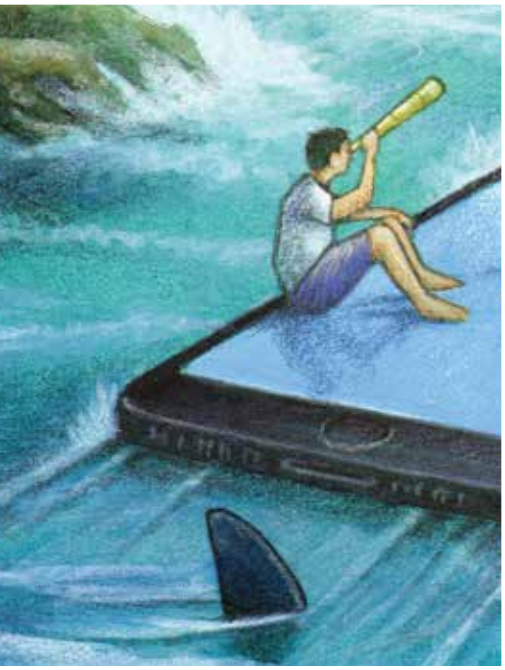
One in a Series of Working Papers from the Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression



The Proposed EU Terrorism Content Regulation:

Analysis and Recommendations with Respect to Freedom of Expression Implications

Joris van Hoboken
Vrije Universiteit Brussels and University of Amsterdam

May 3, 2019

TRANSATLANTIC WORKING GROUP

The Transatlantic Working Group Papers Series

Co-Chairs Reports

Co-Chairs Reports from TWG's Three Sessions: Ditchley Park, Santa Monica, and Bellagio.

Freedom of Expression and Intermediary Liability

Freedom of Expression: A Comparative Summary of United States and European Law
B. Heller & J. van Hoboken, May 3, 2019.

Design Principles for Intermediary Liability Laws
J. van Hoboken & D. Keller, October 8, 2019.

Existing Legislative Initiatives

An Analysis of Germany's NetzDG Law
H. Tworek & P. Leerssen, April 15, 2019.

The Proposed EU Terrorism Content Regulation: Analysis and Recommendations with Respect to Freedom of Expression Implications
J. van Hoboken, May 3, 2019.

Combating Terrorist-Related Content Through AI and Information Sharing
B. Heller, April 26, 2019.

The European Commission's Code of Conduct for Countering Illegal Hate Speech Online: An Analysis of Freedom of Expression Implications
B. Bukovská, May 7, 2019.

The EU Code of Practice on Disinformation: The Difficulty of Regulating a Nebulous Problem
P.H. Chase, August 29, 2019.

A Cycle of Censorship: The UK White Paper on Online Harms and the Dangers of Regulating Disinformation
P. Pomerantsev, October 1, 2019.

U.S. Initiatives to Counter Harmful Speech and Disinformation on Social Media
A. Shahbaz, June 11, 2019.

ABC Framework to Address Disinformation

Actors, Behaviors, Content: A Disinformation ABC: Highlighting Three Vectors of Viral Deception to Guide Industry & Regulatory Responses
C. François, September 20, 2019.

Transparency and Accountability Solutions

Transparency Requirements for Digital Social Media Platforms: Recommendations for Policy Makers and Industry
M. MacCarthy, February 12, 2020.

Dispute Resolution and Content Moderation: Fair, Accountable, Independent, Transparent, and Effective
H. Tworek, R. Ó Fathaigh, L. Bruggeman & C. Tenove, January 14, 2020.

Algorithms and Artificial Intelligence

An Examination of the Algorithmic Accountability Act of 2019
M. MacCarthy, October 24, 2019.

Artificial Intelligence, Content Moderation, and Freedom of Expression
E. Llansó, J. van Hoboken, P. Leerssen & J. Harambam, February 26, 2020.

www.annenbergpublicpolicycenter.org/twg

A project of the Annenberg Public Policy Center of the University of Pennsylvania in partnership with
The Annenberg Foundation Trust at Sunnylands and the Institute for Information Law of the University of Amsterdam,
with support from the Embassy of the Kingdom of the Netherlands in Washington, D.C.

