracy, is privately designed and managed, how can citizens influence the rules which determine the limits of what may be said online? How can platforms enable greater citizen involvement in norm-setting, and what would the consequences of doing so be for platforms, states, individuals and societies? In essence: Who is allowed to define the rules which regulate online spaces? Do rules formulated exclusively by platforms exercising their domiciliary right ipso facto suffer a legitimacy deficit? Can platforms be assigned institutions which help ensure greater accountability to society? Can such “social media councils” platform-proof democracy?

With the establishment of the Facebook Oversight Board³ in 2020, there now exists an example of a social media council which can be analysed – in terms of both its strengths and its weaknesses – as a sample of the institutionalised expression of the desire to integrate external experts in content governance decisions. Do social media councils offer a silver bullet to the challenge of political participation in the digital age? Does establishing such councils offer a convincing way of platform-proofing democracy and making the platforms more democratic?

These are the questions that helped establish the scope of this study, which is made up of six parts. We start with an introduction to the challenges of political participation in the digital age, with a particular focus on the normative design of opportunities for participation in setting and enforcing platform norms. (2.) This is followed by an overview of the history and conceptualisation of social media councils and (3.) an analysis of the Facebook Oversight Board as a paradigmatic social media council (4.). An overview of other social media councils (5.) and a concluding summary and appraisal (6.) round the study off.

2. Introduction

In 2014, a workshop on “Public International Law of the Internet”, hosted in Berlin by the German Foreign Office, among others, concluded that all digital policy stakeholders without exception were dissatisfied with the status quo: “States are frustrated about being unable to enforce the law on the internet. In the absence of clear and applicable regulations, companies don’t know how to deal with (state and private) requests; they are effectively given no choice but to administer justice. Users worry about their data and about violations of their fundamental rights.”⁴ These frustrations represent a considerable challenge both for the 4.4 billion people who have access to the internet and for the 3.3 billion who do not,⁵ as internet governance and access to online content were recognised as being constitutionally relevant topics at an early stage. For example, the United Nations were quick to link democratic constitutionality and development, but also orient the internet towards human development based on constitutional principles. At the UN World Summit on the Information Society (WSIS) (2003, 2005), the states of the world committed themselves to “a people-centred, inclusive and development-oriented Information Society”, to be based on the purposes and principles of the Charter of the United Nations, international law and multilateralism, and “respecting fully and upholding the Universal Declaration of Human Rights”.⁶

1 EGMR 01.02.2015, Nos 48226/10 and 14027/11, Cengiz and Others vs Turkey, section 49.

2 Committee of Ministers of the Council of Europe (2018): Recommendation CM/Rec(2018)2 of the Committee of Ministers to member States on the roles and responsibilities of internet intermediaries, 7 March 2018, https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=0900001680790e14, p. 2.

3 Oversight Board, <https://oversightboard.com>.

4 Alexander von Humboldt Institut für Internet und Gesellschaft (HIIG), Workshop zu „Völkerrecht des Netzes“, 8 September 2014, 7.

5 Kettemann, Die normative Ordnung der Cyber-Sicherheit. Zum Potenzial von Cyber-Sicherheitsnormen, Normative Orders Working Paper 01/2019; Kettemann, Ein Internet für alle Menschen, Tagesspiegel Background Digitalisierung und KI, 5 June 2019, <https://background.tagesspiegel.de/ein-internet-fuer-alle-menschen>.

6 UN Doc. WSIS-05/TUNIS/DOC/7-E, <https://www.un.org/depts/german/conf/wsis-05-tunis-doc7.pdf>.

In and of itself, internet access does not lead to more democracy, although the rule of law and high internet access levels are positively correlated. However, the internet can be deployed as an effective means to strengthen civil society engagement. At the same time, and in addition to protecting spaces where individual freedom can be exercised, we also have to secure the societal prerequisites of social cohesion, which represents a considerable challenge in the face of the privatisation of online communication spaces and the dynamisation of online communication (including the renegotiation of “truths”, the questioning of shared information assets, the changing of communication practices, and the distribution of media portfolios).⁷

A diffuse sense of unease persists, be it regarding the sharing of disinformation related to Covid-19, or, more recently, when the account of a sitting US president was suspended: The measures taken by platform companies are often welcomed in substance, but their impact on democratic discourse processes and democracy per se is considered a challenge. A comment by German Chancellor Angela Merkel was emblematic of this unease. In a statement on the Trump matter, the chancellor said that it was problematic that important decisions regarding communication rules (and the presence of politicians) in communication spaces were no longer being made by “lawmakers”, but by “the managers of social media platforms”.⁸

That all relevant forces should be involved in developing, adopting and enforcing rules in functioning democracies is not a new insight. It helps to counter the concentration of public opinion and the concentration of power in the (communicative) structures in which social innovation is generated. As Germany’s Federal Constitutional Court emphasised in 1986 on the topic of the freedom of broadcasting, it is sufficient to transfer “all significant decisions to an external organ that is independent of the state and which is subject to the influence of the relevant social forces and trends” while putting in place effective legal provisions to prevent a concentration of the power to shape public opinion.⁹

Yet currently many decisions with a considerable impact on online communication are essentially being taken by platforms on their own. While it is true that platforms have increasingly constructed their own normative orders as coherently conceived rule sets equipped with narratives to establish legitimacy,¹⁰ they are generally far removed from the demands of the Committee of Ministers of the Council of Europe in its recommendation on internet intermediaries: “The process of drafting and applying terms of service agreements, com-

munity standards and content-restriction policies should be transparent, accountable and inclusive. Intermediaries should seek to collaborate with (...) organisations representing the interests of users and affected parties (...) before adopting and modifying their policies. Intermediaries should seek to empower their users to engage in processes of evaluating, reviewing and revising, where appropriate, intermediaries’ policies and practices. (...) Internet intermediaries should make available – online and offline – effective remedies and dispute resolution systems that provide prompt and direct redress in cases of user, content provider and affected party grievances.”¹¹

It is without any doubt possible to align these requirements with the core responsibilities of states in this context, namely the protection of fundamental and human rights in the digital environment.¹² States not only have the negative obligation of not violating the right to freedom of expression and other human rights in the digital context, but also the positive obligation to protect human rights while creating a regulatory environment for all, in which everybody can exercise these rights.

As most communication spaces on the internet are privately owned, intermediaries, including social media companies, have become important normative actors. Network effects and acquisitions have led to a situation where a relatively small number of important platform companies dominates the market. These companies have certain obligations under international and domestic law. In accordance with the UN Guiding Principles on Business and Human Rights and the embedded “Protect, Respect and Remedy” framework (UN Guiding Principles, “Ruggie Principles”),¹³ intermediaries should respect the human rights of their users (and other affected parties) in all their activities (including in formulating and applying terms of use) and remedy any negative impacts on human rights directly linked to their business activities.

At the global level, rights-based entitlements of individuals to participate in internet governance are being incorporated by means of the increased inclusion of individuals in governance decisions related to the internet.¹⁴ Everybody – and especially citizens – has a democratic interest in participating in the internet and its regulation, in other words a stake, a value-based interest in the process and outcome of regulation, the operationalisation of which requires involving all stakeholders in all phases and normative processes.¹⁵

7 See Kettemann, Menschenrechte und politische Teilhabe im digitalen Zeitalter. Expert opinion provided in response to a request by the Committee on Human Rights and Humanitarian Assistance of the German Bundestag (Arbeitspapiere des Hans-Bredow-Instituts, Works in Progress # 2), 17 June 2020, <https://leibniz-hbi.de/de/publikationen/menschenrechte-und-politische-teilhabe-im-digitalen-zeitalter>.

8 Tagesspiegel, 11 January 2021, <https://www.tagesspiegel.de/politik/meinungsfreiheit-von-elementarer-bedeutung-merkel-kritisiert-twitter-sperrung-fuer-trump/26786886.html>.

9 BVerfG, 4 November 1986, 1 BvF 1/84 (4. Rundfunkentscheidung), <https://openjur.de/u/175210.html>.

10 Kettemann/Schulz, Setting Rules for 2.7 Billion. A (First) Look into Facebook’s Norm-Making System: Results of a Pilot Study (Hamburg: Working Papers of the Hans-Bredow-Institut, Works in Progress # 1, January 2020), https://leibniz-hbi.de/uploads/media/Publikationen/cms/media/5pz9hwo_AP_WiP001InsideFacebook.pdf.

