Terrorist propaganda in cyberspace: emerging trends and innovation in policy-making

“National law has no place in cyberlaw. Where is cyberspace? If you don’t like banking laws in the United States, set up your machine on the Grand Cayman Islands. Don’t like the copyright laws in the United States? Set up your machine in China. Cyberlaw is global law, which is not going to be easy to handle, since we seemingly cannot even agree on world trade of automobile parts” - Nicholas Negroponte.

Introduction: a curse and blessing

Some 20 years ago, the Bush Administration launched the global ‘War on Terror’. Since then, America and her allies’ military might has achieved the dispersion of al-Qaeda, the reconquest of the territories captured by ISIS, and a tentative peace, however fragile, in Afghanistan. Has the West, then, won the ‘War on Terror’? For most terrorism scholars and security experts, the answer would be to the negative: The military onslaught against terror groups may have done a great deal in shattering the physical and traditional organizational infrastructure of terror groups, but popular sympathy for radical Islamism and support for the jihadi cause is on the rise not only in those areas which have been plunged into chaos in the wake of US invasions, but also in ever-more corners of the world. The pain in the aftermath of 9/11 and the calls for revenge may have blinded policy-makers to the reality of terrorism: terror is merely a weapon in a larger struggle – it is a struggle over the hearts and minds of people. And in this ‘War of Ideas’, the West has been less than victorious.

Much of the jihadi groups’ success in keeping up the ‘battle of ideas’ is owed to the evolution and savvy instrumentalization of Information and Communication Technologies (ICT). On one hand, the internet has sometimes been compared to the ‘agora of Athens’ and the end of the silence of the voiceless majority: a borderless public place where people can discuss ideas and receive information in real-time. Yet the same features which make ICT a tool of popular empowerment - the ease by which information can be disseminated and accessed, its ability to evade central control and to connect people from all over the globe - can render it a menace to the very fabric of democratic societies when destructive ideas gain traction online. Where geographical distance may have prevented the consolidation of fringe movements before the age of communication technology, ICT today provides conspiracy theorists, hate groups and terrorists with the means to connect vast audiences ready to listen and to believe, luring them with narratives about ‘real’ truths, virtuous struggles and enemies worth of destruction. Q-Anon, for example, a conspiracy group wielding wildly implausible theories involving Satan-worshipping pedophiles, sex-trafficking and Donald Trump in the role of the savior, largely owes its proliferation across the globe to social media (Mezzofiore et al, 2020).
Terror groups similarly have found ICT to be a highly effective instrument through which to wage ideational warfare. The Islamic State (IS) in particular has received much attention for the creation of a sophisticated propaganda machine. It has been able to create its own ‘brand’ through which it communicates its message of injustice and the need for violent jihad to address vast audiences across the globe with the aim of attracting supporters, radicalizing and intimidating (Azani & Liv, 2018). Propagandistic messages may take a variety of forms, for example that of video games, magazines and regularly published news items documenting the groups’ successes, instruction documents for building bombs and executing attacks and videos in which jihadis feature as heroes and martyrs. They are published by the IS’ own media department and shared and translated across internet platforms, whereby the Web2.0 revolution has provided terrorist groups an extraordinary opportunity to increase their visibility (Binder & Gluck, 2018). Although the exact role of the internet in radicalization is still a topic of discussion among researchers, there is little doubt that the visibility of terror propaganda in virtual spaces poses a formidable threat to democratic societies. Phenomena which have made their inroads into contemporary terrorism literature – for example those of ‘homegrown terrorists’ and ‘foreign fighters’ – are at least facilitated, if not at all made possible, by ICT establishing the communicative link between recruiters and recruited. Indeed, Clingendael Institute in The Netherlands predicts that the use of technology by terrorists will present one of the most concerning challenges 2020-2025 (Sweijs & Pronk, 2020).

Ridding cyberspace of terror propaganda thus became a pressing issue within a new policy paradigm focused on fighting ideas rather than physical targets. Yet, despite the strategic need to counteract the proliferation of terrorist content in cyberspace by means of a globally coordinated response, regulatory approaches thereto suffer from a lack of an overarching strategy leveraging common understandings (Housen-Couriel, 2019). This condition is the result, among other things, of the need to address two relatively underdeveloped and highly politicized policy themes – internet governance and (counter)terrorism – through an international system founded on the concepts of State sovereignty and territorial integrity, which is only slowly adapting to global security challenges.

**About this paper**

Some scholars have argued that there is no such thing as ‘cyberlaw’. In the words of US Judge Frank Easterbrook, cyberlaw is like the ‘law of the horse’: students would do better to study laws applicable to incidents with horses – tort law, property law, commercial law – than the ‘law of the horse’ (Easterbrook, 1996). In the same sense, ‘cyberlaw’ embraces a number of disciplines ranging from privacy law to criminal law, commercial law or obscenity law, depending on the phenomenon studied – which, when applied to the digital context, differ from real-life application in the manner they are enforced rather than in substance. In the same way, the present paper is predominantly concerned with the enforcement of legal provisions applicable to terrorist content in cyberspace. While we will touch upon the substantive tools available to policy-makers in different jurisdictions to counter terrorist
propaganda both online or offline, we mostly deal with such approaches as having evolved out of the need to effectively enforce rules in digital environments.

