

ARTICLE 19

Social Media Councils

One piece in the puzzle of content moderation
12 October 2021

ARTICLE 19**T:** +44 20 7324 2500**F:** +44 20 7490 0566**E:** info@article19.org**W:** www.article19.org**Tw:** @article19org**Fb:** facebook.com/article19org**© ARTICLE 19, 2021**

This work is provided under the Creative Commons Attribution-Non-Commercial-ShareAlike 3.0 licence.

You are free to copy, distribute and display this work and to make derivative works, provided you:

- 1) give credit to ARTICLE 19
- 2) do not use this work for commercial purposes
- 3) distribute any works derived from this publication under a licence identical to this one.

To access the full legal text of this licence, please visit:
<https://creativecommons.org/licenses/by-sa/3.0/legalcode>

ARTICLE 19 would appreciate receiving a copy of any materials in which information from this report is used. ARTICLE 19 bears the sole responsibility for the content of the document.

Contents

Executive summary	4
One piece in the puzzle of content moderation	5
1. The making of an idea	7
Endorsement from the UN Special Rapporteur on freedom of expression	9
Global versus national	9
Wider views on the SMC	10
2. The first steps towards an Irish SMC	13
Why Ireland?	13
A multi-stakeholder model	15
A governance structure for the Irish SMC	17
Jurisdiction in Ireland	18
Geographical jurisdiction	18
Subject-matter jurisdiction	19
A mechanism for an individual remedy	20
International human rights law and content moderation	22
The SMC and the legal and regulatory framework in Ireland	23
Giving voice to diverse perspectives on complex social issues	23
Individual complaints mechanism	25
Funding	26
Next steps: Creating a working group to develop the Irish SMC	27
Annex: Draft terms of reference for the working group	28
Endnotes	31

Executive summary

Current content moderation practices allow social media companies to wield significant power over people's right to freedom of expression. Despite this, social media platforms are still not accountable for the way they moderate content. Therefore, one of the initiatives of ARTICLE 19 has been to put forward a model for a multi-stakeholder mechanism for the oversight of content moderation on social media, known as the Social Media Council (SMC).

Since our [initial publication](#) on the SMC in 2018, we have consulted a wide range of actors about this mechanism and have taken the first steps towards setting up a pilot SMC in Ireland.

This report brings that information and experience into the public domain.

The key objectives of the SMC are to:

- **Review** individual content moderation decisions made by social media platforms on the basis of international standards on freedom of expression and other fundamental rights.
- **Provide** general guidance on content moderation practices to ensure they follow international standards on freedom of expression and other fundamental rights.
- **Act** as a forum where stakeholders can discuss recommendations.
- **Use** a voluntary-compliance approach to the oversight of content moderation that does not create legal obligations.

Discussions with Irish stakeholders have indicated a real interest for creating a pilot SMC in Ireland. This report summarises these discussions and recommends the creation of a working group to carry the project forward.

One piece in the puzzle of content moderation

The equation of accountability of social media platforms for content moderation has not been resolved yet,¹ and significant changes are taking place that will shape the future legal and regulatory framework.

As of October 2021, the only existing external self-regulatory mechanism for the oversight of content moderation, the Facebook Oversight Board, has rendered 15 decisions. While Facebook is bound by the Oversight Board's decisions on whether a particular piece of content should be allowed on the platform, general recommendations that the Oversight Board makes are not binding. The long-term effectiveness of this model remains to be assessed.² No other social media giant has committed to a similar initiative.³

Recent legislative developments are guided by an ambition to make 'Big Tech' accountable. At the same time, attention is shifting towards a more systemic approach to regulating online content, where rather than ensuring that every single case is handled properly, the regulator must now instead verify that the platform has adopted 'appropriate measures' to moderate problematic content. Among planned legal initiatives that are likely to be globally influential, the UK Draft Online Safety Bill poses a [severe risk of becoming a chokehold for freedom of expression](#). The draft [EU Digital Services Act](#), while retaining the general conditional immunity from liability for hosting providers and a prohibition on general monitoring, raises a number of concerns about the protection of freedom of expression online.⁴

Along with monitoring legal developments for the regulation of social media platforms, ARTICLE 19 has been developing a model for a multi-stakeholder voluntary-compliance mechanism, known as a Social Media Council (SMC). The SMC provides a transparent and independent forum to address content moderation issues on social media platforms on the basis of international human rights standards.

The key objectives of the SMC are to:

- **Review individual content moderation decisions made by social media platforms** on the basis of international standards on freedom of expression and other fundamental

rights. The right of appeal gives the SMC more credibility in the eyes of the public and gives individual users an opportunity to be heard on matters that directly impact them.

- **Provide general guidance on content moderation** guided by international standards on freedom of expression and other fundamental rights. While there is a growing consensus on the relevance of international human rights law for content moderation, this is still an emerging field with many open questions.
- **Act as a forum** where all stakeholders can discuss and adopt recommendations (or the interpretation thereof). This participatory methodology promotes collective adoption and interpretation of guidelines and can help embed international standards into practices of content moderation.
- **Use a voluntary-compliance approach to the oversight of content moderation** where social media platforms and all stakeholders sign up to a model that does not create legal obligations and where they voluntarily implement the SMC's decisions and recommendations. The SMC will be a self-regulatory mechanism where representatives of the various stakeholders come together to regulate the practices of the sector.