11 Council of Europe, Recommendation CM/Rec(2018)2 of the Committee of Ministers to member States on the roles and responsibilities of internet intermediaries, section 11; appendix 2, especially 2.2.2. on enabling users to participate in formulating rules and 2.5 regarding access to effective complaint mechanisms (our emphasis).

12 Kettemann, The Normative Order of the Internet. A Theory of Online Rule and Regulation (Oxford: Oxford University Press, 2020).

13 See “Ruggie Principles”: Report of the Special Representative of the Secretary-General on the issue of human rights and transnational corporations and other business enterprises, John Ruggie, Guiding Principles on Business and Human Rights: Implementing the United Nations “Protect, Respect and Remedy” Framework, UN Doc A/HRC/17/31 dated 21 March 2011 (German version).

14 German Foreign Office, Recommendation 5A/B, Options for the Future of Global Digital Cooperation, https://www.global-cooperation.digital/GCD/Redaktion/EN/Downloads/options-for-the-future-of-global-digital-cooperation.pdf?__blob=publicationFile&v=2. See Kettemann/Kleinwächter/Senges/Schweiger, Comments on Recommendation 5A/B of the UN High Level Panel on Digital Cooperation, How to Build an Enhanced Mechanism for Digital Cooperation. A Multistakeholder Statement from Germany, 27 April 2020, https://www.global-cooperation.digital/GCD/Redaktion/EN/Downloads/kleinwaechter-kettemann.pdf?__blob=publicationFile&v=2.

15 Kettemann, Internet Governance, in Jähnel/Mader/Stauder (eds.), Internetrecht, 4th edition (Vienna: Verlag Österreich, 2020), 47-73.

But how can the participation entitlements of individuals be realised at a smaller scale, too – by platforms? How can these be made “more democratic”?

The concept of social media councils is a valuable starting point in this debate. In Germany, it can be embedded in the fertile legal context of decades of experience with “council-based” governance in the media sector, including through the broadcasting and television councils of public broadcasters, the media councils of the state media authorities responsible for broadcast and telemedia,¹⁶ and sectoral self-regulating bodies such as press¹⁷ and advertising¹⁸ councils.

Properly understood, social media councils are no utopia of self-regulation in the sense of John Perry Barlow’s famous “Declaration of Independence” of cyberspace.¹⁹ They are not meant to *replace* existing models of private and state regulation, but rather *complement* them “to create an independent, accountable, and transparent mechanism that can cooperate with platforms to improve their own systems and eliminate the need for some regulation.”²⁰

Accepting that there are no silver bullets and that incremental improvements are the best one might hope for in the complex regulatory triangle between states, companies and civil society²¹ – not least because of the challenging diversity of regulatory objectives – the question arises: How can social media councils contribute to such improvements?

3. History and design of social media councils

3.1. Overview and definition

We use “social media councils” to mean external governance structures tasked either with *formulating and/or applying rules or determining the discoverability or visibility of content* on social networks in addition to or instead of the platforms; or tasked with *monitoring* the platform’s activities relating thereto. This implies that the membership of “social media councils” can include civil society representatives²² and/or experts with the aim of creating multi-stakeholder governance, although it is not strictly necessary for them to be referred to by that name. The name should therefore not pre-empt the complex question of whether such an institute effectively constrains the influence of a company or – going even further – legitimises the social accountability of the governance system.

3.2. Origins of the concept

Proposals for governance mechanisms that provide affected parties with an independent channel for complaints and/or involve civil society or user representatives in formulating the private rules of platform companies are nothing new. Quasi-judicial private institutions for reviewing company decisions started appearing in the German debate on regulation almost a decade ago, for instance with reference to “Cyber Courts”²³ or in the form of alternative dispute resolution mechanisms.²⁴ In the US, participatory approaches to formulating the private rules of companies have been proposed, such as a “Content Congress”²⁵ or external advisory bodies.²⁶

16 This refers specifically to the decision-making bodies of the state media authorities and commissions, which are referred to as media councils (“Medienräte”) in Baden-Württemberg, Bavaria, Berlin/Brandenburg, Hamburg/Schleswig-Holstein, Saarland and Saxony; for example, see sections 39-47 of the State Media Treaty for Hamburg and Schleswig-Holstein, the “Medienstaatsvertrag HSH”.

17 Deutscher Presserat (German Press Council), <https://www.presserat.de>.

18 Deutscher Werberat (German Advertising Council), <https://www.werberat.de>.

19 See Barlow (1996): “A Declaration of the Independence of Cyberspace” stating that “We believe that from ethics, enlightened self-interest, and the commonweal, our governance will emerge.”, <https://www.eff.org/de/cyberspace-independence>.

20 Donahoe/Hughes/Kaye (2019): “Social Media Councils: From Concept to Reality.” https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/gdpiar_t19_smc_conference_report_wip_2019-05-12_final_1.pdf, p. 8.

21 Gorwa (2019). The platform governance triangle: conceptualizing the informal regulation of online content. *Internet Policy Review*, 8(2). <https://doi.org/10.14763/2019.2.1407>.

22 For an overview of the definition and potential of this approach to expression on the internet, see Strickling/Hill (2018): “Multi-stakeholder Governance Innovations to Protect Free Expression, Diversity and Civility Online”, in: Donahoe/Hampson: “Governance Innovation for a Connected World. Protecting Free Expression, Diversity and Civic Engagement in the Global Digital Ecosystem” (pp. 45-52), <https://www.cigionline.org/sites/default/files/documents/Stanford%20Special%20Report%20web.pdf>.

23 adeur/Gostomzyk: “Der Schutz von Persönlichkeitsrechten gegen Blogs”, *NJW* 2012, 710 (pp. 713); Ladeur: “Neue Institutionen für den Daten- und Persönlichkeitsschutz im Internet: “Cyber-Courts” für die Blogosphäre”, *DUD* 2012, 711 (pp. 712); also see Vesting (2015): *Die Medien des Rechts*. Bd. 4: Computernetzwerke. Weilerswist: Velbrück Wissenschaft, p. 205.

24 Spindler: “Persönlichkeitsschutz im Internet – Anforderungen und Grenzen einer Regulierung” Gutachten F on the occasion of the 69th German Jurists’ Conference, 2012, p. 133; a similar proposal was made by Wagner: “Haftung von Plattformen für Rechtsverletzungen (Teil 2)”, *GRUR* 2020, 447; such procedures appear to hold promise regarding justice in individual cases, but seem less suited to triggering general improvements in the governance systems of the platforms; see the comparison by Brown (2020): “Models of Governance for Online Hate Speech”, Council of Europe <https://rm.coe.int/models-of-governance-of-online-hate-speech/16809e671d> (p. 84).

25 Tomson, D., Morar, D. (2018). A Better Way to Regulate Social Media. *Wall Street Journal*, <https://www.wsj.com/articles/a-better-way-to-regulate-social-media-1534707906>.

26 Ash, Timothy Garton, Robert Gorwa and Danaë Metaxa. 2019. *Glasnost! Nine Ways Facebook Can Make Itself a Better Forum for Free Speech and Democracy*. Reuters Institute for the Study of Journalism, https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2019-01/Garton_Ash_et_al_Facebook_report_FINAL_0.pdf, pp. 19 – 20.

The current debate around social media councils was advanced by proposals from NGOs, such as Global Partners Digital²⁷ and ARTICLE 19²⁸. ARTICLE 19's concept of "Social Media Councils" was mentioned in the 2018 annual report by the then UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, David Kaye.²⁹ His recommendation for platform companies was that they

"must open themselves up to public accountability. Effective and rights-respecting press councils worldwide provide a model for imposing minimum levels of consistency, transparency and accountability to commercial content moderation. (...) All [companies] that moderate content or act as gatekeepers should make the development of industry-wide accountability mechanisms (such as a social media council) a top priority."³⁰

The debate around social media councils is closely related to demands that platforms should align their private rules, which are often international in scope, with international human rights standards. In this context, social media councils as potential external supervisory bodies would play a role by publicly criticising violations, thereby in a sense acting as an institutionalised "trigger" to create societal and political pressure on companies.

