ARTICLE 19

The Social Media Councils:

Consultation Paper

June 2019

# Table of Contents

# Executive Summary

This document outlines a roadmap for the creation of what ARTICLE 19 has called Social Media Councils (SMCs), a model for a multi-stakeholder accountability mechanism for content moderation on social media. SMCs aim to provide an open, transparent, accountable and participatory forum to address content moderation issues on social media platforms on the basis of international standards on human rights. The Social Media Council model puts forward a voluntary approach to the oversight of content moderation: participants (social media platforms and all stakeholders) sign up to a mechanism that does not create legal obligations. Its strength and efficiency rely on voluntary compliance by platforms, whose commitment, when signing up, will be to respect and execute the SMC's decisions (or recommendations) in good faith.

At the current stage of consultation with partners and stakeholders, there are different views on what SMCs should be. These different visions are not mutually exclusive, and are in fact essential to consider in pursuit of the common goal of addressing content moderation issues on the basis of international standards on freedom of expression and other fundamental rights.

With the present document, we present the different options and submit them to a public consultation. The key issues we seek to address through this consultation are:

1) Substantive standards: could SMCs apply international standards directly or should they apply a 'Code of Human Rights Principles for Content Moderation'?

2) Functions of SMCs: should SMCs have a purely advisory role or should they be able to review individual cases?

3) Global or national: should SMCs be created at the national level or should there be one global SMC?

4) Subject-matter jurisdiction: should SMCs deal with all content moderation decisions of social media companies, or should they have a more specialised area of focus, for example a particular type of content?

The consultation also seeks input on a number of technical issues that will be present in any configuration of the SMC, such as:

A. Constitution process

B. Structure

C. Geographic jurisdiction (for a national SMC)

D. Rules of procedure (if the SMC is an appeals mechanism)

E. Funding

An important dimension of the Social Media Council concept is that the proposed structure has no exact precedent: the issue of online content moderation presents a new and challenging area. Only with a certain degree of creativity can the complexity of the issues raised by the creation of this new mechanism be solved.

ARTICLE 19's objective is to ensure that decisions on these core questions and the solutions to practical problems sought by this initiative are compatible with the requirements of international human rights standards, and are shaped by a diverse range of expertise and perspectives.

The consultation survey can be accessed at: https://lime.article19.org/index.php/991112?lang=en

# Introduction

## The context

In today's world, dominant tech companies hold a considerable degree of control over what their users see or hear on a daily basis.[1] Current practices of content moderation on social media offer very little in terms of transparency and virtually no remedy to individual users.[2] The impact that content moderation and distribution (in other words, the composition of users' feeds and the accessibility and visibility of content on social media) has on the public sphere is not yet fully understood, but legitimate concerns have been expressed, especially in relation to platforms that operate at such a level of market dominance that they can exert decisive influence on public debates.

This raises questions in relation to international laws on freedom of expression and has become a major issue for democratic societies. There are legitimate motives of concern that motivate various efforts to address this issue, particularly regarding the capacity of giant social media platforms to influence the public sphere. However, as with many modern communication technologies, the benefits that individuals and societies derive from the existence of these platforms should not be ignored. The responsibilities of the largest social media companies are currently being debated in legislative, policy and academic circles across the globe, but many of the numerous initiatives that are put forward do not sufficiently account for the protection of freedom of expression.

## Regulatory approaches

### Solo-regulation

> Although they are sometimes referred to under the all-encompassing umbrella of 'self-regulation', situations where a private company unilaterally controls content on its own platform according to its own internal rules will be described in this document as '**solo-regulation**'[3] or '**regulating speech by contract**'.

There is a strong consensus among international experts on freedom of expression that 'solo regulation' on social media platforms has failed to provide sufficient transparency and protection for freedom of expression and other human rights.[4]

In November 2018, Facebook declared their intention to create an 'oversight board'. A draft Charter was made public in January 2019 and a series of consultations were ongoing at the time of writing.[5]

### Self-regulation

> '**Self-regulation**' is a mechanism of voluntary compliance at sector or industry level: legislation plays no role in enforcing the relevant standards. Its *raison d'être* is holding members of self-regulatory bodies accountable to the public, promoting knowledge within its membership and developing and respecting ethical standards.

---

[1] Moore, Tambini, Digital Dominance, OUP, 2018.

[2] ARTICLE 19, Side-stepping rights: Regulating speech by contract, June 2018, https://www.article19.org/resources/side-stepping-rights-regulating-speech-by-contract/

[3] The term has been proposed by Marko Milosavljevic and Sally Brougthon Micova in their article 'Banning, Blocking and Boosting: Twitter's solo-regulation of expression', Medijske Studije / Media Studies, 2016 7 (13), 43-58.

[4] ARTICLE 19 has recently analyzed in detail the flaws of 'regulation by contract' through which social media companies sovereignly regulate speech on their own platforms: Side-stepping rights: Regulating speech by contract, op. cit.

[5] ARTICLE 19 has contributed to some of the regional consultations. We have also summarized our views on the draft charter at https://www.article19.org/resources/facebook-oversight-board-recommendations-for-human-rights-focused-oversight/.

Those who commit to self-regulation generally do so for positive reasons such as the desire to further the development and credibility of their sector, although other motivations may also play a part in encouraging actors to get on board a self-regulatory mechanism – such as the desire to avoid statutory regulation. Self-regulation models rely first and foremost on members' common understanding of the values and ethics that underpin their professional conduct – usually in dedicated "codes of conduct" or ethical codes. Meanwhile, members seek to ensure that these voluntary codes correspond to their own internal practices.