ARTICLE 19 believes that the SMC will play a positive and effective role in defending and promoting the right to freedom of expression in the development of accountability for content moderation. The idea has been endorsed by both the previous and the current UN Special Rapporteurs on the promotion and protection of the right to freedom of opinion and expression.

The making of an idea

ARTICLE 19 first put forward the idea of the SMC in a March 2018 publication titled [Self-regulation and 'hate speech' on social media platforms](#). This report aimed at initiating discussions on creating a new self-regulatory model for social media platforms and proposed that:

- An SMC could be created at a national level to ensure sufficient proximity and understanding of the relevant community and context, or at an international level, or a combination of both.
- The SMC would need to be independent from any particular social media company and should include representatives from all relevant stakeholders, such as media associations, media regulatory bodies, freedom of expression experts, academia, and civil society.
- Social media platforms would have to commit to providing an appropriate level of information on their internal content moderation practices to the SMC. While the SMC's decisions would not be legally binding, platforms would also commit publicly to accepting and implementing its decisions.
- The SMC could adopt a Charter of Ethics for social media through transparent and open consultations with all relevant stakeholders. At a minimum, the Charter would include a commitment to comply with international human rights standards on freedom of expression and due process, respecting an individual's legal rights.
- The SMC would receive complaints from individual users only after all possibilities of remedying the issue with the social media company (either through ombudspersons or other flagging procedures) have been exhausted. The SMC's decision could include the possibility of a sanction that promotes rather than restricts speech (such as a right of reply, an apology, or the publication of its decision).
- The SMC would be able to discuss general guidelines for social media platforms.

- The SMC would have to receive a stable and appropriate level of funding to ensure its independence and capacity to operate. Social media platforms would have to commit to providing at least part of this funding on a multi-annual basis. As briefly discussed in Chapter 2 below, additional resources could be provided by public subsidies, other stakeholders, or philanthropic organisations.
- The SMC would be accountable to the public because it would serve the general interest (through such means as an annual report, transparency on governance and funding, and publication of decisions on the website).

[Based on the experience of press councils](#),⁵ ARTICLE 19 suggested that the SMC should:

1. Be independent from government, commercial, and special interests.
2. Be established via a fully consultative and inclusive process – major constitutive elements of their work should be discussed in an open, transparent, and participatory manner that allows for broad public consultation.
3. Be democratic and transparent in their selection of members and decision-making.
4. Include broad representation – it is important that the self-regulatory body includes representatives that reflect the diversity of society (including the representation of minorities and groups in situations of vulnerability or marginalisation).
5. Have a robust complaint mechanism and clear procedural rules to determine if applicable standards were breached in individual cases.
6. Have the power to impose only non-financial sanctions.
7. Work in the public interest and be transparent and accountable to the public.

Comparable ideas for the oversight of content moderation on social media platforms were put forward around the same time.⁶

Endorsement from the UN Special Rapporteur on freedom of expression

A month later, in April 2018, the SMC concept was endorsed by then UN Special Rapporteur on freedom of expression David Kaye in a report that proposed “a framework for the moderation of user-generated online content that puts human rights at the very centre.”⁷

In his report, the Special Rapporteur defended the idea that “Companies should incorporate directly into their terms of service and ‘community standards’ relevant principles of human rights law that ensure content-related actions will be guided by the same standards of legality, necessity and legitimacy that bind State regulation of expression.”⁸ He observed that:

“...companies could work with one another and civil society to explore scalable solutions such as company-specific or industry-wide ombudsman programmes. Among the best ideas for such programmes is an independent ‘social media council’, modelled on the press councils that enable industry-wide complaint mechanisms and the promotion of remedies for violations. This mechanism could hear complaints from individual users that meet certain criteria and gather public feedback on recurrent content moderation problems such as over censorship related to a particular subject area. States should be supportive of scalable appeal mechanisms that operate consistently with human rights standards.”⁹

Global versus national

In February 2019, the Global Digital Policy Incubator at Stanford University, the UN Special Rapporteur, and ARTICLE 19 convened a two-day seminar to further explore the SMC idea with academics, civil society organisations, and Twitter, Facebook, and YouTube. The [report from the seminar](#) also acknowledged the complexity of the challenges inherent to the constitution of an SMC.

While the meeting showed a strong consensus that the regulation of online content should be compatible with international standards on freedom of expression and other

fundamental rights, there was no unanimity on whether efforts should be directed towards the creation of a *global* SMC or if the new model should be considered at a *national level*.

While the Global Digital Policy Incubator advocated for a global SMC, ARTICLE 19 suggested that the SMC should be created at a national level (unless there was a risk that it would be easily captured by the government or other powerful interests). Creating an SMC at a national level would ensure the involvement of local decision-makers who are well-informed of the local context and understand its cultural, linguistic, historical, political, and social nuances.

Since the event, ARTICLE 19 has continued to submit the concept of an SMC to discussions with academics, international experts, representatives from the social media industry, and civil society organisations at various international events such as [World Press Freedom Day](#) (2019) or panels at the [Tunis edition of Rights Con](#) (2019), and an online public consultation.

While the SMC has generally received interest, opinions around it have remained divergent. Considering that the voluntary oversight of content moderation on social media platforms is an entirely new field and inevitably a complex undertaking, it comes as no surprise that there would be a broad range of views on how the SMC should be created or what its precise remit or geographical scope should be.