In addition, social media councils could also be used to verify that national agencies' commands and requests to platforms are in compliance with international human rights standards. In cases where such actions were in violation of applicable human rights standards, the councils could publicly back companies in rejecting the requests; but this is only being proposed in isolated instances.³¹ The debate – and, therefore, this study – is primarily focused on ways of boosting the democratic legitimacy of the platforms' private orders.

27 With a proposal that concentrates on non-binding notes regarding the formulation of private rules by an "Independent Online Platform Standards Oversight Body", see Global Partners Digital (2018): "A Rights-Respecting Model of Online Content Regulation by Platforms", <https://www.gp-digital.org/wp-content/uploads/2018/05/A-rights-respecting-model-of-online-content-regulation-by-platforms.pdf>, pp. 26 - 28.

28 Article 19: "Self-regulation and 'hate speech' on social media platforms" (2018). <https://www.article19.org/resources/self-regulation-hate-speech-social-media-platforms/>

29 UN General Assembly (2018): "Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression", UN A/HRC/38/35, https://ap.ohchr.org/documents/dpage_e.aspx?si=A/HRC/38/35, paragraphs 58, 59, 63, 72.

30 Ibid.

31 The idea was favourably received by some companies in 2019, but was not pursued further, see Donahoe/Hughes/Kaye (2019): "Social Media Councils: From Concept to Reality." https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/gdpart_19_smc_conference_report_wip_2019-05-12_final_1.pdf at the bottom of p. 13; the most powerful illustration of the need for robust, international mechanisms to protect human rights on digital platforms – even against infringements by states – is the Rohingya genocide in Myanmar, which was partly stoked by appeals on social media platforms, see Irving (2019): "Suppressing Atrocity Speech on Social Media", in: AJIL Unbound 113: 256-261, <https://www.cambridge.org/core/journals/american-journal-of-international-law/article/suppressing-atrocity-speech-on-social-media/494334D2936A6A6E7C547C70816714D4>; the most recent example of the potential of such an institution is provided by directives issued by India against Twitter, which likely contravened the guarantees by the Indian Constitution, Mahapatra/Fertmann/Kettemann (2021): Twitter's Modi Operandi: Lessons from India on Social Media's Challenges in Reconciling Terms of Service, National Law and Human Rights Law, [Verfassungsblog, https://verfassungsblog.de/twitters-modi-operandi](https://verfassungsblog.de/twitters-modi-operandi).

3.3. Requirements

Social media councils as a concept do not have a very long history. Insofar as they are used to supervise the discretionary powers of platforms and act in areas in which the platforms are not restricted by applicable national legislation, national law does not provide much in the way of substantive or procedural criteria for them. Human rights are a more important source of law in this context, not least because of the implied supervisory function for private platform law; this applies even more for quasi-judicial social media councils.

The United Nations Guiding Principles for Business and Human Rights (UNGP) represent the most important benchmark for quasi-judicial social media councils. Their "soft law" defines a corporate responsibility to respect human rights. Under Principles 29, 30 and 31, the UNGP also formulate guidelines for creating non-governmental, especially corporate or independent, complaints institutions and procedures.³²

In this sense, social media councils (also) have to ensure that their decisions comply with international human rights norms. Regarding the institutional design and the procedural practices of social media councils, the UNGP require institutions to be suitable to justify the trust of those who use them, based on transparent and independent membership and transparent processes. Furthermore, they have to ensure that proceedings are fair and accessible to all who may potentially be affected. They have to warrant that affected parties dispose of all the information needed to present their cases, and their decision-making methods must be transparent.

In addition to the UNGP, the "Santa Clara Principles", an industry norm,³³ can also be used to develop minimum criteria for social media councils. The principles require that "companies should provide a meaningful opportunity for timely appeal of any content removal or account suspension." The accompanying list of minimum standards for such appeals mechanisms lists some of the key components of a due process, e.g., "Human review by a person or panel of persons that was not involved in the initial decision. An opportunity to present additional information that will be considered in the review. Notification of the results of the review, and a statement of the reasoning sufficient to allow the user to understand the decision. In the long term, independent external review processes may also be an important component. (...)"

Even in the absence of national regulations for social media councils, there are therefore already guidelines that can be used to assess and shape the design and decision-making practice of such institutions.

32 See also Council of Europe: Recommendation CM/Rec(2018)2 of the Committee of Ministers to member States on the roles and responsibilities of internet intermediaries, section 11; appendix 2, especially 2.2.2. on enabling users to participate in formulating rules and 2.5 regarding access to effective complaint mechanisms.

33 "The Santa Clara Principles: On Transparency and Accountability in Content Moderation" are an industry norm developed through collaboration by civil society and academia. They have been adopted by many companies on a voluntary basis; see <https://santaclaraprinciples.org/>; and Crocker et al: Who Has Your Back? Censorship Edition 2019, <https://www.eff.org/de/wp/who-has-your-back-2019>.

3.4. Design decisions

3.4.1. Areas of activity

As comprehensive lists of possible areas of activity of social media councils already exist,³⁴ we will direct our attention to some core questions regarding the design of such institutions here. A transatlantic working group on platform governance accurately described the many possible factors affecting the design of social media councils: "Policy makers and multi-stakeholder groups might consider a wide range of organizational structures and precedents to choose from, with format, purpose, jurisdiction, makeup, member selection, standards, scope of work, and scalability to be determined in line with the underlying mission of the council."³⁵

A council's underlying mission can be manifold. Possibilities include:

- preventive protection against unjustified measures against content; remediation after such measures have been imposed;
- systematic impulses to improve the governance systems of companies beyond individual cases;
- enabling an access to justice for as many affected parties as possible beyond automated and/or internal platform mechanisms;
- greater transparency;
- diversity-oriented supervision of content curation;
- securing the discoverability of certain content in the public interest;
- specific supervision regarding political campaign advertising and political communication; and
- supervision of basic design decisions and potential influences that guide users (so-called persuasive design and dark patterns).

3.4.2. Regulatory implementation: self-regulation or co-regulation

Social media councils can be implemented on the basis of voluntary cooperation between companies and experts and/or civil society (self-regulation), or alternatively in the form of models where social media councils are embedded in a framework defined by the state (co-regulation, regulated self-regulation³⁶).

Statutory implementations of co-regulation are conceivable. But it is unclear to which extent lawmakers can, within a constitutional framework, prescribe procedures and institutions that lie outside the domain of government to platforms that make decisions and formulate rules which lawmakers cannot make and formulate precisely because the opinion-forming process lies outside the jurisdiction of government.

If social media councils are to improve company decision making in areas where lawmakers are unable or unwilling to formulate guidelines, it is apparent that social media councils will – at least initially³⁷ – be implemented by means of self-regulating initiatives. But even such self-regulation could be politically encouraged and collaboratively shaped, for instance when concrete, but non-binding proposals for an institution are formulated and political pressure is exerted on companies to participate in implementing them effectively ("quasi-regulation").³⁸ Examples include initiatives by expert NGOs such as ARTICLE 19, which exerted considerable influence on the process of developing the social media council concept and which is currently involved in introducing a national social media council in Ireland, as well as Ranking Digital Rights, which is already measuring the transparency of large technology companies by means of an annual "Corporate Responsibility Index".³⁹ The decisions made by such a social media council, which had its origins in voluntary self-regulation, could later be taken into consideration by agencies and courts when interpreting existing obligations, thereby helping to solidify them.⁴⁰ It is also conceivable that incentives for participating in such an institution could be set at a later stage, for example by pointing out that the alternative is more stringent regulation.⁴¹

34 See the functions listed by Tworek, (2019) Social Media Councils, pp. 99, https://www.cigionline.org/sites/default/files/documents/Platform-gov-WEB_VERSION.pdf#page=100.

35 Transatlantic High Level Working Group (2020), Freedom and Accountability A Transatlantic Framework for Moderating Speech Online, https://cdn.annenbergpublicpolicycenter.org/wp-content/uploads/2020/07/Freedom_and_Accountability_TWG_Final_Report.pdf, p. 26.

36 An early contribution: Wolfgang Schulz / Thorsten Held: Regulierte Selbstregulierung als Form modernen Regierens. Commissioned by the Federal Commissioner for Cultural and Media Affairs of Germany. Final report. Hamburg: Publisher: Hans-Bredow-Institut, May 2002, p. 5, <https://www.hans-bredow-institut.de/uploads/Wds/media/Publikationen/cms/media/a80e5e6dbc2427639ca0f437fe76d3c4c95634ac.pdf>.