Press councils are the typical example of self-regulatory mechanisms:[6] in the light of international standards on freedom of expression, self-regulation for print media has been considered to be the least restrictive means available through which the press can be effectively regulated and the best system through which high standards in the media can be promoted.[7] Other examples of self-regulatory mechanisms for content regulation include the video games industry[8] or the advertising industry.[9]

Importantly, the existence of an effective self-regulation mechanism can reduce pressure on the courts and the judiciary. As a significant number of cases can be dealt with in a quick and satisfactory manner, and at low cost, legal proceedings need only be initiated in the most severe of cases. Generally, when a problem is effectively managed through self-regulation, the need for state regulation is eliminated.[10]

## Co-regulation

'**Regulated self-regulation**' or '**co-regulation**' refer to a regulatory regime involving private regulation (be it self-regulation or solo-regulation) that is actively encouraged or even supported by the State.

Co-regulation systems can include the recognition of self-regulatory bodies by public authorities, as could for instance be the case under the German law NetzDG.[11] Public authorities generally also have the power to sanction any failure by self-regulatory bodies to perform the functions for which they were established.

---

[6] ARTICLE 19, Freedom and Accountability: Safeguarding Freedom of Expression Through Self-Regulation, 2005; L. Fielden, Regulating the Press, A Comparative Study of International Press Councils, April 2012, Reuters Institute; A. Hulin, Statutory media self-regulation: beneficial or detrimental for media freedom?, European University Institute, Robert Schuman Centre for Advanced Studies, Centre for Media Pluralism and Media Freedom, EUI Working Paper RSCAS 2014/127; A. Hulin, L'autorégulation des médias : glaive ou bouclier pour la liberté ?, Revue des Médias, n°30-31 Printemps - été 2014, http://la-rem.eu/2014/10/18/lautoregulation-des-medias-glaive-ou-bouclier-pour-la-liberte

[7] Although press councils are one of the sources of inspiration for the SMC model, ARTICLE 19 does in no way suggest that (a) it should be the only reference for SMCs, or that (b) existing press councils should extend their jurisdiction or mandate to social media platforms.

[8] PEGI (Pan European Game Information) is a not for profit company whose mission is to operate a rating mechanism for the video games industry, see https://pegi.info/

[9] In the UK, the Advertising Standards Authority (ASA) is a self-regulatory system which applies a code of standards adopted by the advertising and media industries, see https://www.asa.org.uk/

[10] ARTICLE 19 has identified several requirements for effective self-regulation in the print media sector: ARTICLE 19, Freedom and Accountability Safeguarding Free Expression Through Media Self-Regulation, March 2005, available at: http://bit.ly/2EtIvBj.

[11] ARTICLE 19 observes that the NetzDG provides for the recognition, by the Ministry of Justice, of 'regulated self-regulatory agencies'. The role of such bodies, which would be financed by social media companies, would be to determine whether a given message is in violation of the law and should be removed. Granting of recognition by the Ministry depends upon conditions such as the independence of the self-regulatory agency, the expertise of the people making decisions, and on their capacity to reach a decision within 7 days. We fear that the guarantees provided for in the law are not sufficient to ensure neither the independence and effectiveness of self-regulation nor the protection of freedom of expression (see our legal analysis of the law: Germany: Act to Improve Enforcement of the Law on Social Networks undermines free expression, Sept. 2017, https://www.article19.org/resources/germany-act-to-improve-enforcement-of-the-law-on-social-networks-undermines-free-expression/.)

ARTICLE 19 considers that a limited degree of support from public authorities can be useful in supporting the emergence and operation of self-regulatory mechanisms, provided that the public intervention is limited to creating a legal underpinning for self-regulation and does not threaten the independence of the self-regulatory bodies.[12]

ARTICLE 19 has analyzed the flaws of 'co-regulatory' approaches such as the EU Code of Conduct on Hate Speech[13], the EU Code of Conduct on Disinformation[14], or the revision of the EU AVMS Directive.[15] Generally, our concerns are that such mechanisms (a) rely on rules that are not compliant with international standards on freedom of expression and have not been elaborated through a transparent and participatory mechanism, (b) put companies (rather than courts) in the position of making decisions on the legality of content restrictions, and (c) lack transparency and do not offer individual users an effective remedy.

*Legislation*

ARTICLE 19 observes that many of the recent **legislative initiatives** related to the Internet and social media companies tend to give disproportionate censorship powers to the State, whether through prison terms, fines or content blocking powers, chilling free expression, or to outsource regulation to private companies with no proper integration of international standards.

Content laws are often overly broad. Where a regulatory body is charged with applying the laws, the regulator often lacks independence and the law does not always provide for a right of appeal or judicial review of the regulator's decisions. Such systems are inconsistent with international standards on freedom of expression, which mandate that approaches to regulation developed for other means of communication – such as telephony or broadcasting – cannot simply be transferred to the Internet but, rather, need to be specifically designed for it.[16] We also note that sanctions powers including the blocking of entire platforms or hefty fines for failure to comply with domestic legal requirements would, in and of themselves, constitute a disproportionate restriction on freedom of expression.[17]

We also note that legislation is not always easily adaptable to the fast-paced evolution of social media landscapes.

## The proposal

The need for a mechanism capable of ensuring effective accountability of platforms for content moderation is increasingly recognized on all sides.[18] To fulfil this need, ARTICLE 19 has proposed the creation of the Social Media Council – a model for a voluntary accountability mechanism that would provide an open, transparent, accountable and participatory forum to address content moderation issues

---

[12] See ARTICLE 19's response to the recognition of IMPRESS, 23 October 2016, available at: http://bit.ly/2EMIEmF.

[13] ARTICLE 19, EU: European Commission's Code of Conduct for Countering Illegal Hate Speech Online and the Framework Decision, Aug 2016, https://www.article19.org/resources/eu-european-commissions-code-of-conduct-for-countering-illegal-hate-speech-online-and-the-framework-decision/.

[14] ARTICLE 19, EU: New Code of Practice on Disinformation fails to provide clear commitments, or protect fundamental rights, 9 Oct. 2018, https://www.article19.org/es/resources/eu-new-code-of-practice-on-disinformation-fails-to-provide-clear-commitments-or-protect-fundamental-rights/

[15] ARTICLE 19, New EU legislation must not throttle online flows of information and ideas, Sept. 2017, https://www.article19.org/resources/new-eu-legislation-must-not-throttle-online-flows-of-information-and-ideas/;

[16] See The 2011 Joint Declaration on Freedom of Expression.

