



Self-regulation and ‘hate speech’ on social media platforms

2018

ARTICLE 19

Free Word Centre
60 Farringdon Road
London,
EC1R 3GA
United Kingdom
T: +44 20 7324 2500
F: +44 20 7490 0566
E: info@article19.org
W: www.article19.org
Tw: [@article19org](https://twitter.com/article19org)
Fb: facebook.com/article19org

© ARTICLE 19, 2018



This report was produced with financial support from the Rights, Equality and Citizenship (REC) Programme of the European Union. The European Commission does not necessarily share the opinions expressed within the report. ARTICLE 19 bears the sole responsibility for the content.

The report was produced as part of “Media Against Hate”, a Europe-wide campaign initiated by the European Federation of Journalists and a coalition of civil society organisations, including the Media Diversity Institute (MDI), ARTICLE 19, the Croatian Journalists Association (CJA), Community Media Forum Europe (CMFE), Community Medien Institut (COMMIT) and Cooperazione per lo Sviluppo dei Paesi Emergenti (COSPE). For more information about the campaign see: <http://europeanjournalists.org/mediaagainsthate/>

This work is provided under the Creative Commons Attribution-Non-Commercial-ShareAlike 2.5 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: <http://creativecommons.org/licenses/by-nc-sa/2.5/legalcode>.

ARTICLE 19 would appreciate receiving a copy of any materials in which information from this report is used.

The Principles were developed as a part of the Civic Space Initiative financed by the Swedish International Development Cooperation, Sida. Sida does not necessarily share the opinions here within expressed. ARTICLE 19 bears the sole responsibility for the content of the document.

Table of Contents

Executive summary	4
Introduction	6
International human rights standards	8
Protection of fundamental rights	8
Freedom of expression online	8
What is self-regulation?	9
Approaches to media regulation	9
Self-regulation in practice	10
Problems with self-regulation	11
Effective self-regulation	11
Social media in context	13
Current application of self-regulation to social media platforms	15
Voluntary self-regulation by social media companies	15
Self-regulation initiatives adopted as a result of state intervention	17
Expanding existing regulatory and self-regulatory models to social media	17
Regulated self-regulatory agencies	18
Adoption of codes of conduct	19
Self-regulation of social media: the future	20
Conclusions and recommendations	23
Endnotes	25

Executive summary

In this brief, ARTICLE 19 seeks to contribute to discussions on greater regulation of social media platforms, including calls for such platforms to be considered publishers. We do so by exploring a possible model for the independent and effective self-regulation of social media platforms.

ARTICLE 19 recognises that dominant social media companies hold considerable power over the flow of information and ideas online, given the vast quantities of content published on their platforms. The way in which social media companies have dealt with content issues on their platforms, especially around ‘hate speech’, has been of particular concern to many stakeholders. Under international standards on freedom of expression, however, it is a fairly complex task to decide whether a specific message can be identified as unlawful ‘hate speech’, and, as such, whether it should or could legitimately be prohibited. More generally, any restriction on freedom of expression, whatever the objective it seeks to achieve, necessarily raises a series of legal questions.

Traditionally, international freedom of expression standards allow for different types of media regulation, based on the type of media. A model of self-regulation has been the preferred approach to print media. It is considered the least restrictive means available by which the press can be effectively regulated, and the best system for promoting ethical standards in the media. Importantly, the existence of an effective self-regulation mechanism can also reduce pressure on the courts and the judiciary. Generally, when a problem is effectively managed through self-regulation, the need for state regulation is eliminated.

A number of recent legislative initiatives on ‘hate speech’, including most prominently the 2017 German NetzDG law on social media, make reference to some forms of self-regulation. Voluntary mechanisms between digital companies and various public bodies addressing ‘hate speech’ and other issues, such as the EU Code of Conduct on hate speech, also make reference to self-regulatory models. However, our analysis of these mechanisms demonstrates that they fail to comply with the requirements of international human rights law. They rely on vague and overbroad terms to identify unlawful content, they delegate censorship responsibilities to social media companies with no real consideration of the lawfulness of content, and they fail to provide due process guarantees.

ARTICLE 19 therefore suggests exploring a new model of effective self-regulation for social media. This model could take the form of a dedicated “social media council” – inspired by the effective self-regulation models created to support and promote journalistic ethics and high standards in print media. We believe that effective self-regulation could offer an appropriate framework through which to address the current

problems with content moderation by social media companies, including ‘hate speech’ on their platforms, providing it also meets certain conditions of independence, openness to civil society participation, accountability and effectiveness. Such a model could also allow for adoption of tailored remedies without the threat of heavy legal sanctions.

In this brief we explore the issues that would need to be taken into account when considering such a mechanism, in greater detail. The brief also reinforces ARTICLE 19 recommendations on freedom of expression-compliant approaches to intermediary liability models.

In the first section, ARTICLE 19 outlines relevant international standards in this area. In the second section, we look at existing models of media regulation and in particular, models of self-regulation. In order to respond to calls of whether social media companies can be considered publishers, we also explore the role played by social media platforms as compared to traditional or legacy media. We then briefly examine the current problems related to how social media companies deal with content on their platforms. Finally, we provide some suggestions as to what a new model for effective self-regulation of social media companies could look like. This model could be applied to ‘hate speech’ appearing on social media platforms, but could also provide an appropriate forum to elaborate ethical standards for social media platforms in general, provide remedies, and identify ways to enable exercise of the right to freedom of expression on social media platforms.

ARTICLE 19 acknowledges that the realisation of this self-regulation model may raise practical challenges and problems. However, we believe that these problems, albeit complex, should be further debated and explored. In today’s digital societies, there is a pressing social need to establish appropriate forums of this kind. We believe it is important to collectively engage more broadly on these issues, which are of major importance to the protection of the right to freedom of expression and human rights in digital environments.

Introduction

In recent years, there has been an increase in calls globally for greater regulation of social media platforms, including in Europe. In this respect, the way in which social media companies address 'hate speech' on their platforms is a particularly burning issue for governments, policy makers, regulatory bodies, self-regulatory institutions, media, civil society and, the public at large.