37 A different view is put forward by Jarren/Gostomzyk (2020): Facebook's Hausgericht, <https://www.medienpolitik.net/2020/04/facebooks-hausgericht>, according to which co-regulation could already be considered at this stage.

38 Tworek, Heidi (2019) Social Media Councils, (p. 100) https://www.cigionline.org/sites/default/files/documents/Platform-gov-WEB_VERSION.pdf#page=100.

39 The initiative "Who targets me?" in support of transparent political advertising has also expressed an interest in participating in social media councils, see <https://who-targets.me/en/oversight-boards-for-everything>.

40 For example, see the reference to the press codex, initially developed by the press in a self-regulating process, in interpreting statutory media obligations: Begr. zum Medienstaatsvertrag, LT-Drs. NRW 17/9052, 135; Lent, ZUM 2020, 593 (599); Heins/Lefeldt MMR 2021, 126.

41 For certain journalistic/editorial online media, the new German State Media Treaty ("Medienstaatsvertrag") also takes this step in section 19, see Klaus: "Staatlicher Zahnersatz für den Presserat: Der Medienstaatsvertrag macht die Selbstregulierung der Presse zum Auslaufmodell", VerfBlog, 29 March 2021, <https://verfassungsblog.de/staatlicher-zahnersatz-fur-den-presserat>.

3.4.3. Mission: advisory, quasi-legislative, quasi-judicial

A further fundamental design decision is whether social media councils should be involved at the level of formulating rules and designing enforcement practices or whether they should only check individual judgments after the fact, in response to user complaints.

Involvement in rule-setting can in principle be designed as a quasi-legislative user parliament,⁴² but is limited to an advisory role in norm-setting within many approaches.⁴³ Limiting a social media council to a purely advisory function risks restricting its influence. Conversely, binding rule-setting by a social media council creates the risk that a company could lose control over its platforms, which would likely disincentivise participation from a business perspective.

Apart from allowing involvement in developing rules and practices, creating opportunities for involvement in individual decisions regarding actions against user content is also conceivable. Taking into account the considerable volume of decisions that have to be made, a social media council would not be suitable as the first decision level for moderator decisions or even for initial appeals, but only as a later or higher-level review authority.⁴⁴

Here, one needs to take into account that restitution (restoring the previous state) is only possible within certain limits because unjustified measures taken against content imply negative impacts in the form of missed communication opportunities which cannot be restored when such measures are lifted days, weeks or even months later (vice-versa, content removal represents also a very limited restitution when measures against the content have initially been erroneously rejected).⁴⁵

As the vast majority of such measures are by now automated, social media councils have an opportunity to exert influence by contributing to the design of such tools. Beyond that, it is precisely the design of platforms' feed and recommendation algorithms that represents a potential source of power which has to be supervised, implying that a complaints-based model cannot control platform measures which users don't notice (so-called shadowbanning). The potential benefits of a complaints-based social media council therefore are to be found primarily in potential systemic improvements which such a council could initiate based on individual cases.⁴⁶

Existing social media council concepts, such as those proposed by ARTICLE 19⁴⁷ or the Stanford Global Digital Policy Incubator,⁴⁸ agree that a combination of a complaints-based institution (quasi-judiciary) and involvement in designing rules (quasi-legislative) is required. In any case, involvement in analysing content governance techniques which users are unaware of is critical.

3.4.4. Membership: councils of experts or citizens

Social media councils can be composed of experts on technology governance and freedom of expression, representatives of civil society groups or even randomly selected citizens.⁴⁹

Such approaches to deliberative democracy in randomly selected small groups are discussed under the term "mini publics" and are not without controversy.⁵⁰ Keeping in mind current challenges to democracy, such as political polarisation and the spreading of disinformation, there is however something to be said for the development of "new forms of deliberative, collaborative and participative decision making that are evolving worldwide."⁵¹

On the other hand, formulating recommendations, defining binding rules or adjudicating complaints regarding expression on the net requires a certain level of expert knowledge, among other things to avoid unintended consequences. In this sense, models that combine representation with expert knowledge would seem advisable.

3.4.5. Geographic jurisdiction: national, regional or global

Social media councils can have national, regional or global jurisdiction. The Facebook Oversight Board operates at a global level. This offers certain benefits, but conceptualising councils at the national or regional level is also conceivable to help ensure that cultural and language contexts are appropriately reflected. Unified regional jurisdictions (such as Europe with EU law and the European Human Rights Convention) also suggest that unified social media councils are feasible, which could be designed in a way that covers multiple platforms. Creating social media councils at the national level is also an option. They could function as alternative dispute resolution mechanisms for content moderation decisions made by platforms.⁵²

42 Tomson, D., Morar, D. (2018). A Better Way to Regulate Social Media. Wall Street Journal, <https://www.wsj.com/articles/a-better-way-to-regulate-social-media-1534707906>.

43 See the proposal by Bradley/Wingfield (2018): "A Rights-Respecting Model of Online Content Regulation by Platforms", Global Partners Digital, www.gp-digital.org/content-regulation-laws-threaten-our-freedom-of-expression-we-need-a-new-approach.

44 For an overview of such tasks, see Brown (2020): "Models of Governance for Online Hate Speech", Council of Europe, <https://rm.coe.int/models-of-governance-of-online-hate-speech/16809e671d> (78-83).

45 One of the first cases handled by the FOB (2020-004-IG-UA) serves as an illustrative example. At issue was the removal of content because of nudity; the content had been posted in October 2020 in connection with "Pink October", an international campaign to generate awareness for breast cancer. The FOB arrived at its decision four months later, in January 2021, and emphasised that the impossibility of restitution in the face of the expired campaign made it clear that its decisions needed to aim at transcending individual cases: <https://oversightboard.com/decision/IG-7THR3S11>.

46 Brown (2020): "Models of Governance for Online Hate Speech", Europarat, <https://rm.coe.int/models-of-governance-of-online-hate-speech/16809e671d>, p. 133.

47 Donahoe/Hughes/Kaye (2019): "Social Media Councils: From Concept to Reality." https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/gdpart_19_smc_conference_report_wip_2019-05-12_final_1.pdf, pp. 30-32.

48 Ibid., pp. 26-29.

49 For example, supported by MEP Geese (2021): Social Media Councils: Power to the people, <https://alexandrageese.eu/der-dsa-teil-05-social-media-councils-power-to-the-people>.

50 Bächtiger, André, et al., (eds.). The Oxford Handbook of Deliberative Democracy. Oxford University Press, 2018, p. 1.

51 OECD (2017): Recommendation of the Council on Open Government, available at: <https://www.oecd.org/gov/Recommendation-Open-Government-Approved-Council-141217.pdf>; the OECD also administers a database of representative deliberative institutions: <https://airtable.com/shrRYPpTSs9NskHbv/tblfOHuQuKuOpPnHh>.

52 Donahoe/Hughes/Kaye (2019): "Social Media Councils: From Concept to Reality." https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/gdpart_19_smc_conference_report_wip_2019-05-12_final_1.pdf, p. 30.

Again, the design choices made here are not necessarily mutually exclusive, as national councils could be connected through a global association that defines best practices regarding the councils' work and principles.

3.4.6. "Material" jurisdiction: platform-specific or industry-wide

The jurisdiction of social media councils can be limited to a specific platform (platform-specific) or extended to cover many or all platforms or a specified type of platform (industry-wide). A specialised social media council would appear to be easier to implement by comparison, as only one company would need to support the concept if it is implemented through voluntary self-regulation. An industry-wide social media council would encounter greater challenges, not least in interacting with a multitude of different platforms and their diverse governance systems. A challenge for building such industry-wide mechanisms may also lie in applicable national anti-trust law that may restrict such forms of product (policy) related cooperation.