[17] See, e.g., European Court, *Yildirim v Turkey,* App. no. 3111/10, 18 Dec. 2012.

[18] GDPi, ARTICLE 19, David Kaye, 'Social Media Councils: From Concept to Reality', conference report, Feb 2019, pp. 10-11. The report is available at https://cddrl.fsi.stanford.edu/global-digital-policy-incubator//content/social-media-councils-concept-reality-conference-report.

on social media platforms, on the basis of international standards on freedom of expression and other human rights.[19]

The Social Media Councils would be created by all relevant stakeholders, with the goal of applying a set of human rights-based principles to the review of content moderation decisions made by social media platforms.[20] The Social Media Councils would have the power to impose non-pecuniary remedies such as a right of reply, the publication of an apology (if, for instance, some content was removed by mistake), the publication of a decision, or the re-upload of suppressed content. The mechanism would not be in itself legally binding, and the participating social media companies would commit to executing the Council's decisions in good faith.

The idea was endorsed by UN Special Rapporteur on Freedom of Expression David Kaye, who recommended that

> 'all segments of the ICT sector 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.'[21]

In order to test and further refine the model, the Global Digital Policy Incubator at Stanford University, the Special Rapporteur, and ARTICLE 19 convened a 2-day seminar on 1-2 February 2019. The event brought together academics, civil society organisations, and dominant platforms (Twitter, Facebook, YouTube). The discussions enabled a deeper analysis of the complexity of issues raised by the project of creating a Social Media Council, and various options and new ideas were also identified through the exchanges with all participants. The meeting also showed that there is a strong consensus that the regulation of online content should be compatible with international standards on freedom of expression and other fundamental rights.[22]

It should be kept in mind that the proposal of Social Media Councils only seeks to address issues of content moderation on social media platforms. The power and influence of these tech companies in today's digital environment raise various interlinked concerns in relation to the collection of personal data, competition in online advertising markets and the impact of market dominance on users' human rights, the funding of quality journalism in the digital age, or the adequate taxation of in-country profit. These are legitimate concerns that democratic societies need to address, but these are different debates. The voluntary approach we advocate for in the SMC project is only intended to focus on the accessibility, visibility and findability of content on social media platforms.

## The SMCs and national laws

It has been suggested by some that SMCs should review governmental requests that social media platforms remove content: this external evaluation would supposedly encourage public authorities to show more restraint. ARTICLE 19 acknowledges that social media platforms can find themselves under intense pressure from state authorities.[23] However, as a voluntary mechanism created by private actors, it is not the role of the SMC to review governments' requests to moderate content: such injunctions should only be adopted by an independent judicial body. International law provides a specific venue for the adoption and review of public authorities' requests that social media platforms delete

---

[19] ARTICLE 19's initial policy brief on the SMC can be found at https://www.article19.org/resources/self-regulation-hate-speech-social-media-platforms/ See also Global Partners Digital, A Rights-Respecting Model of Online Content Regulation by Platforms, May 2018, at https://www.gp-digital.org/publication/a-rights-respecting-model-of-online-content-regulation-by-platforms/

[20] See A19, Side-Stepping Rights, pages 9-11, for a presentation of human rights responsibilities of the private sector.

[21] Available at https://daccess-ods.un.org/TMP/9810305.23777008.html

[22] Social Media Councils: From Concept to Reality, op. cit., at available at https://cddrl.fsi.stanford.edu/global-digital-policy-incubator//content/social-media-councils-concept-reality-conference-report.

[23] See, by comparison, our views about the telecom and ISP sector: ARTICLE 19, Getting Connected, op. cit.

content: there should be a clear legal basis, a proportionate application of the law, and the decision of an independent tribunal.

While the SMC will look at content moderation through the lens of international standards, national laws are not always aligned with international law. Should the SMC find that a social media platform has not complied with international standards in a specific case, the platform could then choose to either adapt their practices as required by the SMC, or, depending on national circumstances and the margins they have in dealing with public authorities, explain that national laws won't allow them to follow the recommendations of the SMC. Either way, in the context of a voluntary mechanism, there would be no legal risk for social media platforms, which would not be legally bound by the decisions of the SMC.

A separate question is that of content which is protected under international standards but can still be considered harmful ("lawful but harmful"), for example certain forms of hate speech. As is generally the case with the application of international standards, the SMC mechanism will necessarily include a margin of appreciation that will serve to accommodate the differences between platforms and their liberty to promote their own community standards, or between national contexts.

## A participatory approach

Terms such as 'self-regulation' and 'multi-stakeholder', which we have used to describe the SMC project, have become ambiguous for being used in a broad variety of ways and in diverse policy, academic or industry circles (which to some degree all function as bubbles). In various countries or regions, these terms may be received either positively or negatively for a series of reasons (e.g., the experience of having seen 'self-regulatory' bodies easily captured and controlled by governments). Here, we present what we view as the appropriate participatory methodology for the creation and operation of a voluntary mechanism.