Wider views on the SMC

As of October 2021, academics and policy experts have integrated the concept of an SMC into policy debates on platform regulation:¹⁰

- The [Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression](#) noted that “Social media councils offer independent, external oversight from public, peer, or multi-stakeholder sources. A high-level, strictly independent body to make consequential policy recommendations or to review selected appeals from moderation decisions could improve the level of trust between platforms, governments, and the public.” As such, they could “enable greater collaboration and information-sharing by companies, facilitating early detection of new behaviors by bad

actors across multiple platforms; provide a ready-made forum to discuss responses in crisis situations such as the Christchurch massacre, or advanced crisis planning; set codes of conduct or establish baseline standards for content moderation that safeguard freedom of expression; set standards and procedures for independent, vetted researchers to access databases; and make companies more accountable for their actions under their terms of service.”

- A [scholarly article](#) observed that Facebook’s Oversight Board and SMCs have “the potential to serve as a useful mechanism to help platforms navigate the complex terrain of translating international human rights standards to the platform moderation context – whether through reviewing the compatibility of emblematic individual cases with international human rights law and/or providing general guidance on the compliance of platform processes and procedures with international human rights standards.”
- The [Canadian Commission on Democratic Expression](#) recommended the creation of an SMC which would serve as “an institutional forum where industry, civil society, citizens and other interested parties can study and debate emerging issues with a mind to providing policy recommendations to the regulator. The SMC will serve as a policy arm and be empowered to convene parties and advise the regulator.”
- In her April 2021 [report on disinformation](#), the UN Special Rapporteur on freedom of expression, Irene Khan, recommended that companies should explore the creation of external oversight models such as an SMC.
- On 7 July 2021, [ARTICLE 19 submitted observations](#) to the [Oireachtas Committee on Tourism, Culture, Arts, Sport and Media](#). The meeting was part of the Committee’s pre-legislative scrutiny of the Online Safety and Media Regulation Bill.
- In the course of the legislative work on the future EU Digital Services Act, [Members of the European Parliament](#) proposed to include a new provision in the Digital Services Act to create a “European Social Media Council”, which would serve as “an independent advisory group” with the mission of “issuing non-binding guiding principles and recommendations to improve content moderation processes, fostering a participative

and transparent public debate around content moderation processes; and issuing policy and enforcement recommendations to the Commission.”

While opinions obviously differ on the exact purpose of an SMC, there is broad agreement that, as a multi-stakeholder transparent model, it could:

“(…) put the societal back into social media. They could establish fair, reliable, transparent, and non-arbitrary standards for content moderation. At a time when decisions by social media companies increasingly structure our speech, councils could offer a comparatively swift method to coordinate and address pressing problems of democratic accountability. Creating a democratic, equitable and accountable system of platform governance will take time. Councils can be part of the solution.”¹¹

Table 1: Summary of the advantages offered by the SMC

Challenges with current practices of content moderation	Advantages of an SMC
Antagonism between stakeholders	Acts as a forum for cooperation and co-learning
No external oversight of content moderation decisions	External oversight based on international human rights law
No remedy for individual users	Individual users have access to a complaints mechanism
Opacity	Support towards more transparency
Content moderation decisions are taken unilaterally	The whole diversity of society takes part in the oversight of content moderation decisions

The first steps towards an Irish SMC

Since March 2020, ARTICLE 19 has approached civil society organisations, social media platforms, academics, journalists, media regulators, and others to explore setting up a pilot SMC in Ireland. We proposed that this pilot SMC would:

- Be a voluntary-compliance body created by stakeholders
- Operate an external mechanism to review individual content moderation decisions
- Provide content moderation guidelines and recommendations
- Work on the basis of international and European human rights law
- Focus only on Irish content moderation cases.

We held online meetings with a broad range of stakeholders in Ireland between March 2020 and July 2021.¹² These discussions were placed under a virtual Chatham House Rule.¹³ In these exploratory conversations, we did not ask participants for commitment. It became apparent from the discussions that an SMC would have potential in Ireland, and the conversations allowed us to explore what an Irish SMC could be like in practice.

We knew from the start that the venture would be complex:¹⁴ such a model can only come to exist if it is collectively built and owned by interested stakeholders. The creation of an SMC raises difficult questions, and solutions can only be designed through dialogue with all actors, and if adapted to the specificities of the national context. We also made it clear that we would only contribute to the development of the mechanism: ARTICLE 19 has no intention to control the Irish SMC once it comes into existence.

Why Ireland?

In the course of our work on the SMC, it had been suggested that a target country for a pilot SMC should be identified based on the following criteria:

- The existence of a culture of self-regulation and/or multi-stakeholderism (and existence of corresponding institutions).
- The state of the media landscape, including online media and regulatory institutions, and their openness to engage in dialogue with other stakeholders.
- The state of social media platforms, including their market share, establishment in the country, and their openness to engage in dialogue with stakeholders.
- The state of civil society organisations, including digital rights organisations that have a focus on content regulation and their capacity to engage in dialogue towards the creation of an SMC.
- The willingness of political authorities to accept and encourage the creation of an SMC without trying to seize control of it.
- The state of freedom of expression and media freedom in the country (democratic or transition).

Based on desk research and discussions with Irish stakeholders, it became clear that a pilot SMC was likely to thrive in Ireland because:

- The Irish legal and regulatory culture is familiar with concepts and practices of self-regulation and co-regulation.
- There are ongoing debates on the regulation of social media platforms that combine the transposition of the EU's [Audiovisual Media Services Directive](#) (AVMSD) and the broader scope of the [Online Safety and Media Regulation Bill](#) (OSMR).
- The European headquarters of the main social media platforms are based in Ireland (which means that the future Irish legal and regulatory framework will have an impact beyond the borders of the country), and these companies have an influence on various aspects of daily life in the country.