It shall be noted that the subject on which we will focus here is only one of a number of terrorist’s uses of cyberspace. ‘Terrorist propaganda’ – or terrorist contents, which for present purposes may refer to the same – can here be understood as the propagation of a particular extremist worldview that brings individuals to consider and justify violence (European Commission, n.d.). Other than that, terrorists use the internet to individually communicate among each other in order to plan attacks or coordinate logistics, for financial transactions, to gather intelligence or to provide instructions and train recruits, to name but a few examples. Responses to each of these activities entail distinct legal means, institutional stakeholders and sets of problems (internal communications, for example, are usually conducted through encrypted technologies meaning that privacy constitutes a main concern with regard to investigation and prosecution; fund-raising requires the collaboration of financial institutions; and so on). Still, the fault lines discussed throughout this paper broadly apply to any subject lying at the intersection between terrorism and internet governance, and the conclusions drawn toward the end of this paper will allow for some generalization beyond the subject of ‘terrorist propaganda’.

Section I will provide the reader with a theoretical basis to make sense of the legal implications following from the distinct characteristics of digital vis-à-vis physical space. Subsequently, we will trace the emerging policies and practices aiming to address these problems: Section II deals with criminal justice approaches – and more specifically the strengthening of inter- and transnational cooperation on crimes involving cyberspace - and section III is concerned with preventive approaches aiming to reduce the visibility of terrorist contents in cyberspace through privatization of law enforcement and the use of technological means. Lastly, the discussion will synthesize findings and make sense of the implications of the evolving relationship between terrorism and cyber-technologies for the global legal and political system.

I: Internet governance: theoretical considerations

In the early years of ICT development, the internet was thought of as an ungovernable space: its decentralized architecture supposedly thwarted any type of control. This stance is reflected in the famous Declaration of the Independence of Cyberspace of the 1990s by American cyber-libertarian political activist John Barlow, who stated that the internet be ‘inherently extra-national, inherently anti-sovereign and your [States’] sovereignty cannot apply to us’ (Barlow, 1996).

But Barlow was mistaken. As the global network pervaded evermore areas of everyday life and its relevance (as well as potential threats) to security, economic, social and political interests came to be fully understood, the need for some type of regulation increasingly became acute. Harvard’s Jack Goldsmith, in his seminal article titled ‘Against Cyberanarchy’, challenged the libertarians’ claims of the ‘ungovernable internet’, arguing that the regulation of cyberspace be not so different from other transnational transactions, and that traditional
legal tools and technology be well capable of resolving multijurisdictional regulatory problems implicated by cyberspace (Goldsmith, 1998). Goldsmith, along with other digital realists, essentially argues that ‘real space’ laws should (and could) be extended to cyberspace. Indeed, the stance that online laws should be assimilated to offline laws is today a widely shared one (Kaesling, 2018).

Still, reality is not so simple as the digital realist view might suggest; that the internet should be governed does not yet tell us how to govern it. Cyberspace has characteristics distinct from those of physical space, and those characteristics pose a direct challenge to traditional regulatory theories and policy-making practices (Hofman et al, 2017).

For Lessig (1998), one of the most prominent early academics studying cyberlaw, the main difference between the regulation of cyberspace and that of real space was that of anonymity. Lessing argued that the same rules which govern real-world behavior – law (regulating via sanctions), social norms (regulating via understandings and expectations), the market (regulating via price) and architecture (regulating via physical constraints such as walls, doors, stairs and so on) – also govern cyberspace, at least in theory: legal regulation applies in the case of, for example, copyright law or obscenity law; social norms regulate through punishment by the community; the market regulates by pricing access to sites; and architecture regulates through code, i.e. the software and hardware that constitutes cyberspace as it is – a set of rules and protocols, implemented or codified, in the software of cyberspace itself. The architecture of cyberspace can be regulated by, for example, requiring users to enter a password or identifying oneself, or by providing only certain language options, hence setting criteria for who may access a domain and who may not. Yet, the fact that cyberspace makes it easy for people to hide their identity means that laws, norms and the market cannot effectively constrain people: “The default in cyberspace is anonymity. And because it is so easy to hide one’s identity, it is practically impossible for the laws, and norms, to apply in cyberspace” (Lessig, 1998). However – so the argument goes – source codes are written by a handful of private actors, who can make decisions about the design of cyberspace and enforce rules through technological means, for example filters or self-identification requirements.

Again, for others, the prime challenge of internet governance has to do with its trans-jurisdictional character. States do have formal control over information networks and data circulating across and within their borders (which can be implemented by technological means such as content filtering and website blocking) but spill-overs of the effects of foreign cyber-activities and data flows are inevitable unless the legislator is willing to accept severe restrictions in access to information for her citizens. The advent of social media has raised further questions: if someone publishes a controversial comment on Facebook or Twitter, who is responsible for determining whether it is legal or illegal or in line with the law of the respective country? How can people in Chile or Bangladesh enforce their rights vis-à-vis a company that operates under the laws of a foreign country but makes the content available globally? (Jaume-Palasi & Spielkamp, 2016).
A solution to multijurisdictional issues could come in the form of extensive streamlining of national approaches as proposed by a variety of contemporary authors, who argue that international law may provide a medium to harmonize divergent legal trends into a unified theory that can be more effectively applied to create a ‘global internet’ (e.g. Perritt, 1998). This, for the moment, appears to fail: while some regions – primarily the EU – have managed some degree of harmonization in terms of the modalities by which the internet is regulated, such efforts have been by far less successful on a global scale.

