



CONFERENCE REPORT

# SOCIAL MEDIA COUNCILS

## From Concept to Reality

*February 1-2, 2019*

# ACKNOWLEDGMENTS

This document is a reflection of the ideas expressed by stakeholders from government, civil society, academia, and the private sector who participated in an international working meeting convened by Stanford University's Global Digital Policy Incubator (GDPI), ARTICLE 19, and David Kaye, UN Special Rapporteur on the promotion and protection of freedom of opinion and expression. The meeting, and this publication, would not have been possible without the active participation and substantial contributions of these stakeholders.

GDPI would also like to thank ARTICLE 19 and David Kaye for their generous cooperation before, during, and after this event. These partners were instrumental in organizing this meeting and coordinating the drafting of this report.

# Contributing Authors

## Global Digital Policy Incubator

**Eileen Donahoe:** Executive Director

**Larry Diamond:** Principal Investigator

**Megan Metzger:** Associate Director for Research | *Principal author*

**Jan Rydzak:** Associate Director for Program

**Roya Pakzad:** Research Associate, Project Leader, Technology & Human Rights

**Sarahi Zaldumbide:** Program Associate

---

## ARTICLE 19

**Thomas Hughes:** Executive Director

**Barbora Bukovská:** Senior Director for Law and Policy

**Pierre François Docquir:** Head of Media Freedom Programme

**Barbara Dockalova:** Senior Campaigner

---

## United Nations Special Rapporteur on Freedom of Opinion and Expression

**David Kaye:** United Nations Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression

**Amos Toh:** Legal Advisor to the United Nations Special Rapporteur on the right to freedom of opinion and expression

---

# CONTENTS

<b>1. PREFACE</b>	01
<b>2. INTRODUCTION</b>	05
<b>3. A MULTISTAKEHOLDER APPROACH</b>	07
<b>4. AREAS OF CONVERGENCE</b>	10
4.1 A Need for Transparency	
4.2 Human Rights and the Freedom of Expression as Foundational Principles	
<b>5. THE FUNCTIONS OF SMCs</b>	12
5.1 SMCs as Appellate or Case Review Bodies	
5.2 SMCs as Advisory Bodies	
5.3 SMCs as Case Review and Advisory Bodies	
<b>6. GEOGRAPHIC SCOPE</b>	16
6.1 SMCs at the National Level	
6.2 SMCs at the Global Level	
6.3 SMCs at the Regional Level	
6.4 Hybrid Models	
<b>7. FINANCING SMCs</b>	21
<b>8. STRUCTURE, SELECTION PROCESS, CRITERIA</b>	23
<b>9. NEXT STEPS: MAKING SMCs A REALITY</b>	25
<b>10. THE FUTURE OF SMCs: TWO VISIONS</b>	26
10.1 GDPi: A Global Multistakeholder Approach	
10.2 ARTICLE 19: A National-Level Approach	
<b>APPENDIX: CONFERENCE AGENDA</b>	33





# 1. PREFACE

On February 1-2, 2019 Stanford’s Global Digital Policy Incubator, ARTICLE 19, and United Nations Special Rapporteur David Kaye convened an international, multistakeholder working meeting: **Social Media Councils: From Concept to Reality.**

Our goal was to start a process to help devise a feasible model for a multistakeholder content moderation body that is maximally protective of free expression while protecting societies from the harmful effects of controversial digital content, including cross-border information operations, incitement to violence and hatred, propaganda, and terrorist recruitment. The meeting was grounded in a proposal created by ARTICLE 19 recommending the development of a multistakeholder model for content moderation called Social Media Councils (SMCs).

Our original plan was to convene a workshop with a small group of civil society members and academics to assess ARTICLE 19’s proposal for the creation of SMCs before consulting with a broader range of stakeholders.

However, the policy context changed on November 15, 2018 when Mark Zuckerberg announced Facebook’s intention to create an “independent governance and oversight” body to oversee hard cases and appeals of

Facebook's content decisions. Immediately after that announcement, we decided to broaden the scope of the event and to include a much wider range of stakeholders. In addition, the week before the event, Facebook publicly released their draft charter for an external oversight board, which became an important point of reference for our working meeting.

These developments underscore the fluid nature of the subject and the context for this discussion. There is no precedent or existing institutional structure that is perfectly fit for the purpose of making hard choices about content restrictions on private sector global digital platforms. A key conceptual challenge is that global digital platforms facilitate instantaneous cross-border information sharing at an unprecedented scale and now function as public spaces for civic discourse in many societies. A great deal of creativity and flexibility will be required to develop a model for SMCs.

The goal of our working meeting was to provide an opportunity to flesh out various dimensions of the model as well as to help assess which aspects are essential and feasible, and which aspects are contested, impractical, or need to be reconsidered.

At the start, we knew several core elements of the ARTICLE 19 proposal needed to be tested. Most obvious among those were related to scope, function, structure and standards. Key starting questions for the meeting included:

- 1) **Geographic scope:** What geographic level is optimal for SMCs? Is a national model feasible, or is a regional or global model better? Relatedly, if SMCs are constituted at the national level, what factors will be used to determine jurisdiction: the origin of the content, the impact of the content, or something else?
- 2) **Industry scope:** Can an "industry-wide" model work or is a platform-specific model more feasible? If industry-wide, how do we define the boundaries of the industry?
- 3) **Function:** What function should SMCs undertake? Should they be primarily advisory and simply provide guidance on content decisions or should they have broader adjudicatory authority?
- 4) **Standards:** How should we think about the interplay between national laws restricting expression, private sector terms of service and community guidelines (which arguably reflect platforms' own free expression), and international human rights law (IHRL)?

5) **Structure:** What structure makes most sense? Who decides on that structure? Who selects SMC members and on what basis?

Notwithstanding our openness to addressing these fundamental questions, we did start with two foundational points of agreement. First, GDPI, ARTICLE 19, and UN Special Rapporteur David Kaye share a strong normative commitment to protecting freedom of expression in the digital realm. We anchor our commitment to protect free expression in international human rights law, which we see as a set of universally applicable global principles that are particularly well-suited to a global digital information ecosystem.

That said, we fully appreciate the fact that embracing a commitment to free expression doesn't answer the hard question of how to protect it on global digital platforms or where to draw the lines on the legitimate, reasonable, and limited restrictions on speech in democratic societies, which are permissible under international human rights law.

In fact, we have all seen how superficial reliance upon free speech norms can function as a form of duplicitous cover for those who actually seek to undermine free expression and to attack free societies. One recent case in point is the lawsuit filed against Facebook in November 2018 by the Internet Research Agency (IRA) alleging that Facebook violated the IRA's free speech rights by taking down its posts and pages. Policy decisions about how to simultaneously protect free expression and ensure that free expression norms are not being co-opted by malign actors are some of the thorniest questions being addressed in democratic societies today. Accordingly, we believe many of the challenges faced by private sector global platforms with respect to content moderation need to be addressed by the larger society.