- There is a dynamic media environment, including innovative start-up media companies.
- There is also a vibrant, highly educated, and cohesive civil society that is familiar with the role of social media in public debates.
- There is a robust culture of collective deliberation (as exemplified by the Citizens' Assembly which was established in 2016).



(illustrations: Fuschia McAree for ARTICLE 19)

A multi-stakeholder model

Initially, we suggested that the SMC should be made up of representatives from:

- **Social media companies**, starting with the dominant or biggest players first (as they have a considerable impact on public debates). However, the model should remain open and easily accessible to smaller or emerging players to avoid reinforcing the competitive advantage of dominant actors.
- **Media and journalists**, as they have an 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.

- **Media regulators**, for example the Press Council and the broadcast regulatory authority.
- **Advertising industry**, including the self-regulatory mechanism for the industry.
- **Academics and researchers** with expertise in international human rights law, media law and regulation, intermediary liability, data protection, journalism and media, consumer protection, competition law, and content moderation.
- **Civil society organisations** with legal and technical expertise in digital policy, media, and digital literacy, and in freedom of expression and other human rights.
- **Civil society organisations** with experience and expertise in representing the viewpoints and sensibilities of the various components of society, especially minorities and groups in situations of vulnerability or marginalisation. These stakeholders should only be organisations that recognise international standards on human rights.

This list of stakeholders has been generally accepted as relevant and meaningful during our discussions with our Irish interlocutors; however, a minority of participants did not think that the advertising industry should be involved in the creation of the SMC. One participant observed that the scope of the SMC would be beyond the remit of the Press Council, which might not be willing to engage with the new mechanism. Another participant noted that the relationship between social media platforms and media companies is currently quite tense due to the debates on social media companies contributing financially to the production of media content. While it has been part of the discussions on the future sustainability of media companies for a while, this topic has recently come to light again when [Australia adopted a law that forces technology platforms to remunerate publishers for the right to use their content](#), either through the conclusion of voluntary agreements between the platforms and media companies or through a binding arbitration mechanism. In Ireland, comparable ideas can be debated by the [Future of Media Commission](#). This tension could have a negative impact on the capacity of media and social media companies to work together to create an SMC.

A governance structure for the Irish SMC

The Irish SMC will be a non-profit organisation and established as a company limited by guarantee (CLG) pursuant to [Part 18 of the Companies Act 2014](#). Many charitable and professional bodies find the company limited by guarantee suitable because it allows them to secure the benefits of a separate legal identity and of limited liability.

The Irish SMC will be constituted by an initial group of interested stakeholders; membership would remain open to newcomers. Based on our initial conversations, we suggested that there should be five categories of members:

1. Social media companies
2. Media and advertising industries
3. Journalists
4. Academics
5. Civil society organisations

We suggested that the board of directors of the Irish SMC (who would hold responsibility for the management and the missions of the SMC, i.e., complaints mechanism and elaboration of guidelines), should be comprised of 13 individuals. Two of these directors should be nominated by social media companies, one by the media and advertising industries, one by journalists, one by academics, and a further two by civil society organisations.

The remaining six should be independent directors who are appointed by a committee composed of the seven directors who are directly nominated by the members. These six independent directors will be appointed based on a call for applications. The candidates should be suitably qualified persons and representative of the broad diversity of civil society in Ireland. While the profile for the role of directors, including clear rules on conflict of interest, would need to be agreed, the overall composition of the board should be guided by principles of diversity, notably gender parity, and inclusiveness.

The chair will be chosen by the board (of 13 directors) either from the six independent directors or from the three directors nominated by academics or civil society organisations. In addition, certain organisations could delegate an observer. Observers can take part in discussions relating to the complaints mechanism and the elaboration of general guidelines, but they do not have the right to vote. It is suggested that the Broadcasting Authority of Ireland (or the new Media Commission¹⁵), the Press Council of Ireland, and the Advertising Standards Authority for Ireland could send an observer to participate in the Irish SMC.

This governance structure was generally accepted as making sense. One participant noted that a smaller number of 9 or 11 members would be a more effective group of decision-makers and suggested a corresponding reduction in the number of independent directors. Another person considered that the advertising industry and the media should be two different categories, and each appoint one director; they also preferred the public authorities to be full members of the SMC rather than observers because they thought this would reinforce the multi-stakeholder nature of the body. One participant felt that there should be no government involvement in the SMC.

Questions about the staff that would support the work of the SMC have remained open. A number of participants suggested that the Irish SMC should begin with a start-up mentality and a growth-oriented spirit: a small agile team of highly motivated people could achieve a lot in terms of bringing the new model to an operational and future development stage.

Jurisdiction in Ireland

Geographical jurisdiction

ARTICLE 19 suggests that the Irish SMC should have jurisdiction over content moderation disputes that have a substantial connection to Ireland. To decide whether a particular case has a substantial connection to Ireland, consideration should be given to:

- The extent to which the content at stake spread across the country, and whether actual damage has been suffered in the country.

- Whether the content at stake has been published by a user who is established in, and is mainly targeting the audience of, the country.
- Whether the author and/or 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.