**II: Intensification of international cooperation**

For most of history, conventional crimes have shared the attribute of locality – i.e. the criminal and the victim belong to the same geographical location (Chaturvedi et al, 2014). Globalization, and the rapid evolution of ICT in particular, have changed this. Computer networks and data by their very nature disregard physical boundaries; this feature has been of much help for terrorist groups increasingly operating across national borders: the originator of a terrorist message may be located in a different jurisdiction than the ISP on which the message is hosted, and the audience of the message may be dispersed across the globe. An effective strategy to bring those spreading terrorist propaganda to justice thus requires integrated responses. If a video calling for violent jihad is sent from US soil (where this may be perfectly legal) through a US-based HSP but is viewed by citizens in France (where such contents are criminalized), what can French authorities do to ensure that the video is taken down and its originator prosecuted? At an even more basic level, how would it at all be possible for French authorities to gather evidence on the identity and location of the originator of such video without violating the US’ territorial integrity?

A prerequisite for international cooperation is, firstly, some degree of consensus on the legal definition and judicial interpretation of ‘terrorist propaganda’ and secondly, the availability of mechanisms facilitating mutual legal assistance, transfer of criminal proceedings and convicted persons and exchange of data and information (United Nations Office for Drugs and Crime, 2012). However, as the paragraphs below will show, although globalization has acted as a driver of the development of transnational criminal law, official procedures for transnational cooperation are in many ways still unsuitable for the prosecution of terrorist crimes facilitated through the internet; only a diminishingly small number of transnationally operating cyber-criminals are prosecuted successfully, a condition that has been termed the ‘global enforcement gap’ (Peters & Jordan, 2020).

- **Streamlining ‘terrorist propaganda’**

One of the most formidable obstacles to the operationalization of a global counter-terrorism strategy generally is the lack of agreed definition of what terrorism is, and how to respond to it. Without lingering too much on the failure to formulate a common definition of the term ‘terrorism’ *per se* – volumes have been published on this issue - let us take a look at how the failure to set common standards play out with regard to terrorist propaganda.
Most governments around the world recognize the harm of terrorist propaganda and have criminalized what they perceive as harmful terrorist speech. However, what are the criteria to delimit terrorist speech from types of speech protected under the freedom of expression – after all, a fundamental right? Major differences in this regard are the result of vastly diverging conceptions about the scope of the right to free expression. Most jurisdictions acknowledge that there exist circumstances where certain types of expressions may cause harm that justify restrictions on free speech for the sake of values viewed as deserving of stronger protection; in other words, free speech is not an absolute right. Depending on the value, hierarchies underpinning different legal systems, however, there are major discrepancies as to which values are seen as trumping free speech. The scope of curtailment hence varies widely across legislations and ranges from a virtual inviolability of free speech to speech criminalization along mere moral/religious justifications.

A common standard which has evolved over the past 20 years is the framing of terrorist propaganda in terms of incitement. Incitement offences - referring to direct appeals to the public whereby an individual invites or urges an audience to commit certain criminal acts (Timmermann, 2006) - have been present in many criminal codes long before the 9/11 attacks, for example in the form of incitement to hatred, violence or genocide. In 2005 the UN, through Resolution 1624 called upon its member States to “adopt such measures as may be necessary and appropriate and in accordance with their obligations under international law to prohibit by law the incitement to commit a terrorist act or acts” (United Nations Security Council, 2005). Some 135 States have since added the offence of ‘incitement to terrorism’ to their criminal codes, but the vague wording and lack of guidance provided by the Resolution resulted in highly diverging judicial interpretations of the offence by national courts (Counter-Terrorism Committee Executive Directorate, 2016). For example, while most States agree that ‘incitement to terrorism’ is an inchoate offence – i.e. an offence which prioritizes the intent of the offender rather than the consequences of the act – a prosecutor in the US must be able to prove that an utterance poses a ‘clear and present danger’ in order to qualify as incitement (Brandenburg v. Ohio, 1969) - a test which, to date, has never been met with regard to terrorist speech (Leibowitz, 2017).

A geographically less wide-reaching but more concise approach towards streamlining ‘incitement to terrorism’ has been offered by the Council of Europe by means of its Convention on the Prevention of Terrorism from 2005. Article 5 thereof defines “public provocation to commit a terrorist offence” as the “distribution, or otherwise making available to the public, with the intent to incite the commission of a terrorist offence, where such conduct, whether or not directly advocating terrorist offences, causes a danger that one or more such offences may be committed” and obliges its member States to “adopt such measures as may be necessary to establish public provocation to commit a terrorist offence [...] as a criminal offence under its domestic law” (Council of Europe, 2005). The Convention therefore makes clear that publicity, criminal intent of the speaker and some link between the utterance and a potential for the commission of a terrorist act are to be treated as elements of the crime of incitement, hence allowing for a somewhat greater degree of harmonization among member States.
That said, there is a broad consensus, especially within Europe, that direct incitement to terrorism would not be sufficiently comprehensive to cover the many forms terrorist propaganda may take. In order to fill this ‘gap’, a novel offence has been incorporated into various criminal legal codes: ‘glorification’ of or ‘apology’ for terrorism broadly refers to the praising of terrorist groups or acts. The creation of this offence provides a powerful tool to States clamping down on terrorist contents in cyberspace, however, free speech defenders and human rights mechanisms have expressed doubt that the concept of “glorification” of terrorism is sufficiently narrow and precise to serve as a basis for criminal sanctions compliant with the requirements of the principle of legality and the permissible limitations of the right to freedom of expression (see e.g. Amnesty International, 2017). Nonetheless, the EU has obliged her members to criminalize the ‘glorification’ of terrorism and many other non-European States, including India and Pakistan, went along with the trend.