This leads to the second core normative commitment reflected in this gathering: We believe a multistakeholder approach to digital policy development generally, and specifically in the development of SMC models, will yield more effective outcomes than policies developed by governments alone or by the private sector alone. Incorporating a wider and more diverse range of inputs can produce better solutions to long-standing problems in content moderation. In addition, we believe a multistakeholder approach will be essential to building legitimacy for the substantive, procedural and structural elements of any SMC model. Global multistakeholder processes can be extremely challenging, but such input is also extremely valuable and will yield better outcomes.<sup>1</sup>

With this event, we started what we hope will be a constructive, global multistakeholder process of developing SMCs that support free expression in the global digital context.

**Eileen Donahoe**

*Executive Director, Stanford Global Digital Policy Incubator*

**Thomas Hughes**

*Executive Director, ARTICLE 19*

**David Kaye**

*UN Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression*





## 2. INTRODUCTION

In the past decade, social media and other online platforms have rapidly become some of the most important spaces for people to express themselves and share information. Billions of people around the world use these platforms for all forms of expression, from sharing baby pictures with family friends to organizing anti-government protests. These same spaces, however, are also used to incite violence and racial hatred, to recruit people to terrorist organizations, and to intentionally spread disinformation about important issues like elections or medicine. For this reason, moderating content online has come to the fore as a key challenge for increasingly digitized societies.

This is a particularly challenging problem because the spaces where content is created and posted are owned by private companies, which have to this point made decisions about content based on their community guidelines (CGs) or terms of service (TOS). Yet, as use of these platforms has skyrocketed, they have increasingly become a key component of the public square – the place where people share opinions, ideas, goods, services, and so much more. This development has spurred the need for methods to moderate content that comply with the standards applied to public speech.

Government regulation of platforms is one solution, but as with any situation in which governments regulate speech, this brings up serious free speech concerns,

and early attempts at regulation have often put platforms in the position of being the enforcers of criminal laws that restrict content. How can we best balance the responsibility to protect free speech online with the need to prevent harmful effects, while factoring in the challenges posed by the private ownership of the digital spaces where these forms of speech are appearing?

Our working meeting at Stanford on February 1-2, 2019 aimed to tackle this set of challenges. ARTICLE 19's original proposal recommended the creation of councils at the national level that would serve as an appeals body for content moderation decisions made by platforms. These national councils would all be governed by a global code of principles grounded in international human rights standards, but these principles would be applied within a local context. Moreover, the national councils would all be linked through a global association of councils that would set best practices in relation to the principles and work of the councils.

The conversations that took place over two days brought up a number of challenges to this model, but also many converging ideas, particularly on the validity and potential benefits of the model itself. This report is an effort to synthesize what was learned during the meeting about the viability of the SMC model, and to provide insight to help move the concept forward. We will begin by discussing the value of a multistakeholder approach to content moderation online, and why this type of model is an important step towards addressing the challenges posed by online content. We will then highlight the areas of convergence from our discussions and explore some of the most critical outstanding questions in depth, before suggesting opportunities for next steps in the development of SMCs.

### 3. A MULTISTAKEHOLDER APPROACH

Multistakeholder models of content moderation can help to avoid the pitfalls of existing private sector approaches to content regulation as well as the regulation of content by governments. They have substantial advantages for platforms, governments, and users.

At the moment, most decisions about content online are made based on the TOS or CGs of private companies. This is beginning to change, however, as governments respond to what they perceive as the proliferation of harmful content online.

The multistakeholder SMC model proposed by ARTICLE 19 is an attempt to find an approach to content regulation that avoids the greatest pitfalls of the existing private sector and government models for governing content. On the one hand, what has been called “solo regulation”<sup>2</sup> by companies has been opaque and left companies open to criticisms about lack of transparency. Because companies are profit-driven, content moderation decisions have often been framed as being motivated by the company bottom line. Companies are particularly prone to responding to public pressure, which may not always be well-informed or motivated by a commitment to human rights. The lack of transparency and external oversight over how decisions are made about content, even when companies are making decisions for legitimate and important reasons, leaves companies open to considerable public criticism.

On the other hand, government regulation of speech in democratic societies is always problematic, and initial attempts made by governments to regulate content online have imposed worrisome restrictions on free speech. The clearest example of this is the German law known as NetzDG, passed in 2017. The law requires strict timelines for the removal of “illegal” content and imposes steep fines if content is not removed quickly enough. The law has been criticized for endangering free speech for three reasons. First, the categories of speech that are criminalized are more expansive than international free speech principles

would allow. Second, it shifts authority for enforcement of criminal laws on speech to private companies. Third, it creates incentives for these companies to err on the side of removal of any questionable or borderline content to avoid fines. NetzDG and similar laws raise very serious concerns about the protection of free speech and are a product of governments' reactionary response to new challenges without fully considering long-term implications. From the platform perspective, this statutory approach is less than optimal, not only because it may create new burdens for companies concerning the enforcement of content moderation, but also because, as more states implement regulations, platforms will be forced to comply with divergent laws around online content in varying jurisdictions.

A multistakeholder “self- or co-regulation” approach like SMCs can help to address many of the concerns posed by both solo regulation and government regulation. First, one of the primary goals of such a model is the protection of free expression.

Both solo regulation and government regulation pose potential threats to free expression online, though for different reasons. An independent body with international human rights standards as its foundation can help to ensure that as we develop new approaches to content, individuals' ability to freely express themselves online is maintained as a central tenet. Secondly, an external and independent body will create greater transparency in the process of content moderation, which benefits all stakeholders involved. Finally, through the application of human rights principles, SMCs could serve as a check on both overreaching government regulation and companies' practices that are not in line with international standards on freedom of expression. This model is not meant to replace platforms' own systems for content moderation or eliminate all government regulation, but instead to create an independent, accountable, and transparent mechanism that can cooperate with platforms to improve their own systems and eliminate the need for some regulation.

***This model is not meant to replace platforms' own systems for content moderation or eliminate all government regulation, but instead to create an independent, accountable, and transparent mechanism that can cooperate with platforms to improve their own systems and eliminate the need for some regulation.***

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



## 4. AREAS OF CONVERGENCE

While some challenging questions about how to develop a successful multistakeholder approach to online content moderation remain, there was convergence on many foundational issues related to content moderation.

### 4.1 A Need for Transparency

First, there was broad agreement that greater transparency and clarity around how content decisions are made is essential. There is currently an enormous amount of confusion among the general public, and even among experts, about how platforms are making content decisions, including when and whether content is being removed because of platform TOS or CGs or due to government-imposed restrictions. This is bad for platforms as well as those using them. Facebook's recent proposal for its own platform-specific external review board is, in part, an attempt to improve the transparency of their system, but this transparency is needed across the industry. There are clear challenges to transparency, and companies cannot always be fully transparent in every case (for example, in some cases where law enforcement may be involved), but increasing transparency is essential.