Under the country of origin principle, the future regulatory authority in Ireland might act as a de facto regulator for the whole of the EU because the European headquarters for technology platforms are in Ireland, but we suggest that the Irish SMC should focus only on Irish content moderation cases. The objective is to [enable local voices to be represented](#) in debates on content moderation that directly impact them. It would also contribute to limiting the number of potential complaints addressed to the SMC. A minority of participants in our discussions expressed a preference for a regional European SMC, which was advocated for by some [Members of the European Parliament](#) in an amendment to the Digital Services Act.

A minority of participants also expressed concerns about a conflict between a national level accountability model and platforms operating at a global level based on global content rules. The fact that SMCs apply the universal rules of international standards on human rights contributes to mitigating that tension.¹⁶

Subject-matter jurisdiction

In the long term, if not in the short term, the SMC's mandate could cover all 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 the identification and moderation of incitement to violence or hatred; spread of disinformation; protection of privacy and reputation; 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.

One participant noted that the SMC's remit could specifically include disinformation – an area currently not covered by the OSMR Bill – as well as the accessibility of data for scientific research purposes.

A mechanism for an individual remedy

While some envision an SMC with a consultative role of elaborating policy recommendations to the regulator or to social media platforms, many of our participants believe that there is value in the SMC providing an external individual remedy – especially at a time where legislative developments seem to focus on systemic accountability rather than individual cases (indeed, individual cases could be a way to uncover systemic issues).

The key themes for the creation of such individual remedies are:

- Individual users who are directly affected by a content moderation decision should be able to send a complaint to the SMC, which then decides whether, in the circumstances of the case, the decision made by the social media platform conformed to the requirements of international human rights standards. By extension, this allows the SMC to serve the public directly, which gives the model more strength and impact.
- It is generally accepted that users should first seek a solution to their complaints with the social media platform before escalating the case to the council. In other words, platform-level remedies must be exhausted before a case can be referred to the SMC – a time limit might be added to avoid the situation where platform-level remedies take too long.
- Reasonable admissibility conditions may be instituted, such as a requirement that the plaintiffs identify themselves (the possibility of admitting anonymous complaints should be considered); a requirement that the plaintiffs send sufficient evidence of the facts on which their complaint is based; and a requirement that the plaintiffs clearly explain what the problem is (although a requirement that the complaint must include an explicit reference to international standards might be an obstacle for a number of users). The

cost and technicality of the process could become an obstacle for members of minorities or vulnerable or marginalised groups of society. In global consultations held by ARTICLE 19, participants have suggested that legal mechanisms such as public defenders or collective actions could be helpful.

- There is inevitably a tension between an individual right of appeal and the risk for the SMC to be drowned in a huge number of complaints. The limited geographical scope of the Irish SMC would contribute to mitigating this risk to some degree, but it is generally admitted that the SMC should be able to select the cases it would review, and that these should be emblematic cases that raise hard questions and/or that can serve as a precedent for several similar complaints. One participant suggested that certain civil society organisations could play the role of ‘trusted complainants’ to support the SMC in selecting cases for review.
- An interim situation might be needed 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.
- The decision of the SMC could only impose non-financial remedies such as an apology, a right of reply, the publication of the decision in a relevant visible online space of the social media platform, or the removal or the re-upload of content.
- Since this is a voluntary system, the social media companies would need to commit to executing the decision in good faith, which indeed would leave companies some margin of manoeuvre as to what compliance with the decision means in practice. Companies would also need to commit to explaining how they execute the decisions of the SMC and why they have executed an SMC decision in a specific way.

International human rights law and content moderation

ARTICLE 19 recommends that social media companies should ensure that their [terms of service comply with international standards on freedom of expression](#) as a consequence of their responsibility under the UN's [Guiding Principles on Business and Human Rights](#). As a result, we have proposed that the rules informing the work of the SMC should be based on international standards on human rights.

In his [2018 report](#), David Kaye, UN Special Rapporteur on freedom of expression, stated that “Companies should incorporate directly into their terms of service and ‘community standards’ relevant principles of human rights law that ensure content-related actions will be guided by the same standards of legality, necessity and legitimacy that bind State regulation of expression.” As demonstrated by scholarly debates,¹⁷ the application of international human rights law to content moderation leaves many questions open.¹⁸ This is why, as a way to further discuss and clarify what the application of international standards to content moderation would entail, we suggested that one of the first tasks for the SMC could be to adopt a code of human rights principles for content moderation.

As one scholar argues, “One of the key benefits of IHRL [international human rights law] in this context is that it can provide a common vocabulary for content moderation debates so that even as rules are contested and the participants in these debates plainly disagree about which policies promote the public good[,] ...there is value to putting them in conversation with one another. ...But the important caveat is that for argumentative practice to be successful, participants must actually be in conversation. Creating legitimacy and accountability through argumentative practice requires an institutional structure that facilitates exactly this kind of argument and contestation.”¹⁹

This is precisely the role that the SMC would play: provide a participative, transparent forum where stakeholders could develop an agreement on approaches to content moderation that comply with international standards on freedom of expression and other fundamental rights.

A majority of participants in our conversations agreed that the work of the SMC should be based on international human rights standards. One participant noted that the [Global Network Initiative Principles](#) could provide useful guidance. Two people insisted that the

fundamental rights that inform the work of the Irish SMC should be specifically European human rights and values and the Irish Constitution.

