



**IIEA**

Sharing Ideas  
Shaping Policy

# Regulating

# Hate Speech:

The European Digital Services Act  
and Ireland's Hate Crimes Bill



Seamus Allen





# 1. Introduction

Today there is significant anxiety regarding the role that the internet can play in facilitating the dissemination of harmful content, such as disinformation, terrorist content, and child abuse material.<sup>1</sup> One of the most prominent and problematic categories of harmful content is hate speech. Hate speech has always been a subject that has caused controversy and generated debate. However, the growing migration of public discourse to the online world, and the perception that online media can amplify hate speech, is triggering new levels of concern.

There is thus a growing consensus that more action should be taken to address the dissemination of harmful online content, including hate speech.<sup>2</sup> It is against this background that the European Commission proposed the Digital Services Act (DSA) in December 2020<sup>3</sup> which would impose common rules regarding how online platforms should combat illegal and harmful content, including illegal hate speech.

Hate speech is a term that is associated with speech that expresses an attitude of hatred, particularly towards certain demographic groups based on characteristics such as ethnicity, religion, race or sexual orientation among others.<sup>4</sup> However, there is no internationally accepted definition of hate speech<sup>5</sup> and there is currently

no EU definition of the term.<sup>6</sup> It is widely agreed - including amongst policymakers, freedom of speech activists, civil liberties groups and other stakeholders - that direct incitements to violence and the bullying or harassment of individuals should usually be legally prohibited and there is relatively little controversy over the essential principles underpinning such regulations. While these issues are outside the scope of this paper, it is worth noting that an Online Safety Bill has been proposed in Ireland to tackle online bullying and harassment.<sup>7</sup> In contrast, hate speech laws can apply to a much broader range of content, including generalised expressions of opinion. This can potentially pose risks to freedom of expression, and it is here where significant public policy controversies can arise, particularly regarding how hate speech should be defined.

It is therefore significant that the DSA does not provide a definition of hate speech. Instead hate speech will be defined by each individual EU Member State, although it should be noted that the European Commission proposed an initiative to extend the list of EU crimes to include hate speech and hate crime in December 2021.<sup>8</sup> Notably, Ireland's government published the general scheme of a new Hate Crimes Bill (HCB) in April 2021 which would introduce new provisions regarding the crime of incitement to hatred in Ireland.<sup>9</sup> This proposal will replace

<sup>1</sup> Jacob Mchangama, *Free Speech: A Global History From Socrates to Social Media*, (Basic Books, London, 2022), p. 357 - 365

<sup>2</sup> Mchangama, *Free Speech: A Global History From Socrates to Social Media*, p. 357 - 365

<sup>3</sup> The Digital Services Act: ensuring a safe and accountable online environment | European Commission (europa.eu), [https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment\\_en](https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment_en)

<sup>4</sup> Michael Herz and Peter Molnar, (editors), *The Content and Context of Hate Speech: Rethinking Regulation and Responses*, (Cambridge University Press, New York, 2012), p. 1-7

<sup>5</sup> UN Strategy and Plan of Action on Hate Speech 18 June SYNOPSIS.pdf, <https://www.un.org/en/genocideprevention/documents/UN%20Strategy%20and%20Plan%20of%20Action%20on%20Hate%20Speech%2018%20June%20SYNOPSIS.pdf>

<sup>6</sup> The Commission proposes to extend the list of 'EU crimes' (europa.eu), [https://ec.europa.eu/commission/presscorner/detail/en/ip\\_21\\_6561](https://ec.europa.eu/commission/presscorner/detail/en/ip_21_6561)

<sup>7</sup> gov.ie - Publication of the Online Safety and Media Regulation Bill (www.gov.ie), <https://www.gov.ie/en/publication/88404-publication-of-the-online-safety-and-media-regulation-bill/>

<sup>8</sup> The Commission proposes to extend the list of 'EU crimes' (europa.eu), [https://ec.europa.eu/commission/presscorner/detail/en/ip\\_21\\_6561](https://ec.europa.eu/commission/presscorner/detail/en/ip_21_6561)

<sup>9</sup> gov.ie - Tough sentences for hate crimes under new Bill from Minister McEntee (www.gov.ie), <https://www.gov.ie/en/press-release/48cb5-tough-sentences-for-hate-crimes-under-new-bill-from-minister-mcentee/>



Ireland's extant Prohibition of Incitement to Hatred Act (1989). This proposed update is partly prompted by the specific challenges relating to online hate speech, and partly by a desire to broaden the scope of hate speech legislation.<sup>10</sup> Crucially, the HCB will determine how the DSA will be applied in Ireland for the purpose of countering hate speech.

This paper will assess the risks that this legislation – the DSA and the HCB – could unintentionally result in the censoring of legitimate, non-hateful content. This paper thus presents Ireland as a case study of how the DSA can interact with the existing hate speech legislation of individual EU Member States and considers the implications this could have for freedom of expression. It should be noted that while a provisional political agreement was reached between the Council of the European Union and the European Parliament on the DSA on 23 April 2022, the finalised agreement and legal text is not yet publicly available at the time of writing.<sup>11</sup> Furthermore, the general scheme of the HCB proposed in April 2021 will be followed by an updated bill which will then be presented to the Oireachtas – the Irish Houses of Parliament – in summer 2022 and it may yet undergo significant change.<sup>12</sup> As the DSA is close to being finalised, Irish policymakers will have the opportunity to ensure that the HCB is adapted in light of the changing European legislative landscape.

This paper first introduces the main features of the DSA proposal. The following

section outlines the main features of Ireland's proposed HCB general scheme. This paper then explores if the proposed HCB legislation could risk criminalising or censoring legitimate, non-hateful material. Finally, the last section will explore the potential implications of the combination of the HCB and DSA for online content moderation. Ultimately, this paper will assess whether the HCB and DSA could interact in ways that pose risks to freedom of expression, including for important contributions to public discourse and for those trying to speak out against hatred.