'Hate speech' is an emotive concept which has no universally accepted definition in international human rights law. Many would claim they can identify it where they see it, however its characteristics are often elusive or contradictory. ARTICLE 19 has issued a number of policy documents detailing the complexities of international law in this area.¹ Our analysis has shown that making a decision as to whether a specific message can be identified as 'unlawful hate speech' and, as such, should or could legitimately be prohibited, is a fairly complex task. It requires an analysis of all the circumstances of a given instance of 'hate speech', which, in turn, demands a deeper understanding of the local and regional context in its political, economic, social and cultural dimensions, as well as a strong knowledge of international human rights standards and relevant case-law.

The ongoing legal and policy debates related to content moderation by social media companies on their platforms are hindered by the difficulty in appropriately classifying the role played by these companies in the modern media landscape. While it is clear that tech giants have a genuine capacity to influence public debates online,² identifying the actual impact they have on their audiences or determining an appropriate response to their perceived power remains a complex undertaking. Existing institutions, and legal and regulatory frameworks have been elaborated to deal with traditional media actors, and, as such, their application to new actors is at the very least uneasy. At the same time, in a number of countries the regime of limited liability, which has allowed Internet intermediaries (including social media platforms) to act as efficient enablers of freedom of expression, has been put under increasing pressure.³

These debates are nonetheless of vital importance to democratic societies. It is also important to ensure that any measures related to content regulation on social media platforms, including of 'hate speech', are fully compatible with international freedom of expression standards.

References to forms of self-regulation as the appropriate approach to deal with content moderation on social media are abundant in recent initiatives in this area, both in the EU and in many States outside of it. However, as discussed in this brief, the proposed mechanisms do not necessarily offer sufficient guarantees for either the independence or effectiveness of self-regulation, or for the protection of freedom of expression.

In this brief, ARTICLE 19 explores possible approaches for effective self-regulation for social media. This proposed model could be applied to ‘hate speech’ appearing on social media platforms, but could also provide an appropriate forum to elaborate ethical standards for social media platforms in general, provide remedies, and identify ways to enable exercise of the right to freedom of expression on social media platforms.

ARTICLE 19 acknowledges that the realisation of this self-regulation model may raise certain practical challenges and problems. However, we believe that these problems, albeit complex, should be further debated and explored. In today’s digital societies, there is a pressing social need to establish appropriate forums of this kind. We believe it is important to collectively engage on these issues, which are of major importance to the protection of the right to freedom of expression and human rights in digital environment more broadly.

This brief should be read in conjunction with ARTICLE 19’s digital policy series, in particular the policy on intermediary liability, the policy on blocking and filtering⁵ and the policy on freedom of expression and terms and conditions (terms of service) of social media platforms.⁶

International human rights standards

Protection of fundamental rights

International and regional human rights standards provide for protection of the right to freedom of expression⁷ and the right to equality.⁸ These protections must form the backbone of any State responses to 'hate speech', intolerance and, discrimination.

The right to freedom of expression is not absolute. States may, exceptionally, limit the right, but any limitations must conform to the strict requirements of the three-part test. Namely, they must demonstrate that the limitation is:

- **Provided for by law**, so any law or regulation must be formulated with sufficient precision to enable individuals to regulate their conduct accordingly;
- **In pursuit of a legitimate aim**, listed exhaustively as: respect of the rights or reputations of others; or the protection of national security or of public order (*ordre public*), or of public health or morals; and
- **Necessary in a democratic society**, requiring the State to demonstrate in a specific and individualised fashion the precise nature of the threat to a legitimate aim, and the necessity and proportionality of the specific action taken, in particular by establishing a direct and immediate connection between the expression and the threat identified.⁹ When restricting the right to freedom of expression, the least restrictive measure capable of achieving a given legitimate objective should be imposed. This is a particularly important issue in regards to media regulation.

Additionally, States are *required* to prohibit by law “any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence.”¹⁰

States have further, positive obligations to ensure the effective protection of human rights, including in the sphere of relationships between private parties. In other words, States may have to regulate the behaviour of private actors where it is necessary to guarantee the effective exercise of the rights of individuals to freely receive and impart information and ideas.

Freedom of expression online

At the international level, international standards recognise that the “same rights that people have offline must also be protected online”;¹¹ and that “any limitations

on electronic forms of communication or expression disseminated over the Internet must be justified according to the same criteria as non-electronic or offline communications.”¹² International standards also recognise that regulatory approaches in the telecommunications and broadcasting sectors cannot simply be transferred to the Internet,¹³ rather, “the legal framework regulating the mass media should take into account the differences between the print and broadcast media and the internet, while also noting the ways in which media converge.”¹⁴

It has been also recommended that tailored approaches to addressing illegal online content should be adopted;¹⁵ and that self-regulation can be an effective tool in redressing harmful speech, and should, thus, be promoted.¹⁶

What is self-regulation?

Approaches to media regulation

There are different types of media regulatory systems around the world, which reflect the cultural, social, and political traditions of individual States. In general:

- The **statutory regulation** model applies to any measure passed by parliaments to regulate the media, and is characterised by stronger state interference;
- The **co-regulation** model, sometimes also called “regulated self-regulation,” varies. However, it typically contains elements of a self-regulatory mechanism underpinned by legislation;
- The **self-regulation model** is a framework that relies entirely on voluntary compliance: legislation plays no role in enforcing the relevant standards. Its *raison d’être* is holding its members accountable to the public, promoting knowledge within its membership and developing and respecting ethical standards. Those organisations that commit to this type of regulation do so not under threat of legal sanction, but for positive reasons such as the desire to further the development and credibility of their profession and sector. Self-regulation models rely first and foremost on members’ common understanding of the values and ethics that underpin their professional conduct.

Statutory or co-regulation models have traditionally been deemed necessary for the broadcast media, wherein the allocation of a scarce natural resource (spectrum) requires the intervention of public authorities in order to create a diverse and pluralistic broadcasting landscape. Self-regulation has been considered the preferred approach for

print media; press councils are the typical example of such mechanisms. Self-regulation is considered to be the least restrictive means available through which the press can be effectively regulated and the best system through which high standards in the media can be promoted.¹⁷

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

Self-regulation in practice

Self-regulation of the press typically means some form of national or regional press council, complaints commission or ombudsperson (either acting alone or in conjunction with a press council). Press councils may be funded by the publishing industry alone, by journalists alone or by a combination of both, and sometimes with government assistance (for example, financial assistance).

Press councils publish their codes of conduct with the approval of journalistic and media organisations. Crucially, the press outlets of the country that are members of the press council must commit themselves to these codes of conduct. Sometimes broadcasting organisations do so as well.