The Proposed EU Terrorism Content Regulation:

Analysis and Recommendations with Respect to Freedom of Expression Implications†

Joris van Hoboken, Vrije Universiteit Brussels and University of Amsterdam[1]

May 3, 2019

Contents

Introduction and recommendations

In the last two decades, the use of internet communications and related services for terrorism[2] (the live streaming of the Christchurch mosque shooting and subsequent viral distribution through white supremacist networks being the latest high-profile example[3]) has been a major area of concern for government regulation of the internet. A series of European Union legislative and policy initiatives has defined new terrorism-related crimes at the EU level, including policies for law enforcement and the responsibility of online service providers. Most recent is the new proposal for a so-called Terrorism Content Regulation (TERREG). The European Commission proposed this measure in September 2018 and it is currently under debate in the European Parliament and the Council.[4]

---

† One in a series: A working paper of the Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression. Read about the TWG: https://www.ivir.nl/twg/.

The Transatlantic Working Group (TWG) used part of its first meeting at Ditchley, UK, to discuss the strengths and weaknesses of the TERREG proposal on the basis of an earlier version of this document.[5] This document has been updated to reflect crucial insights from these discussions as well as recommendations in light of the ongoing debate about the proposal at the EU level. Taking into account these discussions in the TWG, the central conclusions and recommendations are as follows:

- It is essential, in particular given the difficulty of defining terrorism in the first place, that legislative **definitions of "terrorism content"** strictly follow established rule of law and freedom of expression requirements. The original proposal's definitions are too wide in this respect and, if adopted, can be expected to become a significant source of abuse.
- Content removal laws like the TERREG proposal risk concentrating resources into an area with **limited tangible benefits**. The proposal is not sufficiently integrated in and connected to the broader legal and policy framework with respect to violent extremism and terrorism.
- The proposed content takedown order procedure should offer **independent judicial oversight** over public interferences with freedom of expression. In principle, such prior independent review should be a requirement for content takedown orders to be issued. For emergency situations, which should be adequately, explicitly and strictly defined, such review should still take place as a rule (and not be made dependent on an appeal), but could be started immediately after an emergency removal order is issued.
- The **one-hour removal** time frame is too rigid and simplistic. A more flexible requirement (promptly, without undue delay) would signal similar urgency, while better respecting established freedom of expression and due process values.
- The proposal's **referral procedure**, which requires platforms to handle law enforcement notifications under their terms of service standards, undermines due process as well as public legitimacy and accountability for limitations on freedom of expression.
- Although a full **separation of public and private regulation** may not be feasible, new rules on public enforcement actions with respect to online expression should follow and further develop established legal safeguards.
- The proposal enlists **platforms as de facto regulators** of online speech. This regulation through proxy challenges the legitimacy of subsequent restrictions on freedom of expression, complicates legal action on behalf of users' freedom of expression, and poses a central challenge to protecting free expression online.
- **Public accountability** and reporting on the use of proposed measures and procedures by competent authorities should be significantly enhanced.
- The **proactive monitoring provisions** violate the ban on preventive monitoring from Article 15 of the e-Commerce Directive (ECD), and lack clarity and supporting evidence for the effective and proportionate use of automation in tackling relevant content. They would create significant legal uncertainty and can be expected to cause large-scale removal of legal speech by relevant platforms.
- Automation in content moderation can be helpful in tackling issues at scale, but there are inherent risks and limitations as a result of the current state of the art of **artificial intelligence for content moderation** purposes and the lack of appropriate and functioning safeguards for users.

- The application of automation in online content moderation should not result in a shift from a presumption of legality of online information and ideas to a presumption of illegality. The principle of **freedom of expression by default** should be developed and implemented.
- **New institutions are needed** to support rule of law and fundamental rights safeguards in the development, application and regulation of new forms of AI in online content moderation.

The TERREG proposal and its freedom of expression implications

The core aim of the TERREG proposal is to tackle the availability of "terrorism content" online, thereby preventing potential radicalization and support for terrorism caused by the dissemination of such content.[6] The proposal does so by (1) providing a general definition of terrorism content at the EU level, (2) establishing two mechanisms for public authorities to obtain removal of relevant content by a broad class of service providers (orders and referrals) and (3) imposing new duties of care on relevant service providers to combat the availability of similar content through their services, including through proactive automated means. A key part of the proposal is that content removal orders would require an effective response in as little as one hour.