## 2. The Digital Services Act

The DSA proposal presented by the European Commission in December 2020 sets out the rules regarding how intermediary services, including online platforms, must address illegal and harmful online content, including illegal hate speech.<sup>13</sup> The DSA introduces due diligence obligations for online platforms which must put in place robust systems to moderate illegal online content. Under the DSA, platforms are mandated to deploy "*the necessary means to diligently mitigate*" such risks.<sup>14</sup> The DSA imposes extra obligations on the very large online platforms (i.e. those with more than 45 million users),<sup>15</sup> including annual risk assessments regarding how their platforms may be used for the dissemination of illegal content.<sup>16</sup> Under the DSA, large platforms

<sup>10</sup> gov.ie – Tough sentences for hate crimes under new Bill from Minister McEntee (www.gov.ie), <https://www.gov.ie/en/press-release/48cb5-tough-sentences-for-hate-crimes-under-new-bill-from-minister-mcentee/>

<sup>11</sup> Digital Services Act: Council and European Parliament provisional agreement for making the internet a safer space for European citizens – Consilium (europa.eu), <https://www.consilium.europa.eu/en/press/press-releases/2022/04/23/digital-services-act-council-and-european-parliament-reach-deal-on-a-safer-online-space/>

<sup>12</sup> Proposal for a Council Decision on Hate Speech and Hate Crime: Motion – Seanad Éireann (26th Seanad) – Thursday, Mar 2022 – Houses of the Oireachtas, <https://www.oireachtas.ie/en/debates/debate/seanad/2022-03-10/9/>

<sup>13</sup> The Digital Services Act: ensuring a safe and accountable online environment | European Commission (europa.eu), [https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment\\_en](https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment_en)

<sup>14</sup> Digital Services Act, European Commission, p. 32, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825&from=en>

<sup>15</sup> Digital Services Act, European Commission, p. 31, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825&from=en>

<sup>16</sup> The Digital Services Act: ensuring a safe and accountable online environment | European Commission (europa.eu), [https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment\\_en](https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment_en)





are also mandated to consider measures to mitigate the dissemination of illegal content by adapting their content moderation systems, algorithmic recommender systems, and online interfaces.<sup>17</sup>

Platforms must also introduce processes for notice-and-action mechanisms, by which platforms are legally obliged to respond in a timely fashion to materials that are reported by users, including private citizens, as being illegal.<sup>18</sup> A system of “trusted flaggers” must also be established, consisting of organisations which will be treated with priority when they report content that they deem to be illegal.<sup>19</sup> It is notable that the original DSA proposal mandates that online platforms must act “without undue delay”<sup>20</sup> and “expeditiously” when they become aware of illegal content.<sup>21</sup> Platforms must also regularly report the “average time” it takes the platform to remove illegal content in regularly published transparency reports.<sup>22</sup> More recently, the Council of the European Union has proposed amendments to the draft legislation which would require large platforms to generally react to notifications of illegal hate speech within 24 hours.<sup>23</sup> Large platforms that are deemed to have failed to adequately fulfil their obligations under the DSA could face fines of up to 6% of their annual turnover.<sup>24</sup>

The DSA contains explicit protections for freedom of expression, by establishing, for example, processes through which users can contest the removal of content they have created.<sup>25</sup> However, while the DSA mandates a role for organisations dedicated to reporting material that should be removed (i.e. trusted flaggers) there is no equivalent role for specialist organisations to object to the wrongful removal of material. There is also no explicit protection for content that violates the law of a Member State or a company’s terms and conditions in cases where a citizen may argue that these themselves threaten freedom of expression.

The risk of significant fines and the time pressures involved may mean that companies could err on the side of removing content that could even potentially be deemed to be hate speech, even when such material is arguably legitimate and non-hateful. It also means that platforms will likely rely on automated algorithms to detect and remove hate speech, even though such algorithms can make errors a human would be unlikely to make.

With regards to enforcement, the political provisional agreement on the DSA of April 2022 between the Council and Parliament envisages that for very large online platforms and very large online

<sup>17</sup> *Digital Services Act*, European Commission, p. 32, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825&from=en>

<sup>18</sup> At a glance: Does the EU Digital Services Act protect freedom of expression? - ARTICLE 19, <https://www.article19.org/resources/does-the-digital-services-act-protect-freedom-of-expression/>

<sup>19</sup> *Digital Services Act*, European Commission, p. 56, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825&from=en>

<sup>20</sup> *Digital Services Act*, European Commission, p. 47,51, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825&from=en>

<sup>21</sup> *Digital Services Act*, European Commission, p. 22,46,47, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825&from=en>

<sup>22</sup> *Digital Services Act*, European Commission, p. 50-51, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825&from=en>


<sup>23</sup> Luca Bertuzzi, DSA: EU ambassadors reach agreement to start interinstitutional negotiations – EURACTIV.com, <https://www.euractiv.com/section/digital/news/dsa-eu-ambassadors-reach-agreement-to-start-interinstitutional-negotiations/>

<sup>24</sup> *Digital Services Act*, European Commission, p. 80, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825&from=en>

<sup>25</sup> *Digital Services Act*, European Commission, p. 54, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825&from=en>







search engines the European Commission will have the primary responsibility for enforcement for the obligations specific to these companies.<sup>26</sup> Smaller platforms will be regulated by the Digital Services Coordinator established in their respective EU Member State.

Significantly, the DSA does not define what constitutes “illegal hate speech” and, as noted above, there is currently no official EU definition for the term. The original DSA proposal states that:

*“for the purpose of this Regulation the concept of “illegal content” should be defined broadly and also covers information relating to illegal content, products, services and activities....it is immaterial whether the illegality of the information or activity results from Union law or from national law that is consistent with Union law and what the precise nature or subject matter is of the law in question.”<sup>27</sup>*

As previously mentioned, there is currently no official EU definition of hate speech, although the European Commission has proposed to extend the list of EU crimes to include hate speech and hate crime.<sup>28</sup> This means that the DSA will be applied within EU Member States in accordance with national hate speech laws. Therefore, across the EU, one piece of legislation - the DSA - will establish the rules by which online platforms must address hate speech, while national legislation will define what constitutes hate speech. Thus, the DSA may interact with national hate speech laws in

ways that could have consequences that were not intended by the creators of either piece of legislation. Given this, this paper provides a case study of how the DSA may interact with the hate speech legislation of a particular EU Member State, in this case Ireland’s proposed Hate Crimes Bill (HCB).

### 3. Ireland’s Proposed Hate Crimes Bill

As mentioned, currently the primary piece of hate speech legislation in force in Ireland is the Prohibition of Incitement to Hatred Act (1989).<sup>29</sup> It is intended that this legislation will be replaced by the HCB, the general scheme of which was published in April 2021.<sup>30</sup> An updated version of this bill is due to be published in summer 2022.<sup>31</sup> There are two main parts to the HCB. One part addresses how penalties for already existing crimes will be increased when these crimes are found to be motivated by hatred. The other part of the bill introduces new measures on incitement to hatred, which is the part of the bill that is explored in this paper. The HCB states:<sup>32</sup>

*“(1) A person is guilty of an offence who – communicates to the public or a section of the public by any means, for the purpose of inciting, or being reckless as to whether such communication will incite, hatred against another person or group of people due to their real or perceived association with a protected characteristic.”*

<sup>26</sup> Digital Services Act: Council and European Parliament provisional agreement for making the internet a safer space for European citizens - Consilium (europa.eu), <https://www.consilium.europa.eu/en/press/press-releases/2022/04/23/digital-services-act-council-and-european-parliament-reach-deal-on-a-safer-online-space/>

<sup>27</sup> Digital Services Act, European Commission, p. 20, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825&from=en>

