ISD | Institute for Strategic Dialogue

sponsored by

Federal Foreign Office

Online crisis protocols – Expanding the regulatory toolbox to safeguard democracy during crises

Christian Schwieter

About the Digital Policy Lab

The Digital Policy Lab (DPL) is an inter-governmental working group focused on charting the policy path forward to prevent and counter the spread of disinformation, hate speech, extremist and terrorist content online. It is comprised of representatives of relevant ministries and regulatory bodies from liberal democracies. The DPL aims to foster inter-governmental exchange, provide policymakers and regulators with access to sector-leading expertise and research, and build an international community of practice around key challenges in the digital policy space. We thank the German Federal Foreign Office for their support for this project.

About this Paper

As part of the DPL, the Institute for Strategic Dialogue (ISD) organised two working group meetings on the topic of online crisis protocols in July and September 2022. The working group consisted of DPL members representing national ministries or departments as well as regulators from Canada, New Zealand, Slovakia, Switzerland, the UK, and the US. Participation also included representatives from academia and civil society. While participants contributed to this paper, the views expressed in this paper do not necessarily reflect the views of all participants or any governments involved in this project.

Editorial responsibility:
Huberta von Voss, (Executive Director, ISD Germany)

Author

Christian Schwieter is Project Manager at ISD Germany, working at the intersection of digital analysis and digital policy. Since 2021, he leads the German-language research project on far-right activity on alternative and emerging online platforms, funded by the Ministry of Justice. Previously, Christian worked as a researcher for the Computational Propaganda Project at the Oxford Internet Institute, where he co-authored reports on state-backed information operations relating to the Covid-19 pandemic. In 2019, Christian was the Specialist Adviser on Disinformation Matters for the UK Digital, Culture, Media and Sports Select Committee at the House of Commons. Christian holds an MSc in Social Science of the Internet from the University of Oxford and a BA from Leiden University College The Hague, Netherlands.

Acknowledgements

We would like to thank all members and participants of the working group for their contributions. We would like to give special thanks to the speakers for providing valuable insights, namely Denis Sparas and Julia van Best (Directorate-General for Communications, Networks, Content and Technology, European Commission), Kristina Kirk (Department of the Prime Minister & Cabinet, New Zealand), Dr Erin Saltman (Global Internet Forum to Counter Terrorism, GIFCT), Antonis Samouris (Europol), Diego Naranjo (European Digital Rights, EDRi), Michael Meyer-Resende (Democracy Reporting International, DRI), and Iverna McGowan (Center for Democracy & Technology, CDT).

ISD Institute for Strategic Dialogue

www.isdgermany.org

Contents

Executive Summary 4

Key lessons learned from existing crisis protocols in the CVE space 4

Key concerns and recommendations to safeguard fundamental rights in the context of the Digital Services Act crisis response mechanism 4

Introduction 5

From ad-hoc crisis responses to protocols and mechanisms 6

Crisis protocols designed to counter violent extremism online 7

Lessons learned from existing online crisis protocols 10

Democratic safeguards – definitions, competencies and procedures 11

Conclusion 13

Endnotes 14

Executive Summary

The EU Digital Services Act (DSA) has brought with it a breadth of new regulatory tools seeking to create a "safer digital space" by protecting the fundamental rights of users and combating the proliferation of illegal and harmful content on online platforms. While there have been many discussions around platform liability, algorithmic audits and annual risk assessments, less attention has been paid to the last-minute additions to the DSA – namely mechanisms and protocols aimed to empower DSA regulators during so-called "crisis" events. To shine a light on this under-explored but potentially crucial new regulatory tool, ISD held a working group to a) review the lessons learned from existing online crisis protocols in the countering violent extremism (CVE) space and b) collect recommendations on how future crisis protocols and response mechanisms can be designed and implemented to safeguard, rather than undermine, fundamental rights.

Key lessons learned from existing crisis protocols in the countering violent extremism space:

1. Smaller platforms and services play a key role in the spread of violent extremist content.
2. Evidence for law enforcement must be preserved while ensuring the timely removal of illegal content.
3. Expertise from the CVE space must be shared with the key stakeholders tasked with designing and implementing the broader crisis mechanisms of the DSA to improve capabilities, build capacities and avoid duplicative efforts.
4. Fundamental rights must be safeguarded during and in the aftermath of crisis events via procedural accountability mechanisms, including regular consultations with stakeholders from civil society.