The SMC and the legal and regulatory framework in Ireland

In Ireland, and in the EU more generally, a new legal and regulatory framework for social media platforms is emerging. In the EU, the evolution comes from the revised AVMSD and the future [Digital Services Act](#). [Article 28b of the AVMSD](#), which is currently being transposed in national laws, provides that video-sharing platforms must adopt ‘appropriate measures’ in relation to the protection of minors, hate speech, and incitement to terrorism, and this duty will be enforced by Member States’ regulatory authorities. Under the future [Digital Services Act](#), online platforms could be required to [put in place an internal complaints mechanism](#) (Article 17) and partake in out-of-court dispute settlement mechanisms certified by national authorities (Article 18).

There is an ongoing legislative process in Ireland in relation to the regulation of social media platforms. Beyond the transposition of the revised AVMSD into Irish law, the OSMR Bill²⁰ aims to regulate ‘harmful online content’. While [preventing online harms is a ‘worthy objective’](#), it needs to be balanced with the fact that the ability of governments to regulate speech is limited by international standards on freedom of expression as well as by constitutional provisions. In particular, international standards require that legal provisions that define categories of content as illegal must be clear and precise. However, legislative initiatives that seek to prevent online harms target categories of content that are not prohibited under law and are inherently difficult to define, e.g., speech that is ‘lawful but harmful’.

Giving voice to diverse perspectives on complex social issues

Societal problems – such as bullying, hate speech, or disinformation – are inevitably complex: a range of expertise and perspectives are needed to understand such issues and, as part of a broader approach, to define appropriate technological measures. No law should encourage social media companies to attempt to replace the combination of

diverse viewpoints and expertise that is required to regulate how social harms manifest online.

The complex nature of ‘online harms’ calls for an approach that gives voice to a diversity of perspectives, and the SMC offers a very effective forum for a transparent and open discussion of such multi-faceted and complex questions. This is the first reason why we suggest that a multi-stakeholder forum such as the SMC can play a positive role in the implementation of the future legal and regulatory framework for social media platforms.

Under [Article 28b of the AVMSD](#), Member States should encourage co-regulation when it comes to implementing ‘appropriate measures’ that video-sharing platforms need to adopt in relation to the protection of minors, hate speech, and incitement. Under the same provision, these ‘appropriate measures’ “shall be determined in light of the nature of the content in question, the harm it may cause, the characteristics of the category of persons to be protected as well as the rights and legitimate interests at stake, including those of the video-sharing platform providers and the users having created or uploaded the content as well as the general public interest.” We suggest that, within a framework of co-regulation where the respective roles of a statutory regulator and the multi-stakeholder model could reinforce each other, the SMC has a role to play in providing a forum where all stakeholders can come to an understanding of what the ‘appropriate measures’ should be.

In this configuration, a statutory public authority would set general and systemic objectives (such as the existence of appropriate processes and measures), while the SMC would provide a space where technical and practical mechanisms and innovations towards these objectives can be discussed with all stakeholders and tested for compliance with international standards on freedom of expression and other fundamental rights. Under monitoring by the statutory authority, the SMC provides some breathing room that facilitates the emergence of a consensus on the appropriate approach towards legal requirements, as well as a broader understanding of the complex challenges of content moderation.

The online safety codes of conduct that will be adopted by the new Media Commission could recognise the SMC as a legitimate approach to online harms. The [General Scheme of the OSMR Bill \(Head 50A\)](#) could provide for the online safety codes to acknowledge that

participation in a multi-stakeholder voluntary-compliance body, such as the SMC, is considered an indication that designated online services have taken measures to tackle online harms.

Individual complaints mechanism

The operation of an individual complaints mechanism is the second reason why the SMC could play a positive role in the future legal regime for the regulation of social media platforms.

The [General Scheme of the OSMR Bill \(Head 52B\)](#) provides that the future Media Commission will be able to treat complaints relating to ‘systemic issues’ on social media platforms. Only bodies nominated by the regulatory authority – such as a non-governmental organisation – will have the opportunity to call the attention of the Media Commission on issues they have identified. Under the OSMR Bill, there is no option for individual users to access an external complaints mechanism.

The SMC can serve as the out-of-court redress mechanisms provided for in [Article 28b, para 7, of the AVMSD](#): “Member States shall ensure that out-of-court redress mechanisms are available for the settlement of disputes between users and video-sharing platform providers.”

The SMC could also serve as the out-of-court dispute settlement mechanism required under the current formulation of Article 17 of the [Digital Services Act](#): the complaints mechanism that it will operate will meet the requirements for certification as provided for in the current formulation of Article 18 of the Proposal for the Digital Services Act.

Finally, we suggest that the participatory nature of sector-wide self-regulation would enable the SMC to serve as a co-learning forum where the new legal and regulatory framework can be debated, while the regulatory authority could step in if the self-regulatory approach fails to deliver its promises.

Funding

The SMC needs sufficient and sustainable funding to ensure its operation and guarantee its independence, and no funding would be accepted that could undermine its independence.

There are various possible sources of funding the SMC, from governmental subsidies to contributions from social media companies or grants from foundations or other international or national donors. From our conversations with stakeholders, the idea emerged that the SMC could also enter into a contractual agreement with the new Media Commission to operate an individual complaints mechanism.

Next steps: Creating a working group to develop the Irish SMC

ARTICLE 19 has kicked off discussions on creating an Irish SMC, and some progress has been made through online meetings, as reflected in this report. However, the future development of the project would be more effective if it was led by a working group composed of interested stakeholders. This approach would contribute to a stronger sense of collective ownership that would boost the SMC.