<sup>28</sup> The Commission proposes to extend the list of ‘EU crimes’ (europa.eu), [https://ec.europa.eu/commission/presscorner/detail/en/ip\\_21\\_6561](https://ec.europa.eu/commission/presscorner/detail/en/ip_21_6561)

<sup>29</sup> Prohibition of Incitement To Hatred Act, 1989 (irishstatutebook.ie), <https://www.irishstatutebook.ie/eli/1989/act/19/enacted/en/html>

<sup>30</sup> gov.ie - Tough sentences for hate crimes under new Bill from Minister McEntee (www.gov.ie), <https://www.gov.ie/en/press-release/48cb5-tough-sentences-for-hate-crimes-under-new-bill-from-minister-mcentee/>

<sup>31</sup> gov.ie - Dáil statement by Minister for Justice Helen McEntee on violence against women (www.gov.ie), <https://www.gov.ie/en/speech/4d9d1-dail-statement-by-minister-for-justice-helen-mcentee-on-violence-against-women/>

<sup>32</sup> General Scheme Criminal Justice (Hate Crime) Bill 2021, Department of Justice, p. 4-5 [https://www.justice.ie/en/JELR/General\\_Scheme\\_Criminal\\_Justice\\_\(Hate\\_Crime\)\\_Bill\\_2021.pdf/Files/General\\_Scheme\\_Criminal\\_Justice\\_\(Hate\\_Crime\)\\_Bill\\_2021.pdf](https://www.justice.ie/en/JELR/General_Scheme_Criminal_Justice_(Hate_Crime)_Bill_2021.pdf/Files/General_Scheme_Criminal_Justice_(Hate_Crime)_Bill_2021.pdf)



The bill also states that:<sup>33</sup>

*“hatred” means detestation, significant ill will or hostility, of a magnitude likely to lead to harm or unlawful discrimination against a person or group of people due to their association with a protected characteristic”*

and,

*“protected characteristic” means race; colour; nationality; religion, ethnic or national origin; sexual orientation; gender; or disability”*

A key point to note is that “reckless” incitement to hatred will now be criminalised. Notes in the bill state that recklessness means:

*“where the person is aware that there is a significant risk that the communication will incite hatred.”*

The Department of Justice website elaborates that recklessness means a person must:<sup>34</sup>

*“at the very least have considered whether what they were doing would incite hatred, concluded that it was significantly likely, and decided to press ahead anyway.”*

As will be argued, this approach to understanding what constitutes “reckless” incitement to hatred may have a broader reach than was intended by the bill’s authors. It should be noted that the HCB

also prohibits incitement to hatred against persons who may only have an “association” with a protected characteristic. For example, the Government’s public consultation report states that this would apply to workers associated with the trans community or against persons who may be rightly or wrongly associated with a particular religion due to the clothing they wear.<sup>35</sup> The Government’s report states that “advocating unlawful discrimination” can include the “spreading of derogatory stereotypes or tropes” or “false statements about groups with certain characteristics.”<sup>36</sup>

### 3.1 Disseminating Hateful Content: The Risk of Double Standards

The HCB proposal makes it an offence when a person “publishes or otherwise disseminates, broadcasts or displays to the public”<sup>37</sup> a communication that is deemed an incitement to hatred. However, there will be exceptions to this provision, including situations in which:<sup>38</sup>

*“the material concerned consisted solely of – a reasonable and genuine contribution to literary, artistic, political, scientific, or academic discourse”*

or when it is:

*“material which has a certificate from the authorising body, in the case of a film or book”.*

<sup>33</sup> General Scheme Criminal Justice (Hate Crime) Bill 2021, Department of Justice, p. 4, [https://www.justice.ie/en/JELR/General\\_Scheme\\_Criminal\\_Justice\\_\(Hate\\_Crime\)\\_Bill\\_2021.pdf/Files/General\\_Scheme\\_Criminal\\_Justice\\_\(Hate\\_Crime\)\\_Bill\\_2021.pdf](https://www.justice.ie/en/JELR/General_Scheme_Criminal_Justice_(Hate_Crime)_Bill_2021.pdf/Files/General_Scheme_Criminal_Justice_(Hate_Crime)_Bill_2021.pdf)

<sup>34</sup> gov.ie – Tough sentences for hate crimes under new Bill from Minister McEntee (www.gov.ie), <https://www.gov.ie/en/press-release/48cb5-tough-sentences-for-hate-crimes-under-new-bill-from-minister-mcentee/>

<sup>35</sup> Legislating for Hate Speech and Hate Crime in Ireland: Report on the Public Consultation 2020, Department of Justice, p. 43, [https://www.justice.ie/en/JELR/Legislating\\_for\\_Hate\\_Speech\\_and\\_Hate\\_Crime\\_in\\_Ireland\\_Web.pdf/Files/Legislating\\_for\\_Hate\\_Speech\\_and\\_Hate\\_Crime\\_in\\_Ireland\\_Web.pdf](https://www.justice.ie/en/JELR/Legislating_for_Hate_Speech_and_Hate_Crime_in_Ireland_Web.pdf/Files/Legislating_for_Hate_Speech_and_Hate_Crime_in_Ireland_Web.pdf)

<sup>36</sup> Legislating for Hate Speech and Hate Crime in Ireland: Report on the Public Consultation 2020, Department of Justice, p. 28, [https://www.justice.ie/en/JELR/Legislating\\_for\\_Hate\\_Speech\\_and\\_Hate\\_Crime\\_in\\_Ireland\\_Web.pdf/Files/Legislating\\_for\\_Hate\\_Speech\\_and\\_Hate\\_Crime\\_in\\_Ireland\\_Web.pdf](https://www.justice.ie/en/JELR/Legislating_for_Hate_Speech_and_Hate_Crime_in_Ireland_Web.pdf/Files/Legislating_for_Hate_Speech_and_Hate_Crime_in_Ireland_Web.pdf)

<sup>37</sup> General Scheme Criminal Justice (Hate Crime) Bill 2021, Department of Justice, p. 5, [https://www.justice.ie/en/JELR/General\\_Scheme\\_Criminal\\_Justice\\_\(Hate\\_Crime\)\\_Bill\\_2021.pdf/Files/General\\_Scheme\\_Criminal\\_Justice\\_\(Hate\\_Crime\)\\_Bill\\_2021.pdf](https://www.justice.ie/en/JELR/General_Scheme_Criminal_Justice_(Hate_Crime)_Bill_2021.pdf/Files/General_Scheme_Criminal_Justice_(Hate_Crime)_Bill_2021.pdf)

