

ARTICLE 19

# Germany: The Act to Improve Enforcement of the Law in Social Networks

---

August 2017

Legal analysis

# Executive summary

---

In August 2017, ARTICLE 19 analysed the Act to Improve Enforcement of The Law in Social Networks (the Act), which the Federal Council (Upper Chamber) of the German Parliament approved on 7 July 2017, and is expected to be published in the Gazette shortly. The Act will enter into force on 1 October 2017. ARTICLE 19 analysed the earlier draft of the Act.

ARTICLE 19 is deeply concerned that the Act will severely undermine freedom of expression in Germany, and is already setting a dangerous example to other countries that more vigorously apply criminal provisions to quash dissent and criticism, including against journalists and human rights defenders.

The Act establishes an intermediary liability regime that incentivises, through severe administrative penalties of up to 5 million Euros, the removal and blocking of “clearly violating content” and “violating content”, within time periods of 24 hours and 7 days respectively. As regulatory offences, it is possible for the maximum sanction to be multiplied by ten to 50 million Euros.

Though the Act does not create new content restrictions, it compels content removals on the basis of select provisions from the German Criminal Code. Many of these provisions raise serious freedom of expression concerns in and of themselves, including prohibitions on “defamation of religion”, broad concepts of “hate speech”, and criminal defamation and insult. Deputising private companies to engage in censorship on the basis of these provisions is deeply troubling, as they should not be criminal offences in the first place.

We are particularly concerned that the obligation to remove or block content applies without any prior determination of the legality of the content at issue by a court, and with no guidance to Social Networks on respecting the right to freedom of expression. Such private enterprises are not competent to make these complex factual and legal determinations, and the Act provides no recourse to users whose content is blocked or deleted unfairly. Though the Act creates a system for recognising “self-regulation institutions” to act as secondary review bodies for “unlawful content”, this recognition is conditional and held by an administrative body not insulated from political influence.

The likelihood of Social Networks being over-vigorous in deleting or blocking content is compounded by the legal uncertainty pervading the Act. The threshold at which a Social Network’s content removal and blocking processes will be considered systemic enough to attract administrative liability is unclear, and ambiguity in the definitions of key terms (including of “Social Network”) is likely to create an environment wherein lawful content is routinely blocked or removed as a precaution. The secondary review that would be provided by “self-regulation institutions”, and the limited oversight provided by the Administrative Courts does nothing to address over-blocking, and provides little protection or due process to Social Networks that in good faith refrain from blocking or removing content in the interests of respecting freedom of expression.

## Summary of recommendations

1. The Act should be repealed, with consideration given to retaining Section 2 on reporting requirements in alternative legislation to increase transparency around online content moderation by private actors;
2. The German Criminal Code should be comprehensively revised to remove offences that are not compatible with international human rights law on freedom of expression, including but not limited to those listed in the Act.



# Table of contents

---

- Introduction ..... 4
- International human rights standards ..... 5
  - The right to freedom of expression ..... 5
  - The right to equality ..... 6
  - Limitations on “hate speech” ..... 6
    - Obligation to prohibit ..... 6
    - Permissible limitations ..... 8
    - Lawful expression ..... 8
  - Freedom of expression online and intermediary liability ..... 8
  - Self-regulatory models for social media ..... 11
- Analysis of the Act ..... 12
  - Defining “Social Network” ..... 12
  - “Unlawful content” ..... 13
  - Handling complaints about unlawful content ..... 16
    - Processes for complaining about unlawful content ..... 16
    - Removal and blocking ..... 17
    - “Recognition” for “self-regulatory” institutions ..... 20
  - Reporting requirements ..... 22
  - Individual responsibility for offences committed online ..... 22
  - Amendments to the Telemedia Act ..... 23
- Conclusion ..... 24
- About ARTICLE 19 ..... 25

# Introduction

---

In this legal analysis, ARTICLE 19 reviews the German Act to Improve Enforcement of the Law in Social Networks (the Act)<sup>1</sup> for its compatibility with international human rights law. According to the Act, it will enter into force on 1 October 2017.

ARTICLE 19 previously expressed serious concerns with earlier drafts of the Act.<sup>2</sup> This analysis is an update to our analysis of the previously published Bill, pointing out significant differences that were made in the later stages of the legislative process. Noting the German government's response to those concerns,<sup>3</sup> we reiterate in this analysis why, notwithstanding changes made, the Act violates Germany's obligations on freedom of expression under international human rights law.

ARTICLE 19 has extensive experience in analysing laws pertaining to the right to freedom of expression online,<sup>4</sup> as well as legislation on "hate speech". The issue of intermediary liability for content created by third parties is a recurring issue and an area where we see increasing pressure on private companies to control information flows. Recently, we have expressed concerns at attempts in this regard at the European Union level,<sup>5</sup> which the Act emulates and expands upon. Our policy position on this is set out in "Internet Intermediaries: Dilemma of liability."<sup>6</sup> The types of content limitation foreseen in the Act are also the focus of several of ARTICLE 19's standard setting policies, including on "hate speech",<sup>7</sup> defamation,<sup>8</sup> and national security.<sup>9</sup>

The Act addresses specific forms of online intermediaries, defined as "Social Networks", and imposes duties on them, inter alia, to adopt particular complaints procedures regarding content that violates provisions of the German Criminal Code, to delete or block access to that content, and to report on their practices in this regard periodically to administrative authorities. Failure to abide by these conditions results in the imposition of severe administrative fines.

This analysis sets out relevant international human rights law standards before examining the main freedom of expression concerns raised by the Act. ARTICLE 19 finds that the Act does not comply with international human rights law, and recommends its full repeal.

We also urge the German Government to bring its legislation into full compliance with international freedom of expression standards, in particular the Penal Code, and we stand ready to provide further assistance in this process.

---

<sup>1</sup> This analysis reviews a [consolidated English version of the Act](#). The analysis assumes the differences between the English and German versions of the Law reflect errors in the English version. In the English version, Section 1(3) omits a reference included in the German version to Section 201(a) of the German Penal Code. In numerous provisions, the term in German more accurately translated as "a recognised system of regulated self-regulation" appears in the English version only as "self-regulation" or "recognised self-regulation." ARTICLE 19 takes no responsibility for the accuracy of the translation or for comments made on the basis of any inaccuracies in the translation.

<sup>2</sup> ARTICLE 19, [Germany: draft bill on social networks raises serious free expression concerns](#), 26 April 2017

<sup>3</sup> [The government of Germany responded to a series of criticisms](#) offered by multiple freedom of expression and press freedom organisations to the Council of Europe's Media Freedom Platform.

<sup>4</sup> See, for example, ARTICLE 19, [Myanmar: Telecommunications Law](#), Legal Analysis, March 2017.

<sup>5</sup> ARTICLE 19, [EU Commission Code of Conduct for Countering Illegal Hate Speech Online and the Framework Decision](#), June 2016.

<sup>6</sup> ARTICLE 19, [Internet Intermediaries: Dilemma of Liability](#), 29 August 2013.

<sup>7</sup> ARTICLE 19, [Hate Speech Explained: A Toolkit](#), December 2015.

<sup>8</sup> ARTICLE 19, [Defining Defamation: Principles on Freedom of Expression and Protection of Reputation](#), 2<sup>nd</sup> Edition, February 2017.

<sup>9</sup> ARTICLE 19, [Johannesburg Principles on National Security, Freedom of Expression and Access to Information](#), 1996.

# International human rights standards

---

ARTICLE 19's comments on the Act are informed by international human rights law and standards, in particular regarding the mutually interdependent and reinforcing rights to freedom of expression and equality.

## The right to freedom of expression

The right to freedom of expression is protected by Article 19 of the Universal Declaration of Human Rights (UDHR),<sup>10</sup> and given legal force through Article 19 of the International Covenant on Civil and Political Rights (ICCPR).<sup>11</sup>

The scope of the right to freedom of expression is broad. It requires States to guarantee to all people the freedom to seek, receive or impart information or ideas of any kind, regardless of frontiers, through any media of a person's choice. The UN Human Rights Committee (HR Committee), the treaty body of independent experts monitoring States' compliance with the ICCPR, has affirmed that the scope of the right extends to the expression of opinions and ideas that others may find deeply offensive,<sup>12</sup> and this may encompass discriminatory expression.