Press councils accept complaints from any member of the public who believes that a published article infringes the respective code of conduct. The members of the press council (or, in some cases a complaints committee of the press council) will then adjudicate on complaints received, publish their conclusions and, in some cases, order the publication of their decision or impose a right of reply on the offending outlet. In very few cases, for particularly serious breaches of the code of conduct, press councils can impose financial penalties.

Some press councils comprise only of representatives of its member media organisations, while others give representation to the wider community and include a balanced representation of publishers, journalists and the public.

Many press councils see the task of hearing complaints as part of a wider responsibility to defend media freedom. These bodies often publish an annual review discussing media concerns and sometimes advocate for legislative changes related to the media. Others see their role solely as a complaints body.

Problems with self-regulation

Self-regulation mechanisms are not without problems in ensuring an effective and respected system of media accountability.

Self-regulatory bodies have sometimes been described as “elitist circle of enforcers”;²¹ and some of them find it difficult to gain public trust.²² Councils whose membership is composed only of media representatives are often open to charges of hypocrisy, as they may be reluctant to challenge the decisions of either their own colleagues or the media proprietors. Media proprietors and editors can operate as a powerful lobby, and are often seen as manoeuvring for position and influence, with a vested interest in the way in which press councils operate. In some countries, self-regulation has proven to be ineffective in preventing abuses by the press and prompted outpourings of public criticism.²³ Furthermore, in countries where the state has traditionally played a more intrusive role, and where there is stricter content regulation through legislation, there is likely to be less appetite for a system of self-regulation.

Effective self-regulation

ARTICLE 19 has maintained that self-regulation of print media should always be seen as the preferred model of press regulation. At the same time, we recognise that self-regulation must be meaningful and effective. It must not only provide protection for members of the profession (i.e. journalists) but also hold them accountable to their profession, and ensure that press outlets are held accountable to the public. Moreover, self-regulation can only prosper alongside a legal framework which provides strong guarantees for the fundamental right to freedom of expression and freedom of information.

ARTICLE 19 has previously identified several requirements for effective self-regulation of the media.²⁴ We submit that sector-wide effective self-regulatory bodies should:

- Be independent from government, commercial and special interests;
- Be established via a fully consultative and inclusive process – the major constitutive elements of their work should be elaborated in an open, transparent, and participatory manner that allows for broad public consultation;
- Be democratic and transparent in their selection of members and decision-making;
- Ensure broad representation.²⁵ It is important that the independence of self-regulatory bodies is ensured, with a composition that includes representatives of civil society;

-
- Adopt a code of ethics for the profession or sector it seeks to regulate;
 - Have a robust complaints mechanism and clear procedural rules to determine if ethical standards were breached in individual cases, and have the power to impose only moral sanctions;²⁶ and
 - Work in the service of the public interest, be transparent and accountable to the public.

In addition, self-regulatory bodies can play an important role in promoting knowledge and understanding of ethical rules throughout the sector or profession. They can do so by adopting and disseminating recommendations and guidelines, or offering trainings to their members.

A limited degree of state support can be useful in supporting the creation of effective self-regulatory mechanisms, provided that state intervention is limited to creating a legal underpinning for self-regulation and does not threaten the independence of the self-regulatory bodies.²⁷ By contrast, situations whereby public authorities pressure private companies to define and regulate speech, under the guise of self-regulation or co-regulation, are seriously at odds with international standards on freedom of expression.²⁸

Social media in context

In current debates on regulation of content on social media, frequent calls are made to treat social media platforms as media outlets or to regulate them in the same way as publishers.²⁹

ARTICLE 19 urges that such proposals be approached with extreme caution; they must take compliance with international freedom of expression standards into consideration. As noted above, international law has traditionally differentiated between the permissible approaches to regulation of broadcast media and print media. It is important to recognise that social media platforms fulfil different functions than the traditional or legacy media (broadcasting and print). The *raison d'être* of the traditional/legacy media companies or publishers is to produce content, be it news or entertainment. The role of social media platforms is distinct: in general they perform one or more of the following functions:

- **Hosting** – social media platforms provide space for individuals and (media) companies to make their content available to the public, storing content uploaded by a third party. Hosting providers offer an infrastructure that enables individuals and companies to exercise their fundamental right to freedom of expression, which is of great value in a democratic society. As they are not involved in modifying the content in question, they should in principle be immune from liability for third party content;
- **Online distribution** encompasses all activities that involve making the content of third parties publicly visible, findable or accessible to the general public or to individual users on a given online platform. Online distribution generally relies, although not necessarily exclusively, on automated decision-making processes ('algorithms') for the selection of media content that is pushed towards users. Although users uploading, liking or sharing content influences the visibility of what appears on their own personal feeds, accountability for online distribution lies with the platforms. Relevant activities in this category would also include scenarios whereby platforms impose particular requirements on third-party editors.³⁰ Editors of media outlets who wish to publish content on a particular social media platform (e.g. Facebook, YouTube or Twitter) will often have to conform to specific requirements in terms of formats, types or length of media content. Therefore, a story edited for publication in a newspaper or on the website of the outlet will need to be reformatted to be published on a social media platform. As a result, certain media outlets now have teams dedicated to working on each of the primary platforms;

-
- Only in *very* limited circumstances do some social media platforms fulfil the **role of traditional or legacy media outlets**. This is when they either edit or commission content over which they have editorial control. For example, in 2016 Twitter bought the rights to broadcast NFL official content through its platform.³¹ It has also been reported that Facebook would start purchasing television series to be distributed on its own streaming platform.³² Outside of such clearly defined cases, which still remain exceptional, the activities of social media platforms differ from those of traditional or legacy media outlets or “publishers”.

Separately, it has been argued that social media platforms tend to lock individuals into so-called filter bubbles or echo chambers, where people encounter only information and ideas that they already view favourably. This can create problems in terms of pluralism and diversity of information. While filter bubbles are not unique to social media, there is a limited understanding on their mechanics and potential effects on society.³³ Indeed, the genuine impact of social media platforms on society is uncertain. While they are frequently decried as vehicles for the circulation of ‘hate speech’ or misinformation (‘fake news’), there have also been cases wherein media organisations have only investigated and covered a story of public interest after its emergence from viral dissemination on social media.