There are numerous ways to conceptualize terrorist propaganda other than in terms of incitement or glorification. Germany, for example, has applied the prohibition of the depiction of violence (art 131 StGB) to terrorist propaganda (see e.g. BGH, 2013), in the UK, expressing support for a proscribed terrorist organization is a criminal act and Turkey simply prohibits the ‘dissemination of terrorist propaganda’ – an arguably purposely unconcise offence which is routinely abused to prosecute journalists and political opponents (Zeldin, 2015).

The paragraphs above point to a number of problems in connection with tackling terrorist propaganda by means of legislative tools. Firstly, laws targeting terrorist speech are ambiguous with regard to key legal aspects, which in turn opens doors for abuse; secondly, legislating against terror propaganda comes at the detriment of the fundamental right to free speech; and thirdly, the lack of agreement on what constitutes ‘terrorist speech’ prevents the formulation of a coordinated response to the phenomenon which more often than not has international ramifications. It is unlikely that States will in the near future be able to agree on a shared approach to terrorism incitement as progress on discussions on even the basic definition of terrorism have come to a standstill. This failure has fatal consequences for any law enforcement approach toward internet-facilitated terrorist crimes.

- Transnational legal cooperation: bringing perpetrators to justice

Timely and effective transnational cooperation between law enforcement, intelligence and judicial agencies for the purposes of gathering and sharing intelligence, identifying perpetrators and obtaining data serving as evidence are indispensable to the successful prosecution of terror-related crimes committed through cyberspace. A particularly salient issue in the present context is that of law enforcement’s access to electronic data as ISPs holding such data, and the criminal justice bodies investigating and prosecuting cyber-related crimes are often located in different jurisdictions.

The general principles applicable to mutual legal assistance in cases involving terrorism or transnational organized crime are part of comprehensive mechanisms set out in the universal and regional counter-terrorism instruments and other instruments dealing with transnational organized crime; a counter-terrorism instrument dealing specifically with Internet issues
connected to terrorism does currently not exist (UNODC 2012). Examples of applicable instruments are the UN Convention against Transnational Organized Crime and the European Convention on Mutual Assistance in Criminal Matters. Furthermore, the Council of Europe Convention on Cybercrime of 2001 (‘Budapest Convention’) is highly relevant as it not only has, in relative terms, a broad geographic reach – it has to date been ratified by merely 71 States – but also because it deals specifically with transnational crimes involving the internet (ibid). It sets common procedural rules, for example with regard to the storage and sharing of evidence in electronic format and provides signatories with guidance on mutual assistance procedures. Although the Budapest Convention does not address terrorism in the online sphere in particular, many of its principles can be applied to the topic (Azani & Liv, 2018). States that are not parties to such Conventions often have bilateral cooperation agreements with one another and some – such as China – perform law enforcement cooperation on a reciprocal basis (Du & Yu, 2019).

Typically, law enforcement cooperation is initiated via Mutual Legal Assistance (MLA) requests by the judicial authority of one State to a judicial authority of another State, in which the requested judicial authority is asked to perform one or more specified actions (United Nations Office for Drugs and Crime, 2018). The issuing, validation and execution of MLA requests can be highly bureaucratic and time-consuming especially when requesting and requested State do not have a common mutual legal assistance treaty: it can take weeks, sometimes months, to process. A specific obstacle to collaboration on cooperative investigation and prosecution of terrorism-related cases is the ‘dual criminality’ requirement, which is a prerequisite for mutual legal assistance: it refers to the requirement that the conduct underlying the request for assistance must be considered a crime in both the requesting and requested State (OECD, 2018). As demonstrated above, criminal prohibitions on the dissemination of terrorist propaganda differ in key legal aspects, a condition which has often led to mutual legal assistance or extradition requests being refused when authorities in requested countries considered dual criminality requirements not to have been satisfied (United Nations Office for Drugs and Crime, 2018).

As a consequence of these shortcomings, criminal justice actors of two States are often required to take recourse to flexible, pragmatic approaches. Informal cooperation – such as agency to agency or police to police cooperation - is therefore often employed by authorities, especially in cases of ICT-facilitated crimes where timeliness is critical: ISPs usually store data only for a limited period of time which, once eclipsed, might result in evidence being lost forever if not retrieved swiftly. There are, however, limits to the extent to which informal mechanisms may be employed. It might be acceptable for a law enforcement actor in one country to directly contact an ISP located in another country in order to ask for preservation or retrieval of data, but the search and seizure of data will usually still require authorization, which can only be obtained by formal means (United Nations Office for Drugs and Crime, 2012). Informal channels bypassing the central institutions of another State, in other words, may be more efficient in retrieving information located abroad, however, they also risk violating State sovereignty as States may perceive retrieval of citizens’ data by foreign security or judicial authorities as intrusive when they are not accompanied by official procedures (United Nations Office for Drugs and Crime, 2019).
A number of institutions have evolved out of the need for flexible and unbureaucratic transnational cooperation. One of those is the installation of the post of Liaison Officers (LOs), and more specialized Counter Terrorism Liaison Officers (CTLOs) serving as bi-lateral contact points for practitioner-level cooperation (Swallow, 2013). Liaison officers are utilized to facilitate relationship-building between law enforcement institutions of two States, but also between international police organizations, for example between Interpol and Europol. Many such organizations have installed so-called ‘24/7 contact points’ establishing points of contact to respond to urgent requests involving, for example, preservation of electronic evidence before more formal legal channels are pursued (Peters & Jordan, 2020). Moreover, the Budapest Convention is currently in the process of being supplemented by a second protocol to its main body, which contains provisions on the disclosure of subscriber information: accordingly, it allows law enforcement actors to directly request certain data from foreign-based ISPs without having to go through the mutual legal assistance process (see Cybercrime Convention Committee, 2020).