<sup>38</sup> General Scheme Criminal Justice (Hate Crime) Bill 2021, Department of Justice, p. 6, [https://www.justice.ie/en/JELR/General\\_Scheme\\_Criminal\\_Justice\\_\(Hate\\_Crime\)\\_Bill\\_2021.pdf/Files/General\\_Scheme\\_Criminal\\_Justice\\_\(Hate\\_Crime\)\\_Bill\\_2021.pdf](https://www.justice.ie/en/JELR/General_Scheme_Criminal_Justice_(Hate_Crime)_Bill_2021.pdf/Files/General_Scheme_Criminal_Justice_(Hate_Crime)_Bill_2021.pdf)





It should be noted that these exceptions are not a protection for freedom of expression but are rather a protection for freedom of *dissemination*. Thus, somebody who is deemed to have recklessly incited hatred by producing material that is considered “*a reasonable and genuine*” contribution to one of the above fields will not be protected. Instead, the exception applies to third parties who disseminate any incitement to hatred – for example a bookseller or an online platform. These protections thus create different standards when it comes to freedom of expression. It is not clear why artistic and literary works should be held to one standard, while another standard should be applied to other kinds of work. This double standard may risk privileging certain groups who may have greater opportunities to create artistic, literary or academic material. Assessing whether a contribution is “*reasonable and genuine*” is also highly subjective. Under the DSA and the HCB, this means that powerful online platforms would be making judgements about what constitutes a “*reasonable and genuine*” contribution to the specified fields of public discourse. Thus, this proposal risks providing greater freedom of expression to some actors over others.

### 3.2 The Bigot’s Veto

In the way that the proposed legislation is currently drafted, there is a risk that the HCB could (i) criminalise non-hateful material and (ii) be invoked to justify censoring material that is legal, particularly when combined with the DSA. This is because the HCB would criminalise “reckless” incitement to hatred. This means that, if a communication is “significantly likely” to incite hatred, and if a citizen is aware of this but chooses to proceed with the communication, then the citizen is committing a criminal offence – even if inciting hatred is not the intention of the speaker. This could give rise to what might be called a “bigot’s veto” when it comes to freedom of expression. If a

communication could be deemed hateful simply because it is likely to encourage feelings of hatred in somebody else, then almost any communications about a group with protected characteristics could at least potentially be deemed hateful. This is because, for any such communication, there could be persons of an intolerant mindset who would feel inclined to hatred as a result. Even statements of fact can potentially encourage hatred. For example, an online post discussing variations in crime rates around the world could potentially encourage persons of an intolerant mindset to feel hostility towards particular communities. It would be a significant threat to freedom of expression if the author could face legal consequences as a result of producing such content. Such a scenario would effectively make speakers legally responsible for the potential thoughts of others even when these are thoughts the speaker themselves may not share – or might even abhor.


### 3.3 Risks for Freedom of Expression

It is important for democratic discourse that citizens should be free to express non-hateful views, even if these views may be critical of particular national, ethnic, racial or religious groups for their actions or behaviours. For example, someone might wish to criticise voters in a particular country for how they voted in an election or referendum, or a speaker in a multi-ethnic society may criticise an ethnic group that they deem to be unfairly dominant.

One reason that a speaker may wish to criticise a group is precisely in order to oppose hatred or discrimination. Hatred and discrimination can manifest in actions and not just in speech. Critical speech can play an important role in opposing hatred and discrimination. Crucially, legislation that is designed to regulate controversial forms of speech – such as hate speech – can be at risk of being applied to speech that







itself criticises hatred and discrimination. This is particularly important because hatred can manifest when a group with "protected characteristics" (such as an ethnic or religious group) embraces a hateful attitude towards another group. Alleged persecution and discrimination in places such as South Africa (during Apartheid), Myanmar (against the Rohingya), Turkey (against Kurds), or in Northern Ireland (during the Troubles) provide some examples.

Socio-cultural attitudes can also be responsible for hatred or discrimination. In many countries, there are widespread views or practices related to topics such as homosexuality,<sup>39</sup> femicide honour killings<sup>40</sup> and female genital mutation<sup>41</sup> which may be closely associated with particular ethnicities and cultures. Groups with protected characteristics can also engage in other forms of wrongdoing, such as the sexual abuse scandals which took place in the Catholic Church in Ireland. A citizen criticising an ethnic or religious group for what they deem to be the oppression of another group could risk being censored under Ireland's HCB in keeping with the "bigot's veto" described above. This is because even legitimate criticism could potentially incite hatred against the criticised group amongst

intolerant persons. There is also a risk that such criticism can be misconstrued, or deliberately misrepresented, as hatred.

In Ireland, criticism of Russia following its invasion of Ukraine in February 2022, or against Israel due to that state's treatment of the Palestinian population, have been portrayed by Russian<sup>42</sup> and Israeli<sup>43</sup> authorities respectively as manifestations of Russophobic or antisemitic hatred. It has also been claimed that even non-hateful criticism of these countries can help to stir up hatred against Russian and Jewish people<sup>44</sup> which could be potentially true even for legitimate criticism. Thus, the HCB would provide a legal tool that could plausibly be used to argue that criticism of Israel and Russia promotes hatred and should be prohibited.

As the Irish Department of Justice states, the HCB criminalises those who *"have considered whether what they were doing would incite hatred, concluded that it was significantly likely, and decided to press ahead anyway."*<sup>45</sup> The proposed legislation does not seem to provide protection for non-hateful actors who carefully evaluate the possible implications of their communication and proceed because of what they deem the greater importance of doing so. This may be true even if these actors try to minimise

<sup>39</sup> *The Global Divide on Homosexuality*, Pew Research Center, <https://www.pewresearch.org/global/2013/06/04/the-global-divide-on-homosexuality/>

<sup>40</sup> *The World's Muslims: Religion, Politics and Society*, Pew Research Center, p. 90, available at <https://www.pewforum.org/2013/04/30/the-worlds-muslims-religion-politics-society-overview/>

<sup>41</sup> Sheila Wayman, Female genital mutilation: 'It's happening here, girls are being taken out of Ireland' (irishtimes.com), <https://www.irishtimes.com/life-and-style/health-family/female-genital-mutilation-it-s-happening-here-girls-are-being-taken-out-of-ireland-1.3374033>

<sup>42</sup> Ronan McGreevy, Ireland at forefront of 'anti-Russian events' in EU, Russian ambassador claims (irishtimes.com), <https://www.irishtimes.com/news/ireland/irish-news/ireland-at-forefront-of-anti-russian-events-in-eu-russian-ambassador-claims-1.4818514>

<sup>43</sup> Colm Keena, Israel reports 'rise in anti-Semitism' in Ireland to UN committee (irishtimes.com), <https://www.irishtimes.com/news/crime-and-law/israel-reports-rise-in-anti-semitism-in-ireland-to-un-committee-1.4725024>