The proposal is the first legislative text in Europe, together with the new copyright proposal, to require proactive filtering of illegal content, breaking with the e-Commerce Directive approach to intermediary liability. The proposal builds on earlier policy documents released by the European Commission in the broader area of tackling illegal content online, and the co-regulatory initiatives of the EU Hate Speech Code of Conduct and the EU Internet Forum. As this document was being finalized, the European Parliament adopted the LIBE Committee report coming out of the EP committees. This sets the stage for the "trialogue negotiations" between the Council, the Parliament and the Commission to be initiated.[7] The LIBE Committee report and the Council position diverge significantly on many crucial aspects of the proposal from a freedom of expression perspective.

The implications for freedom of expression of the proposal are varied, significant and widely acknowledged.[8] The proposal itself contains a number of safeguards to address freedom of expression concerns. Most significantly, the proposal puts forward an obligation on service providers to allow users to complain if they believe their content has been removed unjustifiably. And the proposal requires human oversight and verification of automated tools for the removal of terrorism content to prevent unjustified removals.

These proposed safeguards notwithstanding, the draft regulation presents a clear threat to freedom of expression. First, the definitions of terrorism content lack the legal detail and precision that should be required for restrictions on freedom of expression. Second, the proposal targets a broad heterogeneous set of intermediary and online service providers and further enlists them into a project of privatized enforcement without proper human rights accountability. Third, the safeguards in the proposal fall short of European and international freedom of expression standards and best practices.

Definitions of targeted speech and communications

Key findings:

- Weak evidence for targeting speech defined as terrorism content in support of counter-radicalization;
- Definition of terrorism content does not include an intent requirement;
- Proposed definitions are broader than the criminal offences currently defined in EU law;
- Definition of "terrorism content" can easily include protected speech, while being subject to takedown orders and referrals;
- Definitions fail to meet the (freedom of expression) prescribed by law standard.

A first concern is that the proposal provides insufficient evidence demonstrating a causal connection between terrorist actions and "terrorism content" as defined in the proposal. While the covered content will generally be shocking and disturbing,[9] there is no clear evidence linking these particular kinds of content and terrorist radicalization or offenses. There is evidence that the internet allows terrorists to effectively disseminate their motivations for committing their crimes,[10] but evidence shows that radicalization may as well be caused by consumption of daily news (including coverage of terrorist acts). Available evidence also shows that radicalization tends to occur primarily as a result of offline rather than online dynamics.[11] This puts the proposal on a weak footing, including from a freedom of expression perspective.

If one accepts that new legal procedures are needed to tackle certain terrorism-related information and communications on the internet, the regulatory challenge is to define precisely which information should be allowed to be targeted by public authorities,[12] thereby satisfying European and international freedom of expression standards. The weak evidence for a causal link between terrorism offenses and terrorism content should have informed a narrow definition of which material could be targeted, namely material that causes an actual risk and/or imminent harm. Under the broad and seemingly simple notion of "terrorism content" in this proposal, however, lurks a wide variety of targeted terrorism-related offences and activities. Notably, the definitions in the proposal are broader than the speech-related terrorist offences currently provided for under EU law. The new definitions generally lack the precision required by tests under freedom of expression law and do not include an intent requirement.

The European Commission's proposed definition for terrorism content in Article 2(5) is the following:

> 'terrorist content' means meets one or more of the following information:
>
> (a) inciting or advocating, including by glorifying, the commission of terrorist offences, thereby causing a danger that such acts be committed;
>
> (b) encouraging the contribution to terrorist offences;
>
> (c) promoting the activities of a terrorist group, in particular by encouraging the participation in or support to a terrorist group within the meaning of Article 2(3) of Directive (EU) 2017/541;
>
> (d) instructing on methods or techniques for the purpose of committing terrorist offences.

The current Council version amends as follows (edits underscored):

(5) 'terrorist content' means material which may contribute to the commission of the intentional acts, as listed in Article 3(1)(a) to (i) of the Directive 2017/541, by:

(aa) threatening to commit a terrorist offence;

(a) inciting or advocating, such as the glorification of terrorist acts, the commission of terrorist offences, thereby causing a danger that such acts be committed;

(b) soliciting persons or a group of persons to commit or contribute to terrorist offences;

(c) promoting the activities of a terrorist group, in particular by soliciting persons or a group of persons to participate in or support the criminal activities of a terrorist group within the meaning of Article 2(3) of Directive (EU) 2017/541;

(d) instructing on methods or techniques for the purpose of committing terrorist offences

The lead European Parliament Committee's report (LIBE) offers the following definition (edits underscored):