On the other hand, an industry-wide approach is especially promising because it could contribute to the independence of the institution. With industry-wide jurisdiction, such a council would not depend on its relationship with one or just a few companies for its existence and acceptance. As with other design dimensions, an iterative approach might be advisable, i.e., a social media council could be launched as an initiative by one or two companies and evolve over time into an industry-wide institution. It cannot be ruled out that other companies might join the platform-specific social media council of a competitor at a later stage. In the discussion to date, the predominant call has been for the establishment of industry-wide social media councils.⁵³

3.4.7. Sources of inspiration of existing institutions of self-regulation

At this stage, the discussion around the design of social media councils is very much at the initial stages. Nonetheless, there are existing models of private institutions of self-regulation which are mentioned as examples or role models in the context of the social media council debate. They include press councils,⁵⁴ the US Financial Industry Regulatory Authority (FINRA),⁵⁵ the Canadian Broadcast Standards Council (CBSC),⁵⁶ the Internet Corporation for Assigned Names and Numbers (ICANN)⁵⁷ and various institutions of self-regulation established in other industries to ensure that practices comply with human rights, for instance in the resource extraction and manufacturing industries.⁵⁸

From a German perspective, an approach to define the design and composition of a social media council could take its guidance from the broadcasting councils of the German public broadcasters. These councils include socially relevant groups, such as unions, employer associations, churches, environmental groups etc., which are considered "trustees of the interests of the general public"⁵⁹ and monitor compliance with statutory duties in this capacity.⁶⁰ There is no constitutional requirement for representatives to be affiliated with such associations. Instead, lawmakers can also require unaffiliated or weakly organised groups to be represented in some other way.⁶¹ This appears increasingly advisable in light of the constitutionally required⁶² consideration of the equality provisions of Article 3 of Germany's Basic Law in determining the composition of the councils. A promising route to pursue is to develop a model for social platform councils based on the extensive literature and constitutional jurisprudence on broadcasting councils. Such a model should at the same time reduce existing representation deficits, for instance through (partly) random selection of council members.⁶³

54 UN General Assembly (2018): "Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression", UN A/HRC/38/35, https://ap.ohchr.org/documents/dpage_e.aspx?si=A/HRC/38/35, Section 58.

55 Transatlantic High Level Working Group (2020), "Freedom and Accountability: A Transatlantic Framework for Moderating Speech Online", https://cdn.annenbergpublicpolicycenter.org/wp-content/uploads/2020/07/Freedom_and_Accountability_TWG_Final_Report.pdf, pp. 26 - 27.

56 See the proposal for a national (Canadian) co-regulation "Council for Moderation Standards," based on the example of the CBSC, in Tenove, Tworek, McKelvey (2018): "Poisoning Democracy: How Canada Can Address Harmful Speech online", <https://ppforum.ca/wp-content/uploads/2018/11/PoisoningDemocracy-PPF-1.pdf>, pp. 27 - 28.

57 Tenove/Tworek/McKelvey (2018): "Poisoning Democracy: How Canada Can Address Harmful Speech online", <https://ppforum.ca/wp-content/uploads/2018/11/PoisoningDemocracy-PPF-1.pdf>, pp. 27 - 28.

58 Gorwa, R. (2019). The platform governance triangle: conceptualising the informal regulation of online content. *Internet Policy Review*, 8(2). <https://doi.org/10.14763/2019.2.1407>.

59 Federal Constitutional Court judgment of 25 March 2014, 1 BvF 1/11, available at https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/DE/2014/03/fs20140325_1bvf000111.html (margin no. 40).

60 For a short overview see Schulz/Held/Dreyer/Wind (2008): Regulation of Broadcasting and Internet Services in Germany: a brief overview, available at <https://doi.org/10.21241/ssoar.71697>, pp. 11-12.

61 BVerfGE 83, 238 - 6. Rundfunkentscheidung (p. 335 et seq), available at <https://www.servat.unibe.ch/Dfr/bv083238.html>.

62 See BVerfGE 83, 238 - 6. Rundfunkentscheidung (p. 336 et seq), available at <https://www.servat.unibe.ch/Dfr/bv083238.html>.

63 This is also suggested for broadcasting bodies, see Dobusch (2019) for example: Zusammensetzung der Rundfunkgremien, https://www.deutschlandfunk.de/zusammensetzung-der-rundfunkgremien-schoeffen-fuer-mehr.2907.de.html?dram:article_id=448617.

53 *Ibid.*, proposals by GDPi (from p. 26) and ARTICLE 19 (from p. 30); also in favour Kaye, UN General Assembly (2018): "Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression", UN A/HRC/38/35, https://ap.ohchr.org/documents/dpage_e.aspx?si=A/HRC/38/35, Section 72;

Pursuing this route would also be helpful in balancing the need for the social media council's independence and the large number of checks required with the cooperative relationship it needs to have with the company. If the chosen approach is that of a national social media council designed with reference to broadcasting councils, this could also help to create the required high levels of acceptance. Such a national institution could also be integrated into an international network of social media platforms, as indicated above.

4. The Facebook Oversight Board

4.1. Evolution

In 2018, Mark Zuckerberg announced that he wished to create "an independent appeal process" that would function "almost like a Supreme Court".⁶⁴ Following Zuckerberg's announcement, a global consultation process was set in motion, involving stakeholders from civil society, academia and politics. It lasted for over a year and was intended to establish the jurisdiction of such an institution and its requirements in terms of staffing, organisational, legal, financial etc. resources.⁶⁵ The process of establishing the Facebook Oversight Board (FOB) was subject to close academic scrutiny, especially by Evelyn Douek⁶⁶ and Kate Klonick.⁶⁷

The FOB is not an isolated initiative. It is embedded in the context of Facebook's ongoing slow-but-steady development of the private order system of its platforms towards greater transparency and justification of their private decision-making powers. Measures introduced include global transparency reports,⁶⁸ a formalised system for amending community standards,⁶⁹ informal methods to enable civil society participation in formulating the standards,⁷⁰ and commissioning external researchers to publicly assess the company's decision-making processes.⁷¹

All of these initiatives are to be welcomed, but in the absence of robust supervisory and enforcement structures, they have not imposed any substantial limits on the company's power.

4.2. Operating principle

The FOB is a global, currently platform-specific, quasi-judicial and advisory social media council. Its members are representatives of civil society and academia⁷² who act as an appeals body that reviews Facebook's decisions to delete user-generated content on its platforms Facebook and Instagram. Decisions are based on the respective community standards and take into account "international human rights standards."⁷³ For any given case, the FOB's decisions are binding on the company in terms of its voluntary commitment and are intended to be transferable to "identical content with parallel context."⁷⁴ Furthermore, the FOB acts as an advisory body that issues public, non-binding "recommendations" regarding Facebook's general rules and practices.⁷⁵ The company is required to respond publicly within specified time frames.⁷⁶ Facebook is funding the legally independent body through a \$130m trust for the period 2020–2026.⁷⁷ The FOB Charter was completed and published in November 2019. It provides the framework for the establishment of the FOB and is complemented by a set of bylaws⁷⁸ and a rulebook⁷⁹ adopted by its members.

The FOB has pointedly been given a structure that allows other platform companies to join it, opening the way for it to become an industry-wide institution in time. But whether joining would be attractive for other companies is questionable, given the close ties between the FOB and Facebook.⁸⁰

64 Klein (2018): "Mark Zuckerberg on Facebook's hardest year, and what comes next." *Vox*, 2. April, <https://www.vox.com/2018/4/2/17185052/mark-zuckerberg-facebook-interview-fake-news-bots-cambridge>; Tworek, *Social Media Councils*, <https://www.cigionline.org/articles/social-media-councils>.

65 See Facebook (2019): "Global Feedback & Input on the Facebook Oversight Board for Content Decisions", <https://about.fb.com/wp-content/uploads/2019/06/oversight-board-consultation-report-2.pdf>.

66 Douek, Evelyn: "Facebook's 'Oversight Board': Move Fast with Stable Infrastructure and Humility (2019)." 21 N.C. J. L. & Tech. 1 (2019), <https://ssrn.com/abstract=3365358>.

67 Klonick (2020): "The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression" (June 30, 2020). *Yale Law Journal*, Vol. 129, Nr. 2418, <https://ssrn.com/abstract=3639234>; Dies. (2021): "Inside the Making Of Facebook's Supreme Court" *New Yorker*, <https://www.newyorker.com/tech/annals-of-technology/inside-the-making-of-facebooks-supreme-court>.

68 Facebook: Transparency reports, <https://transparency.facebook.com>.

69 Facebook: *Gemeinschaftsstandards*, <https://www.facebook.com/communitystandards/>, German version <https://de-de.facebook.com/communitystandards>.