<sup>44</sup> Ronan McGreevy, Ireland at forefront of 'anti-Russian events' in EU, Russian ambassador claims (irishtimes.com), <https://www.irishtimes.com/news/ireland/irish-news/ireland-at-forefront-of-anti-russian-events-in-eu-russian-ambassador-claims-1.4818514>; *Contemporary Antisemitism in the Political Discourse in Ireland*, The Institute for National Security Studies Tel Aviv University, <https://www.inss.org.il/wp-content/uploads/2021/06/%D7%9E%D7%96%D7%B%D7%A8-214-Contemporary-Antisemitism-in-the-Political-Discourse-of-Five-We...-83-89.pdf>; David Collier, *Ireland Antisemitism Report*, [https://embassies.gov.il/dublin/NewsAndEvents/Documents/Collier%20Antisemitism%20in%20Ireland%20\\_final\\_online.pdf](https://embassies.gov.il/dublin/NewsAndEvents/Documents/Collier%20Antisemitism%20in%20Ireland%20_final_online.pdf)

<sup>45</sup> gov.ie - Tough sentences for hate crimes under new Bill from Minister McEntee (www.gov.ie), <https://www.gov.ie/en/press-release/48cb5-tough-sentences-for-hate-crimes-under-new-bill-from-minister-mcentee/>



the risk that their communication could incite hatred. As the Government's report on its public consultation states: *"There should be no requirement for the material [hate speech] to be threatening, abusive or insulting in itself", and material can constitute hate speech even if it "is couched in polite or coded language."*<sup>46</sup> The HCB also states that material can be counted as hate speech *"whether or not any actual instance of harm or unlawful discrimination is shown to have occurred, or to have been likely to occur, as a result"*.<sup>47</sup> The above cases therefore demonstrate examples of where the HCB could potentially be open to misuse or be inadvertently applied to censor legitimate non-hateful content.

## 4. The Misuse of Hate Speech Laws: Examples and Implications

The possibility that hate speech laws could be misused is not only an abstract possibility but is something that has already occurred, even in democracies. Hate speech laws are also often intended to protect vulnerable groups and yet can be used against the groups they are intended to protect. In the UK the first

person convicted of hate speech under the Race Relations Law was a black man who used a racial insult against a white policeman.<sup>48</sup> Three members of the Universal Coloured People's Association in the UK were also convicted for inciting hatred against white people.<sup>49</sup> In France hate speech laws have been used against critics of Israel's treatment of Palestinians (who were deemed to be antisemitic),<sup>50</sup> critics of allegedly "cruel" animal slaughter (deemed as being anti-Islamic)<sup>51</sup> and against a radical feminist (for incitement of hatred against men).<sup>52</sup> In many countries, hate speech legislation has been used to silence those who have criticised dominant ethnic groups with disproportionate political power, including in Ethiopia, Rwanda and Kenya.<sup>53</sup>

It is also problematic that nothing in the legislation prevents new groups from claiming to be a religious group or ethnicity, including groups that may be controversial. This is also concerning because hate speech law can be used to protect groups that promote hatred. For example, in Ireland in 2021 an activist was arrested by appointment after the religious group, Society of Saint Pius X Resistance (SSPX Resistance), accused her of inciting hatred against them.<sup>54</sup> The SSPX Resistance had held sermons preaching that Jewish

<sup>46</sup> *Legislating for Hate Speech and Hate Crime in Ireland: Report on the Public Consultation 2020*, Department of Justice, p. 44 [https://www.justice.ie/en/JELR/Legislating\\_for\\_Hate\\_Speech\\_and\\_Hate\\_Crime\\_in\\_Ireland\\_Web.pdf/Files/Legislating\\_for\\_Hate\\_Speech\\_and\\_Hate\\_Crime\\_in\\_Ireland\\_Web.pdf](https://www.justice.ie/en/JELR/Legislating_for_Hate_Speech_and_Hate_Crime_in_Ireland_Web.pdf/Files/Legislating_for_Hate_Speech_and_Hate_Crime_in_Ireland_Web.pdf)

<sup>47</sup> *General Scheme Criminal Justice (Hate Crime) Bill 2021*, Department of Justice, p. 7 [https://www.justice.ie/en/JELR/General\\_Scheme\\_Criminal\\_Justice\\_\(Hate\\_Crime\)\\_Bill\\_2021.pdf/Files/General\\_Scheme\\_Criminal\\_Justice\\_\(Hate\\_Crime\)\\_Bill\\_2021.pdf](https://www.justice.ie/en/JELR/General_Scheme_Criminal_Justice_(Hate_Crime)_Bill_2021.pdf/Files/General_Scheme_Criminal_Justice_(Hate_Crime)_Bill_2021.pdf)

<sup>48</sup> Herz and Molnar, (editors), *The Content and Context of Hate Speech: Rethinking Regulation and Responses*, p. 244

<sup>49</sup> Herz and Molnar, (editors), *The Content and Context of Hate Speech: Rethinking Regulation and Responses*, p. 97

<sup>50</sup> Benjamin Dodman, France's criminalisation of Israel boycotts sparks free-speech debate (france24.com), <https://www.france24.com/en/20160120-france-boycott-israel-bds-law-free-speech-antisemitism>; France: Landmark ECtHR judgement finds boycott campaign against Israel cannot be criminalized - Amnesty International, <https://www.amnesty.org/en/latest/news/2020/06/france-landmark-ecthr-judgment-finds-boycott-campaign-against-israel-cannot-be-criminalized/>

<sup>51</sup> Timothy Garton Ash, *Free Speech: Ten Principles for a Connected World*, (Atlantic Books Ltd, London, 2017) p. 219


<sup>52</sup> Alison Flood, I Hate Men: Attempts to ban French pamphlet sends sales skyrocketing (irishtimes.com), <https://www.irishtimes.com/culture/books/i-hate-men-attempts-to-ban-french-pamphlet-sends-sales-skyrocketing-1.4350587>

<sup>53</sup> Herz and Molnar, (editors), *The Content and Context of Hate Speech: Rethinking Regulation and Responses*, p. 76, 352 – 374

<sup>54</sup> Conor Gallagher, Cork activist questioned over allegations of harassment by Christian fundamentalists (irishtimes.com), <https://www.irishtimes.com/news/crime-and-law/cork-activist-questioned-over-allegations-of-harassment-by-christian-fundamentalists-1.4629871>

<sup>55</sup> Conor Gallagher, Cork activist questioned over allegations of harassment by Christian fundamentalists (irishtimes.com), <https://www.irishtimes.com/news/crime-and-law/cork-activist-questioned-over-allegations-of-harassment-by-christian-fundamentalists-1.4629871>





people caused the COVID-19 crisis and were manipulating stock markets in order to cause war.<sup>55</sup> This arrest took place under Ireland's 1989 Prohibition of Incitement to Hatred Act. Given that the HCB is intended to have a broader scope than the 1989 Act,<sup>56</sup> it could potentially give rise to similar cases.