(5) 'Terrorist content' means one or more of the following material:

(a) inciting the commission of one of the intentional offences listed in points (a) to (i) of Article 3(1) of Directive (EU) 2017/541, where such conduct, directly or indirectly, such as by the glorification of terrorist acts, advocates the commission of terrorist offences, thereby causing a danger that one or more such offences may be committed intentionally,

(b) soliciting another person or group of persons to commit or contribute to the commission of one of the offences listed in points (a) to (i) of Article 3(1), of Directive (EU) 2017/541, thereby causing a danger that one of more such offences may be committed intentionally;

(c) soliciting another person or group of persons to participate in the activities of a terrorist group**,** including by supplying information or material resources, or by funding its activities in any way within the meaning of Article 4 of Directive (EU) 2017/541, thereby causing a danger that one of more such offences may be committed intentionally;

(d) providing instruction on the making or use of explosives, firearms or other weapons or noxious or hazardous substances, or on other specific methods or techniques for the purpose of committing or contributing to the commission of one of the terrorist offences listed in points (a) to (i) of Article 3(1) of Directive (EU) 2017/541;

(e) depicting the commission of one or more of the offences listed in points (a) to (i) of Article 3 (1) of Directive (EU) 2017/541, and thereby causing a danger that one or more such offences may be committed intentionally.

For terrorist offences, EU law typically includes strict requirements of intent and likelihood of speech resulting in criminal action. But these are notably absent from the above definitions of "terrorism content" and "terrorism content dissemination" in the proposal.[13] By using relatively weak legal language, such as "causing a danger that" (EC, proposal) or "may contribute to" (Council and EP version), the definitions open up more space for restrictions than is necessary.

In other words, online media posts could be considered "terrorism content" even though they are clearly not intended to support or incite terrorism. To give an example: TERREG might encourage platforms to remove content by (citizen) journalists and opinion leaders who are responding to and quoting from terrorist propaganda. These parties may actually be rebutting terrorist causes, not supporting them, but since the Directive's definitions of "terrorist content" do not include intent, platforms may still find reason to remove this speech. It is worth noting, finally, that the amended definitions of the European Parliament's LIBE report and the Council do include some references to intent, although these are predominantly tautological references to the intent requirements in the underlying terrorist offenses instead of intent requirements related to the posting of content and the intended results thereof.

The proposal's recitals do clarify the scope of the definitions. For instance, recital 9 stipulates that "content disseminated for educational, journalistic, counter-narrative or research purposes should be adequately protected," but these safeguards are lacking in the actual legal text of the proposal.

In addition, the definition of terrorism content is not connected to existing speech-related offences in the area of terrorism, such as recruitment, training, and financing, which were already defined previously in Article 5-12 of Directive 2017/541. The proposal seems to have prioritized a broad and sweeping definition over a precise, legally sound set of definitions. More precise definitions would help to ensure that the Directive's most far-reaching provisions, such as the framework for proactive measures, are only imposed when absolutely necessary.

With the current definitions, the TERREG proposal provides for government-sanctioned removal mechanisms for speech that does not present an actual or imminent risk for terrorist offenses, including communications that are themselves not necessarily criminal under EU law. This raises pertinent questions about the precise legal basis for government-mandated removal of such material, and whether these mechanisms can be considered necessary and proportionate in the first place.

Regulation by proxy and privatized enforcement

Key findings:

- Proposal targets very broad range of online service providers;
- Proposal undermines the existing safe harbor regime in e-Commerce Directive;
- Proposal violates ban on general monitoring;
- Proposal codifies problematic practice of informal referrals by law enforcement and extralegal removal of information online.

The proposal targets a broad range of online service providers ("hosting service providers"), ranging from cloud infrastructure companies to online marketplaces, file storage services, social media and

search engines. The proposal uses the definition of hosting service providers from the e-Commerce Directive (2000/31/EC), which was introduced to limit the responsibility that could be imposed on intermediary services. It now connects to this definition to introduce new obligations to police and remove online speech.[14] Thus, the proposal narrowly focuses on the ability of online service providers to act as control points and censors of expression online, without taking account of the precise role different services play in the online environment and their relationship with expressive activities they help to facilitate. The scope of the proposal is one of the key areas of debate and amendment. Infrastructural service providers, like Amazon Web Services or Microsoft Azure, and other cloud infrastructure or service companies with more remote relations to the actual content are excluded from the regulation in the LIBE report.