70 For greater detail on stakeholder engagement, see https://www.facebook.com/communitystandards/stakeholder_engagement; for information on how it unfolded in practice, see the observations provided in Kettemann/Schulz: "Setting Rules for 2.7 Billion", https://www.hans-bredow-institut.de/uploads/media/Publikationen/cms/media/5pz9hwo_AP_WiP001InsideFacebook.pdf, pp.15; pp. 23.

71 See Reports by the Data Transparency Advisory Group, most recently Bradford, Ben et al.: "Report of the Data Transparency Advisory Group April 2019", https://law.yale.edu/sites/default/files/area/center/justice/document/dtag_report_5.22.2019.pdf.

72 The initial 20 members announced on 6 May 2020 included academics, political (internet) activists, a former prime minister of Denmark and a former judge of the ECHR, see <https://www.oversightboard.com/news/announcing-the-first-members-of-the-oversight-board>; but "some commentators" have criticised the fact that US citizens and people with a legal background are overrepresented if one considers the global Facebook user base, see Article 19 (2020). The Facebook Oversight Board: A significant step for Facebook and a small step for freedom of expression, <https://www.article19.org/resources/facebook-oversight-board-freedom-of-expression>.

73 Facebook: *Community Standards*, available at <https://de-de.facebook.com/communitystandards> (Introduction).

74 Facebook (2019): *Oversight Board Charter*, Article 4, https://about.fb.com/wp-content/uploads/2019/09/oversight_board_charter.pdf.

75 Facebook (2019): *Oversight Board Charter*, Article 3, Section 1 https://about.fb.com/wp-content/uploads/2019/09/oversight_board_charter.pdf.

76 Facebook (2021): *Oversight Board Bylaws*, 2.3.2, <https://www.oversightboard.com/sr/governance/bylaws> (p. 25).

77 Harris: "An Update on Building a Global Oversight Board", available at <https://about.fb.com/news/2019/12/oversight-board-update>.

78 Facebook: *Oversight Board Charter*, https://about.fb.com/wp-content/uploads/2019/09/oversight_board_charter.pdf; *Oversight Board Bylaws*, https://about.fb.com/wp-content/uploads/2020/01/Bylaws_v6.pdf.

79 Oversight Board (2020): *Rulebook for Case Review and Policy Guidance*, <https://oversightboard.com/sr/rulebook-for-case-review-and-policy-guidance>.

80 Lapowsky (2020): How Facebook's oversight board could rewrite the rules of the entire internet, <https://www.protocol.com/facebook-oversight-board-rules-of-the-internet>.

4.3. First decisions

By March 2021, within its first six months of operation, the FOB had published decisions on a dozen cases from a range of countries. That is not even a drop in the ocean; it is a water molecule, considering that Facebook makes around three million⁸¹ decisions regarding content removal every single day.

A closer look at the FOB's decisions shows, however, that it is serious about defining its own position in Facebook's system of norms. In doing so, it refers more to international human rights standards than to the Facebook terms of use.⁸² Even though there is no "hard" mechanism to force Facebook to implement the FOB's decisions, the public pressure generated by its published decisions does seem to have an effect: Facebook has responded in a cooperative way and implemented the recommendations, as far as one can tell. However, there are still significant problems with accessing the data needed to establish to which degree Facebook has really modified its practices.⁸³ The initial decisions indicate that many problems will remain, but that the FOB will gradually improve the platforms' governance systems.

5. Further examples of social media councils

5.1. Other company initiatives

In addition to the FOB, there are several platform-specific advisory bodies which the respective companies use to a greater or lesser extent to market the idea of legitimisation through participation, but which are not designed to be effective "social media councils," as defined here. The bodies in question represent a form of informal cooperation rather than an institution. For instance, the activities of these bodies and/or their influence on the company's practices are opaque, while their members are not protected against being recalled without good cause by the companies in question.

Such initiatives include the "Councils" established by TikTok for the EU,⁸⁴ the Asia-Pacific region⁸⁵ and the US,⁸⁶ with a membership constituted in each case by individuals including (former) politicians and/or civil society representatives. As far as can be established with any degree of certainty, these institutions do not communicate independently with the public, do not report on their activities and are not independent, neither legally nor by means of a voluntary commitment by the company.

The "Trust and Safety Council"⁸⁷ operated by Twitter has a different kind of membership, where the members are NGOs rather than individuals. But this "Council" also serves as little more than a company forum for obtaining non-binding opinions on the company's actions. Because of the absence of an independent organisational structure or public image, it is more akin to an informal cooperation.

Beyond that, there is a range of industry-wide cooperation mechanisms focusing on appropriate ways of dealing with content. The Global Network Initiative is a well-known initiative in this category. It has brought together companies, NGOs and research institutes to develop codes of conduct in support of corporate practices that respect human rights and to which companies which are members commit themselves.⁸⁸ Such initiatives are criticised for not going far enough and not holding companies to their voluntary commitments sufficiently,⁸⁹ but they could represent a promising launchpad to start building social media councils at the international level.

81 Calculation based on global statistics provided by the company on measures taken against content (posts, comments etc.) for "substantive" violations of its community standards, see "Community Standards Enforcement Report" for Q3 2020. "Substantive" means that content belonging to the categories Fake Accounts (1.3 billion items of content) and Spam (1.9 billion items of content) was not counted; if those categories were included, the figure would reach 36 million pieces of content per day. The following categories were included: "Adult Nudity & Sexual Activity: 36,700,00; Bullying & Harassment: 3,500,000, Child Nudity & Sexual Exploitation: 12,400,000, Dangerous Organizations: Organized Hate, 4,000,000, Terrorism: 9,700,000, Hate Speech 22,100,000, Regulated Goods: Drugs, 4,730,000, Regulated Goods: Firearms 1,050,000, Suicide and Self-Injury 1,300,000, Violent & Graphic Content 19,200,000 – totalling 116,700,000 pieces of content for the period July-September 2020. The data can be downloaded from <https://transparency.facebook.com/community-standards-enforcement>.

82 Gradoni, Lorenzo: Constitutional Review via Facebook's Oversight Board: How platform governance had its Marbury v Madison, VerBlog, 2021/2/10, <https://verfassungsblog.de/fob-marbury-v-madison>.

83 Douek (2021): The Oversight Board Moment You Should've Been Waiting For: Facebook Responds to the First Set of Decisions, Lawfare Blog, <https://www.lawfareblog.com/oversight-board-moment-you-shouldve-been-waiting-facebook-responds-first-set-decisions>.

84 "European Safety Advisory Council", See TikTok (2021): Meet TikTok's European Safety Advisory Council, <https://newsroom.tiktok.com/en-be/meet-tiktoks-european-safety-advisory-council>.

85 "TikTok Asia Pacific Safety Advisory Council", See TikTok (2020): Introducing the TikTok Asia Pacific Safety Advisory Council, <https://newsroom.tiktok.com/en-sg/tiktok-apac-safety-advisory-council>.

86 "TikTok Content Advisory Council", See TikTok (2020): Introducing the TikTok Content Advisory Council, <https://newsroom.tiktok.com/en-us/introducing-the-tiktok-content-advisory-council>.

87 Twitter (2021): <https://about.twitter.com/en/our-priorities/healthy-conversations/trust-and-safety-council>

88 See <https://globalnetworkinitiative.org/gni-principles>, members include Facebook and Google.

89 See Labowitz/Meyer (2016), for example: Why We're Leaving the Global Network Initiative, <https://bhr.stern.nyu.edu/blogs/why-were-leaving-the-gni>.