A recent publication co-authored by GDPi[24] proposed that an authentic multi-stakeholder process needs to be driven by the stakeholders, who are actual decision makers rather than being confined to an expert or advisory role, and must be open to the participation of a broad range of stakeholders. Such a process also needs to be transparent and to provide public access to its decisions and documents in order to foster trust and accountability. Multi-stakeholder processes rely on consensus to achieve their outcomes, which creates win-win situations for a broad spectrum of stakeholders. The authors add that: '(m)ulti-stakeholder approaches have repeatedly proven themselves to be exceptionally well-suited to rapidly changing technologies and business practices and to the global environment in which the internet exists.'[25]

The report of the Stanford conference notes that

> '[a] multistakeholder approach to online content moderation has substantial advantages for all sides. For platforms, an independent multistakeholder body can help to provide legitimacy to their internal processes and demonstrate a commitment to free expression, which is valuable for their public profile. It can also serve as a resource, providing outside perspectives from experts to help navigate particularly complex problems. For governments, these bodies promote the democratic principle of transparency and can help to ameliorate societal concerns about content online. They can also take pressure off of courts by creating an accountable body that can moderate many content decisions. For users, the councils will help them better understand the content

---

[24] E. Donahoe, F. O. Hampson, Governance Innovation for a Connected World - Protecting Free Expression, Diversity and Civic Engagement in the Global Digital Ecosystem, 2018.
[25] L. E. Strickling, J. F. Hill, Multi-stakeholder Governance Innovations to Protect Free Expression, Diversity and Civility Online, in E. Donahoe, F.O. Hampson, op. cit., p 47.

*moderation process and create more transparency about what steps are being taken to protect free speech while addressing issues of abuse.'[26]*

On the basis of ARTICLE 19's analysis of the conditions for effective self-regulation in the print media sector,[27] we consider that the adoption of a self-regulatory/multi-stakeholder approach for the development of SMCs means that the SMCs should be:

1. Independent from government, commercial and special interests;
2. Established via a fully consultative and inclusive process – major constitutive elements of their work should be elaborated in an open, transparent and participatory manner that allows for broad public consultation;
3. Democratic and transparent in their selection of members and decision-making;
4. Include broad representation: it is important that the composition of the self-regulatory body includes representatives of the diversity of society;
5. Have a robust complaint mechanism and clear procedural rules to determine if standards[28] were breached in individual cases, and have the power to impose only moral sanctions;
6. Work in the service of the public interest, be transparent and accountable to the public.

We argue that under international standards on human rights, such an approach – a voluntary, participatory, accountable and transparent mechanism – corresponds to the least restrictive measure that is effective for protecting freedom of expression online while achieving the protection of other rights and legitimate interests related to content moderation on social media platforms.

We observe that numerous ongoing or forthcoming initiatives are seeking to establish self-regulatory mechanisms at the national level within a legal framework of 'co-regulation', under the guise of bringing a swift end to the dissemination of often vaguely defined harmful content. There are differences between various iterations of co-regulation (sometimes also called 'regulated self-regulation'), according to whether or not the overseeing public authority is properly kept at arms-length from regulating the substance of issues and restrains itself to assessing the processes put in place by either the social media companies themselves or a self-regulatory institution. These initiatives clearly indicate that co-regulation is an approach that is currently deemed optimal by the industry and public authorities when it comes to dealing with content moderation on social media platforms. We suggest that the Social Media Council offers a model that can deliver a form of co-regulation that fully ensures the protection of the fundamental right to freedom of expression. Moreover, the SMC model offers a stop-gap between the state body charged with overseeing the self-regulatory mechanism and the social media companies, without which companies are likely to apply mechanisms and execute decisions that do not comply with international human rights standards.

In the future development of the project, ARTICLE 19 plans to support the implementation of a pilot Social Media Council in one country, which will be identified on the basis of the following criteria:

- The existence of a culture of self-regulation and/or multi-stakeholder approaches, and of corresponding institutions;
- The state of development of the media landscape, including the online media and the regulatory institutions, and their openness to engage in dialogue with other stakeholders;
- The presence of social media platforms, including their business presence in the country and their openness to engage in dialogue with stakeholders;

---

[26] Social Media Councils: From Concept to Reality, op. cit., p. 8.
[27] See, e.g. ARTICLE 19, Freedom and Accountability Safeguarding Free Expression Through Media Self-Regulation, March 2005, available at: http://bit.ly/2EtIvBj.
[28] In the context of press councils, they are ethical standards (professional standards of journalism); in the context of the SMC, as discussed below, the standards will be based on international human rights law.

- The existence of an active civil society, including digital rights organisations and other civil society organisations (CSOs), which have a focus on content regulation, and their capacity to engage in dialogue towards the creation of a SMC;
- The willingness of public authorities to accept and encourage the initiative of the creation of a SMC, without the risk of them trying to seize control of this new institution;
- The state of freedom of expression and media freedom and pluralism, as well as broader human rights protections in the country.

## The consultation

At this stage of discussing our proposal with CSOs, academics, social media companies and other stakeholders, it appears that there are different visions for the creation of an open, transparent, accountable and participatory forum that applies international standards to content moderation on social media. For instance, some support the idea of a global SMC with an advisory role while others advocate for national SMCs with the power to make decisions on cases submitted by individuals. We consider that these different conceptions, which all pursue the common goal of bringing content moderation in line with international standards, are not mutually exclusive: they can be designed to be complementary. In a similar perspective, SMCs could interact with Facebook's oversight board (pending an assessment of what this board will actually amount to when it comes into existence).

Accordingly, in this document, we present and discuss the core questions around which options may diverge:

1) Substantive standards: could SMCs apply international standards directly or should they apply a Code of Human Rights Principles for Content Moderation?

2) Functions of SMCs: should SMCs have an advisory role or should they be able to review individual cases?

3) Global or national: should SMCs be created at the national level or should there be one global SMC?

4) Subject-matter jurisdiction: should SMCs deal with all content moderation decisions of social media companies, or should they have a more specialised area of focus?

Choices on each of these core questions will lead to different possible configurations for the SMC. In any case, the creation of a SMC will be a complex process that will require the resolution of numerous technical problems. In this document, we also present and discuss proposals to address technical problems such as:

A. Constitution Process

B. Structures

C. Geographic Jurisdiction (for a national SMC)

D. Rules of Procedure (if the SMC is a appeals mechanism)

E. Funding

An important dimension of the Social Media Council project is that the proposed structure has no exact precedent: only with a certain degree of creativity can the complexity of the issues raised by the creation of this new mechanism be solved.