While the right to freedom of expression is fundamental, it is not absolute. A State may, exceptionally, limit the right under Article 19(3) of the ICCPR, provided that the limitation is:

- Provided for by law; 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;
- Necessary in a democratic society, requiring the State to demonstrate in a specific and individualised fashion the precise nature of the threat, 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.<sup>13</sup>

Thus, any limitation imposed by the State on the right to freedom of expression, including to limit "hate speech," must conform to the strict requirements of this three-part test. Further, Article 20(2) ICCPR provides that any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence must be prohibited by law (see below).

At the European level, Article 10 of the European Convention on Human Rights (ECHR)<sup>14</sup> protects the right to freedom of expression in similar terms to Article 19 of the ICCPR, with permissible limitations set out in Article 10(2).<sup>15</sup> Within the European Union, the right to freedom of

---

<sup>10</sup> Through its adoption in a resolution of the UN General Assembly, the UDHR is not strictly binding on states. However, many of its provisions are regarded as having acquired legal force as customary international law since its adoption in 1948; see *Filartiga v. Pena-Irala*, 630 F. 2d 876 (1980) (US Circuit Court of Appeals, 2<sup>nd</sup> circuit).

<sup>11</sup> The ICCPR has 167 States parties, including Germany.

<sup>12</sup> See HR Committee, General Comment No. 34 on Article 19: Freedoms of opinion and expression, CCPR/C/GC/34, 12 September 2011, para 11.

<sup>13</sup> *Ibid.*, para. 22.

<sup>14</sup> Convention for the Protection of Human Rights and Fundamental Freedoms, Rome, 4 September 1950.

<sup>15</sup> Article 10 (1) of the ECHR: Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers. This Article shall not prevent States from requiring the licensing of broadcasting, television

expression and information is guaranteed in Article 11 of the Charter of Fundamental Rights of the European Union.

### The right to equality

Article 1 of the UDHR states that “all human beings are born free and equal in dignity and rights” and Article 2 provides for the equal enjoyment of the rights and freedoms contained in the declaration “without distinction of any kind,” and Article 7 requires protection from discrimination. These guarantees are given legal force in Articles 2(1) and 26 of the ICCPR, obliging States to guarantee equality in the enjoyment of human rights, including the right to freedom of expression, and equal protection of the law.

At the European level, the ECHR prohibits discrimination in Article 14 and more broadly in Protocol No. 12.

### Limitations on “hate speech”

While “hate speech” has no definition under international human rights law, the expression of hatred towards an individual or group on the basis of a protected characteristic can be divided into three categories, distinguished by the response international human rights law requires from States:<sup>16</sup>

- Severe forms of “hate speech” that international law *requires* States to prohibit, including through criminal, civil, and administrative measures, under both international criminal law and Article 20(2) of the ICCPR;
- Other forms of “hate speech” that States *may* prohibit to protect the rights of others under Article 19(3) of the ICCPR, such as discriminatory or bias-motivated threats or harassment;
- “Hate speech” that is lawful but nevertheless raises concerns in terms of intolerance and discrimination, meriting a critical response by the State but which should be protected from restriction under Article 19(3) of the ICCPR.

#### *Obligation to prohibit*

Article 20(2) of the ICCPR obliges States to prohibit by law “any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence.” In General Comment No. 34, the HR Committee stressed that while States are *required* to prohibit such expression, these limitations must nevertheless meet the strict conditions set out in Article 19(3).<sup>17</sup>

The Rabat Plan of Action,<sup>18</sup> adopted by experts following a series of consultations convened by the UN Office of the High Commissioner for Human Rights (OHCHR), advances authoritative conclusions and recommendations for the implementation of Article 20(2) ICCPR:<sup>19</sup>

---

or cinema enterprises; Article 10 (2), The exercise of these freedoms, since it carries with it duties and responsibilities, may be subject to such formalities, conditions, restrictions or penalties as are prescribed by law and are necessary in a democratic society, in the interests of national security, territorial integrity or public safety, for the prevention of disorder or crime, for the protection of health or morals, for the protection of the reputation or rights of others, for preventing the disclosure of information received in confidence, or for maintaining the authority and impartiality of the judiciary.

<sup>16</sup> For a full explanation of ARTICLE 19’s policy on “hate speech”, see *Hate Speech Explained: A Toolkit*, *op.cit.*

<sup>17</sup> HR Committee, General Comment No. 34, 21 June 2011, CCPR/C/GC/34, para. 52.

<sup>18</sup> [The Rabat Plan of Action on the prohibition of advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence](#), A/HRC/22/17/Add.4, Appendix, adopted 5 October 2012.

<sup>19</sup> The Rabat Plan of Action has been endorsed by a wide range of special procedures of the UN Human Rights Council, see, for example: Report of the Special Rapporteur on protecting and promoting the right to freedom of opinion and expression on hate speech and incitement to hatred, A/67/357, 7 September 2012; Report of the

## Germany: The Act to Improve Enforcement of the Law in Social Networks

- Incitement: prohibitions should only focus on the advocacy of discriminatory hatred that constitutes incitement to hostility, discrimination, or violence from the audience, rather than the advocacy of hatred without regard to its tendency to incite action by the audience against a protected group.
- Six-part threshold test: to assist in judicial assessments of whether a speaker intends and is capable of having the effect of inciting their audience to violent or discriminatory action through the advocacy of discriminatory hatred, six factors should be considered:
  - Context: the expression should be considered within the political, economic, and social context prevalent at the time it was communicated, for example the existence or history of conflict, existence or history of institutionalised discrimination, the legal framework and the media landscape;
  - Identity of the speaker: the position of the speaker as it relates to their authority or influence over their audience, in particular if they are a politician, public official, religious or community leader;
  - Intent of the speaker to engage in advocacy to hatred; intent to target a protected group on the basis of a protected characteristic, and knowledge that their conduct will likely incite the audience to discrimination, hostility or violence;
  - Content of the expression: what was said, including the form and the style of the expression, and what the audience understood by this;
  - Extent and magnitude of the expression: the public nature of the expression, the means of the expression and the intensity or magnitude of the expression in terms of its frequency or volume;
  - Likelihood of harm occurring, including its imminence: there must be a reasonable probability of discrimination, hostility or violence occurring as a direct consequence of the incitement.
- Protected characteristics: States' obligations to protect the right to equality more broadly, with an open-ended list of protected characteristics, supports an expansive interpretation of the limited protected characteristics in Article 20(2) of the ICCPR to provide equal protection to other individuals and groups who may similarly be targeted for discrimination or violence on the basis of other recognised protected characteristics.
- Proportionate sanctions: the term "prohibit by law" does not mean criminalisation; the HR Committee has said it only requires States to "provide appropriate sanctions" in cases of incitement.<sup>20</sup> Civil and administrative penalties will in many cases be most appropriate, with criminal sanctions an extreme measure of last resort.

The Committee on the Elimination of all forms of Racial Discrimination (the CERD Committee) has also based their guidance for respecting the obligation to prohibit certain forms of expression under Article 4 of the ICERD on this test.<sup>21</sup>

---

Special Rapporteur on freedom of religion or belief on the need to tackle manifestations of collective religious hatred, Heiner Bielefeldt, A/HRC/25/58, 26 December 2013; Report of the Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance on manifestations of racism, racial discrimination, xenophobia and related intolerance on manifestations of racism on the Internet and social media, Mutuma Ruteere, A/HRC/26/49, 6 May 2014; and the contribution of the UN Special Advisor on the Prevention of Genocide to the expert seminar on ways to curb incitement to violence on ethnic, religious, or racial grounds in situations with imminent risk of atrocity crimes, Geneva, 22 February 2013.

<sup>20</sup> Human Rights Committee, General Comment 11: prohibition of propaganda for war and inciting national, racial or religious hatred (Art. 20), 29 July 1983, para. 2.

<sup>21</sup> UN Committee on the Elimination of Racial Discrimination, General recommendation No. 35: Combating racist hate speech, 26 September 2013, paras. 15 - 16. The CERD Committee specifies that five contextual factors should be taken into account: the content and form of speech; the economic, social and political climate; the position or status of the speaker; the reach of the speech; and the objectives of the speech. The CERD Committee also specifies that States must also consider the intent of the speaker and the imminence and likelihood of harm.

