Online Safety Bill: Distinguishing between public and private communication

Jonathan Hall KC (Independent Reviewer of Terrorism Legislation) and Professor Stuart Macdonald (Swansea University)

1. The focus of this briefing note is the distinction between public and private communication. In the Online Safety Bill, this distinction delineates the scope of OFCOM's power to issue terrorism content notices, to impose a proactive technology requirement, and to include in a code of practice a proactive technology measure.

2. Clause 203 specifies three factors that OFCOM must in particular consider when deciding whether content is communicated publicly or privately by means of a user-to-user services, for the purpose of the powers listed above. These are:
   - The number of individuals in the United Kingdom who are able to access the content by means of the service;
   - Any restrictions on who may access the content by means of the service; and,
   - The ease with which the content may be forwarded to or shared with users of the service other than those who originally encounter it.

3. The stipulation that OFCOM must consider these factors "in particular" indicates that it is a non-exhaustive list of potentially relevant considerations, although the identification of these criteria within the body of the statute accords them especial weight compared to potential other factors such as "reasonable expectation of privacy". This could be because, although reasonable expectations are the touchstone for whether an individual's private life is affected,[1] the function of the public/private divide in the Bill is less the protection of individual rights than delineation of OFCOM's enforcement powers.

4. To consider how clause 203 will operate in practice, in this note we apply it to the propaganda dissemination process of Islamic State (IS). The key features of this process include its multi-stage and cross-platform dimension. We explore the implications of this for clause 203 and propose amendments to the clause's existing wording.

IS propaganda dissemination process

5. A key node in the IS propaganda dissemination process is the platform Telegram. Telegram is a messaging app on which users can share an unlimited number of photos, videos and files, of up to 2 gigabytes each.[2] It has over 500 million active users[3] and is popular for its enhanced privacy and encryption.[4] Its features include: secret chats, with

[1] *Murray v Express Newspapers plc* [2008] EWCA Civ 446 at para 24, per Sir Anthony Clarke MR, summarising the principles stated by Lord Nicholls in *Campbell v MGN Ltd* [2004] UKHL 22; [2004] 2 AC 457.

[2] 'Telegram FAQ', https://telegram.org/faq, accessed February 9, 2023.

[3] ibid.

[4] Dave Johnson, 'What is Telegram? A quick guide to the fast and secure messaging platform' *Business Insider*, March 24, 2021 https://www.businessinsider.com/what-is-telegram?r=US&IR=T

end-to-end encryption; a self-destruct timer that permanently deletes secret messages after a set period of time; groups, which are multi-person chats and can have up to 200,000 members; and, of particular relevance, channels, which are a tool for broadcasting messages to large audiences and can have an unlimited number of subscribers.[5] Channels can be public or private. Public channels have a username, so anyone can find them in Telegram's search function and join, whereas to join a private channel a user must be added by the owner or receive an invite link (known as a joinlink).[6]

6. When a new item of official IS propaganda is produced, it is posted in private Telegram channels.[7] It is then acquired by pro-IS users, following which the dissemination process "becomes rapidly decentralized".[8] These users store each piece of propaganda on multiple file-sharing sites, creating large banks of URLs by generating multiple URLs for each item on each site.[9] Often, these file-sharing sites are small or micro companies.

7. These banks of URLs are then made openly available on public Telegram channels.[10] From here, IS sympathisers can gather the URLs and post them on "beacon" platforms, such as Twitter.[11] These Twitter *ghazwah* (invasions) commonly rely on the use of throwaway accounts, created for the specific purpose of disseminating propaganda and in the expectation that they will be swiftly suspended.[12] The volume of URLs and speed with which they are disseminated are key, often achieved by the use of bots, along with other tactics such as hashtag hijacking and use of the @reply and @mention functions to try and maximise exposure.[13]

[5] 'Channels FAQ', https://telegram.org/faq_channels, accessed February 9, 2023.

[6] ibid.