5.2. Planned “Social Media Council” in Ireland

Regarding projects which have already been launched, the most promising initiative apart from the FOB is the establishment of a national social media council with broad jurisdiction in Ireland. ARTICLE 19, a non-governmental organisation, has proposed the creation of an Irish “Social Media Council” and promoted the idea that future Irish legislation⁹⁰ should support this endeavour. The NGO also offered to manage the institution in the context of a pilot project. At this stage it is not yet clear if this will in fact happen; however, one of the organisers expressed cautious optimism at a presentation in November 2020.⁹¹

5.3. “Institutions of regulated self-regulation” under the German Network Enforcement Act (*Netzwerkdurchsetzungsgesetz, NetzDG*)

In Germany, the recognised institutions of regulated self-regulation provided for under Article 3.2.3b) of the Network Enforcement Act provide a legal framework for quasi-judiciary institutions which make decisions about removing certain types of content covered by the Act. Social networks which are covered by the Act because of their size can opt to join these institutions to submit content for legal review. Currently, only one such institution is active: the NetzDG review panel of the German Association for Voluntary Self-Regulation of Digital Media Service Providers (*Freiwillige Selbstkontrolle Multimedia-Diensteanbieter, FSM*).⁹²

The conceptualisation of these institutions falls short of the potential of social media councils. This is not so much owing to the way it has been implemented in practice by the FSM (the only example to date), but is a consequence of the NetzDG regulatory context. The NetzDG application framework limits the jurisdiction of institutions to the tightly circumscribed area of the (omitted) *deletion* of supposedly *unlawful* content. It does not address questions of content aggregation or restrictions on the visibility of content, nor does it cover the much broader field of removing content that only violates the terms of use of the platform providers.

Within this narrowly limited scope of the law, the NetzDG also fails to leverage these institutions to encourage platforms to implement systemic improvements of their internal procedures and decision-making structures. The institutions do not develop any form of “case law” from their regular reviews of individual cases and do not assess the process leading up to the submission of content to them. Within the NetzDG, any systemic improvements are to be triggered through a traditional approach of imposing fines for regulatory non-compliance (so-called “command and control regulation”).

Even in terms of the extremely limited scope of the NetzDG, this regulatory approach is already highly controversial from a constitutional law perspective. Broadening it to other areas, such as measures against “legal, but harmful” content within the (at least *de facto*) existing discretion of the platforms, would likely overstep these boundaries. This traces back to constitutional requirements relating to the independence of the process of opinion formation from the state and the very limited legislative power of the *Bund* (opposed to the *Länder*) in this context.⁹³ The institutions provided for under the Network Enforcement Act therefore do not offer a suitable framework for the development of social media councils.

90 Online Safety and Media Regulation Bill, regarding the current legislative process, see <https://www.gov.ie/en/publication/d8e4c-online-safety-and-media-regulation-bill>.

91 See the recording of a lecture by Pierre François Docquir (ARTICLE 19), <https://www.youtube.com/watch?v=zYYw7bBg88o&t=1530s>, from time stamp 25:30.

92 Freiwillige Selbstkontrolle Multimedia-Diensteanbieter: NetzDG-Prüfausschuss, <https://www.fsm.de/de/netzdg>.

93 The legislative competence of the federal government regarding the Network Enforcement Act in its current form is mostly rejected in the literature, see Liesching in: Spindler/Schmitz/Liesching, 2. ed. 2018, NetzDG § 1 margin no. 10-12 and also Hoven/Gersdorf in: BeckOK InfoMedienR/ 31. Ed. 1 May 2019, NetzDG § 1 margin no. 5-8 with numerous further citations; in support, see Schwartmann (2020), Stellungnahme im Rahmen des NetzDG-Änderungsgesetzes (BT-Drucksache 19/18792), available at <https://www.bundestag.de/resource/blob/700958/c2132ca5c-bf50c600d04a0df0058c1b8/schwartmann-data.pdf>; as well as, by the same author: (2017) Stellungnahme zum NetzDG-Entwurf (BT-Drucksache 18/12356), available at <https://www.bundestag.de/resource/blob/510886/002a8ce4b15005b96318abacee89199d/schwartmann-data.pdf>.

6. Summary and conclusion: the potential of social media councils

The concept of social media councils and especially the Facebook Oversight Board (FOB) is not without controversy. In response to the establishment of the FOB, for example, a “Real Facebook Oversight Board” was created, an initiative by critics of Facebook who call for government regulation while criticising the lacking effectiveness and slow establishment of the FOB, which – contrary to earlier announcements – only commenced work after the US presidential elections.⁹⁴

With specific reference to the FOB, its structure⁹⁵ is seen to be insufficiently independent while its mandate⁹⁶ is considered to be too limited, among other things because practices not related to deletions can hardly be monitored and assessed.⁹⁷ A further criticism is that the FOB does not have the structural capacity to improve the situation because it can only review a very small number of individual cases.⁹⁸

Some commentators also express a concern that quasi-judicial social media councils could erode the effectiveness of government legal protections,⁹⁹ impede the international standardisation of guarantees such as freedom of expression¹⁰⁰ or hamper the ability to use national courts in other ways.¹⁰¹ These concerns cannot be dismissed entirely if quasi-judicial social media councils were in fact to be established and if they made a great number of decisions. However, considering that the approach to designing such social media councils – presumably the only feasible way – is to focus on reviewing a very small number¹⁰² of “leading cases” to improve general systems, this does not appear to be a particularly significant risk.

A further concern regarding social media councils is that they effectively stabilise private orders without giving rise to real changes.¹⁰³ A counterargument is that (even minimal) iterative improvements through stronger social accountability for decisions, including those made by private companies, can only be a good thing and that – arguing from the point of view of the separation of powers – even a slightly broader distribution of the decision-making power of platforms should not be rejected as a matter of principle (“more could be done”). In this sense, social media councils could resemble institutional role models like press and broadcasting councils, which have existed in Germany for decades. As compromise solutions, they also face inevitable criticism but have never been replaced due to the absence of constitutionally, politically and/or practically viable alternatives.¹⁰⁴

A common perspective on social media councils rooted in legal theory is that such councils form part of the emergent phenomenon of private “constitutions”.¹⁰⁵ This viewpoint illustrates the considerable potential of the internal, functional differentiation of companies towards democratisation – or, at least, a partial implementation of rule of law-principles by the platforms and the creation of internal checks and balances.¹⁰⁶ In this way, social media councils can contribute to generating internal transparency and discussions about rules.¹⁰⁷

On the precision scales that measure the balance of power in the field of digital expression, the relationship between platforms and states is far from being in balance. This may be lamentable, but it offers the advantage that it makes even bold institutional experiments worthwhile. In this sense, social media councils cannot replace private or government regulation.

94 Butcher (2020): ‘The Real Facebook Oversight Board’ launches to counter Facebook’s ‘Oversight Board’, <https://techcrunch.com/2020/09/30/the-real-facebook-oversight-board-launches-to-counter-facebooks-oversight-board/>; <https://the-citizens.com/real-facebook-oversight>.

95 Morar (2019): Facebook’s Oversight Board: A toothless Supreme Court?, <https://www.internetgovernance.org/2019/10/02/facebook-oversight-board-a-judiciary-with-no-constitution>.

96 Reed (2019): Facebook’s Oversight Board needs a broader mandate that integrates human rights principles, <https://rankingdigitalrights.org/2019/05/22/facebook-oversight-board-needs-broader-mandate-that-integrates-human-rights-principles>.

97 Weinzierl (2019), Difficult Times Ahead for the Facebook “Supreme Court”, *VerfBlog*, 2019/9/21 <https://verfassungsblog.de/difficult-times-ahead-for-the-facebook-supreme-court>.

98 McNamee, Roger; Ressa, Maria: Facebook’s “Oversight Board” Is a Sham. The Answer to the Capitol Riot Is Regulating Social Media (2021), <https://time.com/5933989/facebook-oversight-regulating-social-media/>; with similar comment regarding its limited capacity see Ghosh, Dipayan; Hendrix, Justin: Facebook’s Oversight Board Just Announced Its First Cases, but it Already Needs an Overhaul, *VerfBlog*, 2020/12/19, <https://verfassungsblog.de/fob-first-cases>.

99 Jarren/Gostomzyk (2020): Facebooks Hausgericht, <https://www.medienpolitik.net/2020/04/facebook-hausgericht>.

100 Wagner, Haftung von Plattformen für Rechtsverletzungen (Teil 2), GRUR 2020, 329 (332).

101 Weinzierl (2019).

102 For every single FOB decision, there are millions of measures taken by Facebook (see above) that are not subject to review; considering the very low probability of the FOB’s accepting a given case for review, it is unlikely that national courts would regard this as an alternative remedy in civil cases, for example.

103 McSherry (2019) “Social Media Councils: A Better Way Forward, Window Dressing, or Global Speech Police?”, <https://www.eff.org/de/deeplinks/2019/05/social-media-councils-better-way-forward-lipstick-pig-or-global-speech-police>.