The role of this working group would be to:

- Draft and adopt a constitution for the Irish SMC.
- Define the essential principles for the operation of the complaints mechanism and the elaboration of general guidelines.
- Ensure funding for the first year(s) of operation of the Irish SMC.
- Organise a steering committee, once the constitution is adopted, that will be in the spirit of a start-up and will lead the Irish SMC to its full operational capacity.
- Further explore the application of international human rights law to content moderation, by preparing the adoption of a Code of Human Rights Principles for Content Moderation.

The working group would be constituted by individuals from the various categories of stakeholders who agree to work together towards creating an SMC in Ireland.

While the working group would be driven by a core group of stakeholders, not all participants in our conversations so far would be ready to commit to that level of engagement in the project: we suggest that some may take part as observers.

At its first meeting, the working group would elect a chair and co-chair, agree on its terms of reference, and adopt a programme of work. Further suggestions on the creation and operation of the working group are included in the Draft Terms of Reference in the Annex.

Annex: Draft terms of reference for the working group

1. From March 2020 to July 2021, ARTICLE 19 held online conversations with various Irish interlocutors in relation to setting up a multi-stakeholder, voluntary-compliance body for overseeing content moderation on social media platforms based on international standards on human rights. This new body has been called the Social Media Council (SMC). A working group was set up to continue these conversations, and those individuals who join the group commit to working together towards creating an SMC in Ireland, based on this report, which serves as a general roadmap for the SMC development.
2. The members of the working group adhere to the following principles:
 - a. We unreservedly adhere to international and European human rights law.
 - b. We agree that the oversight of content moderation on social media should be based on international and European human rights law.
 - c. We agree that the composition of the SMC should respect gender parity and ensure the representation of the whole diversity of society, including vulnerable or marginalised groups.
 - d. We agree that the working group will work in a spirit of independence from public and private interests and in the objective of setting up an SMC that serves the general interest.
 - e. We agree that the future multi-stakeholder body should be:
 - i. Independent from government, commercial, and special interests.
 - ii. Established via a fully consultative and inclusive process – major constitutive elements of their work should be discussed in an open, transparent, and participatory manner that allows for broad public consultation.
 - iii. Democratic and transparent in their selection of members and decision-making.
 - iv. Include broad representation: it is important that the composition of the self-regulatory body includes representatives of the diversity of society.
 - v. Have a robust complaint mechanism and clear procedural rules to determine if applicable standards were breached in individual cases and have the power to impose only moral sanctions.

- vi. Work in the service of the public interest and be transparent and accountable to the public.

3. The mandate of the working group is:

- a. To discuss and adopt a constitution for the Irish SMC.
- b. To identify possible sources of funding for the first years of operation of the Irish SMC.
- c. To ensure that the Irish SMC will operate smoothly within the future legal and regulatory framework for the regulation of social media platforms in Ireland.
- d. To ensure a smooth transition by appointing a steering committee, once the constitution is adopted, that will be in charge of setting up a company limited by guarantee and the operation of the SMC.

4. Members join the working group on a voluntary basis, in their individual capacity or as representative of a particular stakeholder. With no consideration to the number of its staff or members taking part in the working group, each organisation only has one vote.

5. New members can join the working group by sending an application to the chair. The chair informs the working group of the applications: should there be opposition to the entry of a new member, it shall be decided at the next meeting of the working group as a matter of priority.

6. The meetings of the working group are open to observers from the various categories of stakeholders. Observers are invited to the working group meetings by the chair. Observers do not take part in votes.

7. The working group will make all reasonable efforts to ensure a balanced participation from all categories of stakeholders. The working group will keep track of the representation of the various categories of stakeholders.

8. At its first meeting, the working group shall review and adopt the terms of reference as its mandate and constituting instrument. The working group will elect a chair, a co-chair, and a secretary. The secretary will prepare the agendas of meetings and keep minutes of the meetings.

9. Decisions are made by consensus. If no decision emerges through consensus, the chair can either postpone the decision on the corresponding item of the agenda or call

for a vote. Decisions are adopted by a majority of votes; in case of equality, the chair has the deciding vote.

10. The working group can create sub-working groups to prepare its work on specific points. The working group can decide to hold online meetings and use online asynchronous collaboration tools.

Endnotes

¹ See, for instance, H. Tworek, '[The dangerous inconsistencies of digital platform policies](#),' CIGI, 13 January 2021; B. Wagner et al., '[Reimagining content moderation and safeguarding fundamental rights: A study on community-led platforms](#),' The Greens/EFA, May 2021.

² D. Kayyali and J. C. York, '[The Facebook Oversight Board is making good decisions – but does it matter?](#),' Tech Policy Press, 28 July 2021; F. Coleman et al., '[The promise and pitfalls of the Facebook Oversight Board: A human rights perspective](#),' Carr Center for Human Rights Policy, May 2021.

³ Facebook is reportedly considering the creation of a body similar to its Oversight Board for election-related matters: see '[Facebook said to consider forming an election commission](#),' *New York Times*, 25 August 2021.

⁴ ARTICLE 19, '[At a glance: Does the EU Digital Services Act protect freedom of expression?](#)', 11 February 2021; '[EU: Due diligence obligations in the proposed Digital Services Act](#)', 21 May 2021; '[EU: Regulation of notice and action procedures in the Digital Services Act](#)', 21 May 2021; '[EU: Regulation of recommender systems in the Digital Services Act](#)', 14 May 2021; '[Digital Services Act: IMCO draft report raises freedom of expression concerns](#)', 30 June 2021.