To be clear, there is no doubt that there is a need to find appropriate ways to deal with the power that a limited number of digital technology and social media companies have over the online flow of information and ideas. Some of these companies are in a dominant, even quasi-monopolistic, position, from which they are potentially able to influence the flow of information and ideas, which in turn informs crucial public debates. The influence of these dominant companies within societies has grown and, intentionally or not, they are as such in a position to influence the public agenda, trends in public opinion, and the topics occupying public debates.³⁴

At the same time, ARTICLE 19 submits that the impact of dominant social media companies, and digital technology companies in general, on civic space and media landscapes is both complex and uncertain. These companies operate in a fast-paced, rapidly changing environment. Solutions to the new problems they potentially pose should, therefore, be cautious, adaptable, and innovative, while fully complying with international freedom of expression standards. We suggest that exploring and adapting existing models of effective self-regulation to social media platforms, based on ethical standards and effectively adapted to the online environment, can offer such a solution.

Current application of self-regulation to social media platforms

Voluntary self-regulation by social media companies

In order to access the services provided by social media companies, its users have to accept the respective Terms and Conditions or Terms of Service and Community Guidelines that define what content is acceptable on the platform. These contractual arrangements can generally be modified at will by social media companies. Social media platforms routinely remove content from their platforms on the basis of these Terms of Service or Community – either of their own initiative, in response to external pressures, or based on the complaints of their users. As such, it can be argued that these companies are already self-regulating in a certain way. However, the existing self-regulation mechanisms operated by dominant social media platforms do not meet the conditions for *effective* self-regulation, as outlined above.

ARTICLE 19 finds that in general, content moderation and removal policies by social media companies are problematic for several reasons, including:

- **Lack of respect for human rights standards:** Although social media companies, as private businesses, are not directly bound by international human rights law, they are increasingly encouraged to implement international standards on freedom of expression in all their practices related to content moderation.³⁵ Available information shows that the internal content moderation policies of some companies address complex and varied factors and may include a certain degree of consideration for freedom of expression and other fundamental rights.³⁶ It is, however, also clear that certain decisions to suppress content are in violation of freedom of expression standards;
- **Lack of legal certainty:** The removal of content on social media platforms is entirely unilateral and the process does not respect the requirements of due process of law. ARTICLE 19 has previously recommended that, as a matter of principle, social media companies and all hosting service providers should only be required to remove content following an order issued by an independent and impartial court or other adjudicatory body that has determined that the material at issue is unlawful.³⁷ This is because the courts apply laws that have been democratically adopted, under all the guarantees of due process of law. Through modalities that vary from country to country, they are also bound to apply human rights standards,

on the basis of national, regional, and international law, to the cases they preside over. This provides a much greater degree of legal certainty. We recognise, however, that it may be too burdensome and costly for the courts to examine all applications for content removal, given the high volume of such requests. However, at a minimum, procedures set up by social media companies should fully comply with certain basic due process requirements;

- **Lack of accountability or transparency over decision-making:** The majority of the decision-making processes and practices of social media platforms underpinning their content policies– including the use of automated decision-making processes – remain opaque. Existing procedures do not ensure sufficient accountability for those decisions. Although some progress has been made with regards to transparency reporting over the years,³⁸ there is still too little information available about the way in which social media platforms apply their Terms of Service in various circumstances. It has been acknowledged that this lack of transparency with regards to their decision-making processes can obscure discriminatory practices or political pressure affecting the companies' practices;³⁹
- **Lack of consistency in stakeholder engagement:** Many of the initiatives undertaken by social media platforms in this area need more meaningful participation of civil society organisations and other stakeholders. It has been repeatedly suggested that practices and decision-making regarding content on dominant social media platforms is in the public interest, and it has been highlighted that “the shaping of those policies might be more effective if done through collective knowledge and debate.”⁴⁰ Experience has shown that dominant social media platforms can be sensitive to public outrage and, when faced with large-scale protests, are willing to reconsider their decisions to remove specific content. This has generally occurred in cases where legitimate content has been removed on the basis of community guidelines or Terms of Service, in particular content depicting nudity⁴¹ or violence,⁴² and the removal of the content has generated a significant public backlash. However, there are many instances whereby lawful content, removed by these platforms, is not defended by a massive public mobilisation and is not supported by government representatives or celebrities. The lack of consistency by social media platforms is additionally increasingly being driven by pressures exerted by advertisers, who wish to avoid their brand's image being tarnished through association with certain types of content.⁴³

On a positive note, in response to public concerns the dominant social media platforms are increasingly engaged in a number of initiatives in order to address the problems raised by certain content on their platforms. This includes initiatives to combat ‘hate speech’, ‘terrorism’, identifying and flagging misinformation, and to provide financial support to media companies adapting to the digital context. Representatives of the social media companies appear to be increasingly engaged in debates with public

authorities, academia and civil society, in national and international forums. The publication of yearly transparency reports by digital technology companies is another positive step towards ensuring greater accountability. Some companies have also engaged in multi-stakeholder initiatives, such as the Global Network Initiative, that encourage companies to undertake human rights impact assessments of their decisions and to produce transparency reports when faced with requests to implement decisions that may undermine the rights to freedom of expression and privacy.⁴⁴

ARTICLE 19 appreciates the willingness of social media platforms to engage with public authorities and other stakeholders, and we acknowledge that some of the internal practices by these companies may contribute to the respect and promotion of freedom of expression when they are in full conformity with international standards. At the same time, we consider that internal policies practices of the dominant social media companies or their partners are, at the very least, insufficiently transparent to be fully credible and reliable, and should be reformed.

Self-regulation initiatives adopted as a result of state intervention

Social media platforms are facing increasing pressure from States to remove content with little or no consideration for the protection of freedom of expression. ARTICLE 19 has identified several trends in this respect, as follows.

Expanding existing regulatory and self-regulatory models to social media

There is a growing tendency for States and intergovernmental bodies to challenge existing models of intermediaries' immunity from liability and to put forward proposals that change the existing safe harbour and broad immunity models.⁴⁵ In some countries and regions, there have been attempts to extend the remit of existing regulatory and self-regulatory mechanisms for traditional or legacy media to dominant social media platforms. In particular, this reflects the situation in the EU. For example:

- Media regulatory authorities in the EU could have the power to extend their jurisdiction to the video sections of the websites of print media.⁴⁶ The rules of the audiovisual media services directive (AVMS Directive) could become applicable to these parts of websites and print media would be subject to the intervention of a regulatory authority for broadcast media (even if only for their online audiovisual content), which would mark a radical divergence from the usual practices in this sector.
- The currently ongoing revision of the AVMS Directive includes provisions that would require "video-sharing platforms" – a loose concept, the definition of which

could encompass social media – to adopt “appropriate measures” to remove ‘hate speech’ and content deemed harmful to minors.⁴⁷ The national regulatory authorities would be responsible for deciding whether the measures adopted by video-sharing platforms and social media platforms were appropriate.