Key concerns and recommendations to safeguard fundamental rights in the context of the Digital Services Act crisis response mechanism:

1. The Commission must clarify whether it will interpret a crisis event as defined in the DSA as a "state of emergency" in international human rights law (IHRL).
2. There is a need for a complementary rapid response mechanism during crisis events through which civil society organisations (CSOs) can flag incorrect removals directly to platforms, possibly coordinated via the European Board for Digital Services (the Board).
3. Delegated acts following the DSA must limit the role of the Commission by empowering the Board and improving the Board's ability to act as an independent oversight body.
4. There must be robust and timely data access provisions for independent researchers in order to evaluate the effectiveness and proportionality of the crisis response, potentially as part of a human rights impact assessment.

Introduction

The last two years have seen an unprecedented global public health crisis, followed by the outbreak of a new war in Europe caused by the Russian invasion of Ukraine. Whether it is the spread of health-related misinformation, the radicalisation of anti-lockdown movements, or the strategic dissemination of war propaganda – in the information era, activity on online platforms such as Facebook, Instagram, YouTube, Twitter, TikTok and Telegram is often found to further exacerbate these crises.

While policymakers have begun to develop regulation to rein in the power of 'Big Tech', some legislators have argued that, in times of crises, regulators ought to be given additional emergency powers, including in the regulation of online platforms. The EU's Digital Service's Act exemplifies this approach, reflecting that its drafting was marked by crises; the global pandemic and finalised in the first weeks of the invasion of Ukraine. In it, Article 48 (originally 37 in previous versions) foresees the establishment of voluntary crisis protocols "for addressing crisis situations strictly limited to extraordinary circumstances affecting public security or public health".

Additionally, in March 2022 the draft agreement introduced an additional "crisis response mechanism" under the umbrella of risk assessments (Article 36, 27a in previous versions) that would empower the European Commission to demand additional threat assessments from very large online platforms (VLOPs) in times of crisis. The Covid-19 pandemic and the war in Ukraine are cases in point where such crisis protocols and mechanisms would be applied.

Supporters of crisis protocols and crisis response mechanisms argue that these tools are a necessary backstop to safeguard fundamental rights during crisis events based on threats to public health or public security (for example, by displaying verified information prominently on the front page of the service to counteract the spread of misinformation). However, critics, including from civil society, argue that the lack of clarity around the implementation of these emergency measures threatens the rule of law, particularly when it is unclear what constitutes a crisis event and who decides when an emergency is declared. In a public statement on the crisis response mechanism published by Article 19 and European Digital Rights (EDRi), 23 CSOs warned that "[d]ecisions that affect freedom of expression and access to information, in particular in times of crisis, cannot be legitimately taken through executive power alone".[1]

While the DSA was adopted in July 2022, the specific design and implementation of the crisis protocols and related mechanisms require further formulation, to ensure they safeguard rather than undermine fundamental rights. The Covid-19 pandemic as well as the war in Ukraine and corresponding platform responses provide useful case studies as to how these mechanisms may be enacted. This policy brief seeks to contribute to this process by exploring: a) the current prominence of online crisis response protocols and mechanisms in discussions around platform governance; b) what lessons can be learned from existing protocols and mechanisms; and c) how fundamental rights can be protected in this context. In the summer of 2022, ISD convened a series of working groups to discuss these issues and inform this paper, including policymakers, regulators and civil society experts as part of the Digital Policy Lab (DPL).

From ad-hoc crisis responses to protocols and mechanisms

At its core, any crisis protocol is a set of rules and processes that are designed to mitigate the effects of an unforeseen emergency. Protocols allow stakeholders to clarify roles, responsibilities and procedures before a crisis event, and to minimise the need for ad-hoc responses. A range of stakeholders, such as governments, businesses, schools and CSOs, may be involved at various stages in designing and implementing such protocols, especially when the envisioned crisis would endanger their safety or welfare. Crisis protocols are mainly applied to high-risk situations, such as threats to life.

Over the past two years, the global public health emergency caused by the Covid-19 pandemic has seen different kinds of ad-hoc responses implemented by different types of stakeholders, including social media platforms such as Facebook, YouTube and Twitter. This includes, for example, prioritising reputable or verified sources on topics related to Covid-19 by displaying them prominently on the front page or adding warning labels to posts containing false or misleading information about vaccines.