Furthermore, the proposal undermines the existing safe harbor regime in the e-Commerce Directive, by creating a proactive duty of care for hosting service providers and moving beyond the reactive notice and takedown obligations that follow from the ECD framework. The ECD, adopted in 2000, is currently under pressure from different sides and risks being eroded completely through a combination of this proposal and others (audiovisual regulation, copyright enforcement). The most striking departure from the ECD is the introduction of legal obligations to prevent known "terrorism content" from becoming available (upload filtering) and more general preventive duties to remove terrorism content through automated content recognition tools. These provisions violate the ban on general monitoring in Article 15 ECD, which the Court of Justice of the European Union has found to support the freedom of expression rights of internet users (e.g., the *Scarlet Extended SA v. Sabam* case).

The proposal codifies the already existing problematic practice of informal referrals by law enforcement and the subsequent extralegal removal of information online on the basis of a company's terms of service.[15] The proposed referral mechanism for online content does not entail a determination by an appropriate authority that the content falls within the definition of terrorism content and whether the content is actually illegal. Instead, when receiving referrals, a company must decide upon content removal on the basis of its terms of service, which tend to be much broader and more flexibly enforced than requirements under criminal procedural law. As a result, the proposal undermines existing legal procedures and due process safeguards for internet users. The proposal obliges services to operate a complaint procedure for internet users whose content is removed, but does not create effective avenues to appeal referrals at the source, i.e., the public authority that has made the referral. Finally, the broader law enforcement referral practices anticipated by the proposal could be amplified through existing industry coordination in the GIFCT hash-sharing database initiative, if law enforcement referrals become a significant source for this industry database.[16]

Deficient freedom of expression safeguards

Key findings:

- Inflexible one-hour response deadline for content takedown orders;
- Takedown orders lack independent judicial review;
- Safeguards for affected speakers/audiences;

- Risks related to deficiencies and bias in automated content recognition tools.

The proposal imposes an inflexible one-hour response deadline for content takedown orders. The stated reason for this short response window is that the most intense dissemination of terrorism content tends to take place in the first hours after its posting. This raises a number of concerns. First, if this is the case, how would a one-hour response window help to address this? Typically, it will take time for content that is posted online to be identified as "terrorism content" by relevant authorities.[17] On top of the time that it will take to process the takedown order, it seems unlikely that dissemination in the first hours can be effectively addressed.[18] Second, smaller service providers will likely lack the resources to provide for effective 24-hour staffing to be able to comply with this obligation. Third, the short window to process orders will incentivize service providers to minimize review of such orders. Considering the lack of judicial review on content takedown orders before they are sent to service providers, this presents a big risk for freedom of expression. Overall, a more flexible obligation to act expeditiously, without undue delay, would better support the necessary and proportionate requirement for interferences with online speech.

Another safeguard that is lacking is judicial review on content takedown orders before they are sent to service providers, or as soon as they are sent in the case of emergency situations in which a prior review would cause undue delay.[19] The proposal refers to appeal mechanisms for service providers that it expects to be in place in the Member State but the proposal does not set minimum standards for these appeal mechanisms.[20] This creates a significant risk of abuse. First, it creates the possibility for non-judicial authorities to use the procedure to censor content without due process. This is particularly problematic for European countries with broader rule of law issues.

Existing appeal mechanisms in the Member States are likely to take far longer than the stipulated one-hour response window.[21] The proposal's broad scope in terms of service providers, which would include collaborative journalism platforms and others for public debate, poses real dangers for robust debate on terrorism-related matters of public concern.[22] The lack of judicial review and effective appeal mechanisms for takedown orders is one of the aspects of the proposal that most clearly violates established freedom of expression case law.

The proposal adopts a narrow view of whose freedom of expression rights require protection. It does not recognize that service providers can invoke freedom of expression when confronted with takedown orders and referrals. It also doesn't recognize that there are others besides the users posting the content whose freedom of expression can be curtailed by the takedowns. First, those other internet users who would have wanted to access the content are now prevented from seeing it. Second, the authors of content that is taken down are not always the same as the individuals uploading the content and may see their expression taken down without an effective remedy. For this reason, it could be worth broadening standing in relevant appeal procedures beyond the user (re-)posting particular content to others unduly impacted in their freedom of expression.

