New EU Proposal on the Prevention of Terrorist Content Online

An Important Mutation of the E-Commerce Intermediaries' Regime

Joan Barata I October 2018

The Center for
Internet and Society

NEW EU PROPOSAL ON THE PREVENTION OF TERRORIST CONTENT ONLINE: AN IMPORTANT MUTATION OF THE E-COMMERCE INTERMEDIARIES' REGIME

Joan Barata
Intermediary Liability Fellow
CIS Stanford Law

I. Introduction

In the course of the last years, the European Union (EU) institutions, and the Commission (EC) in particular, have shown a growing concern regarding the use of online intermediary platforms for the dissemination of illegal content, particularly content of terrorist nature. Despite lack of complete certainty and differences between member States about what terrorist content the law prohibits, or even can prohibit consistent with fundamental rights to free expression, the truth is that there is a broad consensus among the national authorities that legislative and regulatory measures should be enacted both at the European and national levels in order to guarantee the swift and almost automatic detection and removal of content related to the commission of acts of terrorism.

The political positions and non-binding documents produced so far have progressively incorporated the notion of "responsibility" for intermediaries, although this could not necessarily be equated to a straightforward intention to impose conventional legal liability obligations on such actors. In particular, the initiatives undertaken so far by the EU institutions basically aimed at promoting platforms' voluntary cooperation with public authorities to detect and remove online illegal content (including terrorist content). Such initiatives include the Code of Conduct on Countering Illegal Hate Speech Online[1], the Recommendation on measures to effectively tackle illegal speech online[2], the Guidelines on Freedom of Expression Online and Offline[3], and the EU Internet Forum[4].

Above all these recommendations, agreements and soft-law standards, the general legally applicable regime has remained so far intact since its approval in 2000: Directive 2000/31/EC, known as the e-commerce Directive, establishes liability exemptions for intermediaries under certain conditions of lack of knowledge of illegal activity or information and expeditious removal and disabling upon knowledge (article

[1] http://ec.europa.eu/newsroom/just/item-detail.cfm?item_id=54300
[2] https://ec.europa.eu/digital-single-market/en/news/commission-recommendation-measures-effectively-tackle-illegal-content-online
[3] https://ec.europa.eu/europeaid/guidelines-freedom-expression-online-and-offline_en
[4] http://europa.eu/rapid/press-release_IP-17-5105_en.htm

14). The Directive also includes an important provision regarding the absence of any legal obligation for providers to monitor content (article 15).

The new proposed Regulation on preventing the dissemination of terrorist content online, which was the object of a first discussion on September 19-20 during the meeting of EU leaders in Salzburg under the Austrian Presidency, may represent a change in the abovementioned approach.

1. Interactions between the proposed Regulation and the e-commerce Directive

For the first time, a true legal instrument (directly binding for member States) will establish obligations that go beyond the promotion of voluntary conduct and cooperation. In fact, the explanatory memorandum of the proposal states that one of the main objectives is “to provide clarity as to the responsibility of hosting service providers in taking all appropriate, reasonable and proportionate actions necessary to ensure the safety of their services and to swiftly and effectively detect and remove terrorist content online”.

Regarding the interaction and consistency between the provisions included in the proposal and the liability regime derived from the e-commerce Directive, the memorandum clarifies that liability exemptions provided in article 14 remain unaltered. Moreover, Recital 5 of the proposal states that “any measures taken by the hosting service provider in compliance with this Regulation, including any proactive measures, should not in themselves lead to that service provider losing the benefit of the liability exemption” provided for in article 14.

However, regarding the provisions contained in article 15, the text takes a far more ambiguous approach by expressing that some of the obligations established in the text vis-à-vis intermediary platforms “should not, in principle, lead to the imposition of a general obligation to monitor”. However, the memorandum also states that “given the particularly grave risks associated with the dissemination of terrorist content, the decisions under this Regulation may exceptionally derogate from this principle under an EU framework”. Such derogation would require, in each case, that the competent

authority strike a fair balance between public security needs and the rest of fundamental rights and interests at stake, particularly the right to freedom of expression.