ARTICLE 19 finds both of the proposals from the EU to be incompatible with freedom of expression standards and incompatible with the limited or conditional liability regime set out under the EU E-commerce directive⁴⁸ The existing regulatory authorities for audiovisual media apply a legal framework that is relevant for audiovisual media, not for the specific and new issues raised by new forms of online distribution. Existing press councils are, similarly, guardians of journalistic ethics, whereas the questions raised by content regulation practices by social media platforms require the adoption of new ethical standards appropriately tailored to the specific circumstances of social media.

Regulated self-regulatory agencies

There have been some initiatives to establish regulated self-regulatory agencies for social media.

This mechanism has been adopted in Germany, under the 2017 Network Enforcement Act (NetzDG). The NetzDG threatens social networks with a fine of up to 50 million EUR if they do not remove “clearly illegal” content within 24 hours of a complaint (or a week when it is not clear that the content is illegal).⁴⁹

The NetzDG also provides for the recognition, by the Ministry of Justice, of “regulated self-regulatory agencies.” The role of such agencies, which would be financed by social media companies, will be to determine whether a given piece of content is in violation of the law and should be removed from the platform. Recognition by the Ministry of Justice is contingent on conditions such as the independence of the self-regulatory agency, the expertise of the agency staff who would act as decision-makers, and the agency’s capacity to reach a decision within seven days. In theory, this mechanism might be a step towards establishing some sort of independent self-regulation for social media under the statute. This might seem like a step towards establishing a self-regulatory mechanism for social media platforms, underpinned by legal statute. ARTICLE 19, however, finds that the guarantees provided for in the NetzDG are insufficient to ensure the independence and effectiveness required for an effective model of self-regulation and for the protection of freedom of expression.⁵⁰

Adoption of codes of conduct

In some instances, specific codes of conduct have been adopted jointly by companies and public institutions. An example of this model is the Code of Conduct on Countering Illegal Hate Speech Online (the Code of Conduct), developed by the European Commission in collaboration with several major digital technology companies (Facebook, Microsoft, Twitter and YouTube).⁵¹

Under the Code of Conduct, when these companies receive a request to remove content from their online platforms, they are required to assess the removal request against their Terms of Service and community guidelines, and, where applicable, national laws on combating racism and xenophobia, which transpose EU law on combating racism and xenophobia. The companies have committed to reviewing the majority of these requests in fewer than 24 hours and to removing the content if necessary. Further, the Code of Conduct looks to strengthen notification processes between the companies and law enforcement authorities by channelling communications between them through national contact points on both sides. The role of civil society organisations (CSOs) as “trusted reporters” of “illegal hate speech” is also highlighted, with the European Commission and Member States helping to ensure access to a representative network of CSO partners and “trusted reporters.”

The Code of Conduct contains further commitments from companies to educate their users about the types of content that are not permitted under their rules and community guidelines, to share best practices between themselves and other social media platforms, and to continue working with the European Commission and CSOs on developing counter-narratives and counter-hate speech campaigns. While the Code of Conduct does not put in place any mechanism to monitor signatories’ compliance – and indeed the Code of Conduct is not binding or otherwise enforceable – the companies and the European Commission have agreed to assess the public commitments contained in the Code on a regular basis. In addition, the European Commission, in coordination with Member States, has committed to promote adherence to the commitments set out in the Code to other relevant platforms and social media companies.

Despite some positive features, ARTICLE 19 considers the Code of Conduct to be problematic for a number of reasons.⁵² It resembles the co-regulatory model of codes of conduct (or broadcasting codes) typically applied to regulate the broadcast media. The Code of Conduct encourages the removal of “illegal hate speech” and the ‘tweaking’ of companies’ Terms of Service by reference to the EU’s Framework Decision. It contains only weak references to the protection of freedom of expression. It is also problematic from a due process perspective. Finally, despite several references to consultations with civil society organisations, no CSOs working to defend freedom of expression were seemingly involved in the drafting of the Code of Conduct.

Self-regulation of social media: the future

ARTICLE 19 believes that considering whether social media platforms are publishers or not, is not the appropriate starting point from which to approach discussions as to their regulation. Instead, it should be acknowledged that social media platforms are “a kind of hybrid beast that does not fit into any of the traditional categories”⁵⁴ and that the situation is both relatively new and still in flux.

At the same time, ARTICLE 19 recognises that current approaches need to be reconsidered. At present, the possibilities presented are:

- The continuation of the current state of affairs, whereby social media platforms self-regulate, and steps are taken in order that problems with their current practices, such as opacity, credibility and accountability are effectively addressed; or
- States continue to adopt (often hastily written) legislation that pressures social media companies to swiftly remove content, often with little or no consideration for freedom of expression.

Both of these possibilities are problematic from a human rights perspective. ARTICLE 19 therefore suggests exploring the possibility of establishing new models of self-regulation for social media, inspired by effective self-regulation models created to support and promote journalistic ethics. With some adjustments, such models could be explored for a variety of issues.

The most ambitious task in this respect would be the creation of an independent self-regulatory body for social media (Social Media Council); it could be created at national level or international level or a combination of both. It would deal with content moderation issues and would be adequately funded by social media companies and relevant stakeholders. The Council could elaborate ethical standards specific to the online distribution of content and cover topics such as terms and conditions, community guidelines and the content regulation practices of social media companies. By making the work transparent to the general public, and through appropriate consultative processes, this mechanism could provide a public forum for important public discussions on the regulation of online content distribution. Through light sanctions, and mainly relying on transparency, peer and public pressure, this body could monitor and promote respect of appropriate ethical standards by social media companies. Transparency and openness, combined with independence, could give this mechanism the needed credibility to gain public trust.