Although it appears that States are increasingly supportive of such mechanisms, it has been noted on various occasions that they often are not sufficiently laid out by means of national legislation, hence presenting a challenge to the rule of law (Osula 2015). Besides the tension of informal law enforcement cooperation with State sovereignty, the efficiency of informal intelligence sharing comes at the cost of privacy protection of citizens when foreign actors can access personal data in absence of a sound legal basis for such transaction, a condition which again highlights that efficiency in counterterrorism often comes at a cost for fundamental democratic rights.

III: Private-sector involvement: the privatization of law enforcement

One of the difficulties in controlling online contents results from the sheer quantity of information circulating in cyberspace, more often than not by anonymous sources, making surveillance as well as investigation and prosecution of each and every potential offence in cyberspace virtually impossible for law enforcement actors. Key to solving this malaise is the outsourcing of judicial (and increasingly surveillance) functions to (private) Internet intermediaries, a type of tech company encompassing the “wide, diverse and rapidly evolving range of service providers that facilitate interactions on the internet between natural and legal persons” (Council of Europe, 2018). Such intermediaries are considered to have better access to data and control over the internet’s architecture and therefore superior technical capacity to manage online contents. In particular those intermediaries that provide platforms for and curate third-party content (‘hosting service providers’ or HSPs, which include social media platforms), are increasingly employed to act as proxies enforcing States’ counter-terror policies (Huszti-Orban, 2018).

While most HSPs had voluntary mechanisms in place to control contents on their services from early onward, the increasing pressure to rid cyberspace of terrorist propaganda compelled governments to regulate the manner in which online intermediaries manage user-generated contents on their products, which came to be perceived as not transparent and
ineffective. The legal tool employed to this end is the imposition of intermediary liability on online services, a type of culpability based on the doctrine of vicarious liability: an innocent third party may be held responsible for the conduct of a primary perpetrator because she has an authoritative legal relationship with another party (for example, when parents are held responsible for the actions of their children or employers for their employees) (Strowel, 2009). Threatening intermediaries with sanctions for user-generated contents at odds with the criminal law provides governments with a convenient and cost-efficient manner of implementing speech regulations.

There are different possible models regarding the role of online intermediaries in society: on one side of the spectrum, States may opt for a pure self-regulatory approach by granting legal immunity to internet intermediaries for third-party content (e.g. in the US) and on the other side, States may subordinate corporate policy to government policy entirely, hence making intermediaries instruments of the State (e.g. in China). Both approaches are controversial as the American approach effectively makes private tech companies the arbiters of free speech in cyberspace, while the Chinese approach interferes with companies’ right to conduct business freely and profitably and ultimately amounts to State censorship. Most democratic States, however, find themselves somewhere in the middle; the EU in particular is currently fashioning herself as the pioneer in the development of a legal framework by which she attempts to reconcile the need to provide safe online environments to citizens and protect their online freedoms with the need to promote innovation and growth in the Digital Single Market.

### Self-regulation

Most of privatized speech regulation – referred to as ‘content control’ – takes place voluntarily as HSPs have strong reputational incentives to keep their services clear from violent, sexually explicit or otherwise objectionable contents – being associated with terrorism in particular will have commercial consequences. Self-regulation works through self-drafted Terms of Service (ToS) outlining which content categories are allowed and which ones not, to which users have to consent to when accessing sites. A ToS is usually not aligned to any particular legal system and uses its own definitions, including on what constitutes ‘terrorist contents’. User violations may be detected and enforced by means of three possible mechanisms: staff member moderators are individuals hired to monitor and moderate content on their employer’s platforms and services; automated tools employ technologies such as artificial intelligence or machine learning; and user moderation refers to a method where content moderation is outsourced to the online community, e.g. by providing users with an option to flag and report objectionable contents (OECD, 2020). Human moderators have the advantage of being able to detect nuances and context, while technological tools tend to be unconcise while, however, being cheaper and faster than human resources (ibid). This constitutes a dilemma as HSP’s – which are, after all, commercial enterprises – tend to prioritize their business interests over fundamental rights, and are therefore likely to employ cheap, over-censoring mechanisms over expensive, more concise ones. YouTube, for example, reports that 7,390,963 out of 7,872,684 videos removed between June and
September of 2020 were flagged by automated tools (YouTube, 2020). Only about half of 50 major social media corporations assessed in a recently published OECD report have some mechanism in place to notify users in case of potential violations of their ToS and give them the opportunity to appeal to removal decisions (OECD, 2020). Problems such as these are more profound when the enforcement of alleged ToS violations by private corporations are more far-reaching than the mere removal of a comment or video, for example when profiles are deleted as a consequence of an alleged violation.

In theoretical terms, privatized law enforcement means that HSPs are transferred quasi-normative functions (being able to draft autonomous codes of conduct), quasi-executive functions (being able to enforce their rules according to own judgment) and quasi-judicial functions (providing own, or no, appeal mechanism) (Huszti-Orban, 2018). Apart from the obvious legitimacy problem this entails from a democratic viewpoint, arguments against a purely self-regulatory approach concern the effectiveness of private speech regulation to diminish the visibility of harmful contents in cyberspace: after all, monitoring and enforcing rules in cyberspace is costly, and as companies are above all driven by economic interests, the absence of negative incentives may render private policing inefficient.