ARTICLE 19's objective is to ensure that choices on core questions and solutions to practical problems all are compatible with the requirements of international standards.

**The consultation is open until 30 September 2019. Please fill in the consultation survey here: https://lime.article19.org/index.php/991112?lang=en**

**You can send also your contributions by email to pierre@article19.org, mentioning your name, job title and employer, or the category of stakeholder you belong to (social media platforms, media, journalism, other social communicators, academics, civil society, human rights experts, media regulatory bodies, advertising industry, government). If you would like to partake in follow-up interviews, calls or meetings, please mention this in your email.**

# I. Substantive Standards

The Social Media Council will need a body of rules to inform and guide its review of content moderation decisions and practices. ARTICLE 19 considers that these rules should reflect international standards on human rights.[29] The reference to international human rights law will ensure the coherence and consistency of decisions made by various SMCs. To be clear, as we are talking of a voluntary mechanism, the application of international standards by SMCs will not be legally binding: the decision of the SMC will only deploy its "authority" in the framework of a private, voluntary commitment.

As is generally the case with the application of international standards, a certain margin of appreciation[30] would be part of the SMC. A degree of variation is inherent to international human rights law (and the same can be said of constitutional law): fundamental rights are general principles, they are standards, and when they are translated into actual detailed rules through (judicial) application, there is unavoidably a certain margin of manoeuvre that comes into play. In the case of the SMC, this margin of appreciation in the application of international standards will allow the application of standards in a specific country context (as is generally the case with international standards), and the differentiation between different companies and their respective products (e.g., Facebook is different from Twitter), including the liberty of a company to adopt stricter restrictions on freedom of expression to accommodate their own editorial choices (although it should be clear that market dominance would result in a narrower margin of appreciation in this respect).

## Option A: Direct Application of International Standards

The Social Media Council could directly apply international standards. The UN Special Rapporteur on freedom of expression David Kaye considered in his 2018 report that

> '(H)uman rights standards, if implemented transparently and consistently with meaningful user and civil society input, provide a framework for holding both States and companies accountable to users across national borders. (…)
>
> The Guiding Principles and their accompanying body of "soft law" provide guidance on how companies should prevent or mitigate government demands for excessive content removals. But they also establish principles of due diligence, transparency, accountability and remediation that limit platform interference with human rights through product and policy development. Companies committed to implementing human rights standards throughout their operations — and not merely when it aligns with their interests — will stand on firmer ground when they seek to hold States accountable to the same standards. Furthermore, when companies align their terms of service more closely with human rights law, States will find it harder to exploit them to censor content.'[31]

Principle 12 of the UN Guiding Principles on Business and Human Rights provides that:[32]

---

[29] We note that there is agreement on that point among the experts and stakeholders that have contributed to preparatory discussions of the SMC model. See GDPi, Social Media Councils, op. cit.

[30] Under international law, states have a 'margin of appreciation' which is a limited capacity to adapt international standards to their specific contexts.

[31] Report of the Special Rapporteur, A/HRC/38/35, 6 Apr. 2018, p. 14.

[32] See also, for further analysis of HR obligations for telecom companies, ARTICLE 19, Getting Connected: FoE, Telcos and ISPs, June 2017, at https://www.article19.org/resources/getting-connected-new-policy-on-freedom-of-expression-telcos-and-isps/, and the Report of the UN Special Rapporteur on FoE on telecommunications and Internet access sector, A/HRC/35/22, 30 March 2017.

> "*The responsibility of business enterprises to respect human rights refers to internationally recognized human rights – understood, at a minimum, as those expressed in the International Bill of Human Rights.*"

The authoritative interpretation by international and regional courts and special mechanisms will provide all necessary guidance to inform the decisions of the SMC. If confronted with an issue that has not yet been decided by legal authorities, the SMC would have to elaborate its own interpretation of international human rights law for its own purposes.

## Option B: Adoption of a Code of Human Rights Principles for Content Moderation

The SMC could apply a Code of Human Rights Principles for Content Moderation. The specific adaptation of international standards to online content moderation through the adoption of a Code would ensure that the SMC operates within a narrower margin of manoeuvre and under stricter guidance than the broad reference to international standards.[33]

This option would require a formal adoption process. An international conference convened by the UN and regional Special Rapporteurs on freedom of expression, in collaboration with GDPi and ARTICLE 19, could serve to discuss and adopt the global Code of Human Rights Principles for Online Content Moderation.[34] This conference should be open to the participation of all stakeholders (provided they respect international human rights standards) and ensure representation from all regions of the world, including the broad diversity of societies, and marginalised groups.

The Code would be organised in a series of chapters that deal with specific areas of content moderation in the online environment, such as hate speech, incitement to terrorism and violence, misinformation, the promotion of the broadest possible diversity in distribution of news and other media content, the protection of minors, defamation and privacy, harassment, a human rights impact assessment of algorithms, and the identification of safeguards that ensure that AI is used in accordance with human rights.

---

**Question 1: Substantive Standards**

**[Please fill in the survey at https://lime.article19.org/index.php/991112?lang=en]**

1.1 (A) Should the SMC directly apply international standards on freedom of expression and human rights **OR** (B) should the SMC apply a Code of Human Rights Principles for Content Moderation?

1.2 If you chose (A) above: what should the scope and limits of the margin of appreciation be? How should they be defined?

1.3 If you chose (B) above: how should the Code be drafted and adopted?

1.4 Additional comments on this topic.

---

[33] The Code of HR Principles for Content Moderation should be a universal set of rules that can be applied by national or regional SMCs (if they are set at a local level).
[34] A formal treaty would not make sense in the context of a self-regulatory/multi-stakeholder mechanism.

# II. Function of the SMCs

The SMC could act as an advisory panel providing guidance to companies, or it could be responsible for reviewing individual cases.

## Option A: An Advisory Body

The Social Media Council could have an advisory role: it would provide general guidance to social media companies on the compatibility of ToS or Community Standards with international standards on human rights. In this configuration, the SMC would be an open forum where stakeholders elaborate recommendations or observations. It would not engage into the discussion of individual cases.