## Germany: The Act to Improve Enforcement of the Law in Social Networks

At the European level, the ECHR 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 of Human Rights (European Court) has recognised that certain forms of harmful expression must necessarily be restricted to uphold the objectives of the Convention as a whole.<sup>22</sup> The Court has also exercised particularly strict supervision in cases where criminal sanctions have been imposed by the State and in many instances it has found that the imposition of a criminal conviction violated the proportionality principle.<sup>23</sup> Recourse to criminal law should therefore not be seen as the default response to instances of harmful expression, if less severe sanctions would achieve the same effect.

At the EU level, the Council's framework decision "on combating certain forms and expressions of racism and xenophobia by means of criminal law" (2008/913/JHA) requires States to sanction racism and xenophobia through "effective, proportionate and dissuasive criminal penalties". It establishes four categories of incitement to violence or hatred offences that States are required to criminalise with penalties of up to 3 years' imprisonment, including condoning, denying or grossly trivialising historical crimes. States are afforded the discretion of choosing to punish only conduct which is carried out in "a manner likely to disturb public order" or "which is threatening, abusive, or insulting", implying that limitations on expression not likely to have these negative impacts can legitimately be restricted. These obligations are broader and more severe in the penalties prescribed than the prohibitions in Article 20(2) of the ICCPR, and do not comply with the requirements of Article 19(3) of the ICCPR.<sup>24</sup>

### *Permissible limitations*

There are forms of "hate speech" that target an identifiable individual, but that do not necessarily advocate hatred to a broader audience with the purpose of inciting discrimination, hostility or violence. This includes discriminatory threats of unlawful conduct, discriminatory harassment, and discriminatory assault. These limitations must still be justified under Article 19(3) of the ICCPR.

### *Lawful expression*

Expression may be inflammatory or offensive, but not meet any of the thresholds described above. This expression may be characterised by prejudice, and raise concerns over intolerance, but does not meet the threshold of severity, at which restrictions on expression are justified. This also includes expression related to denial of historical events, insult of State symbols or institutions and other forms of expression that some individuals and groups might find offensive.

This does not preclude States from taking legal and policy measures to tackle the underlying prejudices of which this category of "hate speech" is symptomatic, or from maximising opportunities for all people, including public officials and institutions, to engage in counter-speech.

## Freedom of expression online and intermediary liability

### *International law*

At the international level, several human rights bodies and mechanisms have developed soft law guidance on intermediary liability.

<sup>22</sup> European Court, *Erbakan v. Turkey*, App. No. 59405/00 (2006), para. 56 or *Gündüz v. Turkey*, App. No. 35071/97 (2004), para. 22.

<sup>23</sup> European Court, *Jersild v. Denmark*, App No 15890/89 (1992), para. 35.

<sup>24</sup> For a full analysis, see ARTICLE 19, [Submission to the Consultations on the European Union's Justice Policy](#), December 2013.



## Germany: The Act to Improve Enforcement of the Law in Social Networks

The UN Human Rights Council (HRC) recognised in 2012 that the “same rights that people have offline must also be protected online”.<sup>25</sup> The HR Committee has also made clear that 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, as set out above.<sup>26</sup>

While international human rights law places obligations on States to protect, promote and respect human rights, it is widely recognised that business enterprises also have a responsibility to respect human rights.<sup>27</sup> Importantly, the UN Special Rapporteur on freedom of opinion and expression (Special Rapporteur on freedom of expression) has long held that censorship measures should never be delegated to private entities.<sup>28</sup> In his June 2016 report to the HRC,<sup>29</sup> the Special Rapporteur on freedom of expression, David Kaye, enjoined States not to require or otherwise pressure the private sector to take steps that unnecessarily or disproportionately interfere with freedom of expression, whether through laws, policies, or extralegal means. He further recognised that “private intermediaries are typically ill-equipped to make determinations of content illegality,”<sup>30</sup> and reiterated criticism of notice and takedown frameworks for “incentivising questionable claims and for failing to provide adequate protection for the intermediaries that seek to apply fair and human rights-sensitive standards to content regulation,” i.e. the danger of “self- or over-removal.”<sup>31</sup>

The Special Rapporteur on freedom of expression recommended that any demands, requests and other measures to take down digital content must be based on validly enacted law, subject to external and independent oversight, and demonstrate a necessary and proportionate means of achieving one or more aims under Article 19 (3) of the ICCPR.<sup>32</sup>

In their 2017 Joint Declaration on freedom of expression, “fake news”, disinformation and propaganda, the four international mandates on freedom of expression expressed concern at “attempts by some governments to suppress dissent and to control public communications through [...] efforts to ‘privatise’ control measures by pressuring intermediaries to take action to restrict content.”<sup>33</sup> The Joint Declaration emphasises that

[I]ntermediaries should never be liable for any third party content relating to those services unless they specifically intervene in that content or refuse to obey an order adopted in accordance with due process guarantees by an independent, impartial, authoritative oversight body (such as a court) to remove it and they have the technical capacity to do that.

They also outlined the responsibilities of intermediaries regarding the transparency of, and need for, due process in their content-removal processes.

<sup>25</sup> HRC Resolution 20/8 on the Internet and Human Rights, A/HRC/RES/20/8, June 2012.

<sup>26</sup> General Comment No. 34, *op cit.*, para 43.

<sup>27</sup> Guiding Principles on Business and Human Rights: Implementing the United Nations “Protect, Respect and Remedy” Framework (The Ruggie Principles), A/HRC/17/31, 21 March 2011, Annex. The UN Human Rights Council endorsed the guiding principles in HRC resolution 17/4, A/HRC/RES/17/14, 16 June 2011.

<sup>28</sup> Report 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, paras. 75-76.

<sup>29</sup> Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, 11 May 2016, A/HRC/32/38; para 40 – 44,

<sup>30</sup> *Ibid.*

<sup>31</sup> *Ibid.*, para. 43.

<sup>32</sup> *Ibid.*

<sup>33</sup> [Joint Declaration on Freedom of Expression and “Fake News”, Disinformation and Propaganda](#), adopted by the United Nations (UN) Special Rapporteur on Freedom of Opinion and Expression, the Organization for Security and Co-operation in Europe (OSCE) Representative on Freedom of the Media, the Organization of American States (OAS) Special Rapporteur on Freedom of Expression and the African Commission on Human and Peoples’ Rights (ACHPR) Special Rapporteur on Freedom of Expression and Access to Information, 3 March 2017.

*European law*

At the European level, the approach to intermediary liability has not always accorded sufficient protections to freedom of expression.

The E-Commerce Directive requires that Member States shield intermediaries from liability for illegal third party content where the intermediary does not have actual knowledge of illegal activity or information and, upon obtaining knowledge, acts expeditiously to remove or disable access to the information.<sup>34</sup> The E-Commerce Directive prohibits States from imposing general obligations on intermediaries to monitor activity on their services.<sup>35</sup> The regulatory scheme under the Directive has given rise to so-called ‘notice-and-takedown’ procedures, which have been sharply criticised by the special mandates on freedom of expression for their lack of clear legal basis and basic procedural fairness.

The limited shield from liability for intermediaries provided by the E-Commerce Directive has been further undermined by the approach of the European Court of Human Rights. In *Delfi AS v. Estonia*, the Grand Chamber of the European Court found no violation of Article 10 of the ECHR where a national court imposed civil liability on an online news portal for failure to remove “clearly unlawful” comments posted to the website by an anonymous third party, even without notice being provided.<sup>36</sup> A joint dissenting opinion highlighted that this “constructive notice” standard contradicts the requirement of actual notice in Article 14(1) of the E-Commerce Directive, necessitating intermediaries to actively monitor all content to avoid liability in relation to specific forms of content, thus additionally contradicting Article 5 of the E-Commerce Directive.<sup>37</sup>

Decisions subsequent to *Delfi AS* appear to confine the reasoning to cases concerning “hate speech”. In *Magyar Tartalomszolgáltatók Egyesülete and/et Index.hu Zrt v. Hungary*,<sup>38</sup> the European Court found a violation of Article 10 of the ECHR where a self-regulatory body of Internet content providers and an owner of an online news portal were held liable for defamatory comments posted by a third party, which the parties removed on receipt of notice (in this instance, the issuance of legal proceedings against them). More recently, the European Court rejected as inadmissible a complaint that the domestic courts had failed to protect the applicant’s right to privacy by refusing to hold a non-profit association liable for defamatory comments posted to their website by a third party. The Court noted that the comments were not “hate speech” or direct threats and were removed upon notice (though a formal notice-and-takedown was not in place).<sup>39</sup> The position and resources of the intermediary were also relevant factors.<sup>40</sup>

Lastly, the European Commission’s “Code of Conduct on Countering Illegal Hate Speech” (May 2016), developed in collaboration with some of the major information technology companies, constitutes a (non-legally binding) commitment to remove “illegal hate speech”, defined on the basis of the “Framework Decision on Combatting Certain Forms and Expressions of Racism and Xenophobia by Means of Criminal Law” (above), within 24 hours. While the Code of Conduct is ostensibly “voluntary”, it is part of a concerning trend whereby States (including through intergovernmental organisations) are increasing pressure on private actors to engage in censorship of content without any independent adjudication on the legality of content in issue.