More recently, the Russian invasion of Ukraine has also led to a variety of emergency measures adopted by both governments and platforms. Most notably, on 2 March 2022, the EU banned the broadcasting of Russian war propaganda, affecting Russian state media accounts on a wide range of online platforms such as YouTube, Facebook and Twitter.[2]

The EU's Digital Services Act – new tools to tackle crisis situations

The provisional agreement of the EU Digital Services Act published in June 2022 (see also September Corrigendum) introduced two additional crisis response tools: the binding crisis mechanism (Article 36, previously 27a) as well as the voluntary crisis protocols (Article 48, previously 37).[3] [4] Article 36 provides a binding crisis response mechanism where, in times of crises, the Commission could mandate Very Large Online Platforms to perform additional, specific ad-hoc risk assessments and oblige mitigation measures. Article 48 outlines how the Commission may initiate the drawing up of voluntary crisis protocols for addressing crisis situations during extraordinary circumstances affecting public security or public health. Both the binding provisions of Article 36 and the voluntary protocols of Article 48 would be triggered by the Commission upon recommendation by the Board, which consists of the national Digital Services Coordinators (DSCs). The binding Article 36 was introduced given the concern among co-legislators that the annual risk assessments already foreseen in the DSA would be insufficient in addressing these crisis situations. The additional ad-hoc risk assessments foreseen by Article 36 would therefore complement and strengthen the annual risk assessments.

As such, various ad-hoc online crisis responses and measures have been applied over the past years. However, they have recently gained new impetus in the context of platform regulation, particularly in light of the EU's Digital Services Act. The goal here is to develop protocols and mechanisms that clarify procedures and stakeholder accountability before a crisis event.

The DSA is not the first time that crisis protocols for social media platforms have been designed and adopted. The CVE community devised crisis protocols designed to combat the spread of terrorist propaganda, providing insights into different approaches designed for different stakeholders and purposes. In the next section, three such crisis protocols are explored in greater depth, providing insights for the design of future online crisis protocols to be used in the space of platform regulation.

Crisis protocols designed to counter violent extremism online

While combating terrorist use of the internet had been high on the political agenda since the emergence of ISIS, the far-right terrorist attack on two mosques on 15 March 2019 in Christchurch, New Zealand was the key driver for the development of online crisis protocols.[5] The terrorist attack demonstrated both the vulnerabilities of many online platforms, as well as the lack of coordination between platforms, governments and law enforcement agencies. The shooting was livestreamed on Facebook and consequently reuploaded across a variety of platforms, including YouTube.

The Christchurch Call was initiated by New Zealand Prime Minister Jacinda Ardern and French President Emmanuel Macron in the aftermath of the attack.[6] It is a set of 24 commitments for supporting governments and tech companies to implement with the goal of eliminating terrorist and violent extremist content online while protecting human rights and a free, open and secure internet. There is a specific commitment to work together on processes to enable a rapid, coordinated and effective response when terrorist and violent extremist content is being disseminated as part of a real-world attack. Call supporters have developed a suite of interlocking, voluntary protocols to give effect to this commitment.

These protocols: a) define what constitutes a crisis and when it is considered over; b) set out the roles of different stakeholders and the actions they will take in a crisis; and c) establish communication channels between those stakeholders to ensure quick and proportionate responses. These protocols include the Christchurch Call Crisis Response Protocol, the Global Internet Forum to Counter Terrorism (GIFCT) Content Incident Protocol (CIP), the EU Crisis Protocol (EUCP), and the Terrorist Content Analytics Platform (TCAP) Crisis Protocol Policy (see Table 1 below for further details on the scope and nature of these various protocols). Complementing these international initiatives, at the national level, there is the New Zealand Online Crisis Response Process and the 'online crisis event' process established by the Australian Online Safety Act.[7]

The Global Internet Forum to Counter Terrorism (GIFCT) Content Incident Protocol (CIP)

Following the Christchurch Call, the already-existing, industry-led GIFCT was adapted as an independent non-profit organisation and its resources and membership significantly expanded.[8] Additionally, the Content Incident Protocol (CIP) was developed. The CIP process consists of three levels: 1) incident, 2) content incident, and 3) content incident protocol. On the first level, there are weekly briefings shared with platforms for awareness which do not necessarily require immediate action. These briefings also feed into the various transparency reports published by GIFCT members. The second, the content incident level, is reached when content by a perpetrator or accomplice of a violent extremist attack is detected.[9] This content is then hashed and added to a database, which can be accessed by GIFCT members to ensure the content is detected and removed on their own platforms.[10] The third level, the CIP, is only activated in case of a livestreamed and ongoing real-world threat.[11] The full CIP was activated after the 2019 shooting in Halle, Germany, and the 2022 shootings in Glendale, Arizona and Buffalo, New York in the US.