[7] Asaad Almohammad and Charlie Winter, *From Battlefront to Cyberspace: Demystifying the Islamic State's Propaganda Machine*, (West Point, NY: Combating Terrorism Center, 2019), https://ctc.usma.edu/wp-content/uploads/2019/05/Battlefront-to-Cyberspace.pdf; Laurence Bindner and Raphael Gluck, "Assessing Europol's Operation Against ISIS' Propaganda: Approach and Impact," accessed February 9, 2023, https://icct.nl/publication/assessing-europols-operation-against-isis-propaganda-approach-and-impact/.

[8] Daniel Milton, *Pulling Back the Curtain: An Inside Look at the Islamic State's Media Organization*, (West Point, NY: Combating Terrorism Center, 2018), https://ctc.usma.edu/wp-content/uploads/2018/08/Pulling-Back-the-Curtain.pdf, 10.

[9] Ahmad Shehabat and Teodor Mitew, "Black-boxing the black flag: anonymous sharing platforms and ISIS content distribution tactics," *Perspectives on Terrorism* 12, no. 1 (2018): 81-99.

[10] Stuart Macdonald, Connor Rees and Joost S, *Remove, Impede, Disrupt, Redirect: Understanding & Combating Pro-Islamic State Use of File-Sharing Platforms*, (Washington DC: RESOLVE Network, 2022), https://doi.org/10.37805/ogrr2022.1.

[11] Ali Fisher, Nico Prucha, and Emily Winterbotham, *Mapping the Jihadist Information Ecosystem: Towards the Next Generation of Disruption Capability*, (London: Royal United Services Institute, 2019), https://static.rusi.org/20190716_grntt_paper_06.pdf.

[12] Daniel Grinnell *et al*., *Who disseminates* Rumiyah? *Examining the relative influence of sympathiser and non-sympathiser Twitter users*, https://www.europol.europa.eu/cms/sites/default/files/documents/dgrinnell_smacdonald_dmair_nlorenzodus_who_disseminates_rumiyah_0.pdf, accessed February 11, 2023.

[13] Mohammed Al Darwish, 'From Telegram to Twitter: The Lifecycle of Daesh Propaganda Material', VOX-Pol Blog, September 11, 2019, https://www.voxpol.eu/from-telegram-to-twitter-the-lifecycle-of-daesh-propaganda-material/, accessed February 11, 2023; Stuart Macdonald, Connor Rees and Joost S, *Remove, Impede, Disrupt, Redirect: Understanding & Combating Pro-Islamic State Use of File-Sharing Platforms*, (Washington DC: RESOLVE Network, 2022), https://doi.org/10.37805/ogrr2022.1.

8. In terms of content moderation, Telegram draws a sharp distinction between public and private channels. Its Terms of Service state that, by signing up to Telegram, users agree not to "Promote violence on *publicly* viewable Telegram channels, bots, etc".[14] Telegram has in the past taken part in Referral Action Days organised by Europol's EU Internet Referral Unit[15] and, in the first four months of 2022, it claimed to have removed 90,349 terrorist bots and channels.[16] While some have nonetheless doubted Telegram's commitment to moderating publicly available content,[17] its stated approach to public channels stands in marked contrast to its refusal to moderate the contents of private channels, undertaking to "ensure that no single government or block of like-minded countries can intrude on people's privacy and freedom of expression".[18] At the same time, Telegram recognises that some users may seek to exploit its public-private dichotomy, stating that "private channels with publicly available invite links will be treated in the same way as public channels, should it come to content disputes".[19]

9. The question that this raises – and which goes unanswered in Telegram's Terms of Service – is when will a joinlink be regarded as publicly available? As noted above, new pieces of official IS propaganda are released in private Telegram channels. While these channels are highly secretive, by scouring the contents of public IS channels it is possible for those with the necessary expertise and patience to locate openly available joinlinks to these private channels. Indeed, it is through this painstaking process that some researchers and investigators manage to gain access to these private channels to monitor the release of IS propaganda. So while, on the one hand, these private channels possess privacy-enabling technical features, on the other hand joinlinks are made openly available (albeit secretively, so that locating them is a laborious task) and, importantly, these channels are being used as part of the propaganda dissemination process, which is in essence a public-facing form of communication.