### 4.2 Human Rights and Freedom of Expression as Foundational Principles

Second, participants in the working meeting converged around the importance of international human rights law and the protection of free expression as the substantive principles that should guide the work of the SMCs. A human rights lens should form the foundation for how we think about all content decisions. Existing human rights standards already provide substantial guidance for how to balance freedom of expression against the risks of certain types of speech

(such as inciting violence). These standards provide a universal anchor point for all stakeholders—governments, companies and users alike. Their practical usefulness lies in resolving disparities between competing values while remaining relatively sensitive to regional differences. Another advantage that human rights principles provide in addressing challenging issues related to speech online is that, through instruments like Article 19 of the International Covenant on Civil and Political Rights, they provide a clear set of guidelines on what the limits of acceptable speech are and where speech can be reasonably restricted. Even in the case of private companies, while not binding, the UN Guiding Principles on Business and Human Rights provide guidance on how companies should approach their obligations related to human rights. Additionally, many companies already endorse the human rights framework and apply elements of human rights due diligence in their operations. Working from the premise that a commitment to human rights must underpin any attempts to create new mechanisms for content moderation online does much of the work in ensuring that different stakeholders are on the same page.

What these areas of convergence across stakeholders signal is a shared commitment to addressing an increasingly visible problem around content online, with a shared set of values from which to address these issues. While the remainder of this report will focus on the key practical questions that need to be resolved to move the concept of SMCs forward, it is important to begin by highlighting that stakeholders were broadly in agreement on these foundational principles. It is in the details that more open questions remain.

## 5. THE FUNCTIONS OF SMCs

The foundational issue identified at the working meeting related to the function of SMCs. The answers to many other questions depend on how we address the primary one: "What function should SMCs fulfill?"

In order to move the SMC concept forward, the precise function of the body must be articulated. The initial model put forward by ARTICLE 19 proposed that SMCs act essentially as an appeals mechanism after platforms' own processes for content moderation have been exhausted. There was an enormous amount of discussion about whether this was the best way for a multistakeholder model to accomplish its goals. Over the course of the meeting, three potential functional models emerged.

### 5.1 SMCs as Appellate or Case Review Bodies

One potential model for SMCs is to act as an appeals or a case review mechanism after platforms have already made content moderation decisions. An SMC system structured in this way would accomplish several goals. First, it would serve to improve transparency in the content review process. It would give users a space outside of platforms themselves to have decisions reviewed, which can provide a greater sense of agency over the process. For companies, it can have positive impacts in cases where their decisions are upheld, as it lends legitimacy to their own internal processes. If, as some platform representatives suggested, TOS and CGs are applied appropriately in most cases, this independent review should be a vehicle for improving trust in the internal processes of companies. In cases where SMCs disagree with the decision of platforms, decision review could provide platforms and users with valuable information about the differential meaning of certain types of content in different national settings as well as the unintended consequences of certain types of removals.

One problem that immediately arises when considering a body that carries out

review of content moderation decisions is scale. During the meeting, platform representatives emphasized that the scale of content review decisions is massive and unlikely to be manageable by a structure such as SMCs. Others pointed out that we do not know for sure how many cases would both go through a full adjudication process at platforms and then apply for a secondary process at SMCs. Additionally, if this were broken down to the national level, the scale might be smaller than anticipated. Broadly speaking, however, there was a sense that the scale of the work was going to be too large for SMCs to serve as an appellate mechanism for individual cases.

A potential solution to this problem is to have SMCs evaluate emblematic cases that represent a type of content moderation issue, or to evaluate cases as groups of content organized by type. This would allow the SMC to provide feedback to companies and guidance to users on the content moderation process and the application of substantive principles to emblematic cases without evaluating hundreds of thousands of individual pieces of content. Such a system would be better described as a case review system than an appellate system, as there would likely not be remedy provided for individual cases.

While much of the public discussion around content moderation online has focused on what platforms remove as violations of CGs or TOS, the more significant concern expressed

***To address the challenge of handling enormous volumes of content, SMCs could evaluate emblematic cases that represent a type of content moderation challenge, or review cases in groups organized by content type.***

by representatives from platforms was government requests to remove content based on a law and/or CGs and TOS. For platforms, this is the primary problem, as they argue that their own misapplication of their CGs or TOS is relatively rare, whereas government requests to remove what they view as legitimate content that should not be removed is relatively common. When government requests are related to enforcing local law, platforms can push back on requests, but in general companies are required to comply. When requests are not based in law, platforms often do push back, but representatives signaled that having support for that push back would be welcome. Although this was not part of ARTICLE 19's initial proposal, some participants in the meeting suggested that social media councils reviewing such requests could serve as an external body putting pressure on governments to minimize the overreach that platforms currently report. One risk here, however, is that it could result in draconian measures from governments — including the complete shutdown of users' access to the platforms — if they perceived that their authority was being undermined. One question that arose about the appellate or case review model is whether

it addresses the most pressing issue: Is the primary concern that needs to be addressed in this space about the misapplication of existing CGs or TOS, or about whether the CGs and TOS themselves are problematic? A case review model could be valuable in addressing misapplications of CGs and TOS or providing local context for unintended impacts of CGs and TOS, in addition to providing a potential method to push back on government attempts to restrict content online. However, the model does relatively little to address the concern that there should be more input from experts outside of companies on the development of CGs and TOS. This is one reason why the second model suggested during the meeting envisioned SMCs serving in an advisory capacity.

## 5.2 SMCs as Advisory Bodies

The second model discussed during the meeting envisioned SMCs serving an advisory function, helping companies to develop and adjust CGs and TOS in line with human rights law and changing circumstances. Those in favor of this model argue that, by and large, companies do a good job of enforcing their own CGs and TOS as written and of correcting mistakes once they are brought to the company's attention.<sup>3</sup> Therefore, some suggested that the place where a multistakeholder organization might actually have the most impact is in serving in an advisory capacity. This body could provide feedback to companies to improve the development and adjustment of their TOS and CGs, with a particular focus on ensuring that they are in line with international human rights standards. They could also provide insight on how applications of CGs and TOS in different contexts may have different and unforeseen impacts. This approach would greatly help in solving the scale issue discussed above, because SMCs would not directly engage in case reviews.

Some meeting participants critical of SMCs as purely advisory bodies pointed out that the advice could lack legitimacy and authority if it were coming from people who were not engaging directly with content moderation decisions. That is, advisory opinions that may seem reasonable on a broad level might make less sense when applied to the complexities of actually evaluating content. Others expressed the concern that SMCs would lack impact as a purely advisory body, questioning whether the model can truly address the problems associated with content moderation. Specifically, it was noted that advice that is not grounded in specific cases can lead to ambiguity in how and when to implement the SMCs suggestions, calling into question whether the SMCs are providing sufficient impact. For these reasons, a number of people at the event proposed that the best system for SMCs might instead consist of a hybrid function as an appeals body and an advisory body.