<sup>34</sup> Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market, Article 14(1).

<sup>35</sup> *Ibid.*, Article 5.

<sup>36</sup> Grand Chamber of the European Court, *Delfi AS v. Estonia*, No. 64569/09, 16 June 2015.

<sup>37</sup> *Ibid.*, Judges Sajó and Tsotsoria’s Joint Dissent Opinion. See also Gabrielle Guillemin, [Delfi AS v Estonia Case Comment: Strasbourg Undermines Freedom of Expression](#), ARTICLE 19 Join the Debate, 1 October 2015.

<sup>38</sup> European Court of Human Rights, *Magyar Tartalomszolgáltatók Egyesülete and/et Index.hu Zrt v. Hungary*, No. 22947/13, 2 February 2016

<sup>39</sup> European Court of Human Rights, *Pihl v. Sweden*, No. 74742/13, 9 March 2017

<sup>40</sup> For further analysis, see: Dirk Voorhoof, [Pihl v. Sweden: non-profit blog operator is not liable for defamatory users’ comments in case of prompt removal upon notice](#), Strasbourg Observers, 20 March 2017.

## Germany: The Act to Improve Enforcement of the Law in Social Networks

In short, the law on intermediary liability remains legally uncertain in Europe, with tensions between the European Court's jurisprudence and the protections of the E-Commerce Directive, as well as with the guidance of the international freedom of expression mandates.

### Self-regulatory models for social media

ARTICLE 19 considers "self-regulation" to refer to a voluntary agreement between the stakeholders of a sector, civil society and the public, to set up an external, independent body, tasked with supporting and promoting ethical standards through recommendations and the application of light, non-pecuniary or "moral" remedies. In the media sector, for example, press councils are a typical example of such a mechanism.

Based on ARTICLE 19's experience with press councils, we consider that self-regulation could provide an interesting model to approach the issue of content moderation on social media platforms, but only where such bodies meet certain requirements.<sup>41</sup> These requirements aim at ensuring that self-regulation is meaningful. As such, self-regulatory mechanisms should act as both a defender of the rights of members of the profession, as well as a guiding force for their conduct and an adjudicator for complaints received from members of the public.

Self-regulatory bodies should therefore:

- Be independent from government, commercial and special interests;
- Be established via a fully consultative and inclusive process – major constitutive elements of their work should be elaborated in an open, transparent and participatory manner that allows for broad public consultation;
- Democratic and transparent in their selection of members and decision-making;
- Be composed of a membership with broad representation,<sup>42</sup> including of civil society, to ensure independence;
- Have the power to impose only non-pecuniary or "moral" sanctions;<sup>43</sup>
- Work in the service of the public interest, be transparent and accountable to the public.

Any self-regulatory body needs to adopt a code of ethics for the profession or sector it seeks to regulate. It also needs a robust complaint mechanism and clear procedural rules to determine if ethical standards were breached in individual cases. In addition, such a body can also play an important role in promoting the 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.

While a certain form of legal underpinning of self-regulation may be compatible with international human rights law on freedom of expression, this must not undermine the independence and effectiveness of self-regulatory bodies.

---

<sup>41</sup> See, e.g. ARTICLE 19, [Freedom And Accountability Safeguarding Free Expression Through Media Self-Regulation](#), March 2005.

<sup>42</sup> In case of press self-regulation, this should include journalists, media owners and members of the public.

<sup>43</sup> 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.

# Analysis of the Act

---

ARTICLE 19 finds that the Act contains various provisions that do not meet international standards on the right to freedom of expression, and create a serious danger of further incentivising the over-removal of expression that should be considered lawful under international human rights law.

## Defining “Social Network”

Section 1(1) of the Act provides that “Social Networks” are “telemedia service providers which, for profit-making purposes, operate Internet platforms which are designed to enable users to share any content with other users or to make such content available to the public.”

The same provision excludes from the definition of “Social Networks” for “[p]latforms offering journalistic or editorial content, the responsibility for which lies with the service provider itself”. Section 1(2) of the Act provides that obligations in relation to reporting (Section 2) and handling complaints (Section 3) apply only to Social Networks with more than 2 million users in the Federal Republic of Germany.

ARTICLE 19 finds that the definition of “Social Network” is too ambiguous and creates significant uncertainty about who the duties in the Act apply to. This is in violation of the requirement of legal certainty under the first prong of the three-part test under Article 19(3) of the ICCPR and Article 10(2) of the ECHR (see above).

In particular:

- **“For profit-making purposes”:** it is not clear how “for-profit” platforms are distinguished from not-for-profit platforms, and what criteria apply in making this determination.
- **“Sharing” or making content “available”** to the public: these terms seemingly bring a range of platforms that allow the posting of third party contributions, or that connect to such content, within the scope of the law. There is no requirement for the content to be publicly available. Various types of platforms not generally considered to be “Social Networks” are captured by this definition. Gaming platforms that enable interaction between users, websites for instant messaging or email, sites with customer/user reviews of goods or services, platforms for collaborative working, and platforms that enable file-sharing or remote storage, may fall within the scope of the Act. What “mak[ing] such content available” constitutes is also unclear, for example whether it covers hyperlinking to such content;
- **“Journalistic content”:** this term, which determines which tele media providers are exempt from duties as “Social Networks”, is not defined and therefore potentially subject to various interpretations. International human rights law advances an expansive and functional (rather than formal) understanding of “journalism”, which would bring many online platforms partly or wholly within the potential scope of this exclusion.<sup>44</sup> It also is unclear whether this exclusion benefits platforms that contain some “journalistic content” for which “responsibility” (also not defined) is accepted, alongside other third-party content for which “responsibility” is not automatically accepted (such as comments). This legal uncertainty may also incentivise platforms with journalist content to remove comment and discussion

---

<sup>44</sup> HR Committee, General Comment No.34, *op. cit.*, para 44: Journalism is a function shared by a wide range of actors, including professional full-time reporters and analysts, as well as bloggers and others who engage in forms of self publication in print, on the internet or elsewhere [...]”

functions or subject them to restrictive prior screening, limiting space for freedom of expression online.

- Two million registered users: the Act gives no technical definition of a “registered user” and no guidance on how their number will be measured. Section 1(2) is slightly clearer than in previous versions, as it now specifies that users must be “registered”, and be “in” the Federal Republic of Germany. Nevertheless, technical questions about how to count “inactive accounts”, accounts of individuals who only periodically use Germany-based IPs or use VPNs, mean that Social Networks face uncertainty as to whether the Act applies to them. Additionally, this exemption does not apply to obligations under Section 5 (requiring the appointment of representatives in Germany), bringing Social Networks of all sizes within the scope of punitive measures under Section 4(1)(7).

The ambiguity in the definition of Social Network within Sections 1(1) and 1(2) of the Act makes uncertain the scope of duties provided for in subsequent sections. For those provisions that have the effect of limiting freedom of expression through requirements aimed at “Social Networks”, ARTICLE 19 considers that they are not “provided for by law”, as business enterprises cannot easily predict whether or not these provisions apply to them.

### “Unlawful content”

Section 1(3) of the Act defines “unlawful content” as content that contravenes a number of provisions selected from the German Criminal Code (GCC).<sup>45</sup>

This provision now excludes from the scope of the Act two criminal offences that had been included in previous drafts, namely: criminal defamation of the President of the Federation (Section 90 GCC); and criminal defamation of the state and its symbols (Section 90a GCC). These two provisions were against international human rights law, and we welcome their removal.

Nevertheless, ARTICLE 19 is disappointed that references to several of the most egregious criminal restrictions on freedom of expression in the Penal Code are retained in the Act. While noting that Article 5 of the Basic Law for the Federal Republic of Germany gives no protection to criminal content,<sup>46</sup> that in itself does not mean that the definition of what constitutes “criminal content” complies with Germany’s obligations under international human rights law.