Clause 203

10. Consideration of how the existing clause 203 might be applied to the IS propaganda dissemination process yields important insights. We examine each of the three factors stipulated by the clause in turn, offering suggestions for how each might be usefully refined.

(a) The number of individuals in the United Kingdom who are able access the content by means of the service.

[14] 'Terms of Service', https://telegram.org/tos, accessed February 9, 2023 (emphasis added).

[15] "Europol and Telegram take on terrorist propaganda online," Europol, https://www.europol.europa.eu/media-press/newsroom/news/europol-and-telegram-take-terrorist-propaganda-online, accessed February 9, 2023.

[16] 'ISIS Watch', https://t.me/s/ISISwatch, accessed February 9, 2023.

[17] Hannah Gais and Megan Squire, 'How an Encrypted Messaging Platform is Changing Extremist Movements', Southern Poverty Law Center, February 16, 2021, https://www.splcenter.org/news/2021/02/16/how-encrypted-messaging-platform-changing-extremist-movements, accessed February 9, 2023.

[18] 'Telegram FAQ', https://telegram.org/faq, accessed February 9, 2023.

[19] 'Channels FAQ', https://telegram.org/faq_channels, accessed February 9, 2023.

11. By referring to the potential rather than an intended or actual number of receivers of the content, the implication appears to be that completely open communications on a platform such as Twitter are bound to be considered public. Indeed, content that is available *to the public generally* could be said to meet the very definition of content communicated publicly.

12. Less obvious is content posted to a service which has a limited capacity to accommodate concurrent users, and which is removed after a period: any member of the public could in principle encounter the content but owing to the technical design or deficiencies of the service only a limited number of individuals will in practice be able to access it. Also, less obvious is content posted to a public Telegram channel with an obscure name, or a private Telegram channel whose joinlinks are made publicly available. In these cases, any determined user may access the content, but the number of actual views may be minimal.

13. The difficulty arises from the qualifying words "by means of *the* service" (emphasis added). The words suggest that the numerical focus is the service's own user-base rather than on the number of individuals who might eventually access the same content on the internet generally, even though the service plays a vital role in its wider dissemination.

14. This could arise where content designed for wide consumption, such as IS propaganda, is placed on one service as part of a wider dissemination strategy using multiple services. The strategy might use different services (Telegram, Instagram, JustPaste.it) simultaneously or deploy one service (for example Telegram) to advertise the presence of material in the hope that it will be picked up and taken viral using other services. The number of individuals on one service ought not to be the primary factor in such circumstances.

15. The qualifying words "by means of the service" are no doubt included to limit the responsibility of services. However, shorn of the wider dissemination context, the harm risked and the need for greater OFCOM intervention may be overlooked.

16. The nature of the online environment and of the content are relevant. In a trusted environment, a person might share personal information in the expectation that others will not violate that trust and share the information with others. A small number of users may indicate that the environment is one in which this relationship of trust exists. But this is not necessarily the case. In private IS Telegram channels, usernames are anonymised and identities are unknown. This is how investigators and researchers manage to obtain access.

17. Moreover, those posting the propaganda in these channels do so in the expectation that the materials will be disseminated more widely, and the content is expressly designed for wide dissemination. As Maura Conway has pointed out, terrorist propaganda that is

designed to be shared widely appears, by its very nature, to be incontestably public.[20] The presence of public terrorism content ought to be relevant to OFCOM's consideration of whether content on that service is being communicated publicly or privately.

18. It follows that criterion (a) should be amended to make it clear that accessing the content "by means of the service" includes simultaneous or later access on other services, having regard to the type of content and the relationship between the users of the service in question.

(b) Any restrictions on who may access the content by means of the service (for example, a requirement for approval or permission from a user, or the provider, of the service).

19. Restrictions on access are a factor, but not determinative, and taken together with the previous factor squarely raise the case of terrorist content on a private Telegram channel with hundreds or perhaps thousands of members. Under the Bill, restrictions on access do not include a requirement to log in or register, to make a payment or take out a subscription, or to access the service using a particular technology (such as the TOR browser) or device, so long as generally available.[21] But all other restrictions on access are in scope including permission from a human administrator who could be a member of a proscribed terrorist organisation and conditional entry based on acceptable answers to bot-administered questionnaires ("Do you hate Jews?") so long as they amount to more than simple registration. The implication of this factor is that restrictions on access are a feature tending towards private communication even though the numbers of those accessing the content may be significant.