The EU Crisis Protocol

The EU Crisis Protocol (EUCP) was adopted by the EU Internet Forum in 2019 as a voluntary framework to facilitate a rapid and coordinated cross-border response mechanism to combat the spread of terrorist content online.[12] Importantly, the protocol is not an everyday tool to address terrorist content online as it requires specific and high-threshold criteria linked to the nature of the terrorist attack to be met before it can be activated. Therefore, since its adoption, the EUCP has only been activated once, following the Islamic terrorist attack on school teacher Samuel Paty in a Paris suburb on 16 October 2020. In contrast to the Global Internet Forum to Counter Terrorism (GIFCT) Content Incident Protocol (CIP), the EUCP is not only to limit the virality of terrorist online content, but to actively support law enforcement investigations. The process is not automated, but instead relies heavily on the coordination between national authorities, private companies, and Europol. In light of this, a new platform is currently being developed to improve the coordination between all stakeholders and consolidate relevant information and communication channels. The entire process on behalf of Europol is subject to the supervision by the European Data Protection Supervisor to ensure the protection of the fundamental right to effective data protection.[13]

Geography	Name	Type	Primary bodies involved	Platforms covered	Type of content covered
Global	Global Internet Forum to Counter Terrorism, (GIFCT) Content Incident Protocol	Voluntary, industry-led	GIFCT	Airbnb, Amazon, Discord, Dropbox, Facebook, Instagram, JustPaste.it, LinkedIn, Mailchimp, Mega.nz Microsoft, Pinterest, Tumblr, Twitter, WhatsApp, WordPress.com & YouTube	Terrorist and violent extremist content
Global	Terrorist Content Analytics Platform (TCAP) Crisis Protocol Policy	Voluntary, aims to particularly support smaller platforms	Tech Against Terrorism, funded by Public Safety Canada	Available to all platforms	Terrorist and violent extremist content
EU	EU Crisis Protocol	Voluntary, but see Regulation 2021/784 for binding obligations on take-down and preservation of content	Europol	Meta, Twitter, Google, Microsoft, Dropbox, JustPaste.it, Dailymotion, Telegram, TikTok, Yubo, Discord, Vimeo & Snap	Terrorist and violent extremist content
Global	Christchurch Call Online Crisis Response Protocol	Voluntary intergovernmental coordination	Christchurch Call governments and industry supporters, led by the governments of France and New Zealand. Also Civil Society, Advisory Network	Tech companies that support the Call are Amazon, Meta, Google, YouTube, Zoom, Dailymotion, Microsoft, Qwant, JV, LINE, Twitter, Roblox, Mega & Clubhouse	Terrorist and violent extremist content
NZ	Online Crisis Response Process	Voluntary, but see Films, Videos, and Publications Classification Act 1993 (2019 update) for binding take-down obligations	Department of Internal Affairs, New Zealand	All, including internet service providers	Terrorist and violent extremist content
Australia	Abhorrent Violent Conduct Powers in an online crisis event	Non-statutory, but see Online Safety Act 2021 & Criminal Code Amendment Act 2019	eSafety Commissioner	All, including internet service providers	Abhorrent violent conduct material
UK	Crisis Response Protocol	Voluntary arrangement, supported by Terrorism Act 2006, which provides for information-sharing between the government and industry, and take-down notices for internet service providers.	Home Office and Counter Terrorism Policing, including the Counter Terrorism Internet Referral Unit (CTIRU)	All	Online content linked to a terrorist act

Table 1: Overview of existing online crisis protocols (non-exhaustive).

Lessons learned from existing online crisis protocols

After consultation with key stakeholders involved in the design and implementation of the crisis protocols, a consensus emerged around four key areas that should inform policymakers working to develop new protocols or improve existing ones. While the range of risks identified in the DSA is much larger than those covered by existing online crisis protocols, all these lessons still apply.