It is also worth noting that hatred can be mobilised by governments, including governments who use hate speech laws to silence critics. The world's two largest democracies have in recent years witnessed the election of leaders who have been widely accused of inciting hatred. In the US, President Donald Trump has been accused of inciting hatred against Muslims and Mexicans.<sup>57</sup> However, strong protections for freedom of speech in the US means that President Trump was unable to prevent his views from being challenged in the media and online.<sup>58</sup> India presents an important contrast. In India, since the election of Prime Minister Narendra Modi in 2014, there has been a significant upsurge in violence against Muslims, Christians and Dalits (formerly known as Untouchables) with the governing right-wing Hindu Nationalist Bharatiya Janata Party (BJP) widely accused of inciting hatred against these groups.<sup>59</sup> India's hate speech laws make this problem worse because, as Human Rights Watch has warned, India's hate speech laws:<sup>60</sup>

*'are intended to protect minorities and the powerless, [but] in practice they are often used at the behest of powerful individuals*

*or groups, who claim that they have been offended, to silence speech. Government officials too often pursue such complaints, leaving members of minority groups, writers, artists, and scholars facing threats of violence and legal action'.*

Thus, the above examples demonstrate how hate speech laws can be applied in ways that are not originally intended by their creators and show how such laws can be misused to silence legitimate criticism, including criticism against hatred. Ireland's HCB is arguably too vague to prevent it from being misused in similar ways.

## 5. The Online Implications

In a possible safeguard for freedom of speech, Ireland's HCB states that it "prevents any prosecutions taking place for incitement to hatred without the consent of the Director of Public Prosecutions."<sup>61</sup> This would not necessarily stop the law from being invoked by other actors, such as publishers, who could refuse to disseminate certain controversial but legal materials out of a sense of risk aversion. However, a significant proportion of public discourse now takes place online, mostly on a small number of large platforms, and this is where the application of the HCB will be determined by the DSA. Notably, it will be the online platform companies who will interpret in the first instance whether or not online material constitutes hate speech. The

<sup>56</sup> gov.ie - Tough sentences for hate crimes under new Bill from Minister McEntee ([www.gov.ie](http://www.gov.ie)), <https://www.gov.ie/en/press-release/48cb5-tough-sentences-for-hate-crimes-under-new-bill-from-minister-mcentee/>

<sup>57</sup> Nicquel Terry Ellis, 'Hate speech' a common theme of Trump's presidency, critics say ([usatoday.com](http://usatoday.com)), <https://eu.usatoday.com/story/news/politics/elections/2020/10/13/hate-speech-common-theme-trumps-presidency/5873238002/>; Fabiola Cineas, Trump's history of inciting violence in words and tweets: A timeline from 2015 through the Capitol attack - Vox, <https://www.vox.com/21506029/trump-violence-tweets-racist-hate-speech>

<sup>58</sup> Mchangama, *Free Speech: A Global History From Socrates to Social Media*, p. 340; Sara Morrison, The Trump administration attacks Section 230 and free speech online - Vox, <https://www.vox.com/recode/2020/6/18/21294331/section-230-bill-barr-josh-hawley-trump-internet-free-speech>

<sup>59</sup> John Harriss, Craig Jeffrey and Trent Brown, *India: Continuity and Change in the Twenty-First Century*, (Polity Press, Cambridge, 2020), p. 111 – 135; p. 245-246

<sup>60</sup> *Violent Cow Protection in India: Vigilante Groups Attack Minorities*, Human Rights Watch, <https://www.hrw.org/report/2019/02/18/violent-cow-protection-india/vigilante-groups-attack-minorities#>

<sup>61</sup> *General Scheme Criminal Justice (Hate Crime) Bill 2021*, Department of Justice, p. 9 [https://www.justice.ie/en/JELR/General\\_Scheme\\_Criminal\\_Justice\\_\(Hate\\_Crime\)\\_Bill\\_2021.pdf/Files/General\\_Scheme\\_Criminal\\_Justice\\_\(Hate\\_Crime\)\\_Bill\\_2021.pdf](https://www.justice.ie/en/JELR/General_Scheme_Criminal_Justice_(Hate_Crime)_Bill_2021.pdf/Files/General_Scheme_Criminal_Justice_(Hate_Crime)_Bill_2021.pdf)



HCB specifically states that corporations *"including social media companies"* must have *"reasonable and effective measures to prevent the dissemination of communications inciting hatred generally."*<sup>62</sup> The HCB further states that, if a company fails to remove an incitement to hatred, the company must have been: *"unaware and had no reason to suspect that this particular content was intended or likely to incite hatred."*<sup>63</sup> This suggests that even if a company has only a very minor suspicion that content could be deemed to be hate speech, it may feel compelled to remove such content.

Some features of the HCB might suggest that there are limits to the risks it poses to online freedom of expression. First, the size of fines proposed by the HCB will be modest, at least from the perspective of large social media platforms. These are "Class A" fines, (the highest value class of fines for summary conviction) which are currently set to a maximum of €5000.<sup>64</sup> Secondly, a prosecution against a social media company under the HCB would have to be approved by the the Director of Public Prosecutions (DPP), potentially making it less likely that the company would face frivolous or vexatious charges. Thirdly, while the HCB prescribes "reasonable and effective" measures to prevent the dissemination of hateful content, it is not prescriptive regarding what these measures should be. However, under the DSA, platforms must respond when anyone flags material as hate speech, and therefore, the veto role

that Ireland's DPP might otherwise play will not apply. Furthermore, while fines proposed under the HCB are relatively modest, the DSA proposes fines of up to 6% of a company's annual turnover, which could prompt companies to avoid even the low possibility of incurring major fines by removing content even in cases where such content is unlikely to be deemed to be hate speech.<sup>65</sup> Hence, companies will be required to determine what constitutes hate speech in accordance with one piece of legislation, while a separate piece of legislation determines how broadly and rapidly they would need to remove such content in order to avoid being fined. Thus, the HCB and the DSA applied together could pose greater risks to online freedom of expression than either piece of legislation would separately. This could exacerbate an already existing problem whereby online material is improperly removed as hate speech, as is discussed in the next section.