To ensure effectiveness, the advisory SMC should have the broad mandate of 'providing guidance to social media companies on how to ensure that content moderation complies with international standards.' It should also be able to undertake the elaboration of general recommendations on any matter it deems necessary on the suggestion of any of its members.

## Option B: An Appeals Mechanism

The Social Media Council could have adjudicatory powers:[35] it will review decisions made by social media platforms on individual cases. Individual users who are directly affected by a content moderation decision will be able to send a complaint to the Social Media Council, which will then decide whether in the particular circumstances of the case the decision made by the social media platform conformed to the requirements of international human rights standards. This would allow the SMC to serve the public directly, which would give the mechanism more strength and impact, but the mechanism would have to deal with the problem of scale (the potential number of cases).

Such a mechanism should be accessible to all. There should also be clear and precise rules of procedure on questions such as admissibility conditions, the time limits, the admissibility of evidence, elements that would be covered by confidentiality, the exchange of arguments and views, elements of publicity or the adoption and publication of decisions (this is further discussed in section 5).

## Option C: An Advisory and Appeals Mechanism

The Social Media Council could combine both the roles of an advisory body and of an appeals mechanism.

---

**Question 2: Functions**

2.1 (A) Should the SMC be an advisory body, and why? **OR** (B) Should the SMC have the power to review individual cases, and why? **OR** (C) Should the SMC should have both advisory and adjudicatory functions, and why?

2.2 Additional comments on this topic.

---

[35] (aka, the power to make a formal judgement on a disputed matter (Oxford dictionary), that is, in other words, the mandate to deal with individual complaints from users).

# III. Global or national?

Should the SMC be an international body, or should SMCs operate on the national or regional level?

## Option A: The Global SMC

Installing the SMC at the global level would ensure a uniform application of international human rights standards to content moderation. Social media platforms operate on a global scale and they would likely more easily partake in an international initiative, instead of having to multiply engagement efforts at national level. While this option might give the impression of creating distance between the SMC and the national or regional context where content moderation issues arise, this gap could be bridged through the composition of the SMC, third-party interventions or other forms of participation in advisory processes, or by inviting additional expertise on an ad hoc basis. It has also been suggested that an international body would not be subjected to capture by the state, although one could also note that existing at the international level won't automatically preserve the SMC from undue influence by public or private interests.

## Option B: The National SMC

Creating the SMC at national level ensures that the new mechanism will be inhabited by decision makers who are well informed of the complexities of the national context in its cultural, political, and social dimensions. In the hypothesis where SMCs are appeals mechanisms, this option would contribute to dealing with the issue of scale, as fewer cases would fall under the jurisdiction of a national SMC than under the broad umbrella of a global institution. A form of network gathering all national SMCs could serve to avoid divergence in decision-making between these SMCs. The presence of international experts on HR in the SMC Board would help nationally based SMCs integrate international standards in their decisions.

## Option C: The Regional SMC

In certain countries, national civil society actors may lack the capacity to fully engage into the creation and operation of such a mechanism; in oppressive regimes, civil society organisations often find themselves under direct control or close monitoring of public authorities. Setting the SMC at the regional level would allow civil society to join forces. The regional level could also alleviate the risk of capture at national level. CSO voices from the Global South have expressed an interest for the creation of national or regional SMCs, as these could have a positive impact on civil society actors and contribute to enlarging the margins of freedom within which civil society operates. Even in national contexts where freedom of expression faces increased restrictions, the actors to be represented in a multi-stakeholder SMC are often those that are at the forefront of defending freedom of expression: as such, giving them greater voice regarding the online sphere may provide more protection of online rights.

---

**Question 3 – Global or National**

3.1 (A) Should one SMC be created at the global level, and why? **OR** (B) should SMCs be created at the national level, and why? **OR** (C) should SMCs be created at the regional level, and why?

3.2 How could the complementarity between national SMCs and a global institution be organised?

3.3. Additional comments on this topic.

---

# IV. Subject-Matter Jurisdiction

In ARTICLE 19's original vision, the Social Media Councils would be responsible for all issues related to content moderation and distribution on social media platforms, including all types of content (user-generated content, media content, etc.). As this represents a wide scope of activities, it may make sense to start with a narrower area of responsibility.

## Option A: General Jurisdiction

The Social Media Councils' mandate will cover all issues areas related to content moderation and distribution on social media platforms, including all types of content (user-generated content, media content, etc.). This could include areas such as for instance the identification and moderation of incitement to violence or hatred, the protection of privacy and reputation, the visibility, accessibility and promotion of accurate and reliable information, exposure to a broad diversity of information and ideas, or the use of automated decision-making processes and artificial intelligence in content moderation and content distribution.

## Option B: A Specialised SMC

In consideration of complexity and scale, the SMC should focus on a narrower area of content moderation. It could for instance focus on the dissemination of disinformation, or any other content issue. The specialised SMC could also be seen as a starting point for the development of an SMC with a broader remit.

---

**Question 4 – Subject-Matter Jurisdiction**

4.1 (A) Should SMCs be responsible for all content moderation decisions of platforms? **OR** (B) Should SMCs have a narrower field of competence, at least in the beginning? If so, what should it be?

4.2 Other comments on this topic.

---

# V. Technical Questions

## A: The Constitution Process

The SMC will be created by a collection of relevant stakeholders, that is, all parties interested in contributing to bringing content moderation practises in line with international standards on freedom of expression and other human rights. A fundamental dimension of the creation of the SMC is taking an important question of democratic interest – the moderation of online speech – to bring it back (where it belongs) into an open, transparent forum: in order to fulfil that role, the SMC needs to be composed by representatives of the broad diversity of society.