104 Most recently, for some useful insights on the criticisms levelled against the German Press Council in the context of the new State Media Treaty, see Klaus: “Staatlicher Zahnersatz für den Presserat: Der Medienstaatsvertrag macht die Selbstregulierung der Presse zum Auslaufmodell”, *VerfBlog*, 2021/3/29, <https://verfassungsblog.de/staatlicher-zahnersatz-fur-den-presserat>, DOI: 10.17176/20210329-195047-0; for general observations on the criticism of press councils, see Pöttker, in: Baum/Langenbucher/Pöttker/Schicha: “Handbuch Medienselbstkontrolle”, https://link.springer.com/chapter/10.1007/978-3-322-80808-0_12, pp. 125-131; similarly, regarding broadcasting councils, see Hahn: “Der Rundfunkrat – ein verzichtbares Kontrollinstrument?” in the same volume, https://link.springer.com/chapter/10.1007/978-3-322-80808-0_18, pp. 159-174.

105 As a representative example of many similar contributions, see douek: “Facebook’s ‘Oversight Board’: Move Fast with Stable Infrastructure and Humility”, *North Carolina Journal of Law & Technology* Volume 21, Issue 1 (2019), p. 4, which describes the social media council of Facebook as “one of the most ambitious constitutional projects of the modern era”; Pozen: “Authoritarian Constitutionalism in Facebookland” (2018), who argued (before the establishment of the Oversight Board) that Facebook was at best an authoritarian constitutional state, <https://balkin.blogspot.com/2018/10/authoritarian-constitutionalism-in.html>.

106 Maroni, Marta: “Some reflections on the announced Facebook Oversight Board”, available at <https://cmpf.eui.eu/some-reflections-on-the-announced-facebook-oversight-board>.

107 Jarren/Gostomzyk (2020): Facebook’s Hausgericht, <https://www.medienpolitik.net/2020/04/facebook-hausgericht>.

They can only partly relieve the regulatory pressure resulting from the manifold challenges of online communication; but they enable interesting models for the necessary redemocratisation of the normative orders of the hybrid¹⁰⁸ communication spaces of the public sphere in the present era.

For the further debate around social media councils, it is important to note that the attempt by Facebook remains only *one* version of such an institution. The “Oversight Board” should not be elevated to an archetype of a social media council, and the concept of social media councils should not be monopolised by Facebook.¹⁰⁹ Still, the Oversight Board provides a useful opportunity to discuss many of the core challenges and design decisions.

An open marketplace of ideas requires competition. Now, it is becoming increasingly clear that competition is also needed among different concepts to institutionally protect this market on the internet.

Further discussions on this topic could be based on the following considerations:

- Despite current uncertainty around their exact design, social media councils represent a good opportunity to increase the legitimacy of the normative orders of platforms, strengthen the protection of individual rights, and promote social cohesion.
- At the same time, it must be recognised that social media councils are no silver bullet. As independent mechanisms, they can complement existing models of private and state regulation of social networks and initiate improvements, but they cannot replace such models. In this context, the main emphasis is on improving company rules and the corresponding enforcement approaches.
- If not properly configured, social media councils may conceal actual power structures and fail to initiate real change. Therefore, such councils have to meet the highest standards of transparency regarding their own operations, while being equipped with appropriate rights to information and data access to ensure that diverse stakeholders can verify which systemic improvements they initiated, if any.

- In the absence of the national regulation of social media councils, there are already guidelines in the field of international human rights standards and industry agreements which would allow the configuration and decision-making practice of social media councils to be designed and measured.
- Politics, civil society, companies and multi-stakeholder groups provide a broad palette of configuration options for social media councils, depending on their underlying goals.
- A promising objective for social media councils is to initiate systemic improvements in the governance systems of companies beyond just individual cases. On the other hand, social media councils do not appear suited to contribute to remedy violations that have already occurred or to provide a type of legal hearing for as many affected parties as possible.
- The most important configuration dimensions of social media councils include their jurisdiction in various respects (industry-wide or platform-specific; national, regional or global; quasi-legislative or quasi-judicial design), their composition and the selection of members, their decision-making tools and the question of how to implement them, either through self-regulation or co-regulation.
- With specific reference to the idea of pursuing a national social media council for Germany, a promising approach would be to develop a model for social platform councils based on the extensive literature and constitutional jurisprudence on broadcasting councils. Such a basis in constitutional law could also contribute to creating the greater acceptance for social media councils needed in the absence of “hard” enforcement mechanisms.

The debate around the potential of social media councils to reimport democratic values into the private orders of public communication has only just begun. Private orders can (and should be) oriented towards public values via laws or legal judgments.¹¹⁰

However, in light of the increase of multiple public spheres (“many publics”)¹¹¹ and the growing recognition of the inner complexity or multi-faceted nature of the platforms themselves,¹¹² the time has come to reconceptualise the democratisation of platforms. In this context, social media councils could be the starting signal in the necessary race to establish a fairer configuration of the normative order of the digital world.¹¹³

¹⁰⁸ For an early example, see Ladeur, *Neue Institutionen für den Daten- und Persönlichkeitsschutz im Internet: “Cyber-Courts” für die Blogosphere*, DUD 2012, 711 (pp. 713): “It is conceivable that one could respond to the hybrid [authors’ note: private-public] nature of the new media (...) by outlining a model of a legal order that is also hybrid in nature. It could start with an attempt to encourage the self-organisation of social rules which form the infrastructure of traditional media law (...) through national law, especially through jurisprudence as a kind of ‘regulator’ of the new media, but without replacing such self-organisation.”

¹⁰⁹ In this context, we should also avoid using “oversight boards” as a generic term for social media councils, which is already happening in some instances. For example, see *Who Targets Me, Oversight boards for everything* (2021), <https://whotargets.me/en/oversight-boards-for-everything>.

¹¹⁰ Kettemann/Tiedeke, *Back up: can users sue platforms to reinstate deleted content?* *Internet Policy Review* 9 (2020) 2, <https://policyreview.info/articles/analysis/back-can-users-sue-platforms-reinstate-deleted-content>, DOI: 10.14763/2020.2.1484.

¹¹¹ Kettemann/Tiedeke, *Online order in the age of many publics*, *Kybernetes* 13 (2020), <https://www.emerald.com/insight/content/doi/10.1108/K-07-2020-0423/full/html?skipTracking=true>.

¹¹² Arun (2021): *Facebook’s Faces*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3805210.

¹¹³ Kettemann, *Deontology of the Digital: The Normative Order of the Internet*, in Kettemann (ed.), *Navigating Normative Orders. Interdisciplinary Perspectives* (Frankfurt/ New York: Campus, 2020), 76-91.

About the authors



PD Dr Matthias C. Kettemann, LL.M. (Harvard)

Matthias C. Kettemann is a *Privatdozent* in international law, internet law and legal theory at the Institute for Public Law at the Goethe University Frankfurt, research programme head at the Leibniz Institute for Media Research | Hans-Bredow-Institut (HBI) and heads research groups at the Humboldt Institute for Internet and Society, Berlin (Global Constitutionalism and the Internet), and the Sustainable Computing Lab at the Vienna University of Economics and Business (*Platform and Content Governance*). After studying law in Graz, Geneva and at the Harvard Law School, he was postdoctoral fellow at the University of Frankfurt. He was called to give evidence on digital rights to parliamentary committees of the German Bundestag, and has acted as reviewer for DAX companies and as an adviser to the German foreign and economics ministries, OSCE, the Fundamental Rights Agency and the Council of Europe in the fields of digital human rights, platform regulation and algorithm governance.

Martin Fertmann



Martin Fertmann is a research associate at the Leibniz Institute for Media Research | Hans-Bredow-Institut in Hamburg, a fellow at the Centre for Law in Digital Transformation at the University of Hamburg and a doctoral candidate at the Albrecht Mendelssohn Bartholdy Graduate School of Law at the same university. After completing his law studies in Hamburg and Beijing, majoring in media law, he worked as a research associate at various corporate law firms and participated in providing pro bono legal advice at the Cyber Law Clinic of the University of Hamburg for several years. He is currently working on his doctoral thesis, focussing on institutional and procedural requirements for social media councils under international human rights standards.