1. **Smaller platforms and 'alt-tech' services play a key role in the spread of violent extremist content**. This has been particularly evident in the aftermath of the far-right terrorist attack in Buffalo, USA in May 2022, when parts of the livestream were viewed by millions on smaller or fringe platforms such as Streamable.[14] These types of platforms are often not members of the GIFCT (or signatories to the Christchurch Call), and are unlikely to fall under the DSA's very large online platform threshold. There are two dimensions to this issue: firstly, many smaller platforms lack both the capacities and capabilities to swiftly respond to a crisis event, even though they do not wish to host this kind of content; secondly, there are platforms (of varying sizes) that are actively adversarial and are therefore not willing to remove violent extremist material (often based on arguments of freedom of speech).[15] Crisis protocols must account for this, for example by ensuring links to such content on fringe platforms are removed from the larger platforms.

2. **Evidence must be preserved while ensuring timely removal of content from platforms.** Here, mechanisms need to be in place to ensure content is preserved for legitimate purposes including law enforcement, international investigations, judicial processes, journalism and research. Platforms that are often the first to encounter the material must implement processes that retain and safely store evidence to be shared with relevant stakeholders in line with local laws, while blocking public circulation of such content.[16] Already in 2017, in the context of atrocities perpetrated during the Syrian civil war, CSOs criticised YouTube for deleting videos on its platform which "could be used in potential war crime prosecutions."[17] All stakeholders involved in the implementation of an online crisis protocol must ensure that their actions do not impede the ability of law enforcement and prosecutors to hold perpetrators to account.[18]

3. **Expertise from the CVE space must be shared with the key stakeholders tasked with designing and implementing the broader crisis mechanisms of the DSA, including the Digital Services Coordinators (DSCs), to improve capabilities, build capacities and avoid duplicative efforts.** The DSCs will likely come from different backgrounds (such as media regulation) and may therefore require additional expertise concerning the (technical) implementation of crisis responses. It is crucial that the crisis protocols and crisis mechanisms introduced as part of this legislation seek to complement and build on the experience of existing protocols and processes. Duplicative or parallel communication channels and processes may hinder swift coordination during a crisis event.

4. **Fundamental rights must be safeguarded during and in the aftermath of crisis events.** All participants agreed that crisis protocols and mechanisms must be designed and implemented with the goal to protect fundamental rights, and that effective safeguards must be enshrined in the process, including access to remedy. This also includes engaging with a wide network of stakeholders from the very beginning, including civil society (as envisioned by the DSA), and ensure there are regular review processes and impact assessments in place (see for example GIFCT working groups).[19] As the need for swift action may hinder real-time transparency, transparency in the design and evaluation phase of these protocols are ever more crucial.

Democratic safeguards – definitions, competencies and procedures

The second meeting of the working group focused on safeguarding fundamental rights in the context of the online crisis protocols and mechanisms envisioned by the DSA. As described in a public statement by Article 19, EDRi, Access Now and 20 additional signatories, a key concern was the lack of transparency during the trialogue process that led to the additional crisis protocols and mechanisms in the DSA, which some working group members perceived as lacking in democratic legitimacy.[20] Beyond these procedural concerns, the key issues raised were:

1. **The Commission must clarify whether it will interpret a crisis event as defined in the DSA as a "state of emergency" in international human rights law.** IHRL experts raised the issue of ambiguous language in the current DSA provisions, which do not clarify whether Article 36 constitutes such an emergency. Consequentially, it is unclear whether the mechanism foresees a potential derogation of fundamental rights under the Siracusa Principles, or whether the article simply formalises the expectations placed on platforms during times of crisis, without declaring a full state of emergency.[21] In other words, does the state of crisis change anything in terms of human rights protections? According to IHRL, a derogation of fundamental rights can only be invoked when there is a "threat to the life of a nation".[22] Such an emergency can only be declared by member states, not the Commission. In a similar vein, it should be clarified how the DSA crisis response mechanism complements the existing EU Integrated Crisis Response (IPCR) led by the presidency of the Council, as well as the existing EUCP led by Europol.
2. **There is a need for a complementary rapid response mechanism during crisis events through which CSOs can flag incorrect removals directly to platforms.** Particularly during crisis events, the incentives for platforms within the scope of the DSA to act in a risk-averse manner may lead to the illegitimate removal of speech. However, it is precisely those situations in which freedom of expression and the need to access (accurate) information becomes most crucial. So, a crisis mechanism seeking to safeguard fundamental rights must also include avenues for swift redress to curb the risk of over-blocking, which could lead to fundamental rights violations.
3. **Delegated acts following the DSA must limit the role of the Commission by empowering the Board and improving the Board's ability to act as an independent oversight body.** In its current shape, there are few checks and balances regarding executive power during online crisis events contained within the DSA.[23] Delegated acts must elaborate on the role of the DSCs Board, and ensure the Board has sufficient expertise and capacities to operate independently from the Commission as an oversight body.
4. **There must be robust and timely data access provisions for independent researchers in order to evaluate the effectiveness and proportionality of the crisis response, potentially as part of a human rights impact assessment.** These insights can help evaluate whether Article 37 is a proportionate tool to deal with crisis events, or whether the general risk management processes contained in the DSA are robust enough to deal with these extraordinary situations. Additionally, such data can provide clarity as to what fundamental rights were affected by the crisis as well as how the response either mitigated these effects or undermined rights. Guidance for the design of such data access provisions may be found in the recently published report of the European Digital Media Observatory's working group on platform-to-researcher data access.[24]