## 5.1 Problematic Removal of Online Content

Online platforms are responsible for moderating enormous quantities of material. YouTube removed over 7.9 million videos in a three-month period in 2020.<sup>66</sup> In the third quarter of 2021, Facebook and Instagram took action against more than 22.3 million pieces of content that were deemed to constitute hate speech.<sup>67</sup> More than 96% of this material was proactively

<sup>62</sup> *General Scheme Criminal Justice (Hate Crime) Bill 2021*, Department of Justice, p. 6 [https://www.justice.ie/en/JELR/General\\_Scheme\\_Criminal\\_Justice\\_\(Hate\\_Crime\)\\_Bill\\_2021.pdf/Files/General\\_Scheme\\_Criminal\\_Justice\\_\(Hate\\_Crime\)\\_Bill\\_2021.pdf](https://www.justice.ie/en/JELR/General_Scheme_Criminal_Justice_(Hate_Crime)_Bill_2021.pdf/Files/General_Scheme_Criminal_Justice_(Hate_Crime)_Bill_2021.pdf)

<sup>63</sup> *General Scheme Criminal Justice (Hate Crime) Bill 2021*, Department of Justice, p. 6 [https://www.justice.ie/en/JELR/General\\_Scheme\\_Criminal\\_Justice\\_\(Hate\\_Crime\)\\_Bill\\_2021.pdf/Files/General\\_Scheme\\_Criminal\\_Justice\\_\(Hate\\_Crime\)\\_Bill\\_2021.pdf](https://www.justice.ie/en/JELR/General_Scheme_Criminal_Justice_(Hate_Crime)_Bill_2021.pdf/Files/General_Scheme_Criminal_Justice_(Hate_Crime)_Bill_2021.pdf)

<sup>64</sup> Fines for criminal offences (citizensinformation.ie), [https://www.citizensinformation.ie/en/justice/criminal\\_law/criminal\\_trial/maximum\\_fines\\_on\\_summary\\_conviction.html](https://www.citizensinformation.ie/en/justice/criminal_law/criminal_trial/maximum_fines_on_summary_conviction.html)


<sup>65</sup> *Digital Services Act*, European Commission, p. 80, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825&from=en>

<sup>66</sup> Fredrik Erixon, "Too Big to Care" or "Too Big to Share": *The Digital Services Act and the Consequences of Reforming Intermediary Liability Rules*, European Centre for International Political Economy, 2021, <https://ecipe.org/publications/digital-services-act-reforming-intermediary-liability-rules/>

<sup>67</sup> Community Standards Enforcement | Transparency Center (fb.com), <https://transparency.fb.com/data/community-standards-enforcement/hate-speech/facebook/#content-actioned>







found by these platforms themselves, with less than 4% of this content being reported by users.<sup>68</sup> Companies have found that it is not practical to rely solely on human moderators and increasingly rely on automated algorithms to remove illegal content. Facebook reported that, in the fourth quarter of 2020, more than 97% of all hate speech removed from its platform was removed by algorithms.<sup>69</sup>

Given the enormous volume of content errors will inevitably occur and, in particular, algorithms can make mistakes that a human would be unlikely to make.<sup>70</sup> For example, videos of cars being washed were removed by algorithms which misidentified them as videos of shootings.<sup>71</sup> This can also apply to hate speech in ways that can have implications for public discourse. For example, Facebook algorithms removed posts quoting the American Declaration of Independence because of the way it referred to Native Americans.<sup>72</sup> A post by a Danish citizen of Iranian heritage

who posted criticism of anti-immigrant sentiment in Denmark found his post incorrectly removed by Facebook on the grounds that it was deemed to be hate speech.<sup>73</sup> Some social media users who have posted about their experiences of racial harassment have found their own content deleted as hate speech, including prominent US activists such as Ijeoma Oluo<sup>74</sup> and Shaun King<sup>75</sup> among others.<sup>76</sup>

As has been noted, misuse of hate speech rules can be particularly concerning when it inadvertently protects groups responsible for hatred from criticism. This can undermine those seeking to challenge the viewpoints promoted by such groups. For example, Facebook and YouTube removed posts and videos about China's persecution of ethnic minorities in Xinjiang as hate speech.<sup>77</sup> Meanwhile, Facebook ran Chinese government advertisements which positively portrayed government-run camps in Xinjiang province.<sup>78</sup>

<sup>68</sup> Community Standards Enforcement | Transparency Center (fb.com), <https://transparency.fb.com/data/community-standards-enforcement/hate-speech/facebook/#content-actioned>

<sup>69</sup> Update on Our Progress on AI and Hate Speech Detection | Meta (fb.com), <https://about.fb.com/news/2021/02/update-on-our-progress-on-ai-and-hate-speech-detection/>

<sup>70</sup> Elizabeth Dwoskin and Nitasha Tiku, Facebook sent home thousands of human moderators due to the coronavirus. Now the algorithms are in charge - The Washington Post, <https://www.washingtonpost.com/technology/2020/03/23/facebook-moderators-coronavirus/>; Mark Scott and Laura Kayali, What happened when humans stopped managing social media content - POLITICO, <https://www.politico.eu/article/facebook-content-moderation-automation/>

<sup>71</sup> Deepa Seetharaman, Jeff Horwitz and Justin Scheck, Facebook Says AI Will Clean Up the Platform. Its Own Engineers Have Doubts. - WSJ, <https://www.wsj.com/articles/facebook-ai-enforce-rules-engineers-doubtful-artificial-intelligence-11634338184?mod=djemalertNEWS>

<sup>72</sup> Sam Wolfson, Facebook labels declaration of independence as 'hate speech' | Censorship | The Guardian, <https://www.theguardian.com/world/2018/jul/05/facebook-declaration-of-independence-hate-speech>

<sup>73</sup> Frederik Stjernfel and Anne Mette Lauritzen, Your Post has been Removed, (Springer Nature, Copenhagen, 2020), p. 125-126

<sup>74</sup> Jessica Guynn, These are Facebook's secret rules for removing posts (usatoday.com), <https://eu.usatoday.com/story/tech/news/2018/04/24/facebook-discloses-secret-guidelines-policing-content-introduces-appeals/544046002/>

<sup>75</sup> Hard Questions: Who Should Decide What Is Hate Speech in an Online Global Community? | Meta (fb.com), <https://about.fb.com/news/2017/06/hard-questions-hate-speech/>

<sup>76</sup> Jessica Guynn, Facebook apologizes to black activist who was censored for calling out racism (usatoday.com), <https://eu.usatoday.com/story/tech/2017/08/03/facebook-ijeoma-oluo-hate-speech/537682001/>

<sup>77</sup> Elizabeth Culliford, Here's what Facebook's oversight board says the company should do | Reuters, <https://www.reuters.com/article/facebook-oversight-factbox-idINKBN29X2AQ>; KAZAKHSTAN - CHINA Activist says that Facebook trying to hide human rights violations of Kazakhs in Xinjiang (asianews.it), PIME AsiaNews, <http://asianews.it/news-en/Activist-says-that-Facebook-trying-to-hide-human-rights-violations-of-Kazakhs-in-Xinjiang-52060.html>; Victoria Waldersee and Parash Dave, EXCLUSIVE-YouTube takes down Xinjiang videos, forces rights group to seek alternative | Reuters, <https://www.reuters.com/article/alphabet-china-idCNL2N2O71H1>; Eileen Guo, How YouTube's rules are used to silence human rights activists | MIT Technology Review, <https://www.technologyreview.com/2021/06/24/1027048/youtube-xinjiang-censorship-human-rights-atajurt/>

