Introduction to Special Issue: The Practicalities and Complexities of (Regulating) Online Terrorist Content Moderation

Maura Conway and Stuart Macdonald

Maura Conway (maura.conway@dcu.ie) (ORCiD: 0000-0003-4216-8592) (Twitter: @galwaygrrl), Dublin City University, Swansea University

Stuart Macdonald (ORCID: 0000-0002-7483-9023) (Twitter: @CTProject_SM), Swansea University

According to the UK's Independent Reviewer of Terrorism Legislation, "most terrorism arrestees are profoundly engaged in expressing and consuming violent and hateful material online, and that online encouragement can be troublingly effective at promoting violence in others."[i] This has also been the experience of counterterrorism police.[ii] A recent study of individuals convicted of extremism offences in the UK provides empirical support for this view, concluding that the internet is playing an increasingly prominent role in radicalisation processes and that radicalisation now takes place primarily online.[iii] In light of these findings, it is unsurprising that the moderation of online terrorist and violent extremist content (TVEC) is a pressing concern for policymakers and practitioners alike. The European Union's (EU) Terrorist Content Online Regulation (TCO) imposes obligations on in-scope platforms, including to remove terrorist content within one hour of receiving a removal order from a competent authority.[iv] National legislatures too have enacted regulatory regimes, such as Germany's *Netzwerkduchsetzungsgesetz* ('Network Enforcement Act' or 'NetzDG' for short),[v] and the UK's currently in-process Online Safety Bill.[vi] As the articles in this special issue demonstrate, both content moderation and the enactment of accompanying regulatory regimes are complex tasks.

The special issue's opening article is by Jonathan Hall, the UK's Independent Reviewer of Terrorism Legislation,[vii] whose annual report for 2021 focused on the online realm.[viii] His article addresses the thorny issues of rights and values in online counter-terrorism. In the physical world, the principal means of deterring violence is generally via arrest and detention. Given physical violence does not take place online however, this article begins by reflecting on why content moderation is a less direct means of public protection than its 'real world' equivalent. It follows-up by elucidating how moderation of terrorism content encroaches on free expression and privacy. Hall concludes by arguing that the difficulty of adopting a rights-based approach to adjudicating between the latter and counterterrorism imperatives is due not just to the unique challenges presented by the internet to protection and exercise of fundamental rights, but also a continued lack of clarity around the compromises necessary for a free and functioning internet.

The University of Hamburg's Reem Ahmed takes up the issue of fundamental rights in her article too. 'Negotiating Fundamental Rights: Civil Society and the EU Regulation on Addressing the Dissemination of Terrorist Content Online' takes the evolution of the EU's TCO regulation as the basis for exploring the dynamics of politicization at play in some contemporary counterterrorism policy-making. To do this, she traces the key discourses that emerged from digital and human rights advocates, on the one hand, and the EU institutions, on the other, during the TCO's negotiation. The article thus draws together different perspectives apparent in the literature on counterterrorism's securitization versus its politicization, as well as the role of digital and human rights non-governmental organizations (NGOs) in contestation and norm diffusion in the area. Ahmed's study aims to demonstrate that whilst securitizing discourses remain important, some EU counterterrorism law is open to scrutiny at the policy-formation and negotiation stages.

Regulatory regimes and their development are also at issue in the special issue's third article by the University of the West of Scotland's Amy-Louise Watkin, titled 'Developing a Responsive Regulatory Approach to Online Terrorist Content on Tech Platforms.' In her article, Watkin draws attention to three major compliance issues likely to arise when tech platforms are required to implement regulations to counter online terrorist content: 1) some tech companies' lack of awareness and/or the necessary expertise required to comply; 2) some tech companies' lack of the necessary capacity and resources to comply; and 3) some tech companies' unwillingness to comply. Watkin's article argues insufficient consideration of these three issues could result in regulation unfairly penalizing tech platforms — particularly smaller platforms — and incentivizing actions that could jeopardize the rights and interests that regulation seeks to protect. For example, over-blocking and infringing on free speech. It argues, instead, for a responsive regulation framework and proposes four regulatory tracks that could be taken to try to minimize compliance issues.

The Global Internet Forum to Counter Terrorism (GIFCT)[ix] is a non-governmental organisation (NGO) founded by Facebook, Microsoft, Twitter, and YouTube in 2017 with the goal of preventing terrorists and violent extremists from exploiting their platforms. Amongst other things, GIFCT develops and shares technology, best practices, and other resources to improve the detection and removal of online terrorist content. Having an eye toward the practical steps necessary to meet a variety of regulatory requirements, GIFCT's Tom Thorley and Erin Saltman provide an overview of some of the algorithmic tools presently deployed by tech companies in their counter-extremism and terrorism content moderation efforts; describe and discuss some of the ethical and human rights issues arising from the deployment of such tools; and report the findings from a GIFCT trial testing a methodology for proactively surfacing content related to credible ongoing TVEC threats.