To initiate discussions on this new self-regulatory model for social media platforms, ARTICLE 19 suggests that the following issues should be considered when exploring this mechanism:

- **Remit:** A Social Media Council (Council) could either be tasked with dealing with a specific issue (such as 'hate speech') or be given general jurisdiction over content issues on the social media platforms that are members of the Council;
- **Scope:** It could be created on a national level to ensure a sufficient level of proximity and understanding of the relevant community and context, or on an international level or a combination of both;
- **Independence:** The Council would have 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. In order to avoid an excessive number of representatives, its composition could vary according to areas of intervention;
- **Commitments:** Social media platforms would have to commit to providing an appropriate level of information on their internal content moderation practices to the Council of which they are a member. They would also have to commit to accepting the decisions of their Council as binding;
- **Charter of ethics/Code of conduct:** As a fundamental part of its remit, Councils would have to adopt a Charter of Ethics for social media. This document would have to be adopted through a transparent and open process, including broad consultations with all relevant stakeholders, including civil society organisations. At a minimum, a Charter of Ethics would include a commitment to comply with international human rights standards, including on freedom of expression and due process;
- **Decision-making:** The Council could adopt recommendations, – either of their own initiative or at the request of its members – to further clarify the interpretation and application of ethical standards in given areas. Such recommendations would have to be adopted through a transparent process, open to participation from all relevant stakeholders and civil society. For instance, Councils could adopt a recommendation on how to include robust notice and counter-notice procedures in social media platforms' terms and conditions;
- **Complaints procedures:** The Council could be empowered to receive complaints from individual users, provided that all possibilities of remedying the issue with the social media company (either through ombudspersons or other flagging procedures) have already been exhausted. The Council would hold a hearing and reach a decision, including the possibility of a sanction that seeks to promote rather than restrict speech (such as a right of reply, an apology or the publication of its decision);

-
- **Other functions:** The Council could also be tasked with providing advice on ethical standards to social media platforms' own ombudspersons, staff, and departments in charge of content regulation;
 - **Funding:** The Council would have to benefit from 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 its income on a multi-annual basis, while additional resources could be provided by other stakeholders or philanthropic organisations; and
 - **Accountability:** The Council would have to ensure its accountability to the public. In particular, it would have to make its work and decisions readily available to the public – including, of course, through social media.

As with the establishment of any new system, the creation of a self-regulatory mechanism for social media platforms raises a number of difficult questions. Additionally, as the experience of establishing press councils shows, it can be a lengthy and complicated process as all relevant stakeholders need to agree on a system that they can all make their own. ARTICLE 19 believes that any new system can only come into existence and prove its effectiveness if all participants are willing to make it work. By shifting the focus towards the process of developing a new mechanism, rather than trying to impose a solution, a self-regulatory mechanism could allow for the adoption of tailored and adaptable remedies unhindered by the threat of heavy legal sanctions.

Developing a new system of independent self-regulation could provide a solid frame of reference through which to assess the initiatives undertaken so far by the dominant social media companies and their partners. In this way, it would prompt questions to be asked about these existing initiatives, such as whether they include all the relevant stakeholders; whether they are purely internal mechanism or if they benefit from a form of external, independent review of decisions; whether they are accountable to the public; whether they work in the public interest; and whether they are captured by private or special interests.

Conclusions and recommendations

In conclusion, ARTICLE 19 reiterates that it is important to acknowledge the role of dominant social media companies in the digital environment. These companies now play a major role in the way in which a growing proportion of the global population interacts with the world. They are an important conduit and source of news for many people. For media outlets that are undergoing complex digital transformations, social media platforms are the place to be in order to reach their audience. Various studies conclusively show that the dominant digital technology companies have a significant role in the public sphere.⁵⁵ The right of individuals to freedom of expression online is, therefore, to a large extent, determined and influenced by the actions of social media platforms.

In this context, ARTICLE 19 therefore believes that, with reference to the positive duty of States to intervene and take active steps to remove obstacles to the exercise of the right to freedom expression, the debate needs to move beyond companies' freedom to conduct their business. Under international law, States have positive obligations to ensure the effective protection of human rights, including in the sphere of relationships between private parties, and to foster an enabling environment for freedom of expression. In other words, States may have to regulate the behaviour of private actors where it is necessary to guarantee the effectiveness of the rights of individuals to freely receive and impart information and ideas.

ARTICLE 19 therefore calls on **States** to:

- Reaffirm and further clarify the principles of limited liability for the hosting of third-party content, in order to preserve the positive impact that social media and other intermediaries have upon the free flow of information and ideas online. ARTICLE 19 has previously produced a detailed proposal on freedom of expression compliant models of intermediary liability and we encourage States to adopt our recommendations;
- Refrain from adopting legislation on content regulation, including on 'hate speech,' that does not comply with the requirements of international standards on freedom of expression;

ARTICLE 19 also calls on **social media platforms** to, at a minimum:

- Include international standards on freedom of expression and due process in their terms and conditions and community guidelines;

-
- Provide for greater clarity and transparency regarding decision-making processes on content removals on their platforms;
 - Ensure that all content related policies and practices, including those on ‘hate speech,’ are elaborated and implemented through transparent processes, which are open to meaningful participation from civil society organisations and all relevant stakeholders;
 - Together with civil society organisations, media actors, academics and other relevant stakeholders, explore possibilities for creation of new mechanisms of independent self-regulation and the development of a charter of ethics for social media (following the suggestions outlined in the previous section). This must be done through a process that is transparent and open to stakeholders’ participation, including civil society.

Endnotes

¹ See, e.g., ARTICLE 19, Hate Speech explained: a toolkit, 2015 Edition; The Camden Principles on Freedom of Expression and Equality (Camden Principles), 2009; Prohibiting Incitement to Discrimination, Hostility or Violence, 2012; or Should ‘hate speech’ be free?, 29 August 2017.

² See, e.g., M. Moore, Tech Giants and Civic Power, April 2016.

³ See ARTICLE 19, New EU legislation must not throttle online flows of information and ideas, 12 September 2017.

⁴ ARTICLE 19, Internet intermediaries: Dilemma of liability, 2013.

⁵ ARTICLE 19, Freedom of Expression Unfiltered: How blocking and filtering affect free speech, 2016.

⁶ Forthcoming, June 2018.

⁷ The right to freedom of expression is guaranteed in Article 19 of the International Covenant on Civil and Political Rights (ICCPR) and Article 10 of the Convention for the Protection of Human Rights and Fundamental Freedoms (European Convention) and in other regional treaties. Within the EU, the right to freedom of expression and information is guaranteed in Article 11 of the Charter of Fundamental Rights of the European Union.