### 5.3 SMCs as Case Review and Advisory Bodies

The final model that received serious consideration during the event combined aspects of both models discussed above. As in the case review model, SMCs would review either emblematic cases or groups of cases of a given type. These might either be “hard cases” forwarded from platforms, brought to the attention of the SMC by civil society groups, or curated by ombudspersons or other SMC members. The SMC would review the cases and provide feedback to companies as well as write and publish publicly available explanations for how content decisions were made, in addition to the SMC’s evaluation of these decisions. The SMCs could also evaluate government requests for content removal, make public comments to help publicize inappropriate requests or engage in other activities to help platforms push back on government overreach, which is both a high priority for the platforms and an important component of protecting free speech online.

***SMCs could evaluate requests for content removal by governments and other actors, help publicize inappropriate requests, or engage in other activities to help platforms push back on government overreach. This is a high priority for the platforms and an important component of protecting free speech online.***

Under this model or other similar hybrid-type models, SMCs would engage directly with content and evaluate decisions, but the outcome and role would be explicitly to provide guidance and advice as well as advocacy on behalf of human rights

and freedom of expression online. This brings together many of the advantages of both models while solving many of the challenges such as scale and lack of impact. It also combines much of what platforms might want from an external body, in that it would help them advocate alongside governments and provide guidance on hard cases while also promoting transparency, human rights and free expression – some of the key goals of the SMC concept.

On the other hand, if a hybrid model is chosen, it is important that the goals and functions of the council are made extremely clear. While there are advantages to such a model, if the mandate is not well-defined at the outset, it could also lead to a lack of clarity in the focus of SMCs. Clearly defining the function of the SMCs, regardless of which model is chosen, was unanimously agreed upon as a priority in moving the idea forward.

## 6. GEOGRAPHIC SCOPE

SMCs can be structured on multiple geographical levels, ranging from a network of national-level councils to an overarching global body. The choice of geographic scope will have a strong impact on their ability to operate effectively.

The original proposal by ARTICLE 19 suggested a system of national SMCs, governed by a globally agreed-upon code of principles. During the event, however, there was significant disagreement about what geographic level was the most appropriate one on which to build SMCs.

Many participants feared that SMCs at the national level were unworkable for a number of reasons, while others argued that without national-level councils SMCs would not be able to deal with the scale of content in question, have the necessary relevance to local stakeholders, or be able to understand local contexts.

### 6.1 SMCs at the National Level

***While SMCs would all adhere to a universally applicable code of principles, the implementation of this code of principles in different settings would be sensitive to variations in regional, cultural, and local contexts.***

There are several advantages to the national-level model proposed in ARTICLE 19's original proposal. Such a model helps to guarantee that local contexts are considered

in review decisions, and that local expertise is readily available. Some at the event suggested that one important role SMCs might play is in facilitating communication between local experts and platforms around complex issues. While SMCs would all comply with the same code of principles, this code of principles could be applied somewhat differently in different regional or cultural contexts.

A national-level SMC would also allow for better responses to local needs. Two content situations that appear similar may actually have very different implications in different countries. Additionally, some governments are already discussing the development of their own content review boards at the national level. By creating independent review boards at the national level, we can potentially preempt government overreach and help to better protect freedom of expression online.

On the other hand, many concerns were expressed about the risk of capture by oppressive states that could use SMCs as a vehicle to accomplish state goals. Those expressing this view emphasized that these concerns were particularly acute in parts of the world that need independent review the most, and that, if captured, SMCs could potentially be co-opted into supporting repressive state activities. In this case, SMCs wouldn't be able to be as impactful if structured at the national level. However, it was also noted that platforms are already beholden to national legislation in the jurisdictions in which they operate.

Additionally, in many of these more challenging contexts, even if state capture could be avoided, there may be legal and physical risks associated with participation in a council that makes any attempt to challenge the government.

In this context, it was proposed that risks faced by persons participating in press councils be reviewed as a possible benchmark regarding risks to SMC members, although it should be noted that press councils do not exist

***Concerns were expressed about the risk of capture by repressive states that could use SMCs as a vehicle to accomplish their goals. These concerns are likely to be particularly acute in parts of the world that need independent review the most.***

in all countries, and the risks of participating in a press council may differ from the risks of participating in SMCs. However, some argued that even in more repressive states, though perhaps not in the very most repressive, the creation of a multistakeholder body might be able to put pressure on the government. Participants argued that structures in these states are already co-opted, but that that does not eliminate the space for agency on the part of civil society and other actors in these countries who could be supported by national SMCs. Nevertheless, a number of those at the conference believed that the risk of state capture made the idea of SMCs functioning at the national level seem unrealistic.

There are also concerns about determining jurisdiction in the case of national councils. Because content produced on social media is not shared only within

a given country, complications could arise when determining which state has jurisdiction. What happens if two national councils believe that they should have jurisdiction over content? Certainly, there would need to be a clear process established for adjudicating such disagreements.

Finally, there would need to be clarification on how to deal with situations in which two national councils decide similar cases differently. How are disagreements between national councils going to be adjudicated, especially cases that could be used to establish an international precedent? This is a particular concern if SMCs are going to evaluate emblematic cases for the purpose of setting precedents for evaluating similar content in the future. Does precedent differ by national context? How should platforms integrate conflicting feedback from national-level councils?

## 6.2 SMCs at the Global Level

***Constituting an SMC at the international rather than the national level avoids the largest risks of direct government interference in SMCs and also makes it easier to firmly ground the council in human rights principles.***

Because of concerns about the national-level model — particularly concerns over state capture of national-level SMCs — a number of participants at the event proposed that an SMC should be developed at the global

level, rather than the national level. Constituting an SMC at the international rather than the national level avoids the largest risks of direct government interference in SMCs and also makes it easier to firmly ground the council in human rights principles. Furthermore, it creates one unified body that can be expected to produce more uniform decisions. The problem of how to understand conflicting decisions by different national councils is resolved if only one council exists at the international level. Some of the platform representatives also argued that because their platforms are global and the content on their platforms is shared globally, constituting councils at the national level is not a good fit. Additionally, the overall costs of one international body would certainly be less than that of several national bodies.

One criticism of this model, however, is that it loses the advantages of local knowledge and expertise that are inherent to a national-level model. How will sufficient representation of local knowledge be ensured when cases are reviewed at a global level? Constructing a global SMC can make it much more challenging to achieve the nuance that would be provided by national-level

councils. Additionally, ensuring truly diverse representation is hard even at a national level, but much more challenging at a global level. Some participants suggested that if the council is not sufficiently diverse, the SMC will be vulnerable to claims of digital imperialism, thus losing legitimacy if it is seen as representing only a small group of elites attempting to dictate to the rest of the world.