The provisions of concern include:

---

<sup>45</sup> [Criminal Code, 13 November 1998](#), Federal Law Gazette [Bundesgesetzblatt] I p. 3322, last amended by Article 1 of the Law of 24 September 2013, Federal Law Gazette I p. 3671 and with the text of Article 6(18) of the Law of 10 October 2013, Federal Law Gazette I p 3799. The listed provisions in the Act are: Section 86 “dissemination of propaganda material of unconstitutional organisations” and Section 86a “using symbols of unconstitutional organisations”, Section 89a “preparation of a serious violence offence endangering the state”, Section 91 “encouraging the commission of a serious violence offence endangering the state”, Section 100(a) “treasonous forgery”, Section 111 “public incitement to crime”, Section 126 “breach of the public peace by threatening to commit offences”, Section 129 “forming criminal organisations”, Section 129a “forming terrorist organisations”, Section 129b “criminal and terrorist organisations abroad”, Section 130 “incitement to hatred”, Section 131 “dissemination of depictions of violence”, Section 140 “rewarding and approving of offences”, Section 166 “defamation of religions, religious and ideological associations”, Section 184b “distribution, acquisition and possession of child pornography” in conjunction with Section 184(d) “distribution of pornographic performances by broadcasting, media services or telecommunications services”, Section 185 “insult”, Section 186 “defamation”, Section 187 “intentional defamation”, Section 201(a) “violation of intimate privacy by taking photographs”, Section 241 “threatening the commission of a felony”, Section 269 “forgery of data intended to provide proof.”

<sup>46</sup> Cf. Foot Note 3, at point 3 of the German government’s response. Point 2 of the German government’s response to criticism of earlier drafts of the Act also seeks to distinguish “hate crime” from “hate speech”. Contrary to this, ARTICLE 19 considers that “hate speech” may also be “hate crime” where the objective element or *actus reus* of an offence is expressive conduct characterised or motivated by hatred. Further explanation of this point is provided in “Hate Speech Explained: a tool kit”, *op. cit.*, at pages 24 – 27.

- Defamation of religions, religious and ideological associations in a manner that is capable of disturbing the public peace (Section 166 GCC): international human rights law protects people, and not abstract concepts, such as religion or belief systems. It is not a legitimate aim under Article 19(3) of the ICCPR to limit freedom of expression to protect religions or belief from criticism, or to shield followers of a religion or belief from offence. The repeal of blasphemy laws is supported by the HR Committee,<sup>47</sup> as well as by the Rabat Plan of Action,<sup>48</sup> numerous special procedures of the HRC,<sup>49</sup> the Venice Commission,<sup>50</sup> the Parliamentary Assembly of the Council of Europe,<sup>51</sup> and the European Union's Guidelines on Freedom of Religion or Belief.<sup>52</sup>

ARTICLE 19 notes that in 2011, HRC Resolution 16/18 moved past the divisive concept of “defamation of religions” to agree a package of measures to address religious intolerance without reference to this concept, with limitations on expression restricted to the incitement of imminent violence on the basis of religion or belief.<sup>53</sup> Various countries in Europe have decided to repeal their blasphemy laws.<sup>54</sup> We do not consider that the limitation of Section 166 GCC to expressions made “in a manner likely to breach the peace” brings it into compliance with international human rights law. This qualification merely privileges persons likely to respond to legitimate expression with violence or other unlawful conduct, at the expense of the person engaged in legitimate expression (the so-called “assassin's veto”). There is no requirement for intent on the part of the speaker to advocate hatred against persons on the basis of a protected characteristic in a way that constitutes incitement to discrimination, hostility or violence, as per Article 20(2) of the ICCPR.<sup>55</sup>

- Criminal defamation and insult, including insult against a group (Sections 185 – 187 and Section 130(1)(2) of the GCC): ARTICLE 19 has long advocated for the repeal of criminal defamation laws and their replacement with appropriate civil defamation laws, as criminal penalties are a disproportionate means for protecting the reputation rights of individuals from injury.<sup>56</sup> The HR Committee has also called for decriminalisation, and affirmed that imprisonment is never an appropriate penalty for defamation.<sup>57</sup> ARTICLE 19 considers that the right to freedom of expression cannot be limited solely on the basis of protecting an individual or group's subjective feelings, and that insult laws therefore do not pursue a legitimate aim.<sup>58</sup> Furthermore, insult or defamation laws that allow individuals to sue on behalf of a group, which does not itself have legal status, as with Section 130(1)(2), are not justifiable.<sup>59</sup>

<sup>47</sup> HR Committee, General Comment No. 34, *op. cit.*, para. 48.

<sup>48</sup> *Ibid.*

<sup>49</sup> Report of the Special Rapporteur on freedom of religion or belief, A/HRC/34/50, 17 January 2017; Report of the Special Rapporteur on freedom of expression, A/71/33, 6 September 2016; Report of the Special Rapporteur on minority issues, A/HRC/28/64, 2 January 2015; UN Working Group on Arbitrary Detention, Opinion No. 35/2008 (Egypt), 6 December 2008, para. 38. The UN Human Rights Council special procedures have also been supported by regional mandates, see: [Joint Declaration on defamation of religions, and anti-terrorism, and anti-extremism legislation](#), 9 December 2008.

<sup>50</sup> The European Commission for Democracy through Law (the Venice Commission), [The relationship between freedom of expression and freedom of religion: the issue of regulation and prosecution of blasphemy, religious insult and incitement to religious hatred](#), October 2008, para. 89.

<sup>51</sup> Council of Europe Recommendation 1805 (2007), [Blasphemy, religious insults, and hate speech against persons on grounds of their religion](#), 29 June 2007.

<sup>52</sup> EU Guidelines on the Promotion and Protection of Freedom of Religion or Belief, 2013.

<sup>53</sup> ARTICLE 19, [New Guide on Implementing UN HRC Resolution 16-18](#), 27 February 2017.

<sup>54</sup> For example, within the European context, the United Kingdom, Iceland, Norway, and Malta, have all recently repealed criminal prohibitions on blasphemy.

<sup>55</sup> For further analysis of these types of limitations, see Hate Speech Explained: A Tool Kit, *op. cit.*, p. 29 – 31.

<sup>56</sup> ARTICLE 19, Defining Defamation, *op. cit.*, Principle 4.

<sup>57</sup> HR Committee, General Comment No. 34, *op. cit.*, para. 47.

<sup>58</sup> Defining Defamation, *op. cit.*, Principle 2

<sup>59</sup> *Ibid.*, Principle 2(b)(v).

## Germany: The Act to Improve Enforcement of the Law in Social Networks

- Publicly approving of, denying or downplaying international crimes committed under the rule of National Socialism, or approving of, glorifying or justifying National Socialist rule of arbitrary force (Sections 130(3) and (4) GCC): international human rights law does not permit limitations on freedom of expression to protect truth claims around historical events or to limit the expression of opinions on those events, even if they are deeply offensive.<sup>60</sup> As the Special Rapporteur on freedom of expression has found: “by demanding that writers, journalists and citizens give only a version of events that is approved by the Government, States are enabled to subjugate freedom of expression to official versions of events.”<sup>61</sup> For the same reasons as in relation to Section 166 GCC, the qualification that such expression must be “capable of disturbing the public peace”, or “violates the dignity of the victims”, does not bring these provisions into conformity with international human rights law.<sup>62</sup> The Grand Chamber of the European Court of Human Rights has found, in relation to historical memory laws, that criminal convictions for expression where there was no intent to incite violence or hatred, in a context where violence or hatred as a consequence of the expression was also unlikely, meant that the limitation was not necessary and therefore violated Article 10 of the ECHR.<sup>63</sup> In the same decision, however, the Grand Chamber affirmed that the denial or gross trivialisation of the holocaust “must invariably be seen as connoting an antidemocratic ideology and anti-Semitism”, and therefore legitimately subject to restriction, with the historical context of a State seemingly being determinative.<sup>64</sup> Nevertheless, ARTICLE 19 advocates, in line with General Comment No. 34, that absent an express requirement showing that denialism or glorification also constitutes incitement as per Article 20(2) of the ICCPR, such provisions do not meet the necessity requirements of Article 19(3) of the ICCPR.
- Other content restrictions incompatible with international human rights law include prohibitions on: disseminating propaganda or using symbols of unconstitutional organisations, (Sections 86 and 86a GCC); “treasonous forgery”, (Section 100(a) GCC); and “dissemination of depictions of violence” (Section 131). These provisions either do not pursue a legitimate national security interest, and/or do not require a sufficiently direct and immediate connection between the expression and the likelihood or occurrence of violence.<sup>65</sup>