Finally, the proposal clearly relies upon the efficacy of automation to identify and take down illegal content, and does so without weighing the evidence and research on these tools. The evidence on automation shows significant problems of false positives (and negatives), and bias with respect to the expression of different viewpoints and groups. The proposal stipulates that any proactive measures

"shall provide effective and appropriate safeguards to ensure that decisions taken concerning that content, in particular decisions to remove or disable content considered to be terrorist content, are accurate and well-founded," but doesn't say what that means. It merely states that safeguards should entail "human oversight and verifications where appropriate and, in any event, where a detailed assessment of the relevant context is required in order to determine whether or not the content is to be considered terrorist content." The proposals for automation appear to connect to the existing industry initiative of a shared database of hashes for violent extremist content that is used to proactively remove such content from their services.

Conclusion

The use of the internet for recruitment and the dissemination of violent extremist materials raises significant policy challenges for public authorities and internet services alike. Freedom of expression has an important role to play in shaping regulation and industry policies. It is clear from the above that the TERREG proposal creates substantial risks with respect to freedom of expression that should be addressed before its adoption.

Official documents

- European Commission proposal
- European Union's Legislative Observatory link with relevant links to the proposal and official texts adopted in the Parliament
- Latest public draft text of the Council, December 2018

Notes

[1] Prof. dr. Joris V. J. van Hoboken, Professor of Law at the Vrije Universiteit Brussels (VUB) and a Senior Researcher at the Institute for Information Law (IViR), University of Amsterdam. At VUB, I am appointed to the Chair "Fundamental Rights and the Digital Transformation," which is established at the Interdisciplinary Research Group on Law Science Technology & Society (LSTS), with the support of Microsoft.

[2] It's worth noting at the outset that there is no uniform definition of the term terrorism, and what is understood as terrorism has also changed significantly over time. This lack of clarity can be a cause for too broad application and misuse of relevant legislation adopted at the national level.

[3] For a detailed overview see Wikipedia, Christchurch mosque shootings: https://en.wikipedia.org/wiki/Christchurch_mosque_shootings#Video_distribution.

[4] The European Parliament's committee rapporteurs for TERREG are: Daniel Dalton, UK (LIBE Committee, lead); Julia Reda, DE (IMCO Committee); Julie Ward, UK (CULT Rapporteur). The Council's presidency is currently held by Romania.

[5] In addition, the hash-sharing database industry initiative (GIFCT) was discussed.

[6] For an overview of counter radicalization strategies in media and communications, see Ferguson, Countering violent extremism through media and communication strategies: A review of the evidence, 2016 available at http://www.paccsresearch.org.uk/wp-content/uploads/2016/03/Countering-Violent-Extremism-Through-Media-and-Communication-Strategies-.pdf. See also Daphne Keller, 'Internet Platforms: Observations on Speech, Danger and Money, A Hoover Institution Essay, 2018. Available at https://www.hoover.org/sites/default/files/research/docs/keller_webreadypdf_final.pdf. Keller stresses how little is

known about the impact of content removal practices on people at risk of radicalization and the dangers of well-intentioned campaigns against violent extremist content backfiring.

[7] Trialogues are expected to begin after the summer.

[8] See, e.g., Aleksandra Kuczerawy, "The Proposed Regulation on Preventing the Dissemination of Terrorist Content Online: Safeguards and Risks for Freedom of Expression," 2018. Available at http://dx.doi.org/10.2139/ssrn.3296864; Letter of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, the Special Rapporteur on the right to privacy and the Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism, 7 December 2018, available at https://spcommreports.ohchr.org/TMResultsBase/DownLoadPublicCommunicationFile?gId=24234; Faiza Patel, "EU 'Terrorist Content' Proposal Sets Dire Example for Free Speech Online," Just Security, 5 March, 2019, available at https://www.justsecurity.org/62857/eu-terrorist-content-proposal-sets-dire-free-speech-online/. EDRi, FRA and EDPS: Terrorist Content Regulation requires improvement for fundamental rights, 20 February 2019, https://edri.org/fra-edps-terrorist-content-regulation-fundamental-rights-terreg/.

[9] The fact that content is shocking and disturbing doesn't mean it won't be protected speech in Europe (*Handyside v. UK*).