Another issue is that some governments are already starting to move forward with regulation, some with their own national-level review bodies. Participants expressed concern that if the SMCs are not developed at the national level, national-level regulation by states will continue, and SMCs will not have a clear voice in national regulatory decisions. National-level SMCs also have the potential to prevent overzealous regulation by satisfying governments outside of regulatory mechanisms. Some at the conference therefore viewed the creation of national-level councils as central to the success of the SMC concept as a whole.

### 6.3 SMCs at the Regional Level

Regional-level councils were proposed as a potential middle ground between global councils and national councils. Those advocating this position argued that it mitigates some of the concerns with both of the

***Regional SMCs might mitigate some of the concerns with both of the other models while maintaining many of their advantages. However, they would still deal with problems of representation, diversity, and government influence.***

other models while maintaining many of their advantages. Regional councils would, for example, go a long way in alleviating concerns about state capture, would still be grounded in contexts that are closer to local, and would have an easier time facilitating the participation of local experts. Because each body would be represented by people from the region, the model would also help alleviate the concerns about digital imperialism expressed in relation to the global model, lending more local legitimacy to the SMCs. A regional body might also be better positioned to advocate as a group with governments in their region than a single national body would, particularly in cases where governments are not fully functioning democracies. Finally, jurisdictional issues, while not fully resolved by a regional model, would be less complex than under a national model.

Nonetheless, regional-level councils would still need to deal with concerns about representation and diversity. While this would be easier than at the global level, depending on how regions are drawn, it can be quite complicated



to ensure regional diversity while also representing people with a range of expertise and experiences. Similarly, to the extent that regional councils are advocating with governments, representatives from certain countries might still face safety concerns that would be challenging to address. It's also unclear whether a regional body would help to stave off national-level regulation. Additionally, given national animosities in some regions, regional councils may be undermined by intra-regional dynamics.

## 6.4. Hybrid Models

In addition to regional models, another option that takes advantage of some of the benefits of both the national and international models is to develop a hybrid between an international and a national-level model. One suggestion was the development of an international SMC which oversees a set of global standards – perhaps responsible for reviewing the most high-profile cases or those with the most substantial reach – and the development of national-level councils as needed to help address local issues. Under this model, states with the highest risk of capture or safety concerns probably would not have national councils, while others might.

Another option is a global SMC model with an established and well-developed network of regional and national-level experts who participate in cases that call upon their local and topic area expertise. This would ensure better access to local knowledge and help allay concerns around an elite, global institution while avoiding many of the pitfalls of a predominantly national-level model.

## 7. FINANCING SMCs

One major question in establishing SMCs is figuring out the best way to fund them. This involves balancing concerns around having sufficient resources for SMCs to operate successfully while also ensuring their ability to operate independently.

It is clear that a financial needs assessment is necessary to understand and estimate the full cost of SMCs, but a number of participants pointed out that they are likely to be more expensive than press councils, which partially inspired the concept, because of the nature of the problems they tackle. Some pointed out, for example, that SMCs will need state of the art data security infrastructure and highly trained data security experts on staff to maintain that security, as the councils will potentially be dealing with private data. Additionally, depending on the structure that is chosen, some SMCs may need significant physical security infrastructure.

An additional foreseeable cost is for participation in the councils themselves. Some representatives with a background in multistakeholder governance pointed out that if participation is on a voluntary basis, there will be limitations on who is able to participate. In particular, a voluntary system would make it harder for those working at NGOs or smaller organizations to be involved because their organizations cannot afford to lose their time. It is harder for participants with fewer economic resources, or from lower-income countries and regions to participate. This amplifies the risk of having a council that lacks in representation. Yet, if participants are paid, salaries would add dramatically to the costs of councils.

The three main sources of funding suggested were platforms, governments and third-party funders. In each case, significant concerns were raised about the legitimacy and independence of the SMCs.

The original ARTICLE 19 proposal, drawing inspiration from the press councils that are used for media self-regulation, suggested that funding should come from participating social media platforms. Concerns were expressed that this approach would undermine the legitimacy and perceived independence of the SMCs. One counter argument is that different platforms actually have quite different incentives, and that, as with media self-regulation, no single funder would have undue influence. Managing influence would be somewhat complicated, however, if funding by platforms was proportionate to their size, as this might be perceived to give larger companies disproportionate influence.

Another option is for governments to provide funding, but this dramatically increases concerns about state capture. This could be an option in some countries with particularly strong democratic

***The three main sources of funding suggested were platforms, governments, and third-party funders. In each case, significant concerns were raised about the legitimacy and independence of the SMCs.***

institutions, but not others. If SMCs are global, a coalition of governments could potentially provide funding, but this would likely amplify concerns about a small group of elites dictating standards to the rest of the world and fail to resolve the problem of governments vying for influence in a multi-stakeholder body.

Finally, some suggested seeking external funding from a third-party donor. Many raised concerns, however, that any donor sufficiently well-situated to provide funding would likely be seen as having a particular agenda and therefore raise questions about legitimacy if SMCs want to be a core component of content moderation online.

## 8. STRUCTURE, SELECTION PROCESS, CRITERIA

A significant set of procedural and structural questions must be addressed in order to bring the SMC concept to fruition, including number of members, the selection process, and the appropriate mix of stakeholders.

The original proposal from ARTICLE 19 outlines a wide range of stakeholders that must be included. These include representatives from the media, civil society, academics, representatives from platforms, and representatives from the advertising industry. In addition, the proposal emphasizes the importance of representing the full diversity of the national context when it comes to religious backgrounds, minority groups, and so forth.

A recurring point of discussion during the event involved the challenge of ensuring diversity, or in some contexts of even understanding what full diverse representation would mean. It will also be critical to carefully vet participants, especially given that similar institutions may have varying levels of trust or social capital in different contexts. For example, there are countries where civil society, or a portion of civil society, suffers from a lack of trust and is commonly seen as subservient to more powerful interests. Further, in addition to being diverse, participants must be sufficiently expert in order to be effective in their role, and balancing diversity and expertise will add another layer of challenge. If SMCs are constituted at the global level, this is an even harder problem to solve, as ensuring diverse representation in a global body is quite complex. The key question, consistently raised throughout the discussion, was how to ensure a range of views without enlarging the council(s) to an unwieldy and unmanageable size, while ensuring expertise.