⁵ On press councils, see ARTICLE 19, '[freedom and accountability safeguarding free expression through media self-regulation](#)', March 2005.

⁶ Global Partners Digital, '[A rights-respecting model of online content regulation by platforms](#)', May 2018; C. Tenove et al., '[Poisoning democracy: How Canada can address harmful speech online](#)', Public Policy Forum 8 November 2018.

⁷ Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, '[A/HRC/38/35](#)', 6 April 2018, section 72.

⁸ *ibid*, section 45.

⁹ *ibid*, section 58.

¹⁰ See, for instance, European Parliament ThinkTank, '[Online platforms' moderation of illegal content online](#)', 23 June 2020; CIGI, '[Models for platforms governance](#)', October 2019; E. Douek, '[The limits of international law in content moderation](#),' *UC Irvine Journal of International, Transnational, and Comparative Law*, 4 December 2020; D. Kayyali and J. C. York, '[The Facebook Oversight Board is making good decisions – but does it matter?](#),' Tech Policy Press, 28 July 2021.

¹¹ H. Tworek, '[Social media councils](#),' CIGI, 28 October 2019.

¹² From December 2020 to March 2021, ARTICLE 19 ran an information campaign on disinformation, regulation of social media platforms, and freedom of expression in Ireland. The campaign aimed to explain to young adults (18–25-year-olds) how responses to disinformation and the regulation of social media platforms should meet the requirements of the protection of freedom of expression. See ARTICLE 19, '[Keep it real](#)'.

¹³ Where anyone at these meetings could use the information from the discussions but were not allowed to reveal sources.

¹⁴ By way of comparison, on the creation of the Facebook Oversight Board, see K. Klonick, '[Inside the making of Facebook's Supreme Court](#),' *New Yorker*, 12 February 2021; E. Douek, '[What kind of oversight board have you given us?](#)' *The University of Chicago Law Review Online*, May 2020; Facebook, '[Global feedback & input on the Facebook Oversight Board for content decisions](#)'.

¹⁵ Under the Online Safety and Media Regulation Bill, the Broadcast Authority of Ireland (BAI) will be replaced by a new regulatory authority, the Media Commission.

¹⁶ In the longer term, to facilitate coordination and harmonisation in the application of international standards to content moderation, national SMCs could organise into a regional or international network or develop relationships with global initiatives such as the Facebook Oversight Board.

¹⁷ J. P. Howell, '[The Lawfare Podcast: The arrival of international human rights law in content moderation](#),' 27 May 2021; E. Aswad, '[The future of freedom of expression online](#),' *Duke Law & Technology Review*, 19 September 2018; M. K. Land, '[Against privatized censorship: Proposals for responsible delegation](#),' *Virginia Journal of International Law*, 24 August 2019; E. Celeste, '[Digital constitutionalism: a new systematic theorization](#),' *International Review of Law, Computers and Technology*, 3 January 2019; M. K. Land, 'Regulating private harms online: content regulation under human rights law,' in *Human rights in the age of platforms*, edited by R. F. Jørgensen (MIT Press, 2019); E. Aswad, '[To protect freedom of expression, why not steal victory from the jaws of defeat?](#)' *Washington & Lee Law Review*, 5 May 2020; B. Sander, '[Freedom of expression in the age of online platforms: The promise and pitfalls of a human rights-based approach to content moderation](#),' *Fordham International Law Journal*, 26 May 2020; S. Benesch, '[But Facebook's not a country: How to interpret human rights law for social media companies](#),' *Yale Journal on Regulation*, 14 September 2020; M. Karanicolos, '[Squaring the circle between freedom of expression and platform law](#),' *Pittsburgh Journal of Technology Law & Policy*, 7 May 2020; E. Douek, '[The limits of international law in content moderation](#),' *UC Irvine Journal of International, Transnational, and Comparative Law*, 4 December 2020.

¹⁸ One scholar notes that: "One benefit of the online environment is that platforms can develop far more nuanced remedies than have traditionally been available to governments. Far beyond the false binary of choosing to leave content up or take it down, they can choose to label it, amplify it, suppress it, demonetize it, or engage in counter-messaging. IHRL encourages if not mandates platforms explore these options by requiring any measure restricting expression be necessary, in the sense that it is the least intrusive instrument. Again, however, to date it does not offer any guidance beyond this general command, but essentially all content moderation decisions will be within this grey zone. A jurisprudence can develop over time, but the platform lawyer charged with making these determinations today is left on their own." (Douek, [The limits of international law in content moderation](#)).

¹⁹ *ibid.*

²⁰ ARTICLE 19 has not conducted a detailed legal analysis of the OSMR bill, and our observations in this document are limited to the relationship between the SMC and the future legal and regulatory framework in Ireland. Other civil society organisations have submitted observations to the Oireachtas Joint Committee on Media, Tourism, Arts, Culture and Sport on the OSMR bill and freedom of expression. See Irish Council for Civil Liberties, '[Free expression concerns with Online Safety and Media Regulation Bill](#)', 11 March 2021; FuJo and Anti-Bullying Centre, '[FuJo-ABC submission on the Online Safety and Media Bill](#)', 9 March 2021; Global Partners Digital, '[Written submission on the general scheme of the Online Safety and Media Regulation Bill](#)', 23 March 2021.