These issues are the point of departure of one of the most salient contemporary debates surrounding cyberspace regulation, namely the scope of private power over online contents. In other words, who should ultimately manage free speech online: tech companies or the government? The former view is taken by the US, where the First Amendment to the US Constitution reads that “Congress shall make no law … abridging the freedom of speech” (U.S. Const. amend. I). In addition, Section 230 of the Communications Decency Act of 1996, called the ‘safe harbor clause’, stipulates that no provider of an interactive computer service shall be treated as the publisher and speaker of third-party content and that companies are allowed to pursue “any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected” (Communications Decency Act, 1996). These stipulations reflect a view of internet intermediaries as business endeavors that should no act in the pursuit of public utilities, and the strong presumption against government interference with individual freedoms, which characterize US legal-political culture.

- **Steering private behavior: regulated self-regulation**

In order to limit the leeway of private companies in deciding about the contents to be accessible on their services, most States today have institutionalized an intermediary liability regime. There are important externalities to keep in mind in connection to intermediary liability. Firstly, an argument against this type of liability is that it may stifle innovation and cripple the digital economy as costly online policing renders private internet services unprofitable; and secondly, critics point out that as intermediary liability entails hefty fines for the non-deletion of often vaguely defined content categories, broad over-censorship will be the result as companies will remove even those contents the illegality of which is doubtful as
a precautionary measure. Consequently, some safeguards must be put in place in order to keep the burden of policing manageable for online service providers and to prevent over-censorship.

In the EU context, the e-Commerce Directive of 2000 sets the baseline criteria for the intermediary liability regime in the EU. While the Directive still imposes on Member States the obligation to enact ‘safe harbor’ clauses for a range of online intermediaries from liability, the Directive specifies that such protections shall only extend to a service provider whose activity is “mere technical, automatic and passive” and where it has “neither knowledge (…) nor control” over content transmitted through or stored on their site (Hornik & Villa Llera, 2017). With regard to hosting services, intermediaries are exempted from liability when; a) the provider does not have actual knowledge of illegal activity or information and is not aware of facts or circumstances from which the illegal activity or information is apparent; or b) the provider, upon obtaining such knowledge or awareness, acts expeditiously to remove or to disable access to the information. Furthermore, art. 15 specifies that intermediaries “may not be subject to any general obligation to monitor the information which they transmit or store” nor “actively to seek facts or circumstances indicating illegal activity” (European Council, 2000).

This regime effectively describes a notice-and-take-down principle that characterizes the intermediary liability system as negligence-based – hosting service providers may only incur liability when they fail to remove contents once they are made aware of it but are not obliged to proactively monitor for such contents (Frosio, 2018). The non-monitoring obligation, as clarified in the landmark case Sabam v Netlog, effectively prevents the imposition of an obligation on companies to make use of ex-ante filtering systems, which would contradict the right to conduct business (since such monitoring is complicated and expensive) and imposes disproportional burdens on the right to free speech as upload filters tend to censor much more comprehensively than ex post content take-downs (Case C-360/10, 2012).

Since 2000, a number of developments – rapid technological evolution especially with regard to HSPs in the context of the Web2.0 revolution, the increasing incidence of terror attacks on European soil and the salience of terrorist propaganda in cyberspace - have led the EU to adopt a number of non-binding instruments and initiatives aiming specifically to reduce terrorist contents on hosting platforms. These, however, did not appear sufficiently effective: as of 2018, 20 major social media platforms surveyed by the Commission on average only removed 63% of terrorist contents upon notification of such (European Commission, 2018).

The principles thus developed in a number of recommendations and communications were eventually adopted as binding law in the form of the Regulation on the Prevention on the Dissemination of Terrorist Propaganda, proposed in 2018, and currently in the final stages of negotiation. The Regulation obliges EU Member States to impose markedly stricter and more concise obligations on HSPs regarding terrorist propaganda. The measures stipulated therein include an obligation for HSPs to remove terrorist contents, once notified of such by competent authorities, within one hour, whereby systematic failure to do so may be punished by up to 4% of the global turnover of the last business year of the HSP in question. Furthermore, the Regulation provides a mandatory definition of ‘terrorist content’, which
formerly had been left to HSP’s own discretion in line with their ToS (European Commission, 2018).

Aware of the aforementioned problem of over-censorship occurring as a by-product of intermediary liability, the proposed Regulation provides a number of safeguards. Firstly, it obliges all service providers to establish effective complaint mechanisms for citizens to challenge the removal of allegedly illicit messages; secondly, it mandates that platforms use human oversight besides automated detection tools; and thirdly, it demands that platforms must publish annual transparency reports (European Commission, 2018).

On the other hand, the regulation also stipulated that service providers shall take “proactive measures to protect their services against the dissemination of terrorist content” (art 6) and explains that, in light of the “particularly grave risks associated with the dissemination of terrorist content” the decisions adopted on the basis of the Regulation could derogate from the ‘non-monitoring obligation’ stipulated in art. 15 of the e-Commerce directive. Ultimately, it was decided to remove the requirement for proactive monitoring (Kucyerawy, 2019); however, the fact that EU legislators have shown willingness to override a core principle by which the EU’s intermediary liability regime upholds the balance between efficiency and speech protection is evidence of the EU’s own tendency to water down the right to free speech.

In short, the EU has adopted a relatively complex approach to intermediary liability which aims at balancing HSP’s business interests (conducting a profitable business) with public interests (diminishing the accessibility of terrorist materials online and safeguarding free speech). It has done so by installing an intermediary regime which precludes HSP’s from being subjected to a general monitoring obligation and by mandating the institutionalization of safeguards by which citizens can challenge take-down decisions by HSPs. Other countries, especially those where political power is concentrated centrally, have taken much stricter measures to ensure alignment of private interests with State policy. China for example, in adopting a regime which entirely subordinates private to public policy, exerts tight control over the activities of internet intermediaries, with fatal consequences for citizens’ right to free speech and access to information.