Much of the discussion focused on who should serve on SMCs rather than how they should be selected, but these questions are deeply intertwined. One

important and recurring question was how to prevent institutional inertia, particularly in the selection of the first cohort of representatives on the SMCs. This grew, in part, out of a conversation about the fact that Facebook's first board will be appointed by the company. Several people pointed out that the initial cohort sets an important precedent that can influence who is selected in subsequent cohorts and how they are selected. This makes it all the more important to have a well-considered process from the outset. One idea suggested to address this was to have a public consultation process as a part of the selection of the board, which would both provide greater transparency and increase the likelihood of the board being broadly representative. Wikipedia, for example, has democratized its process for selecting admins. Admins must meet a loose set of criteria, but then there is public voting for selection. Such a direct democracy approach may be unfeasible for SMCs and can also be gamed, but a public consultation process could provide greater opportunity for input.

Additionally, mirroring the question of funding, questions arose about whether government representatives should be included. Obviously, this heightens concerns about

***Much of the discussion focused on who should serve on SMCs rather than how they should be selected, but these questions are deeply intertwined.***

potential state capture, and may put pressure on SMCs to follow government views, particularly in more restrictive contexts. Some others suggested, however, that government involvement might be essential in order to ensure buy-in, although perhaps this could be achieved without explicit participation on the council itself. For example, some governments might provide some start-up assistance in establishing councils but divest themselves from the actual content moderation process.



## 9. NEXT STEPS: MAKING SMCs A REALITY

In order to develop the optimal SMC model, it will be critical to solicit input from a wide range of global stakeholders.

The discussions that took place over two days left many questions still to be answered. Nonetheless, the exercise validated the potential value of an institution that can serve as an independent voice in protecting free expression online as we struggle to prevent some of the troubling societal repercussions stemming from online content in recent years. Such a model can yield advantages for all the major stakeholders and work in tandem with platforms' own processes, increasing legitimacy and transparency around decisions. It can also help prevent the need for government regulation by providing a clearly independent body to help address content moderation challenges.

The most important next steps are to integrate the lessons learned in this first meeting, and to gather more input to find answers to the outstanding questions posed here. We will first develop an updated and more detailed proposal that integrates the feedback from this initial meeting. Once this proposal is complete, it will be made available to the public as a consultation document to be used as the foundation for gathering further input. We will begin that consultation process virtually, soliciting feedback from stakeholders from a range of regions and backgrounds, including the participants from the first meeting. Then, we will organize several presentations of the proposal in front of a broad range of stakeholders so as to generate as diverse a spectrum of input as possible. This includes panels at UNESCO's World Press Freedom Day in Addis Ababa (May 2019) and at RightsCon in Tunis (June 2019). We will then organize a co-creation event that will bring together a variety of stakeholders to finalize the SMC model.

# 10. THE FUTURE OF SMCs: TWO VISIONS

Following the working meeting, GDPi and ARTICLE 19 each drafted short position papers that outlined potential next steps from their respective vantage points.

## 10.1 GDPi: A Global Multistakeholder Cross-Platform Social Media Council

The call for a multistakeholder social media council model to help address challenges associated with online content moderation is grounded in the interplay between several increasingly troubling trends:

1. **Governments** are considering or have passed regulations to address harmful content online that undermine freedom of expression;
2. **Private sector companies** are exercising extraordinary control over what content is and is not allowed on their platforms, which in many settings constitute a substantial proportion of the public square, with inadequate transparency or accountability in their processes;
3. **Assessments** about how to combat harmful online content while also adhering to fundamental free expression commitments have become increasingly complex, require nuanced contextual understanding, and are difficult to scale.

Platforms have different community values and need their own processes and internal standards for content assessment. Governments have a legitimate interest in protecting their citizens' privacy, freedom and security. But GDPi believes that there is an urgent need for an independent body that can advise both platforms and governments as they confront these challenges. In this

advisory capacity, the body can develop a baseline set of guidelines that serves as the foundation for how both states and platforms handle content moderation online, consistent with free expression principles. This body should bring together a mix of stakeholders from across the globe, with a wide variety of relevant expertise. It must have buy-in from platforms and from governments, but should have independence from both. It would serve not in an adjudicatory role, but simply as an advisory body that develops guidelines for online content moderation, reviews emblematic cases and gives guidance to companies and governments. The core function of the global SMC would be to mediate the relationship between government and industry around challenging issues, and to help provide transparency and education to the public.

**SCOPE:** GDPi believes there is a need for a social media council that is global in scope. Online content does not fall neatly within national boundaries. The problem is global, and solutions must address and reflect those global dimensions. This does not preclude more targeted, national councils or other national-level activity where appropriate, but these must be connected to a coordinated, multistakeholder body at the global level and to a foundational commitment to universally applicable human rights principles.

**FUNCTION:** GDPi advocates a model in which the multistakeholder social media council will not directly adjudicate individual cases or serve as an appeals body. Instead, it should serve several key functions:

1. The global SMC should develop a set of **core guidelines** grounded in international human rights principles for how to approach content moderation online. These guidelines would form the best practices around which both platforms and governments develop policies, and would emphasize the protection of free expression as well as clearer elucidation of legitimate restrictions online. These guidelines would apply across platforms, but would serve as a floor on top of which platforms can construct their own individual terms of service and community guidelines
2. When new or challenging types of cases arise, the global SMC would evaluate **emblematic cases** in order to provide advice to platforms on how to address them. This can lead to the development of precedent that can then be applied in future cases.
3. The global SMCs should serve to help **advise platforms** in the development or implementation of their terms of service and community guidelines to ensure that they are in line with core principles.
4. Global SMCs could serve as a conduit to **connect platforms with experts**

**and civil society representatives** to help provide local context, especially in places where it may be challenging for platforms to develop their own teams of experts.

5. A global SMC should help **mediate the interaction between governments and platforms**, in particular by providing a normative basis for the platforms to resist government demands that violate core principles or infringe on freedom of expression. The global SMC can also provide advice, recommendations and expertise to governments as they develop regulation around these issues.

**VALUE ADDED:** The global SMC will add value for all stakeholders.

1. **For Platforms:** Relying upon guidance from an independent body that articulates foundational principles for content moderation standards will help to add transparency and legitimacy to platforms' own processes, as well as an additional layer of accountability. A global SMC can also serve as a key resource in helping platforms push back against overreaching government demands around content removal. Finally, a global SMC can serve as a resource helping connect platforms with civil society to gain local context and expertise when needed, in places where they currently lack such local knowledge.
2. **For Governments:** A global SMC can be a vehicle for addressing many of the most urgent concerns around content online in an open transparent way. It would possess greater expertise around content restriction standards than most governments currently possess, and help governments avoid regulations that may not effectively address challenges, but may in practice violate human rights standards. They can be a resource for governments as problems arise and help to facilitate better platform/government relationships.
3. **For Civil Society:** A global SMC can give civil society a seat at the table and a voice in how standards around online content are developed and implemented. It can provide a new vehicle for connecting civil society with platforms, particularly in cases where they have relevant insight but haven't had a means for sharing that insight. It can also help them advocate for policy that protects their rights with governments.
4. **For Everybody:** A global SMC will increase the transparency about and accountability for online content moderation decisions. It can help ordinary people better understand these processes and provide greater legitimacy by ensuring that decisions are guided and influenced by an independent,

multistakeholder body. It can help to provide education and resources for the people impacted by these decisions.