Therefore, the new proposed Regulation declares that under its provisions State and Union bodies may be able, in certain cases, to provide for specific derogations of the regime established in article 15 of the e-commerce Directive. This is a major step with regards to the alteration of a legal principle which had remained accepted and untouched for almost two decades, and which was part of a series of rules that were publicly presented and considered as the frontispiece of EU regulation on online platforms. Considering the implications (beyond the scope of the proposal) of this derogation, it is a bit surprising to find it acknowledged and declared in such a nuanced manner and by using the *back door* of the explanatory memorandum.

In order to properly discern the extent of these derogations and the way the new responsibilities of intermediary platforms are defined by the proposal, it is important to take a look at some relevant provisions of the text.

2. Definitions

Article 2 includes a series of definitions for the purposes of the Regulation. An important definition to look at is the one on “terrorist content” (a concept that will be used in very important articles of the proposal, as will be noted). Among other things, it covers “inciting or advocating, including by glorifying, the commission of terrorist offences, thereby causing a danger that such acts be committed”.

This definition needs a close critical analysis, as it amends, with regards to online content, the definition included in the current and recent Directive of 15 March 2017 on combatting terrorism[5]. According to this Directive, the “public provocation to commit a terrorist offence” is understood as “the distribution, or otherwise making available by any means, whether online or offline, of a message to the public, with the intent to incite the commission of (terrorist offences) where such conduct, directly or indirectly,

[5] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32017L0541

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," and such public provocation "is punishable as a criminal offence when committed intentionally".

Both definitions look quite similar. However, the Directive requires member States to criminalize the distribution of messages that cause a danger that a terrorist act may be committed, and which advocate for such actions, whereas the proposed regulation additionally refers to incitement. The incitement of terrorist acts is a much broader, vaguer and more general legal notion than the advocacy of their commission. The Regulation thus appears to give a wider and more discretionary power to State authorities, and creates the possible risk of limiting the expression of certain extreme, but fully legal, ideas or the publication of journalistic work related to terrorism.

Another relevant problem is the fact that the Directive also requires that the acts of provocation that it identifies must be punished by member States as criminal offences. Therefore, an important question here, which raises a relevant level of legal uncertainty, is whether the Regulation needs to be interpreted as an amendment to the Directive (and therefore as the imposition of an obligation for member States to criminalize any forms of incitement to terrorism, beyond advocacy), or instead introduces a form of "terrorist content" that can be removed but is not necessarily punishable according to national law. It is important to note, in this sense, that the other two items included by the Regulation in the definition of terrorist content, that is "encouraging the contribution to terrorist offences" and "promoting the activities of a terrorist group, in particular by encouraging the participation in or support to a terrorist group" can be linked to punishable terrorist offences within the provisions of the Directive (particularly according to article 6, 8, 10, and 11).

3. Measures to prevent the dissemination of terrorist content online

Article 3 incorporates a general obligation for hosting services to "take appropriate, reasonable and proportionate actions (…), against the dissemination of terrorist content and to protect users from terrorist content". When doing so, these actors shall act in a "diligent, proportionate and non-discriminatory manner, and with due regard

to the fundamental rights of the users and take into account the fundamental importance of the freedom of expression and information in an open and democratic society". Hosting services must also include in their terms and conditions the rules necessary to prevent the dissemination of terrorist content. This first group of legal duties is specifically labeled by the Regulation as "duties of care". According to article 18 of the proposal, States shall establish penalties for hosts that fail to incorporate such rules into the terms and conditions.

Despite the fact that basically all hosting services have already incorporated into their terms and conditions specific provisions aimed at tackling online terrorist content, the establishment of such legal obligation puts on the table important issues. First, given the abovementioned issue with the definition of "terrorist content," Article 3 appears to require hosts to contractually prohibit (and protect users from) expression that is currently legal – but that will subsequently be banned by specific national criminal legislation. Second, the draft does not provide clear criteria that would give hosting services the reassurance that they are properly meeting the "duties of care" the legal text imposes under the threat of penalties. The nature of the penalties is to be determined by member States, but considering the thematic area of the Regulation they will probably be severe. The general principles mentioned in the first paragraph of article 3 do not seem to be precise enough in this sense, therefore creating a very clear pressure on the side of private companies, even as recitals 12 and 16 nominally exclude the existence of a general monitoring obligation on service providers. In other words, hosting services are obliged to adopt an extensive set of rules in a complex legal matter as well as to directly enforce them and limit the exercise of the right to freedom of expression by their users. Moreover, these actions will thus not be the mere result of the establishment and enforcement of internal conduct rules, but the exercise of a responsibility enshrined in a legal instrument and supervised by State authorities. As service providers' obligations are unclear, they will naturally be incentivized to err on the side of over-enforcement, which will create obvious detrimental effects for freedom of expression.