Content moderation is by nature limited in terms of what it can achieve – as research has consistently shown (i.e., displacement, migration, etc.).[x] These limitations mean that extremist and hateful actors are, the variety of above-described tools and activity notwithstanding, able to maintain an online foothold thereby leaving space for (digital) conflict with others that oppose them. The final special issue article departs somewhat from the other contributions therefore with its focus on digital antifascist activists' informal modes of disrupting online hate, which fill the gap left by – and stand in contrast to – more formal content moderation practices. Having said this, Michael Loadenthal's (University of

Cincinnati) 'We Protect Us: Cyber Persistent Digital Antifascism and Dual Use Knowledge' combines description and discussion of these grassroots efforts with comparison to US government cyber defense architects' doctrinal strategy of Persistent Engagement (PE) — the "dual use knowledge" of his title.

The five articles composing this special issue were selected from those delivered at the third international conference on Terrorism and Social Media (#TASMConf), hosted at Swansea University on 28–29 June 2022. Organized by the University's Cyber Threats Research Centre (CYTREC), the conference registered 201 delegates from 15 countries. In addition to academic researchers, these delegates included representatives from a wide range of non-academic stakeholders, including policymakers, law enforcement, social media companies, and think tanks. The keynote speakers were the UK's Independent Reviewer of Terrorism Legislation, Jonathan Hall KC, and Professor Maura Conway, Paddy Moriarty Professor of Government and International Studies at Dublin City University and founder of VOX-Pol.[xi] The conference concluded with the session 'In Conversation with the Tech Sector,' which featured Dina Hussein (Head of Counterterrorism and Dangerous Organizations Policy for Europe, the Middle East and Africa at Meta), Lucy Calladine (YouTube's Product Public Policy lead in EMEA) and Lisa McInerney (Twitter's then Global Policy Area Lead for Violent Organizations). In addition, a total of 97 others presented their research into extremists and terrorists' use of the Internet and allied issues across 24 breakout panels over the two days.[xii] The articles included herein showcase some of that research.

---

[i] Jonathan Hall, *The Terrorism Acts in 2021: Report of the Independent Reviewer of Terrorism Legislation on the Operation of the Terrorism Acts 2000 and 2006, and the Terrorism Prevention and Investigation Measures Act 2011* (His Majesty's Stationery Office, 2023), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1140911/E02876111_Terrorism_Acts_in_2021_Accessible.pdf, accessed March 18, 2023, 160.

[ii] Stuart Macdonald and Andrew Staniforth, *Tackling Online Terrorist Content Together: Counterterrorism Law Enforcement and Tech Company Cooperation*, (London: Global Network on Extremism and Technology, 2023), https://gnet-research.org/wp-content/uploads/2023/01/31-Tackling-Online-Terrorist-Content-Together_web.pdf, accessed March 19, 2023.

[iii] Jonathan Keynon, Jens Binder and Christopher Baker-Beall, *The internet and radicalisation pathways: technological advances, relevance of mental health and role of attackers* (HM Prison & Probation Service, 2022), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1121985/internet-radicalisation-report.pdf, accessed March 19, 2023.

[iv] The full text of Regulation (EU) 2021/784 of the European Parliament and of the Council of 29 April 2021 on Addressing the Dissemination of Terrorist Content Online is available at https://eur-lex.europa.eu/eli/reg/2021/784/oj, accessed March 31, 2023.

[v] The full text of the Act, in English, is available at https://germanlawarchive.iuscomp.org/?p=1245, accessed March 31, 2023.

[vi] For the most up-to-date version of the Bill and its progress, go to https://bills.parliament.uk/bills/3137, accessed March 31, 2023.

[vii] The Independent Reviewer's website is at https://terrorismlegislationreviewer.independent.gov.uk, accessed 31 March, 2023.

[viii] Jonathan Hall, *The Terrorism Acts in 2021: Report of the Independent Reviewer of Terrorism Legislation on the Operation of the Terrorism Acts 2000 and 2006, and the Terrorism Prevention and Investigation Measures Act 2011* (London: His Majesty's Stationery Office, 2023), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1140911/E02876111_Terrorism_Acts_in_2021_Accessible.pdf.

[ix] For more on the GIFCT and its activity, go to https://gifct.org.

[x] See, for example, Amarnath Amarasingam, Shiraz Maher, and Charlie Winter. 2021. *How Telegram Disruption Impacts Jihadist Platform Migration*. Lancaster, UK: CREST, https://crestresearch.ac.uk/resources/how-telegram-disruption-impacts-jihadist-platform-migration/, accessed April 1, 2023; Stuart Macdonald, Sara Giro Correia and Amy-Louise Watkin. 2019. 'Regulating Terrorist Content on Social Media: Automation and the Rule of Law.' *International Journal of Law in Context* 15(2).

[xi] For more on VOX-Pol, see https://www.voxpol.eu/.

[xii] Recordings of the keynote session and selected other presentations are available at https://youtube.com/playlist?list=PLDYmf88ufmvPtq5HbCXAKA2KW8RFga3tF.