**TIME FRAME:** GDPi believes that the time to develop and implement this model is now. Facebook is moving forward with their own platform-specific model, with plans to establish in late 2019 or early next year. It is critical that we emphasize right now that platform-specific models are necessary, but on their own they are not sufficient. There must be accountability from outside of platforms themselves. Microsoft has called for greater cross-platform collaboration to develop effective means of combatting harmful content while protecting fundamental human rights principles. The sense of urgency within governments to protect societies from the harmful effects of online content has also risen to new levels within the past year. This need can only be met through a global multistakeholder collaborative effort that is rooted in a deep commitment to free expression. We should aim to have fully developed the global SMC concept and a clear strategic plan for implementation by the end of the first quarter of 2020.

**Global SMC as Complement to Platform-Specific and National Councils:**

With its cross-platform perspective, the global SMC would complement platform-specific structures, like Facebook's independent oversight board, and national-level SMCs. A global SMC would not limit platforms' own review processes for content decisions, but work in conjunction with and add value to platform-specific bodies by providing baseline guidelines to inform content moderation processes as well as an outside structure to help evaluate hard cases. It would also serve as a resource for connecting platforms to local civil society actors around the world who can provide context-specific insight. A global, cross-platform SMC can also provide essential guidance to national-level SMCs when there are tensions between national legislation and universal human rights principles. National SMCs will gain increased legitimacy if checked and grounded in global principles, particularly in cases of national restrictions on content in hard cases. In addition, a global SMC can provide important support and guidance in cases where content crosses national boundaries and jurisdictional tensions arise.

## 10.2 ARTICLE 19: A National-Level Approach

We first imagined the original model for a Social Media Council with an ambitious scope: we have envisioned a network of national or regional councils entrusted with providing general guidance to social media platforms and deciding individual complaints brought by individual users, operating on the basis of international standards on human rights and coordinating through the mediation of an international structure. We believe that these multi-stakeholder, transparent, accountable and independent forums can weave freedom of expression within all aspects of online content moderation and distribution, from integrating international standards in decisions to delete or demote content to ensuring exposure to the broadest possible diversity of information and ideas through a form of ‘human rights optimized’ algorithmic distribution. We see these new forum as emerging at the exact point of convergence between the goals and interests of human rights groups and social media platforms: avoiding the pitfalls of harsh legislative approaches that often come with disproportionate sanctions, contributing to restoring trust from users through transparency and accountability, providing an effective yet adaptable form of regulation that can easily accommodate the constant evolution of tech platforms, and ensuring that moderation of speech is done on the universal grounds of international law. In addition, this model also takes the burden of complex public interest questions from the shoulders of private companies and brings them back into the public debate.

ARTICLE 19 views the SMC as necessary to ensure that social media platforms that exert considerable influence over media landscapes and public debates can effectively serve as facilitator of social discourse and social movements and support the online circulation of a broad range of information and ideas. Confronting an ideal, if slightly utopian, broad original vision with the real world, and seeking to identify the best possible concrete SMC in the current context, ARTICLE 19 considers that the SMC offers a model of self-regulation (or multi-stakeholderism) that appropriately responds to national-level challenges.

The national level offers an excellent entry point for the identification of the stakeholders that would take part in the creation and operation of the SMC. While the SMC should represent a broad diversity of the living forces of society, these would be representatives of groups that recognise international standards on human rights. 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 local context in its cultural, political, and social dimensions. In certain countries, national civil society actors may lack the



capacity to fully engage into the operation of such a mechanism; in oppressive regimes, civil society organizations often find themselves under direct control or close monitoring of public authorities. The regional level can be an option that would allow civil society to put forces together. The regional level, possibly combined with a working relationship with a global body, could also serve to alleviate the risk of capture at national level. We also note that CSOs from the Global South have expressed an interest for the creation of national or regional SMCs, as these could have positive impact on civil society actors and contribute to enlarging the margins of freedom within which civil society operates. It is also noted that even in national contexts where freedom of expression faces increased restrictions, the actors to be represented in a multistakeholder 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, in a context where without an SMC these would be directly exposed to the same forces that impinge expression offline. In this regard, the national and regional SMCs would apply international standards on human rights, either directly or on the basis of a universal Code of Principles adopted at the global level.

We observe that numerous ongoing or forthcoming initiatives are seeking to establish self-regulatory mechanisms at the national level, often within a legal framework of ‘regulated self-regulation’, under the guise of bringing an end to the dissemination of often vaguely defined harmful content. There are differences between various iterations of ‘regulated self-regulation’, according to whether or not the overseeing public authority is properly kept at arm’s 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. However, we are concerned that these developments fail to properly protect freedom of expression. We suggest that in the context of current or projected legislative frameworks of ‘regulated self-regulation’, the Social Media Council offers a model that can deliver a form of ‘self-regulation under 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 ‘regulated self-regulation’ 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 perspective of international standards on freedom of expression, self-regulation and multistakeholderism are the least restrictive method to approach content moderation. Moreover, such a model fits well with the fast-changing contexts of social media companies, operating in complex and proprietorial technology environments. In order to be compatible with

international standards and deliver an effective solution to the challenges raised by the circulation of harmful content on social media platforms, such bodies need to respond to strict conditions of independence, transparency, accountability and participation.

Independence means that the Social Media Council, a voluntary mechanism created through contract by all relevant stakeholders, must be immune to undue influences from public, private or other special interests. Its funding structure must be preserved from capture and should ensure sustainability on a pluri-annual basis.

This is a mechanism that serves the general interest. As such, it needs to be accountable to the general public, and not to any particular interest. Accountability to the public translates into the creation of a website and open communication channels with the public, the publication of decisions, and the publication of annual reports (included audited accounts). In order to be credible and demonstrate both independence and accountability, the SMC has to be transparent: its internal rules, its processes, its procedures need to be available to the public.

A fundamental dimension of the creation of the SMC consists in taking an important question of general democratic interest - the moderation of online speech – to bring it back (where it belongs) in 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 considers that the SMC should have an adjudicatory function in order to be able to fulfil its role, to build a relationship of trust with the public and to build its own credibility. While there would be value in an advisory mechanism that would for instance offer guidance to platforms on how to adapt their ToS or Community Standards to bring them in compliance with international standards, the right of individual users to bring a complaint before the SMC will be decisive in creating trust and delivering evidence that the new mechanism can hold platforms accountable. However, we suggest that users' complaints should first be dealt with at platform level before they can be referred to the SMC, whilst in order to avoid the risk of overload, the SMC should have full control of its docket and be able to filter the cases it wants to review.