Article 4 contains a series of obligations that will also count on the penalty powers of public authorities. According to this provision "competent authorities" will have the power to order a hosting service provider to remove "terrorist content" or disable

access to it within one hour from the receipt of the order. Hosts must also report their compliance to the relevant authorities without undue delay.

This appears to be the most delicate part of the Regulation, as it enables a so far undefined set of State bodies to take decisions on content hosted by intermediaries, and gives those same State bodies the power to remove or disable access to it in one hour using a private proxy. A platform that is notified that a particular piece of content violates law in one State must then choose between taking the content down EU-wide or even globally on the one hand, and carrying out State-by State legal analysis on the other. Any other response would expose the platform to risk of being charged with knowledge, and thus liability under Article 14, for hosting illegal content in other States. The clear incentive in this case is to accept the most conservative interpretation of the law as a global standard, in contravention of the EU's reservation of power to Member States to define and enforce their own free expression protections.

An element that will raise controversy is the very short period of one hour given to service providers to take down content. This is obviously a very tight timeframe that would require investing resources in order to guarantee full compliance at all times and in all circumstances. This short timeframe is also delicate from the point of view of possible appeal or redress mechanisms in the hands of hosting services. The Regulation obliges the competent authority to provide "a statement of reasons" that would justify the removal order, although a "detailed statement of reasons" would only be provided if requested by the service provider and without prejudice to its duty to execute the order within the legal timespan. This means that intermediaries do not only need to act in an extremely rapid manner, but also that both they and the users whose expression and information rights are affected are practically deprived of any right or chance to question the decision. The Regulation only mentions a few and very limited exceptions, including force majeure, de facto impossibility, manifest errors or lack of sufficient information. In other cases, as has just been said, hosting providers will have to act expeditiously unless they want to assume a very high risk of being seriously punished under article 18. In this sense, paragraph 2 of article 18 establishes that a systematic failure to comply with obligations pursuant to Article 4(2) (removal orders) "is subject to financial penalties of up to 4% of the hosting service provider's global turnover of the last business year". For these reasons, it is obvious that any

doubt or concern from the service provider's side regarding the legality of the measure, and/or the competence or jurisdiction of the authority in question will be resolved by executing the order, with only the remedy of a subsequent appeal according to the correspondent national legislation.

Article 5 has similarities to article 4 but here the legal approach is slightly different. Beyond the power to send the removal or access disabling orders, competent authorities "or the relevant Union body" (which basically refers to Europol) are also given the competence to send a referral to hosting service providers for their expeditious "voluntary consideration". Following to this referral, intermediaries "shall, as a matter of priority, assess the content identified in the referral against its own terms and conditions and decide whether to remove that content or to disable access to it" and "expeditiously inform the competent authority or relevant Union body of the outcome of the assessment and the timing of any action taken". This "invitation" to consider taking down content or information on grounds of terrorism basically implies that hosting service providers will need to take a decision upon request from authorities, based on their own terms of use. This basically enables public authorities to tackle a legal issue via the enforcement of hosting service providers' terms and conditions by such companies themselves. It goes without saying that in case hosting service providers understand that there are no grounds for such action, competent authorities would be able to activate the mechanism established in article 4, or even examine whether the duties of care established in article 3 are properly fulfilled, in the sense that service providers must enforce an appropriate set of terms and conditions that enables them to properly tackle terrorist content. In such scenario, State bodies could also assess whether the platform faces liability under Art 14 of the e-commerce Directive because it could be argued that already knows about the illegal content but did not take it down.

For this reason, despite the appearance of a "voluntary" measure (or an induced voluntary one), these referral powers may become a mechanism to delegate to private entities the responsibility to decide and enforce measures that otherwise would need to be adopted by public bodies with proper opportunity for judicial review. This may prove a particularly useful option when those bodies have doubts about the legality of

the qualification of certain content or information as terrorist and therefore prefer not to make use of the powers granted by article 4.