- **Controlling private behavior: networked authoritarianism**

   The idea of a citizenry that can freely exchange ideas naturally sits uncomfortably with leaders of authoritarian-style governments. However, in today’s digital world, depriving citizens from using the internet would be accompanied by massive economic setbacks. Some States, notably Xi Jinping’s China, have solved this dilemma to their advantage by instrumentalizing internet services to improve the efficiency of intelligence and surveillance accompanied by heavy censorship, hence stabilizing their authoritarian regimes. This approach toward internet governance has been termed ‘networked authoritarianism’ (Burgers & Robinson, 2016). It stands on the other side of the spectrum, as it requires tight government control of internet intermediaries.
China’s approach toward internet governance is broadly laid out in its White Paper on Internet Policy of 2010, which outlined a vast range of vaguely defined prohibited topics, including such ones “endangering State security, divulging State secrets, subverting State power and jeopardizing national unification; damaging State honor and interests; instigating ethnic hatred or discrimination and jeopardizing ethnic unity; jeopardizing State religious policy, propagating heretical or superstitious ideas; spreading rumors, disrupting social order and stability; disseminating obscenity, pornography, gambling, violence, brutality and terror or abetting crime” (Information Office of the State Council of the People’s Republic of China, 2010). Internet companies operating in China are expected to conform to these requirements and to assume responsibility for the content published on their platforms; failure to do so is punished through fines and revocation of licenses (Ruan, 2019). Foreign service providers refusing or failing to abide by China’s strict censorship rules may simply be denied access to the Chinese digital market through what has been termed the ‘Great Firewall of China’, a number of technological measures enabling the government to inspect any data being received or sent, and to block destination IP addresses and domain names (Economy, 2018). Most international social media and messaging platforms are completely blocked in China, including Facebook, Whatsapp and Twitter, the websites of Amnesty International and Human Rights Watch, and a large number of international news websites (Ministry of Foreign Affairs of the Netherlands, 2020).

In absence of foreign competition, Chinese hosting service providers are able to flourish – given that they agree to comply with the government’s internet policy which has been outlined in a large number of regulations and ordinances formulated since 2010. These include the obligation to proactively monitor services for above-mentioned categories of ‘harmful information’, to install real-name online registration systems aimed at preventing users from hiding behind pseudonyms, to grant authorities access to business premises to secure information and to transfer of data (including encryption keys) to State actors when requested (Ministry of Foreign Affairs of the Netherlands, 2020).

The Chinese Communist Party – and more specifically the Cyberspace Administration of China (CAC) – ultimately controls not only contents circulating on internet sites, but also the entire digital infrastructure and is therefore able to not only filter whatever contents it deems inappropriate or dangerous to its regime, but also to shut the internet down for entire regions. Such shutdowns are becoming increasingly common not only in China but in various Asian countries featuring repressive (cyber) regimes: Myanmar, for example, regularly orders Internet Service Providers to shut down access to internet in conflict-riddled Rakhine state, and India – a democratic country – had suspended the availability of high speed 4G Internet services (2G services were still operational) in Indian-administered Jammu & Kashmir for 165 days straight (Shastry, 2020). In Pakistan, there are regions, especially in the Balochistan province where internet has been totally suspended for ‘security reasons’, in some cases for years now. Governments typically use national security and the prevention of insurgency as go-to justifications for internet black-outs, however, human rights groups warn that shutting down internet services in entire regions deprives affected citizens not only of some abstract rights, but effectively puts their health and lives in danger as they may not be able to, for
example, inform themselves of the development and available protection of the Covid-19 pandemic.

IV: Discussion and Recommendations

The above sections have outlined two pathways by which novel legal challenges linked to globalizing factors – in the present case, the use of cyberspace by terrorists – have been approached. Firstly, within the criminal justice paradigm, the internationalization of crimes precipitates efficient cooperation between the judicial and law enforcement bodies of sovereign States. We have seen that legal harmonization on the subject of terrorist propaganda has been only minimally successful, which directly affects the applicability of transnational legal enforcement mechanisms; this, in turn, has led to the development of various informal cooperation mechanisms. Secondly, within policy paradigms following the logic of prevention, the privatization of law enforcement is increasingly viewed as a means of mitigating the visibility of terrorist propaganda in virtual space, whereby jurisdictions take different views on the scope of public control over private speech regulation.

The proliferation of ICT in societies all over the globe, and terrorists’ cunningness in exploiting cyberspace, represents a formidable challenge to traditional policy-making and law enforcement. The failure to harmonize counter-terrorism frameworks and to cooperate in law enforcement and prosecution of crimes involving terrorism and cyberspace allows terror groups to conduct cyber-operations – propagandistic or otherwise – in relative safety: statistics on successful cybercrime prosecutions, especially where transnational components are involved, do not exist but are assumed to be so bleak as to amount to virtual impunity for perpetrators (Peters & Jordan, 2020).

The reasons for this failure are manifold. At the core the problem lies within the organization of the international system on the basis of the Westphalian model of sovereign nation-States with distinct legal-political systems, inherited from the enlightenment and in many ways ill-equipped to deal with the complex, borderless threats of the 21st century. The informalization of law enforcement cooperation described earlier is a manifestation of the recognition of this shortcoming and, from the perspective of efficient prosecution and law enforcement of transnational crimes, a desirable trend.