ARTICLE 19 finds that requiring private enterprises to censor content that violates these provisions of the GCC may bring those private enterprises into direct contravention of their responsibility under international human rights law to respect the right to freedom of expression.<sup>66</sup>

Other provisions are too broad, and do not fully comply with the requirements of the ICCPR and the ECHR:

- Disturbing the public peace: Section 130(1) GCC, criminalises anyone who, “in a manner capable of disturbing the public peace”, “incites hatred against a national, racial, religious group or a group defined by their ethnic origins, against segments of the population or individuals because of their belonging to one of the aforementioned groups or segments of the population or calls for violent or arbitrary measures against them.” Since this criminalises conduct that falls short of advocacy of hatred constituting incitement to discrimination, hostility or violence, the criminal sanctions foreseen do not meet the requirements of necessity and proportionality.
- Violations of intimate privacy: While omitted from the English language translation, the Gazetted version of the Act (in German) includes in its listing Section 201a GCC, which

<sup>60</sup> General Comment No. 34, *op. cit.*, para. 49.

<sup>61</sup> Report of the Special Rapporteur on freedom of expression, 7 September 2012, A/67/357, para. 55.

<sup>62</sup> For further analysis of these types of limitations, see Hate Speech Explained: A Tool Kit, *op. cit.*, p. 32 – 34.

<sup>63</sup> Grand Chamber of the European Court, *Perinçek v. Switzerland*, No. 27510/08, 15 October 2015.

<sup>64</sup> *Ibid.*, para. 243.

<sup>65</sup> Johannesburg Principles, *op. cit.*; Principles 2 and 6.

<sup>66</sup> The Ruggie Principles, *op. cit.*

makes “violations of intimate privacy by taking photographs” an offence. ARTICLE 19 considers that privacy offences may constitute legitimate restrictions on the right to freedom of expression if they protect individuals from substantial harm, such as the non-consensual disclosure or distribution of private sexual content, provided that they are narrowly drawn, contain sufficient defences for the protection of freedom of expression, and do not impose disproportionate sanctions.<sup>67</sup> However, there are circumstances, particularly in regards to public figures, where the disclosure and distribution of intimate private photographs may be justified by an overriding public interest.<sup>68</sup> Since the German Criminal Code does not provide defences in that regard, delegating enforcement of this provision raises potential freedom of expression concerns for cases where disclosure or distribution is justified in the circumstances.

ARTICLE 19 also recalls that an explanatory memorandum accompanying an earlier draft version of the Act explained that private enterprises determining the legality of content need only consider the objective elements (*actus reus*) of an offence, and not the mental elements (*mens rea*), or defences/justifications, even though all three would otherwise be determinative of criminal responsibility. It is not clear from the Act whether this is still how private enterprises should approach “unlawful content”, since where the mental element of an offence is not present, or a complete defence applies, that content would in fact be lawful. This ambiguity is likely to further contribute to an over-cautious approach to content moderation by Social Networks.

### Handling complaints about unlawful content

Section 3 of the Act prescribes how Social Networks (as defined in Section 1(1)) will handle complaints about “unlawful content” (as defined in Section 1(3)). Only Social Networks with more than 2 million users are required to comply with this provision.

Section 3 was substantially revised between the draft Bill stage and the adoption of the finalised Act. Those changes include limited exemptions to the 7-day limit for Social Networks to remove “unlawful content”; the specification that any required retention of user-data should respect data protection obligations; and, significantly, the creation of powers for the Federal Office of Justice (*Bundesamt für Justiz*) as an administrative body to “recognise” “self-regulatory institutions” for the review of content moderation decisions. The Federal Office of Justice is the central service authority of the Federal German judiciary.<sup>69</sup>

While the following analysis of Section 3 focuses on process, this does not change the fact that much of the content likely to be at issue under the Act cannot legitimately be restricted by the government under international human rights law, even if this is deputised to private enterprises.

### *Processes for complaining about unlawful content*

Section 3(1) of the Act requires that Social Networks “maintain an effective and transparent procedure for handling complaints about unlawful content” in accordance with the subsequent sub-sections; the procedure must be “easily recognisable, directly accessible and permanently available”.

Encouraging the institution of accessible, transparent and usable complaints procedures integrated to websites regarding unlawful content is a positive aim. These are spelled out more clearly in Section 3(2); they include:

- The requirement that Social Networks acknowledge without undue delay receipt of a complaint (Section 3(2)(1));

<sup>67</sup> ARTICLE 19, [Global Principles on Freedom of Expression and Privacy](#), March 2017; at Principle 3(b) .

<sup>68</sup> *Ibid.* at Principles 12 and 13.

<sup>69</sup> [Federal Office of Justice](#) is also port of call for international legal transactions and performs other functions.



## Germany: The Act to Improve Enforcement of the Law in Social Networks

- The requirement that Social Networks inform both the complainant and the author of the content in issue of any decision reached, as well as the reasons for the decision made (Section 3(2)(5));
- That training and counselling be provided to content moderators in German (Section 3(4)).

Online intermediaries, in particular Social Networks, too infrequently provide this level of information and transparency to complainants or to users whose content is subject to removal. However, it is potentially disproportionate to compel these practices through the threat of severe administrative penalties (Section 4). It is also concerning that the Act does not require Social Networks to provide to users the means to challenge content-removal decisions, for example on the basis that a removal violated the user's right to freedom of expression.

ARTICLE 19 further notes that the vast majority of intermediaries that moderate content shared by users on their platforms make removal decisions on the basis of their terms of service, and not the national laws where a complainant is situated (unless a request is received by a national law enforcement authority requesting removal on this basis). It is unclear whether to comply with Section 3(1), a Social Network would need to amend their reporting mechanisms and supplement their terms of service to provide their users in Germany new categories for “violating content” expressly tied to the specified provisions of the GCC. Several of these provisions, for example on “defamation of religions”, would not necessarily find corresponding restrictions in a Social Network's existing terms of service. How Social Networks should ensure the “accessibility” of complex matters of German Criminal law within amended or supplementary terms of service is unclear.

Moreover, the process prescribed in the Act does not acknowledge more innovative and alternative approaches to dealing with harmful content online that some Social Networks have advanced in recent years. These include, for example, functions that allow a complainant to initiate a conversation with the author on why content is harmful and should be removed voluntarily, or the availability of end-user controls for a user to block or silence content they find upsetting (but should not necessarily be blocked or removed). Prescribing processes that rely solely on take-down or blocking under threat of administrative sanctions, provides little flexibility for other forms of dispute resolution, and may even inhibit innovation in this space. These alternatives should be encouraged: they can be less intrusive to fundamental rights; they may provide users with protective mechanisms more tailored to their needs; and, they may provide a forum for more effectively challenging (and changing) unlawful or harmful online behaviour.

The impact of Section 1(3) of the Act is therefore still likely to discourage intermediaries from allowing user engagement on their platforms, and therefore chill freedom of expression, rather than incentivise more user-friendly complaint procedures.

### *Removal and blocking*

The Act requires Social Networks to adopt a particular approach to removal and blocking of unlawful content in response to user-generated complaints or complaints sent by other bodies; there is slightly more detail built into these requirements than in previous draft versions of the law, as well as some notable differences.

Section 3(2)(2) and Section 3(2)(3) of the Act require Social Networks to block or delete “content that is manifestly unlawful” within 24 hours of receipt (or a longer term to be determined between a Social Network and relevant law enforcement authority), and “unlawful content” within 7 days of receipt (describing this, somewhat incongruously, as “immediate” in the English language version).

Two requirements included in earlier draft versions of the Act have been removed, including that the Social Network should block or delete “copies” of unlawful content (previously Section

## Germany: The Act to Improve Enforcement of the Law in Social Networks

3(2)(6)), and the obligation to institute measures to prevent future uploads of the same content (previously Section 3(2)(7)). This is welcome, as these provisions would likely have been incompatible with the European Union's E-Commerce Directive, Articles 14(1) and 5.