[10] Most recently, for instance, CNN reports, that the San Diego synagogue shooter posted a letter on 8chan ("The letter writer talks about planning the attack and references other attacks on houses of worship, including the attack on the Tree of Life Synagogue in Pittsburgh and the Christchurch mosque shootings in New Zealand"). See Ray Sanchez and Artemis Moshtaghian, "Mayor says synagogue shooting in California that left 1 dead and 3 wounded was a 'hate crime,'" CNN, 28 Paril 2019, available at https://www.cnn.com/2019/04/27/us/san-diego-synagogue/index.html. Leading voices have called for the blocking of 8chan by dominant internet companies.

[11] See Ferguson 2016.

[12] The question of what is the right definition for internet companies to use when targeting terrorism and violent extremism is a different one. For a detailed discussion of relevant considerations for industry policies in this area, see Brian Fishman, "Crossroads: Counter-terrorism and the Internet," Texas National Security Review, Vol 2, Issue 2 (April 2019), available at https://tnsr.org/2019/04/crossroads-counter-terrorism-and-the-internet/ (arguing that public authorities may only see a tip of the iceberg of what a company like Facebook is doing with respect to terrorist content). Fishman leads efforts against terrorist and hate organizations at Facebook.

[13] The terrorist offences provided for at the EU level in Directive 2017/541 generally require that these acts, "given their nature or context, may seriously damage a country or an international organization." In addition, they are only defined as "terrorist offences" where committed with the aim of "seriously intimidating a population," "unduly compelling a government or an international organisation to perform or abstain from performing any act" and/or "seriously destabilising or destroying the fundamental political, constitutional, economic or social structures of a country or an international organization."

[14] The scope of the hosting service provider definition is legally contested and hotly debated. For a discussion, see Van Hoboken et al., Hosting Intermediary Services and Illegal Content Online, Study for the European Commission, (forthcoming).

[15] For a discussion of referrals from a human rights perspective, see Jason Pielemeier and Chris Sheehy, "Understanding the Human Rights Risks Associated with Internet Referral Units," The GNI Blog, 25 February 2019, available at https://medium.com/global-network-initiative-collection/understanding-the-human-rights-risks-associated-with-internet-referal-units-by-jason-pielemeier-b0b3feeb95c9.

[16] Due to the lack of transparency about the details of the GIFCT initiative, the question whether this could be the case is difficult to answer.

[17] It would have strengthened the proposal if better data was (made) available on the time it takes relevant authorities to become aware of relevant material online and in what ways that timeframe could be minimized effectively.

[18] The dynamics around the Christchurch mosque shooting show how difficult it is to effectively stop dissemination, as groups mobilize to circumvent removal mechanisms deployed by internet companies. See, e.g., Kate Klonick, "Inside the Team at Facebook That Dealt with the Christchurch Shooting," The New Yorker, 25 April 2019, available at https://www.newyorker.com/news/news-desk/inside-the-team-at-facebook-that-dealt-with-the-christchurch-shooting. See also Brian Fishman, "Crossroads: Counter-terrorism and the Internet," Texas National Security Review, Vol 2, Issue 2 (April 2019), available at https://tnsr.org/2019/04/crossroads-counter-terrorism-and-the-internet/.

[19] The proposal currently does not stipulate an emergency procedure, but seems to build on the assumption that for terrorism content as defined such an emergency always exists. Considering the issues with the definitions discussed earlier, it's not clear that this assumption is correct.

[20] The proposal does contain a mechanism for service providers to ask for clarification in case of missing information in the order or manifest errors.

[21] The proposal lacks documentation and analysis of existing laws and procedures in the Member States.

[22] The EC proposal only contains a (weak) reference to the need to protect journalistic coverage of terrorism content in the preamble (recital 9). In support of journalism and freedom of expression, the Council position adds some more detail to this recital and adds a provision that the "Regulation shall not have the effect of modifying the obligation to respect fundamental rights and fundamental legal principles as enshrined in Article 6 of the Treaty on the European Union" (Article 1(3)). This provision is purely declaratory: the Regulation has to comply with the Treaty of the EU and the Charter of Fundamental Rights regardless of this text. The EP has the clearest exception for journalism in its 1st reading of the proposal, with a new provision in Article 1, paragraph 2 a stating that the "Regulation shall not apply to content which is disseminated for educational, artistic, journalistic or research purposes, or for awareness raising purposes against terrorist activity, nor to content which represents an expression of polemic or controversial views in the course of public debate." All of these provisions leave important questions as to what will be effectively covered by these exceptions open, with a particular danger that the concept of journalism and who can claim to be engaged in journalistic activities, will be narrowly construed.