Conclusion

This policy paper set out to contextualise current discussions around the role of online crisis response mechanisms to combat illegal and harmful content. This builds on the debate of online crisis response protocols implemented in the aftermath of the 2019 far-right terrorist attack in Christchurch. While many lessons can be learned from existing online crisis protocols designed and implemented to combat terrorism and violent extremism online, the key task is now to translate these findings into tangible policies that can be applied beyond the CVE space, while ensuring the protection of fundamental rights remains the overarching goal. Policymakers must ensure that existing inter- and intragovernmental coordination mechanisms, as well as industry-led initiatives, are taken into consideration to avoid duplicative efforts that could hamper effective responses. There must also be robust transparency requirements to ensure independent scrutiny of the effectiveness and proportionality of measures taken under the crisis response mechanisms, with a special emphasis on the impact of the free exercise of fundamental rights online. This includes public and legislative oversight of actions taken during a crisis event by both platforms as well as governments, regulators and international organisations.

Endnotes

1 Article 19. (13 April 2022). *EU: Digital Services Act crisis response mechanism must honour human rights.* https://www.article19.org/resources/eu-digital-services-act-crisis-response-must-respect-human-rights/.

2 An ISD study has shown how the EU ban is being circumvented. See Kata Balint, Jordan Wildon, Francesca Arcostanzo and Kevin D. Reyes (October 2022). *Effectiveness of the Sanctions on Russian State-Affiliated Media in the EU – An investigation into website traffic & possible circumvention methods.* https://www.isdglobal.org/isd-publications/effectiveness-of-the-sanctions-on-russian-state-affiliated-media-in-the-eu-an-investigation-into-website-traffic-possible-circumvention-methods-2/. See also Sara Bundtzen and Mauritius Dorn (5 April 2022). *Banning RT and Sputnik Across Europe: What Does it Hold for the Future of Platform Regulation?* https://www.isdglobal.org/digital_dispatches/banning-rt-and-sputnik-across-europe-what-does-it-hold-for-the-future-of-platform-regulation/.

3 European Parliament. (October 2022). *REGULATION (EU) 2022/... OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of ... on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC.* https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/IMCO/DV/2022/06-15/DSA_2020_0361COD_EN.pdf.

4 European Parliament. (September 2022). *Corrigendum.* https://www.europarl.europa.eu/doceo/document/TA-9-2022-0269-FNL-COR01_EN.pdf.

5 In 2017, the UK government established one of the first of such online crisis protocols, in response to a series of domestic terrorist attacks. For further information see Global Internet Forum to Counter Terrorism. (2022). *Introducing 2022 GIFCT working group outputs* (p10-11). https://gifct.org/wp-content/uploads/2022/07/GIFCT-22WG-CRP-MapGap-1.1.pdf/.

6 Christchurch Call. *Christchurch Call text.* https://www.christchurchcall.com/about/christchurch-call-text/.

7 eSafety Commissioner. (December 2021). *Abhorrent violent conduct powers: regulatory guidance (eSC RG 5).* https://www.esafety.gov.au/sites/default/files/2022-03/Abhorrent%20Violent%20Conduct%20Powers%20Regulatory%20Guidance.pdf.