At the same time, the Act introduces two new exemptions to the application of the 7-day time limit for “unlawful content” (as opposed to “manifestly unlawful content”, which must always be removed within 24 hours):

- Section 3(2)(3)(a) allows for a longer period (of unspecified duration) where the determination of legality “depends on the falsity of a factual allegation or is clearly dependent on other factual circumstances.” It requires the user (presumably the author of the content) to respond to the complaint before the decision is made.
- Section 3(2)(3)(b) allows the decision regarding unlawfulness to be referred to a “self-regulatory institution” that is recognised by the Federal Office for Justice in accordance with the requirements set out in the Act (see below). To receive and retain its recognition, any “self-regulatory institution” is required to guarantee that referrals will be considered within 7 days, and the Social Network is obliged to adhere to their decision to benefit from the exemption.

Sections 4(1)(2) and 4(1)(3) of the Act make it an administrative offence for a Social Network to fail to “provide, to provide correctly or to provide completely” a procedure in conformity with the requirements of Section 3(1). Section 4(2) retains administrative penalties of up to 5 million Euros for such violations. However, the potential application of the Act on Regulatory Offences enables the fine to be multiplied ten-fold to 500 million Euros. Section 3(5) provides that a Social Network's procedure under Section 3(1) may be subject to “monitoring” by an agency tasked to do so by the Federal Office for Justice.

To impose liability for failures to remove or block content, the Federal Office of Justice must first seek a decision of the Administrative Court which deals with objections to regulatory fines (Section 4(5)). This process requires the administrative body to furnish the Administrative Court with a statement by the Social Network, but specifically excludes the possibility for any oral hearing, and specifies that the decision of the Administrative Court “shall not be contestable” (i.e. there is no availability of appeal). The procedure does not allow for the inclusion as parties of either the user(s) whose “unlawful content” is at issue or the complainant.

Additionally, Section 5 requires Social Networks to name a person authorised to receive service in relation to procedures under section 4 or in judicial proceedings, and this person is obliged to respond to requests from law enforcement within 48 hours of receipt.

ARTICLE 19 finds that the requirement in the Act that Social Networks remove “unlawful content” on the basis of notice provided by any party, without a court order, is against international freedom of expression standards. This type of duty has been criticised by the UN and regional freedom of expression mandates as an attempt to privatise controls over public communications.<sup>70</sup>

The following criticisms exemplify why this approach to intermediary liability should be avoided, and why the Act is so fundamentally flawed from a freedom of expression perspective:

- The threshold for liability is unclear: it is difficult to determine how a Social Network will be assessed to have “not provided”, “not provided correctly” or “not completely provided” a procedure in compliance with the Act's requirements. If a small number of failures to remove unlawful content while still having the correct procedures in place does not attract liability, then the proportion of failures in relation to well-founded complaints that must be shown is not clear. While statistics would be available through quarterly reports furnished by Social Networks under Section 2, these would not illuminate the proportion of complaints

<sup>70</sup> Joint Declaration on Freedom of Expression and “Fake News”, Disinformation and Propaganda, 2017, *op. cit.*

that concerned content actually found by an independent court to be illegal, but instead would only show the determinations made by the Social Network through their own assessment. There is consequently considerable legal uncertainty about the point at which liability is engaged.

- Private enterprises should not be arbiters of legality in the exercise of fundamental rights: Section 3(2)(2) and Section 3(2)(3) require private actors to make their own assessment of the legality of content following notification from parties who similarly aren't competent to make authoritative determinations on the legality of content. International standards require that liability only be imposed where an intermediary refused to comply with an order from an independent, impartial, authoritative oversight body, such as a court. The threat of sanctions incentivises intermediaries to be over-cautious and err towards removal or blocking (or remove functions for third party comments entirely) to avoid liability. While the addition of Sections 3(2)(3)(a) and (b) seek to provide further safeguards in this respect, they also reveal how complex and unrealistic the requests on Social Networks to determine legality are. As explained below, the “regulated” “self-regulatory” institutions are also insufficient.
- **The distinction between “manifestly unlawful” and “unlawful” is not clear, and difficult for a private enterprise to accurately predict:** the Act provides no guidance to Social Networks on how they should differentiate “manifestly unlawful content” from “unlawful content”, and contains no duty on the part of Social Networks to consider user’s rights to freedom of expression when making these determinations (even though law enforcement authorities are required to consider this when acting to restrict freedom of expression pursuant to the GCC). Private companies are clearly not well placed to be making complex legal and factual determinations on the interpretation of the German Criminal Code, which require taking into account the content of expression, the mental state of the person publishing the content, the context in which that information is shared, and considerations of the impact of that expression on other individual’s rights or broader State interests in national security or public order. The addition of Section 3(2)(3)(a) acknowledges that complexity, but does not provide a solution to the core problem. The majority of these factors cannot be determined by a facial assessment of the content at issue (with the exception of child abuse images), requiring further investigatory work that are beyond the capabilities or competence of any intermediary to do accurately, notwithstanding the addition of Section 3(2)(3)(a) and the requirement for Social Networks to engage in investigatory work themselves.
- Process for administrative review lacks adequate safeguards: while the enforcement of fines for failure to remove or block content requires judicial authorisation under Section 4(4), the due process guarantees therein are inadequate, and provide no protection for good faith decisions to not remove or block content (including where the intermediary considers their responsibility to respect freedom of expression requires such constraint). Similarly, there is no exemption from liability where a Social Network moves promptly to remove content following the decision of the Administrative Court. These factors are likely to increase the tendency among Social Networks to be over-cautious when making content-removal decisions.
- No mechanisms for users whose content is removed to assert their rights: Section 3(2) and Section 4 provide no recourse for users whose lawful content is removed pursuant to the Act. Administrative authorities are only required to intervene in relation to failures to remove or block content, and no process is foreseen for users to complain about or seek reviews of illegitimate removal or blocking decisions taken on the basis of the Act.
- Proportionality of sanctions: the sanctions foreseen under Section 4(2) of up to 5 million Euros, potentially multiplying by ten to 500 million Euros through the application of the Act on Regulatory Offences, are grossly disproportionate. The independence of the Administrative Body decision-making processes from political influence is a particular

concern, and is not safeguarded through the limited judicial supervision provided by the Administrative Court in Section 4(5).

- Arbitrary audits: the specification in Section 3(5) that a Social Network's procedure under Section 3(1) may be subject to "monitoring" by an (unspecified) agency tasked with this responsibility by the Federal Office for Justice is unclear. On what basis monitoring can be ordered (e.g. in response to an alleged breach of the Act, rather than as a fishing expedition), and what the scope and duration of that audit might be, is unclear.

### ***"Recognition" for "self-regulatory" institutions***

The Act provides the Federal Office for Justice with powers to recognise "self-regulatory institutions" (Section 3(5)(7)). While the Act does not require Social Networks to create, join or fund "self-regulatory institutions", the framework the Act establishes incentives for Social Networks to do so, and to seek recognition for these institutions from the Federal Office for Justice.

On terminology, there are notable differences between the English translation of the Act and the official German Act: in the English, what is translated as "self-regulation" should more accurately be translated from the German as "regulated self-regulation."<sup>71</sup> The official Act is therefore more transparent in revealing the framework for public supervision of content moderation that the Act advances. While described as "self-regulation", the framework of supervised content moderation advanced in the Act departs quite significantly from what ARTICLE 19 considers to be effective "self-regulation" compatible with international human rights law (see above, at pages 11 – 12).

Section 3(5)(6) of the Act sets out five criteria for the Federal Office for Justice to consider when conferring recognition for the purposes of Act. These include: (i) the independence and expertise of its analysts; (ii) that it can guarantee a decision on referred questions of law within seven days; (iii) that it provides rules and procedures for the scope and structure of analyses, for submissions requirements, and the possibility of review; (iv) the establishment of a "complaints service"; and, (v) a funding stream from "several" Social Networks or "establishments", with a membership open to the admission of new providers.

The Federal Office for Justice not only has the power to recognise the "self-regulatory institutions", but also to wholly or partially withdraw that recognition where the listed criteria are no longer met (Section 3(5)(8)), or to temporarily bar Social Networks from making referrals to a "self-regulatory institution".

Section 3(2)(3)(b) of the Act reveals the main incentive for Social Networks to establish, join and seek recognition for a "self-regulatory institution". Social Networks may defer to these institutions decisions regarding the take-down of "unlawful content" (but not "manifestly unlawful content") and thereby extend the required deadline for removals by 7 days. The Social Networks are then required to adhere to the decision of the "self-regulatory" institution.