20. It may be useful to distinguish between restrictions in terms of (a) features of the platform and (b) in practice. At the platform level, Telegram private channels are designed to restrict users to those approved by the channel administrator. But in practice, making joinlinks openly available (albeit difficult to locate) undercuts the raison d'etre of the privacy-enabling feature. Clause 203 could provide that OFCOM must have regard to the practical operation of the restriction on access, as well as the nature of the restriction itself.

21. The purpose of the restriction is also fundamental. In most instances, restrictions on access will be designed to limit the availability of the content (e.g., members of a research institute sharing materials for a security-sensitive research project). But there are some circumstances in which the purpose of restrictions on access will be to facilitate increased dissemination of materials. A clear example would be restrictions designed to prevent law enforcement from accessing the content and taking steps to have it removed. The restrictions on access to private IS Telegram channels are designed to safeguard the initial stages of the propaganda dissemination process, in order to enable wider subsequent circulation of the materials. Here the restriction serves the purpose of wider dissemination and making it more and not less publicly available.

[20] Conway, M., 'Online Extremism and Terrorism Research Ethics: Researcher Safety, Informed Consent, and the Need for Tailored Guidelines', *Terrorism and Political Violence*, Vol.33, No.2 (2021): 367-380.
[21] Clause 203(3).

22. At the very least, therefore, criterion (b) should only apply to restrictions that are designed to limit the number of users who may ultimately access the content (whether on that or another service).

(c) The ease with which the content may be forwarded to or shared with users of the service other than those who originally encounter it.

23. This factor addresses the channel's affordances (specifically, whether users can promote or highlight terrorist content to other users within a channel) rather than the circumstances of the initial communication and recognises that the numerical reach of content does not depend on the intention of the source or nature of the message but upon the reaction of recipients. Whether the message is itself public or private at the point of communication may therefore be subordinated to capacity, of which the original sender may be ignorant, of the channel to allow internal amplification within the service. To this extent it appears to expand the ordinary meaning of publicly communicated. Consistency with the first factor suggests that some assessment should be made of the *number* of other users who may subsequently encounter it.

24. Once again, the qualifying words "with users of *the* service" (emphasis added) reduce the likelihood of capturing wider dissemination strategies. The IS propaganda dissemination process illustrates the importance of an "ecosystems" approach. Different platforms are used for different purposes, in a combined and interdependent way. The difficulty with this requirement is that it focuses on the specific service on which the content was originally shared.

25. To capture the cross-platform dimension of contemporary propaganda dissemination, this factor should read "… with users of the same or another service …"

Proposed amendments

26. To address the issues identified above, we propose the following four amendments to subsection (2) of clause 203:

> *(2) The factors are—*
>
> *(a) the **type of content and the** number of individuals in the United Kingdom who are **or are intended to be** able to access the content by means of the **same or another** service;*
> *(b) any restrictions **that are designed to limit** who may access the content by means of the **same or another** service (for example, a requirement for approval or permission from a user, or the provider, of the service);*
> ***(c) the practical operation of any such restrictions;***
> *(d) the ease with which the content may be forwarded to or shared with users of the **same or another** service other than those who originally encounter it.*

27. First, the words "the service" have been amended to "the same or another service" throughout, to reflect the cross-platform nature of propaganda dissemination.

28. Second, the words “type of content” and “are intended to be” have been added to paragraph (a), to encompass instances where content that is designed for wide dissemination is made available at the initial stages of a chain dissemination process.

29. Third, the words “that are designed to limit” have been added to paragraph (b). This is to make it clear that paragraph (b) does not encompass restrictions on access that are designed to safeguard the initial stages of a chain dissemination process in order to enable wider subsequent circulation.

30. Fourth, a new paragraph (c) has been added, which invites consideration of how any restrictions on access to content operate in practice. This complements paragraph (b), which focuses on the nature of the restrictions themselves.

1 March 2023