Provided that these conditions are met, the national SMCs will provide a solution to the online dissemination of harmful content while also ensuring the application of international standards on freedom of expression, privacy and other fundamental rights.

# APPENDIX: CONFERENCE AGENDA

## Social Media Councils: From Concept to Reality Stanford University

Friday, February 1, 2019

### **Purpose of the Working Meeting**

- Introduction to purpose of program: to discuss challenges associated with protection of free expression and content moderation on global social media platforms; engage in an assessment of ARTICLE 19's proposal for Social Media Councils.
- Our goal: to find a basis on which to build multistakeholder governance model(s), particularly with respect to hard questions related to illegal or harmful content on social media platforms.
- Considerations related to Protection of the International Human Right to Freedom of Opinion and Expression
- Discussion of UDHR/ICCPR ARTICLE 19 and commitment to use international law as normative framework for technology

### **Introduction to ARTICLE 19 Social Media Council concept**

High-Level Reactions to the Concept: What are the challenges?

### **Working Session I: Substantive Standards and Scope for Content Moderation**

Breakout Groups:

- What substantive standards and principles should SMCs be grounded in?
- How do/should universal human rights standards and platform community standards/TOS interact?
- Should substantive standards be the same for different types of platforms?
- What type of content should be covered by SMCs?
- Should rules about substantive standards be equally applied to cases of both deletion and demotion of content?

Report Back from Break-out Groups

## **Working Session II: Jurisdictional Scope and Challenges**

Breakout Groups:

- Should social media councils be local, national, regional, global?
- What are the implications of a national v. global SMC model?
- How should SMCs deal with the difference between democratic and authoritarian governments?
- How should SMCs resolve inconsistencies between national law and international human rights standards?
- How would the SMC decisions interact with national law and the judiciary?
- Should SMCs be platform-specific, applicable to all platforms, or applicable to platforms that meet a minimum threshold number of users in the relevant jurisdiction?

Report Back from Break-out Groups

## **Working Session III: Transparency, Accountability, Procedural Standards**

Breakout Groups:

- How will cases be referred to the SMCs? (E.g., appeals process, remanding by platforms)
- What procedural standards should SMCs be grounded in?
- Should decisions be illustrative, binding or advisory?
- What will be the preceding content moderation process at the platform level before a case is referred to an SMC?
- How will SMCs balance transparency and confidentiality in the decision-making and communication process?
- How will evidence be gathered?
- What would independence, transparency and accountability of the SMC mean in practice?
- What remedy will be available as an outcome of the SMC process?

Report Back from Break-out Groups

\*\*\*\*\*

Saturday, February 2, 2019

## **Insights from Day One and lead into Day Two**

Responses from Stakeholders

## **Working Session IV: Structure of SMCs**

Breakout Groups:

- How will social media councils be structured?

- What combination of stakeholders should sit on such an SMC?
- How should members be selected?
- What requirements should there be in terms of conflicts of interest and independence?
- How many members will sit on the council?
- How long will their mandate be?
- Who decides what the structure and mandate of SMCs will be?
- What will be the process by which the SMC's mandate will be determined?

Report Back from Break-out Groups

## **Working Session V: Risks and Opportunities SMCs**

Breakout Groups:

- How will we ensure that SMCs are not hijacked/coopted by malign actors?
- How will we ensure that SMCs do not stifle innovation?
- How would SMCs co-exist and interact with internal oversight within platforms?
- How will automated content moderation and SMC review interact?

Report Back from Break-out Groups

## **Developing a Feasible Multistakeholder Model for Social Media Councils**

Discussion: How do we work together? How can we ensure that a multistakeholder model of content regulation on global digital platforms is feasible?

- How can we make the SMC model ready for implementation, what is missing?
- How should we prepare a pilot program?
- What minimal conditions would need to be met in order to launch of a pilot?
- What country contexts are most promising for piloting the idea?

## **Closing: Next Steps**

## Endnotes

1. For more on multistakeholder approaches to protecting free expression online see: L. E. Strickling, J. F. Hill. (2018) "Multi-stakeholder Governance Innovations to Protect Free Expression, Diversity and Civility Online." In E. Donahoe, F. O. Hampson. Governance Innovation for a Connected World - Protecting Free Expression, Diversity and Civic Engagement in the Global Digital Ecosystem. Retrieved from <https://www.cigionline.org/publications/governance-innovation-connected-world-protecting-free-expression-diversity-and-civic-0>.
2. Milosavljevic, Marko and Broughton Micova, Sally. (2016). "Banning, Blocking and Boosting: Twitter's solo-regulation of expression". *Medijske Studije / Media Studies*. Vol. 7, No. 13, 43-58.
3. It is worth noting that even if erroneously removed content is reposted, depending on the type of content its removal in the first place may have had substantial impacts. This, however, is beyond the scope of SMCs to fix.



## About GDPi

The **Global Digital Policy Incubator (GDPi)** is a multistakeholder collaboration hub for the development of norms and policies to protect human rights, civic space and democratic processes in digital society.

We convene cross-regional, cross-disciplinary, multistakeholder programs to help government, private sector and civil society actors work together to develop norms and craft policies that protect human rights in the digital context.

We evaluate the human-rights impacts of digital technologies themselves, as well as the impacts of policy and regulatory responses to the technology, with particular emphasis on risks to free expression, privacy, security, and democratic engagement. We also explore practical and conceptual governance challenges related to content moderation on global digital platforms, and address threats to democratic processes and civic space from digital disinformation. We also advocate for use of the human rights framework in governance of artificial intelligence.

Our aim is to help policy-makers and other stakeholders effectively address the risks associated with digital technology in ways that are consistent with universal human rights so that society can capitalize on the benefits of digital technology.

## About the UN Special Rapporteur on freedom of opinion and expression

**David Kaye** was appointed UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression in August 2014. In 1993, the United Nations Commission on Human Rights established the mandate of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression. The Commission's successor body, the Human Rights Council, extended the mandate for another three years in its March 2008. The mandate has since been renewed three more times - in 2011, 2014, and 2017 - for a period of three years in each instance.

## About ARTICLE 19

**ARTICLE 19** works for a world where all people everywhere can freely express themselves and actively engage in public life without fear of discrimination.

We do this by working on two interlocking freedoms which set the foundation for all our work:

1. The Freedom to Speak concerns everyone's right to express and disseminate opinions, ideas and information through any means, as well as to disagree with, and question power-holders.
2. The Freedom to Know concerns the right to demand and receive information by power-holders for transparency, good governance and sustainable development.

When either of these freedoms comes under threat, either by the failure of power-holders to adequately protect them, ARTICLE 19 speaks with one voice, through courts of law, through global and regional organisations, and through civil society wherever we are present.