Other functions of the "self-regulatory" institutions, revealed by the criteria that they must meet to receive official recognition, are less clear. They are required to have in place "the possibility to review decisions" (Section 3(6)(3)), but it is not apparent which decisions this refers to, and on the basis of a request from whom (e.g. from the administrative authority, the Social Network itself, or the original complainant or author). They are also required to have in place a "complaints service" (Section 3(6)(4)), but it is not apparent what complaints the "complaints service" will deal with, i.e. who may submit complaints, and in relation to what conduct or decision by whom.

ARTICLE 19 considers that the framework for "recognition" of "self-regulated" or "regulated self-regulated" institutions does not comply with international human rights law.

<sup>71</sup> See: Section 2(7); Section 3(2)(3)(b); Section 3(5)(6); Section 3(7); Section 3(9).

## Germany: The Act to Improve Enforcement of the Law in Social Networks

There is nothing in principle that is concerning about Social Networks establishing external and independent self-regulatory mechanisms, to pool resources and expertise in how they moderate their platforms in accordance with their own terms of service or in response to lawful requests from States. Assuming that any self-regulatory mechanism created by Social Networks would (a) be open to participation from civil society, (b) be transparent, and (c) be accountable to the public, this may provide opportunities for achieving greater fairness and transparency in Social Networks' operations, while also building public confidence in them as platforms for sharing and receiving information. A truly self-regulatory approach may also allow for a greater emphasis on how Social Networks fulfil their responsibilities to respect human rights, including both the rights to equality and freedom of expression, and therefore include civil society and academic experts on human rights and freedom of expression. This may lead to the development of alternatives to content takedown that are more effective and satisfactory for complainants and authors of unlawful content.

However, ARTICLE 19 believes that the approach in the Act is unlikely to work in this direction. The framework of “regulated self-regulation” was clearly added as an afterthought, seemingly in response to criticism of earlier drafts, but does not remedy the weaknesses identified above.

There are several general issues with the framework as adopted:

- It does not provide an alternative model for resolving disputes, but instead merely delays by a matter of days, for a narrow category of “unlawful content”, the engagement of the punitive administrative law process. Though this provides time for a secondary examination of legality, content is still removed without a determination of legality by a judicial body, and the deadlines provided are clearly unreasonable.
- The problems with the system for removal and blocking listed in the previous subsection therefore all still apply, with the regulatory system still grounded in the German Criminal Code (rather than an ethical code defined and agreed to by the self-regulatory body), and the harsh sanctions regime still creating an incentive structure that favours content removals without safeguards for the protection of freedom of expression.
- The power of “recognition” is a significant one, and is vested in a body that is not independent from political influence but potentially beholden to it. The power for recognition to be revoked or to suspend the privileges Social Networks receive by using this mechanism, in particular, further incentivises any “regulated self-regulatory” institution to err on the side of censorship rather than protect freedom of expression. Any recognition body must be wholly independent from the State, and should include representation from all relevant stakeholders from within industry and civil society.

The powers of recognition as set out in the Act would have to be fundamentally altered in order to comply with international human rights law. In particular, the framework would need to specify:

- Detailed and objective criteria for “independence and expertise” that members of a self-regulatory mechanism should be qualified with, in order to prevent the abuse of ambiguous criteria;
- That analysts be appointed through a fully consultative and inclusive process, which is democratic and transparent;
- That the self-regulatory mechanism's membership be comprised of individuals from within industry and civil society;
- That the self-regulatory mechanism be tasked with identifying and setting out its own ethical standards, and seek through its procedures for its industry members to adhere with those ethical standards, rather than for a self-regulatory mechanism seek to enforce compliance with the law;

## Germany: The Act to Improve Enforcement of the Law in Social Networks

- That any rules and procedures that analysts or self-regulatory body members are required to follow must comply with basic principles of due process;
- That adequate time be given to consider fully any individual complaint, which should be more than seven days;
- Powers for the self-regulatory mechanism to require its members abide with their decisions, including non-pecuniary remedies;
- That the self-regulatory body is adequately and sustainably funded, to ensure that they are able to achieve their mission;
- That the power of recognition be given to a body insulated from political and other interferences, rather than entrusted to an administrative body.

Given that the framework proposed in the Act is so fundamentally at odds with these requirements and safeguards, ARTICLE 19 considers that its openness to unfair decisions and political influence, while failing to provide true accountability to the public, makes it incompatible with international human rights law.

### Reporting requirements

Section 2(1) of the Act requires Social Networks that receive more than 100 complaints per calendar year about “unlawful content” to provide quarterly reports on their handling of complaints.

Section 2(2) details the information these reports must contain, including general information about the processes in place and specific resources allocated to content moderation, numbers of complaints received (disaggregated on basis of complaint, and whether received from a public authority or private party), and on the outcome of decisions (disaggregated in the same way, with the time taken to delete or block), whether both complainant and author were informed of the outcome, as well as whether a referral was made to a “regulated self-regulatory” institution.

Failure to meet reporting requirements is an administrative offence under Section 4 of the Act, punishable with fines up to 500,000 Euros. Again, the Act on Regulatory Offences allows this to be multiplied ten-fold to 50 million Euros.

ARTICLE 19 considers that intermediaries should disclose these types of information, in particular in relation to removal or blocking requests made by States, and opposes any law that seeks to limit the disclosure of such information. Compelling disclosure is in the public interest, and in ARTICLE 19's view should extend to content removal and blocking practices broadly (for example, in relation to content removed for violating terms and conditions which is otherwise lawful in the country concerned).

### Individual responsibility for offences committed online

ARTICLE 19 is concerned that the Act focuses almost exclusively on the administrative liability of Social Networks seemingly as an alternative to seeking accountability of the authors of “unlawful content”. The removal of such content is unlikely to affect users' behaviour or deter the posting of further unlawful content, and will not impact the underlying causes of criminal behaviour, in particular that which is motivated by hatred or bias against particular groups. Given that there are alternative and more effective means of achieving this policy objective, the necessity of the Act, with its dominant focus on coercive administrative sanctions is questionable.

## Germany: The Act to Improve Enforcement of the Law in Social Networks

### Amendments to the Telemedia Act

Article 2 of the Act amends the Telemedia Act (Telemediengesetz) (which transposes the E-Commerce Directive into German law) to expand the list of grounds on which law enforcement can request users' inventory data from "service providers". This new ground includes inventory data "necessary for the enforcement of civil law claims arising from the violation of absolutely protected rights by unlawful content as defined in the Network Enforcement Act." Addressing previous criticism, any such disclosure must be on the basis of a court order requested by the injured party through a regional court; it allows for the Social Network to be involved as an interested party in the proceedings.

## Conclusion

---

ARTICLE 19 finds the Act, taken overall, to be dangerous to the protection of freedom of expression in Germany, and we are particularly concerned that countries with much weaker institutional and legal safeguards for the protection of human rights are looking at this Act as a model for increasing intermediary liability.

For these reasons, we recommend that the German Parliament immediately consider the repeal of the Act, while simultaneously taking measures to repeal provisions of the German Criminal Code that do not comply with international human rights law on freedom of expression.



## About ARTICLE 19

---

ARTICLE 19 advocates for the development of progressive standards on freedom of expression and freedom of information at the international and regional levels, and their implementation in domestic legal systems. The Law Programme has produced a number of standard-setting publications which outline international and comparative law and best practice in areas such as defamation law, freedom of expression and equality, access to information and broadcast regulation.

On the basis of these publications and ARTICLE 19's overall legal expertise, the organisation publishes a number of legal analyses each year, comments on legislative proposals as well as existing laws that affect the right to freedom of expression. This analytical work, carried out since 1998 as a means of supporting positive law reform efforts worldwide, frequently leads to substantial improvements in proposed or existing domestic legislation. All of our analyses are available at <http://www.article19.org/resources.php/legal>.

If you would like to discuss this analysis further, or if you have a matter you would like to bring to the attention of the ARTICLE 19 Law Programme, you can contact us by e-mail at [legal@article19.org](mailto:legal@article19.org).

For more information about the ARTICLE 19's work in Europe and Central Asia, please contact Katie Morris, Head of Europe and Central Asia at ARTICLE 19, at [katie@article19.org](mailto:katie@article19.org).