ARTICLE 19 has suggested that the SMC would be created conjointly by:

- The social media companies. It may be an option to start with the dominant players first (as they have a considerable impact upon public debates), but the mechanism should remain easily accessible to smaller or emerging players (to avoid reinforcing the competitive advantage of dominant actors);

- Media, as they have a vital business interest in the distribution of news on social media platforms, the professional expertise in the production of information, and the experience of regulating content through self-regulatory mechanisms. As a balanced media landscape comprises public service media, private media, and local and community media, the participation of media in the creation of the SMC should reflect these 3 categories of media actors;

- Journalists, as they have the professional expertise in the production of news, the knowledge of professional ethics in that area, and a vital interest in the work conditions of the profession;

- Other social communicators such as bloggers or non-professional contributors to public debates;

- Media regulators such as press councils and broadcast regulatory authorities;

- The advertising industry, as advertising is a key element of the business model of social media platforms;

- Academics and researchers with expertise in international human rights law, media law and regulation, intermediary liability, data protection, consumer protection, competition law and the online circulation of media content or content moderation;

- CSOs with legal and technical expertise in digital policy and freedom of expression (as the integration of international standards into the core of the SMC is a prerequisite to its effectiveness and success);

- CSOs with experience and expertise in representing the viewpoints and sensibilities of the various components of society, especially the CSOs that represent minorities and vulnerable groups in society. While the SMC should represent a broad diversity of the living forces of society, these should only be groups that recognise international standards on human rights.

Governments might be associated with the creation of the SMC as observers, with no decision-making power. This participation might contribute to alleviating concerns among policymakers.

These stakeholders would elaborate a Charter of the Social Media Council during a 'co-creation work-shop' or a 'chartering conference'. The signature of the Charter will create the new voluntary mechanism. The Charter would describe:

- The mandate of the SMC;

- The composition, powers and roles of its organs;

- Detailed rules of procedure for all decision-making processes and all aspects of the creation, composition and operation of the SMC (e.g., how to appoint members in internal organs);

- The substantive standards that the SMC will apply;

- The functions the SMC will fulfil;

- The commitment of participants (i.e. that SMC decisions will be executed in good faith).

This may require the creation (and registration) of a not-for-profit legal entity, as per national law.

---

**Question 5. Constitution Process**

5.1 On the basis of your answers to previous questions, who are the stakeholders that should take part in the creation and operation of the SMC?

5.2 Who in your view are the key stakeholders that should absolutely be part of the mechanism in order to ensure the feasibility and credibility of the project?

5.3 Should the SMC only concern dominant players or should it also be open to smaller, emerging players? Would specific modalities be necessary for smaller players?

5.4 Other comments on this topic.

---

## B. Structure of the SMC

The SMC will probably take the legal form of a not-for-profit association (as per the requirements of relevant national law). The constituting parties will have to agree on the respective weight of each category of stakeholders in the decision-making body of the SMC, and on the corresponding appointment process, which will also have to ensure gender parity and a proper representation of the whole diversity of society, including minorities and vulnerable groups.

The Board is the main decision-making body of the SMC. It adopts decisions on advisory/adjudicatory processes, votes budget, hire staff, adopts annual reports, and other such tasks.

In situations where the SMC is organised as a legal entity with a general assembly, the GA would elect the Board on the basis of a list of names put forward by the stakeholders. Another option would be for the Board to appoint new members after applications have gone through a public consultation. The first Board would be appointed by the signatories of the Charter.

Candidates would have to meet a series of expectations clearly set out in the Charter. While being generally competent on the basis of relevant expertise and experience, and well informed of the state of online and media landscapes and corresponding legal and regulatory framework, Board members should also be constrained by rules of incompatibility that preserve their independence and protects the SMC against undue interference from public or private interests.

The Board would be renewed every two years in part (to ensure continuity), with members staying for a maximum of three mandates (six years). The Charter should also make clear that Board members can only be dismissed for a limited number of reasons set out in the Charter, and through a decision of the Board itself made by a special majority vote.

The Board and staff of the SMC should have to abide by a Code of Conduct that would detail relevant ethical principles, such as reasonable rules on conflicts of interest, remuneration, etc. The Code of Conduct should be public.

It only makes sense to discuss the number of representatives for each category in the light of the specific circumstances of a given context (depending on whether the SMC is created at national, regional or global level), but generally the composition of the Board should achieve a balanced representation of the various categories of participant stakeholders. In certain countries, for instance, the main religious groups or certain linguistic or ethnic minorities might need to be represented, while it would not necessarily make sense in a different national context. In certain cases, bloggers are organised in an association but it's not necessarily the case everywhere – or in certain cases, the unions of journalists would be the relevant interlocutors, but in other contexts professional associations would be in a better position to play that role. A Board could – *and this is just an example* – include:

- 3 representatives of social media companies (that have significative presence in the country in the hypothesis of a national SMC);

- 3 representatives of media, including 1 representative of public service media, 1 representative of private media and 1 representative of community media;

- 3 representatives of journalists;

- 1 representative from bloggers and other social communicators;

- 2 representatives from freedom of expression and digital rights CSOs;

- 2 representatives from the academic sector;

- 4 representatives from various philosophical, religious, political trends (as makes sense in the specific context of the country);

- 4 representatives from minority and vulnerable groups;

- 1 representative from the advertising industry.

It is expected that the Board would hire a secretary general and a small staff to ensure daily operation and prepare its meetings. While participation in the Board meetings would not be remunerated, members could expect expenses to be reimbursed on the basis of clear rules in the Charter. While the General Secretary and staff would be full-time employees, the Board would only meet twice a month, for the duration of a day.

---

**Question 6: Structure of the SMC**

6.1 Based on your preference for a national or global SMC, how should the Board be composed? That is, which weight should each category of stakeholders have? What is the ideal size for the Board? How should the representatives of each category be appointed?

6.2 What would the main ethical principles that would ensure the independence and effectiveness of the Board's members?

6.3. Other comments on this topic.

---

## C. Geographical Jurisdiction

The creation of national or regional SMCs would require robust criteria to determine the scope of cases they would be allowed to consider (geographical jurisdiction). Variations between various national SMCs could also cause a certain degree of fragmentation that would lead social media platforms to engage in additional geofencing.