Article 6 contains a group of additional “proactive measures” that definitely change the role of intermediaries in this area, particularly regarding their content monitoring responsibilities. According to the first paragraph, service providers “shall, where appropriate, take proactive measures to protect their services against the dissemination of terrorist content. The measures shall be effective and proportionate, taking into account the risk and level of exposure to terrorist content, the fundamental rights of the users, and the fundamental importance of the freedom of expression and information in an open and democratic society”. The fulfillment of this principle is not subject to clear control or penalties according to the proposal, and may look closer to a *desideratum* than an actual legal obligation. It also needs to be noted that it contains an indication or directive vis-à-vis private service providers by requiring the respect for fundamental rights, particularly the right to freedom of expression. Paragraph 2 refers to cases of prior application of the mechanisms set in article 4 (in particular, when a hosting service provider has already been obliged to take removal or disabling measures with an order that became final). Competent authorities will have the power to “request the hosting service provider to submit a report, within three months after receipt of the request and thereafter at least on an annual basis, on the specific proactive measures it has taken” with a view to “preventing the re-upload of content” and “detecting, identifying and expeditiously removing or disabling access to terrorist content”. Such reporting obligations may be particularly burdensome for platforms, not only because of the resources that may be required, but also because they may become a *de facto* or indirect supervisory mechanism for authorities and thus an additional pressure for service providers to take down doubtful content.

State bodies will also have the power to impose “specific additional necessary and proportionate proactive measures”, which can be subject to review at the request of the service provider. It also needs to be particularly noted that failure to adopt and report on the implementation of imposed measured is subject to penalties as well.

The establishment of the need to adopt what are generally described as proactive measures puts intermediary platforms in a new status where (even if not called by this

name) content monitoring becomes also part of their role. Moreover, such role goes beyond a merely supervisory one as they also will need to promptly and properly adopt the necessary measures in cases when terrorist content is detected as the result of this proactive behavior. In addition to this, according to article 13.4 if "hosting service providers become aware of any evidence of terrorist offences, they shall promptly inform authorities competent for the investigation and prosecution". Although this cannot be considered as a monitoring obligation *per se*, the combination of the ensemble of responsibilities enshrined in the proposal creates an overarching new regime of private content regulation vis-à-vis terrorist content.

As was outlined earlier, this expansion of platforms' monitoring activities is a major change affecting the way intermediary liability and responsibilities will be understood within the European Union from now on. Even if the proposed Regulation refers only to the area of combatting the dissemination of terrorist content (however this is specifically defined), this legal reconsideration can also be detected in the new proposals regarding copyright or the regulation of audiovisual media services. Therefore, the European legislators seem to move towards a progressive delegation on private companies of true law enforcement powers, depriving Internet users (and hosting service providers themselves) of the legal and procedural safeguards applicable to this kind of decision until now. Moreover, intermediary platforms may be progressively put in a position where cautiously overbroad decisions will be taken, as this will be the only way to avoid the high and somewhat vaguely defined penalties that may be imposed on them. Obviously, this is not good news for freedom of expression and due process in Europe.

About the Author

Joan Barata is an international expert in freedom of expression, freedom of information and media regulation. As a scholar, he has spoken and done extensive research in these areas, working and collaborating with various universities and academic centers, from Asia to Africa and America, authoring papers, articles and books, and addressing specialized Parliament committees. He was Principal Adviser to the Representative on Freedom of the Media at the Organization for Security and Cooperation in Europe (OSCE), as well as Secretary General of the Catalan Audio-Visual Council in Spain, while also being a member of the Secretariat of the Mediterranean Network of Regulatory Authorities. As an international expert, Joan has provided advice to international organizations, NGOs, Governments, legislators and regulators in most regions of the world, including Africa, Asia and Latin America.

About the Center for Internet and Society

The Center for Internet and Society (CIS) is a public interest technology law and policy program at Stanford Law School and a part of Law, Science and Technology Program at Stanford Law School. CIS brings together scholars, academics, legislators, students, programmers, security researchers, and scientists to study the interaction of new technologies and the law and to examine how the synergy between the two can either promote or harm public goods like free speech, innovation, privacy, public commons, diversity, and scientific inquiry. CIS strives to improve both technology and law, encouraging decision makers to design both as a means to further democratic values. CIS provides law students and the general public with educational resources and analyses of policy issues arising at the intersection of law, technology and the public interest. CIS also sponsors a range of public events including a speakers series, conferences and workshops. CIS was founded by Lawrence Lessig in 2000.