⁸ Articles 2(1), 26 and 27 of the ICCPR provide for equality in the enjoyment of human rights and the equal protection of the law. An expansive range of “protected characteristics” are recognised under international human rights law, including race, ethnicity, national or social origin, religion or belief, disability, migrant or refugee status, sex, sexual orientation, and gender identity. At the European level, the European Convention prohibits discrimination in Article 14 and more broadly in Protocol No. 12.

⁹ See, Human Rights Committee, General Comment No. 34 on Article 19: Freedoms of opinion and expression, CCPR/C/GC/34, 12 September 2011, para 22. See, also jurisprudence of the European Court of Human Rights on (European Court) permissible limitations set out in Article 10(2) of the European Convention.

¹⁰ *C.f.* Article 20(2) of the ICCPR. At the European level, the European Convention does not contain any *obligation* on States to prohibit any form of expression, as under Article 20(2) of the ICCPR. However, the European Court has recognised that certain forms of harmful expression must necessarily be restricted to uphold the objectives of the European Convention as a whole; see, e.g. European Court, *Erbakan v. Turkey*, App. No. 59405/00 (2006), para 56; or *Gündüz v. Turkey*, App. No. 35071/97 (2004), para 22. The obligations under Article 20(2) of the ICCPR is unpacked in the Rabat Plan of Action on the prohibition of advocacy

of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence (Rabat Plan), A/HRC/22/17/Add.4, Appendix, 5 October 2012..

¹¹ HRC Resolution 20/8 on the Internet and Human Rights, A/HRC/RES/20/8, June 2012.

¹² General Comment No. 34, *op cit.*, para 43.

¹³ The 2011 Joint Declaration on Freedom of Expression and the Internet, June 2011.

¹⁴ General Comment No. 34, *op.cit.*, para. 39.

¹⁵ 2011 Joint Declaration, *op.cit.*

¹⁶ *Ibid.* See also UN Special Rapporteur on Freedom of Expression, A/66/290, 10 August 2011.

¹⁷ In practice, there is no uniform definition of “self-regulation,” however. Models labelled “self-regulation” in one country may qualify as “co-regulation” elsewhere.

¹⁸ *C.f.*, Resolution on the Declaration of Principles on Freedom of Expression in Africa, adopted at the 32nd Session, October 2002, Principle IX.

¹⁹ For example, in Norway broadcasting standards have been upheld by an entirely self-regulatory press council since 1997.

²⁰ The rationale behind restricting these councils to media representatives is

partly to protect the independence of the journalistic process.

²¹ See, e.g. D.C. Koene, Press Councils in Western Europe, Netherlands Press Fund, 2009.

²² See, e.g. I. Bartle & P. Vass, Self-Regulation and the Regulatory State - A Survey of Policy and Practice, Research Report No. 17, Centre for the Study of Regulated Industries, University of Bath, 2005, p. 23.

²³ For example, so called “Leveson Inquiry” in the UK concluded that the self-regulatory body, the Press and Complaints Commission, was “not independent from the industry it was overseeing, causing problems both of substance and of perception. The way in which it and the self-regulatory system more generally conducted itself in public was often unhelpful. The purported investigations into press misconduct... were ineffectual and inadequate; and their conclusions ... gave false comfort to policy-makers and the public. Taken together these factors caused the self-regulatory system to fail.” See Leveson Inquiry, The report, An Inquiry into the culture, practices and ethics of the press, November 2012.

²⁴ See, e.g. ARTICLE 19, Freedom and Accountability Safeguarding Free Expression Through Media Self-Regulation, March 2005.

²⁵ In case of press self-regulation, this should include journalists, media owners and members of the public.

²⁶ For press self-regulatory bodies, these should include sanctions such as the publication of a correction or an apology. They should not be entitled to fine or ban media outlets or exclude individual members from the profession.

²⁷ See ARTICLE 19, ARTICLE 19's response to recognition of IMPRESS, 23 October 2016.

²⁸ See ARTICLE 19, Legal analysis, EU: European Commission's Code of Conduct for Countering Illegal Hate Speech Online and the Framework Decision, 20 June 2016.

²⁹ For example, at a hearing of the digital, culture, media and sport committee in the UK, Patricia Hodgson, the chairperson of the media regulator Ofcom, said that she believed internet businesses such as Google and Facebook were publishers; see Digital, Culture, Media and Sport Committee, Oral evidence: The Work of Ofcom, HC 407, 10 October 2017. See also, E. Bell, Who owns the news consumer: Social media platforms or publishers?, *The Columbia Journalism Review*, 21 June 2016.

³⁰ Distribution is a classic term of media policy. For instance, it has been used to define the "traditional" activity of a cable television provider: through its infrastructure, the cable company allowed television channels to reach the audience, and the audience to receive media content. The cable company chose which channels should be distributed through its cable network; it also affected the findability or visibility of the television channels by assigning them a certain

number on the remote (for example, it is easier to reach channel with assigned number 1 than the number 265). In certain jurisdictions, legal duties of must-carry (i.e., the cable operators were obliged to carry the signal of a particular television channel, such as the public service media channels) were imposed upon cable distributors in order to protect pluralism.

³¹ See, e.g. The Business Insider, Twitter has won the rights to stream NFL Thursday Night games, 5 April 2016; C. O'Brien, Twitter plans to launch 24/7 sports network Stadium this fall, *Venture Beat*, 27 July 2017.

³² N. Lesage, Soon TV shows on Facebook? The social network relies on original programming, *Numerama*, 26 June 2017.

³³ For example, the 2017 edition of the Digital News Report found that whilst filter bubbles affect a certain proportion of users, "on average, users of social media, aggregators, and search engines experience more diversity than non-users." According to the report, people are exposed to a significantly wider range of sources of information (in the sense of brands or outlets) through social media than they would normally have used in the analogue context. And while the algorithmic processes used to distribute information on social media affect the visibility, findability or accessibility of information and ideas, the choices made by individual users in terms of refining their newsfeed on social media also play a part. See, Reuters Institute for the Study of Journalism, *Digital News Report*,

available at: <http://www.digitalnewsreport.org/>

³⁴ See, e.g. Z. Tufekci, What Happens to #Ferguson Affects Ferguson: Net Neutrality, Algorithmic Filtering and Ferguson, The Message, 14 August 2014; or Z. Tufekci, How Facebook's Algorithm Suppresses Content Diversity (Modestly) and How the Newsfeed Rules Your Clicks, The Message, 7 May 2015. See also the research project of the Data & Society, Algorithms and Publics, or Who controls the public sphere? at https://datasociety.net_

³⁵ See, e.g., reports of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, 16 May 2011, A/HRC/17/27; and, 11 May 2016, A/HRC/32/38.