ARTICLE 19 proposes that the national SMC holds jurisdiction over content that has a substantial connection to the country, and only if the jurisdiction of the SMC can be considered the most appropriate to consider an individual complaint. To that end, consideration should be given to a combination of the following elements:

- The extent to which the content actually spread in the country and whether an actual damage has been suffered in the State;

- Whether the content at stake has been published by a user that is established in, and mainly targeting the audience of, the country;

- Whether the social media platform that has been used to disseminate the content has significant presence in the country (as could be deduced from either the SMP having an office in country, or by looking at the importance of usage, users base, etc.);

- Whether the author and victim are established in the country;

- Whether the content was uploaded in the country;

- Whether the content is in an official language of the country.

In the longer term, a network of national SMCs would coordinate to resolve cross-jurisdictional issues and harmonise the understanding of international standards. Inspiration will be drawn from cooperative practices of national human rights institutions, press councils or independent broadcast regulatory authorities.

A certain degree of localisation of international standards is inherent to the application of international law on the national scenes and does not threaten the universality and coherence of international law. For instance, privacy laws in France and Germany are different, but both are compatible with international standards.

As to the impact on social media platforms of SMC decisions that would require geofencing (to adapt to a variety of national interpretations of international standards), the platforms would be under no legal obligation to engage with that path, as the voluntary mechanism creates no such obligation.

---

**Question 7: Geographical Jurisdiction**

7.1 What is the most effective manner to delineate the geographical jurisdiction of national SMCs? Do you agree with the suggested criteria?

7.2 Other comments on this topic.

---

## D. Rules of procedure for the appeals mechanism

In consideration of the volume of content on social media platforms, it is clear that algorithms will necessarily be a part of first-line content moderation practices.[36] It is also generally accepted that users should first seek a solution to their complaints with the social media platform before escalating the case to the Social Media Council: in other words, platform-level remedies must be exhausted before a case can be referred to the Social Media Council (a time-limit might be added here to avoid the situation where platform-level remedies take too long).

Reasonable admissibility conditions might be instituted, such as:

- a requirement that the plaintiffs identify themselves (although there should be consideration for granting anonymity);

- a requirement that the plaintiffs send sufficient evidence of the facts on which their complaint is based;

- a requirement that the plaintiffs clearly explain "what the problem is" (although it is important to note that a requirement that the complaint includes an explicit reference to international standards or the Code of Principles might be an obstacle for a number of users, especially from marginalised groups).

The SMC will need to be able to filter and choose amongst the cases that will be referred to it by platforms and individual users. It is absolutely necessary that the SMC remains the master of its docket in order to avoid being overloaded with cases, which would lead to paralysis of the system. The criteria and process for selecting cases will have to be specified, but it is expected that the SMC would select emblematic/hard cases.

The SMC will need to have access to the complete 'case file' of the social media platform (that is, the documented succession of facts that preceded the referral of the case to the SMC). There may be important privacy interests at stake (including for third parties not present in the procedure) in relation of the transmission of such information by a tech platform, which need to be balanced through clear rules.

The cost and technicality of the process could become an obstacle for members of minorities or vulnerable groups. It has been suggested, in preparatory discussions, that solutions inspired from public defenders and/or collective actions might be helpful to remedy the issue of universal accessibility.

Some have suggested that panels (that is, sub-sections of the SMC) could decide on individual cases, but the existence of a variety of panels within a SMC might lead to fragmentation, with each panel developing its own 'case law'. Others have suggested that panels might be in charge of preparatory analysis of complaints, preparing the case file for the SMC to examine and decide – another suggestion is to install an 'attorney general' at the SMC, who would be in charge of selecting and submitting cases to the SMC.

An interim situation might need to be organised for the duration of proceedings before the SMC. It has been suggested that the default interim decision should be to not delete or demote online content, unless the seriousness of likely harm is such that it could not be compensated or redressed at a later stage.

---

[36] See ARTICLE 19, Governance with teeth: How human rights can strengthen FAT and ethics initiatives on artificial intelligence, Apr. 2019, at https://www.article19.org/resources/governance-with-teeth-how-human-rights-can-strengthen-fat-and-ethics-initiative-on-artificial-intelligence/

Some have suggested that CSOs, human rights organisations and other stakeholders could have the option to submit third-party interventions to the SMC, which could also serve as a manner of bringing additional diversity of expertise and viewpoints to a global SMC.

The decision of the SMC could only impose non-pecuniary remedies such as excuses, a right of reply, the publication of the decision in a relevant visible online space of the social media platform, or the re-upload of content. Since this is a voluntary system, the commitment of social media platforms would be to execute the decision in good faith, which indeed would leave them some margin of manoeuvre as to what compliance with the decision means in practice. It could also be part of the commitment of a social media company to explain how and why they have executed a decision of the SMC in a specific way.

---

**Question 8: Appeals Mechanism**

8.1 What objectives should rules of procedure pursue? How to achieve these objectives?

8.2 Other comments on this topic.

---

## E. Funding

The SMC will need sufficient and sustainable funding to ensure its operation. Relying on a sole source of income – be it social media platforms, states or private donors – could endanger its independence or its credibility.

There are various possible sources of income to finance the operation of a Social Media Council, from governmental subsidies to contributions from social media companies or grants from independent foundations or other donors. The creation of a buffer institution, possibly a trust or foundation set at the international level, could serve to pool all available resources and allocate resources to SMCs. The Charter of the SMC should also include guarantees that no funding would be accepted that could undermine its independence or credibility.

---

**Question 9: Funding**

9.1 What is in your view the most effective way to ensure the financial sustainability of SMCs without jeopardising their independence?

9.2 Other comments

---

**Question 10: Additional comments.**

Other suggestions, links to online material you would like to share.

---