There is, however, a problem, which can be traced throughout this paper: efficiency in counterterror tends to come at a cost for the rule of law and fundamental rights protection. The theoretical question to be asked, then, is to what extend the increasing importance of transnational cooperation justifies a weakening of procedural safeguards for citizens. Take, for instance, Germany-based journalist Deniz Yucel, a Turkish national sentenced to almost 3 years in jail in absentia for spreading ‘terrorist propaganda’ (Kucukgcmen, 2020). Behind the backdrop of Turkey’s crackdown on regime-critical journalists, should Germany assist in Yucel’s prosecution without adequately reviewing his case, hence risking conspiring with a highly repressive regime?
The privatization of law enforcement, in the same way as the creation of informal criminal cooperation channels, may be understood within a broader shift toward pragmatic policy-making resulting from a failure to address novel threats through established paradigms. Terrorists’ ability to exploit emerging technologies has allowed them to engage the international community in a cat-and-mouse game where the latter is doomed to lose; by the time States have formulated and coordinated a counter-strategy, malicious actors will already have taken recourse to new tactics. Private regulation by those actors who have direct access to the internet’s ‘architecture’, or code, provides the flexibility, informality and de-centrality required to deal with the fast-changing technological environment and concurrently evolving threats. Yet here too, the tension between efficiency and fundamental rights protection is apparent; handing over responsibility to manage fundamental rights to private actors motivated by commercial interests evidently comes at a cost for democratic rights.

In other words, from increasing restriction of free speech resulting from the creation of new, vaguely defined offences such as ‘glorification of terrorism’, the move away from ‘non-monitoring obligations’ of ISPs and the erosion of procedural safeguards in light of the informalization of law enforcement cooperation, the tendency to prioritize efficiency in counter-terrorism over fundamental rights protection is a well-documented trend which has been observed with concern by human rights organizations, academics and journalists alike (see e.g. Guild & Bigo, 2018). On one hand, in the sense that that law reflects the needs of society, restrictions to fundamental rights may be justifiable by reference to societal developments (and more specifically the emergence of new threats). On the other hand, there needs to be some appreciation of the fact that the increasing limitation of fundamental rights in the name of the ‘fight against terror’ may well play into the hands of terrorists who, after all, typically place the destruction of democracy at the apex of their agendas.

In practice, this means that while the legal trends sketched above may in principle constitute proportionate and necessary responses to the complex, amorphous threats created as by-products of globalization, it should be ensured that the impact such measures may have on fundamental rights is limited as far as possible. Firstly, undue intrusion into citizens’ rights may be prevented through clear definition of terrorism-related offences as the criminalization of vague offence categories such as the ‘dissemination of terrorist propaganda’ opens doors to arbitrary application and the silencing of legitimate opposition. Secondly, the proportionality of a measure may be restored by equipping it with safeguards: the externalities of private speech regulation may, for example, be mitigated when companies are obliged to disclose their speech practices and to provide users with the ability to challenge take-down decisions by ISPs (as the EU has done). Similarly, by clearly defining the boundaries within which law enforcement agencies of two States can interact in absence of formal procedures, arbitrariness in transnational law enforcement may be prevented.

Thirdly, policy-makers should explore less restrictive options before taking recourse to more restrictive ones. Case studies show, for example, that practical cooperation under MLA procedures often suffer from lack of funding and personnel trained in digital forensics (Carrera & Stefan, 2020). Instead of entirely by-passing official MLA procedures (and hence endangering procedural safeguards), States may consider installing bodies specialized in
facilitating transnational legal transactions, and systematically provide training to law enforcement and judicial actors so as to enhance their preparedness to the challenges connected to transnational cyber enforcement. Likewise, governments could invest in the development of new, more efficient content detection tools or stimulate knowledge and skill-sharing between private actors. Such measures would constitute low-hanging fruits by which the efficiency of counterterrorism can be improved without impacting civil rights.

On the long run, the harmonization of legal standards with regard to, amongst others, data protection, speech protection, conceptualization of terrorism and responses thereto remains a prerogative in the absence of which major progress in counter-terrorism is unlikely to materialize. Both transnational terrorism and the proliferation of ICT-facilitated crimes are unlikely to disappear in the near future; yet the most promising proposals to address either, or both in conjunction – for example the creation of an international tribunal to adjudicate terrorist offences or the adoption of a universal instrument bridging cyber-crime and terrorism – hinge on the ability of governments to develop some shared understanding of the threat in question and responses thereto.

Although the prospects for developing such consensus currently appear a far cry, international organizations must continue providing platforms for governments to negotiate, and States must recognize that when harmonization efforts come to a halt, so will any major progress in the fight against terrorism.
Bibliography


BHG, Beschluss vom 22.08.2013 – 3 StR 244/13


Brandenburg v. Ohio, 295 U.S. 444 (1969)


Case C-360/10, Belgische Vereniging van Auteurs, Componisten en Uitgevers SVBA (SABAM) v Netlog NV [2012] OJ C 98, pp 6


European Commission, 2018. *Commission Staff Working Document Impact Assessment Accompanying the document Proposal for a Regulation of the European Parliament and of the Council on preventing the dissemination of terrorist content online*. Available at [https://www.parlementairemonitor.nl/9353000/1/j9vij5epmj1ey0/vkrs7i664lz7](https://www.parlementairemonitor.nl/9353000/1/j9vij5epmj1ey0/vkrs7i664lz7)


K. Huszti-Orban, "Internet intermediaries and counter-terrorism: Between self-regulation and outsourcing law enforcement1," *2018 10th International Conference on Cyber Conflict (CyCon)*, Tallinn, 2018, pp. 227-244