<sup>78</sup> Newley Purnell, Facebook Staff Fret Over China's Ads Portraying Happy Muslims in Xinjiang - WSJ, <https://www.wsj.com/articles/facebook-staff-fret-over-chinas-ads-portraying-happy-muslims-in-xinjiang-11617366096>; Sigal Samuel, China paid Facebook and Twitter to help sow anti-Muslim misinformation - Vox, <https://www.vox.com/future-perfect/2019/8/22/20826971/facebook-twitter-china-misinformation-ughur-muslim-internment-camps>



According to Human Rights Watch, legitimate content reporting human rights abuses in Israel-Palestine is frequently and wrongly removed by Facebook and Instagram on the grounds that it constitutes hate speech.<sup>79</sup> A collection of activist organisations – including Article19 and Jewish Voice for Peace – accused Twitter and Facebook of “systematically silencing users protesting and documenting the evictions of Palestinian families from their homes.”<sup>80</sup> In Myanmar, military officials and state-backed paramilitaries were able to use Facebook to spread hatred against the Rohingya minority.<sup>81</sup> In contrast, Facebook’s hate speech moderation system frequently deleted legitimate content, including posts that were criticising the spread of hatred in Myanmar.<sup>82</sup> Moreover, Rohingya accounts and Rohingya activists who posted about real atrocities against Rohingya had their posts deleted or their accounts blocked by Facebook.<sup>83</sup>

The above examples show how hate speech rules can inadvertently benefit actors who perpetrate hatred, by protecting such actors from criticism. More generally, hate speech rules can often be misapplied, triggering the removal of legitimate content. The combination of the HCB and DSA makes it more likely that platforms will remove contentious material, thus exacerbating the problem whereby legitimate material is wrongly removed as hate speech – including material that is politically important or created to oppose hatred.

## 6. Conclusion

Ireland’s HCB proposal and the EU’s DSA proposal are a welcome sign that policymakers are paying serious attention to the dangers of hate speech. As mentioned, some measures that restrict freedom of speech are relatively uncontroversial, for instance, the prohibition of direct incitement to violence, or bans on bullying and harassment.<sup>84</sup> It is when the law goes beyond this, to ban general expressions of opinion that might be deemed hateful, that caution is needed. As has been demonstrated, care is required to ensure that hate speech laws do not have unintended consequences and negative implications for freedom of expression. In some cases, hate speech laws can lead to the censorship of non-hateful material, and even of those who are trying to speak out against hatred. This does not mean that hate speech laws should not be adopted. Rather, hate speech laws should avoid language that is vague, broad and subjective, which could potentially lead to excessive moderation. This is especially important today where a handful of companies and social media platforms play a disproportionately large role in determining what citizens will and will not see.

In conclusion, while there are positive features in both the HCB and DSA, arguably a more complex and nuanced


<sup>79</sup> Israel/Palestine: Facebook Censors Discussion of Rights Issues | Human Rights Watch (hrw.org), <https://www.hrw.org/news/2021/10/08/israel/palestine-facebook-censors-discussion-rights-issues>; see also: Systematic Efforts to Silence Palestinian Content On Social Media, 7amleh - The Arab Center for the Advancement of Social Media, <https://7amleh.org/2020/06/07/systematic-efforts-to-silence-palestinian-content-on-social-media>

<sup>80</sup> Sheikh Jarrah: Facebook and Twitter silencing protests, deleting evidence - ARTICLE 19, <https://www.article19.org/resources/sheikh-jarrah-facebook-and-twitter-silencing-protests-deleting-evidence/>

<sup>81</sup> David Kaye, *Speech Police: The Global Struggle to Govern the Internet*, (Columbia Global Reports, New York, 2019), p. 29

<sup>82</sup> Thant Sin, Facebook bans racist word ‘Kalar’ in Myanmar, triggers censorship | Business Standard News (business-standard.com), [https://www.business-standard.com/article/international/facebook-bans-racist-word-kalar-in-myanmar-triggers-censorship-117060300423\\_1.html](https://www.business-standard.com/article/international/facebook-bans-racist-word-kalar-in-myanmar-triggers-censorship-117060300423_1.html); Tin Htet Paing, Kyaw Phyo Tha and Thet Ko Ko, Facebook Ban of Racial Slur Sparks Debate in Myanmar (irrawaddy.com), <https://www.irrawaddy.com/news/burma/facebook-ban-of-racial-slur-sparks-debate-in-burma.html>

<sup>83</sup> Kaye, *Speech Police: The Global Struggle to Govern the Internet*, p. 29; Betsy Swan, Exclusive: Facebook Silences Rohingya Reports of Ethnic Cleansing (thedailybeast.com), <http://www.thedailybeast.com/exclusive-rohingya-activists-say-facebook-silences-them>; BBC Trending, Why are posts by Rohingya activists getting deleted? - BBC News, <https://www.bbc.com/news/blogs-trending-41364633>; Julia Carrie Wong, Michael Safi and Shaikh Azizur Rahman, Facebook bans Rohingya group’s posts as minority faces ‘ethnic cleansing’ | Facebook | The Guardian, <https://www.theguardian.com/technology/2017/sep/20/facebook-rohingya-muslims-myanmar>



approach is required to address hate speech. Particularly, policymakers should ensure that there are adequate protections for freedom of expression in this legislation. Otherwise, there is a risk that this legislation could incentivise online platforms to remove non-hateful material, including vital contributions to public discourse and material which speaks out against hatred. In some cases, this could risk exacerbating the very problem that such laws are designed to address.

<sup>84</sup> As earlier noted, in Ireland these topics will be addressed by the proposed Online Safety and Media Regulation bill: gov.ie - Online Safety and Media Regulation Bill (www.gov.ie), <https://www.gov.ie/en/publication/d8e4c-online-safety-and-media-regulation-bill/>

The Institute of International and European Affairs (IIEA) is Ireland's leading international affairs think tank. Founded in 1991, its mission is to foster and shape political, policy and public discourse in order to broaden awareness of international and European issues in Ireland and contribute to more informed strategic decisions by political, business and civil society leaders.

The IIEA is independent of government and all political parties and is a not-for profit organisation with charitable status. In January 2021, the Global Go To Think Tank Index ranked the IIEA as Ireland's top think tank.

© Institute of International and European Affairs, June 2022

Creative Commons License

This is a human-readable summary of (and not a substitute for) the license.

[https://creativecommons.org/licenses/Attribution-NonCommercial-ShareAlike 4.0 International \(CC BY-NC-SA 4.0\)](https://creativecommons.org/licenses/Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0))

You are free to:

- Share - copy and redistribute the material in any medium or format
- Adapt - remix, transform, and build upon the material
- The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

**Attribution** — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

**NonCommercial** — You may not use the material for commercial purposes.

**ShareAlike** — If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

**No additional restrictions** — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.



The IIEA acknowledges the support of the Europe for Citizens Programme of the European Union.