³⁶ See, e.g., J. C. Wong, Facebook blocks Pulitzer-winning reporter over Malta government exposé, The Guardian, 19 May 2017; Global Voices, Facebook 'Mistakenly' Censored Tiananmen Massacre Tribute, 30 May 2017. <https://advox.globalvoices.org/2017/05/30/facebook-mistakenly-censored-tiananmen-massacre-tribute/>

³⁷ Dilemma of liability, *op.cit.* For repository of publications on trends in this area, see, e.g. The Center for the Internet and Society, Stanford, Intermediary Liability at: <http://stanford.io/2F73pqP>.

³⁸ For example, Twitter now publishes information about government removal requests based on the company's Terms of Service, see <https://transparency.twitter.com/en/gov-tos-reports.html>; Google has

overhauled its Transparency Report, which provides additional contextual information or data e.g. relating to traffic disruption or National Security Letters, see <https://www.blog.google/topics/public-policy/new-look-our-transparency-report>.

³⁹ *Ibid.*, 2011 Report, para 42.

⁴⁰ E. Bell, Facebook's moderation is of public interest. It should be public knowledge, Columbia Journalism Review, 23 May 2017; or J. Angwin, Facebook's Secret Censorship Rules Protect White Men from Hate Speech But Not Black Children, ProPublica, 28 June 2017. See also The Guardian Series on "Facebook Files" at: <https://www.theguardian.com/news/series/facebook-files>.

⁴¹ An example of this problem is a worldwide criticism, including by many state officials, of Facebook removing an iconic historical picture of "Napalm girl" on the basis of the child nudity. In Brazil, the suppression of a poster showing naked indigenous people to announce a photography exhibition was a cause for similar debate in April 2015.

⁴² See, ARTICLE 19, Facebook vs Norway: Learning how to protect freedom of expression in the face of social media giants, 14 September 2016.

Ibid. In July 2016, the removal of a video that showed an African American dying from violence inflicted by US policemen caused an outcry that led Facebook to quickly reinstate the clip.

⁴³ For example, in 2018, Unilever announced that it will withdraw its advertising from online platforms such

as Facebook and Google if they fail to eradicate content; see Unilever, Unilever will not invest in online platforms that create division, 12 February 2018. In 2013, Facebook was criticised for allowing images glorifying rape, and subsequently advertisers such as Dove, Nationwide and Nissan UK pulled their advertising from the platform over the issue. Facebook eventually removed the images from its platform and amended its community guidelines to include provisions stipulating that Facebook removes content “that threatens or promotes sexual violence.” See, Facebook vs Norway, op.cit. See, also, e.g. L. Stamper, Facebook Will Block Photos Celebrating Rape Following Ad Boycott, The Business Insider, 28 May 2013.

⁴⁴ For Global Network Initiative work, see: <https://www.globalnetworkinitiative.org>.

⁴⁵ There are three distinct models of liability for intermediaries, including social media: a) the strict liability model under which companies are liable for third party content and are effectively required to monitor content in order to comply with the law; b) the safe harbour model grants immunity, provided they comply with certain requirements. This model is at the heart of the so called ‘notice and take down’ procedures and can be sub-divided into vertical and horizontal approach; and c) the broad immunity model grants internet intermediaries broad or conditional immunity from liability for third-party content and exempts them from any general requirement to monitor content. See, Dilemma of Liability, op.cit., p. 7.

⁴⁶ This is a result of the decision of the European Court of Justice in NewsMedia Online GmbH (C-347/14).

⁴⁷ See, e.g. EDRI, AVMS Directive reform: Document pool, May 2017.

⁴⁸ As for the first problem, ARTICLE 19 believes that the extension of a regulatory authority’s jurisdiction to audiovisual content published by print media could potentially be conducive to stricter control of the press and would not be justified under international standards on freedom of expression. As for the AVMS Directive proposal, we believe this would be incompatible with the EU E-commerce directive, which states that Internet intermediaries (including social media) are not liable for third-party content that is stored on their servers, unless they have edited such content or have been made aware of its unlawfulness. See, e.g., ARTICLE 19, New EU legislation must not throttle online flows of information and ideas, 12 September 2017.

⁴⁹ The Network Enforcement Act, 30 June 2017, in force of 1 October 2017, Federal Law Gazette 2017 I, Nr. 61, page 3352. The term “clearly illegal content” refers to the criminal law provisions on prohibition of incitement to terrorism and incitement to hatred.

⁵⁰ For the critique of NetzDG, see, e.g. ARTICLE 19, Germany: Act to Improve Enforcement of the Law on Social Networks undermines free expression, 1 September 2017.

⁵¹ European Commission, the Code of Conduct on countering illegal hate speech online, 31 May 2016.

⁵² For the critique of the Code of Conduct, see, ARTICLE 19, Legal Analysis, op.cit.

⁵³ The Code of Conduct specifically refers to the Framework Decision on Combatting Certain Forms and Expressions of Racism and Xenophobia by Means of Criminal Law (Framework Decision) as the legal basis for defining illegal hate speech under the Code; see the Council Framework Decision 2008/913/JHA of 28 November 2008 on combating certain forms and expressions of racism and xenophobia by means of criminal law. For a full analysis of the Framework Decision, see ARTICLE 19, Submission to the Consultations on the EU Justice Policy, December 2013.

⁵⁴ See, e.g. The Guardian view on moderating Facebook: we need to talk: Editorial, The Guardian, 22 May 2017.

⁵⁵ See, e.g. M. Moore, Tech Giants and Civic Power, The King's College, 2016..

⁵⁶ For more details, see Dilemma of Liability, op.cit.

DEFENDING FREEDOM OF EXPRESSION AND INFORMATION

ARTICLE 19 Free Word Centre 60 Farringdon Road London EC1R 3GA

T +44 20 7324 2500 F +44 20 7490 0566

E info@article19.org W www.article19.org Tw [@article19org](https://twitter.com/article19org) facebook.com/article19org

© ARTICLE 19