8 Global Internet Forum to Counter Terrorism. *Membership.* https://gifct.org/membership/.

9 Global Internet Forum to Counter Terrorism. *Content incident protocol.* https://gifct.org/content-incident-protocol/.

10 Global Internet Forum to Counter Terrorism. *Tech innovation.* https://gifct.org/tech-innovation/.

11 Global Internet Forum to Counter Terrorism. *Content incident protocol.* https://gifct.org/content-incident-protocol/.

12 European Commission. (October 2019). *A Europe that protects EU Crisis Protocol: responding to terrorist content online.* https://home-affairs.ec.europa.eu/system/files/2019-10/20191007_agenda-security-factsheet-eu-crisis-protocol_en.pdf.

13 Europol. (November 2021). *Data protection and transparency.* https://www.europol.europa.eu/about-europol/data-protection-transparency.

14 Kellen Browning and Ryan Mac. (May 2022). *After Buffalo shooting video spreads, social platforms face questions.* New York Times https://www.nytimes.com/2022/05/15/business/buffalo-shooting-social-media.html.

15 See Tech Against Terrorism. https://www.techagainstterrorism.org/ for resources and training available for smaller tech companies.

16 WG members acknowledged the potential abuse of such mechanism in authoritarian contexts. So these measures should be complemented by sufficient democratic safeguards. See Global Internet Forum to Counter Terrorism. (2022). *Introducing 2022 GIFCT working group outputs* (p15). https://gifct.org/working-groups/.

17 Malachy Browne. (August 2017). *YouTube removes videos showing atrocities in Syria.* New York Times. https://www.nytimes.com/2017/08/22/world/middleeast/syria-youtube-videos-isis.html.

18 Business for Social Responsibility. (July 2021). *Human rights impact assessment: Global Internet Forum to Counter Terrorism.* https://www.bsr.org/en/our-insights/report-view/human-rights-impact-assessment-global-internet-forum-to-counter-terrorism. This notes the importance of all these uses of content in ensuring that victims of human rights abuses can access effective remedies.

19 Global Internet Forum to Counter Terrorism. *Working groups 2022.* https://gifct.org/working-groups/.

20 Article 19. (13 April 2022). *EU: Digital Services Act crisis response mechanism must honour human rights.* https://www.article19.org/resources/eu-digital-services-act-crisis-response-must-respect-human-rights/.

21 International Commission of Jurists. (1 July 1984). *Siracusa Principles on the Limitation and Derogation Provisions in the International Covenant on Civil and Political Rights.* https://www.icj.org/siracusa-principles-on-the-limitation-and-derogation-provisions-in-the-international-covenant-on-civil-and-political-rights/.

22 It is noted that some working group members raised the concern that the IHRL definition of an "emergency" in itself is broader and less clear than the definition of a crisis contained in the DSA.

23 Specifically, current checks and balances on the powers of the Commission for the binding crisis response mechanism (Article 36) state that the Commission must act on the recommendation of the Board in requiring crisis-specific actions of VLOPs and VLOSE. Additionally, the Commission is tasked with ensuring VLOP and VLOSE action is "necessary, justified and proportionate", that the measures respect fundamental rights, and are limited to a maximum of three months. The Commission must also make its decision "publicly available" and must "inform the Board". After the implementation of measures on behalf of the VLOP or VLOSE, the Commission must report to the Board in its VLOP and VLOSE monitoring "at least on a monthly basis". The Commission may also amend its original decision (for example, by revoking the decision or extending the crisis response period for an additional three months), again only upon recommendation of the Board. Lastly, the Commission must report on the application of specific VLOP and VLOSE crisis measures to the European Parliament and the Council "on a yearly basis [...] and, in any event, three months after the end of the crisis". Checks and balances on Commission powers for voluntary crisis protocols (Article 48) include that the Board "may recommend that the Commission initiate the drawing up [...] of voluntary crisis protocols" and that the Commission "shall, as appropriate, involve Member States' authorities, and may also involve Union bodies". Additionally, the Commission "may, where necessary and appropriate, also involve civil society organisations".

24 European Digital Media Observatory. (May 2022). *Report of the European Digital Media Observatory's Working Group on Platform-to-Researcher Data Access*. https://edmo.eu/wp-content/uploads/2022/02/Report-of-the-European-Digital-Media-Observatorys-Working-Group-on-Platform-to-Researcher-Data-Access-2022.pdf.

ISD | Institute for Strategic Dialogue

Amman | Berlin | London | Paris | Washington DC

www.isdgermany.org

sponsored by

Federal Foreign Office