**Counter-terrorism policies and human rights: Lessons from a digital rights perspective**

16 July 2019

*Submission of Access Now regarding the call for submissions on “Soft law and informal lawmaking in the global counter-terrorism architecture: Assessing implications on the promotion and protection of human rights and fundamental freedoms”[[1]](#footnote-1)*

*To the Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism*

Authors: Christopher Andreou, Peter Micek, Isedua Oribhabor

**TABLE OF CONTENTS**

**Introduction** 1

1. **PROBLEMATIC APPROACHES TO COUNTER-TERRORISM** 2
2. Recent Government Initiatives 2
3. Internet Shutdowns 4
4. Encryption, Data Localization and Retention 5
5. Content Removal 6
6. **RIGHTS-RESPECTING PRINCIPLES AND RECOMMENDATIONS** 7
7. For Governments 7
8. For Technology Companies and Platforms 8

**Introduction**

In an effort to combat the propagation of terrorist content online, governments around the world are enacting policies that violate users’ rights to freedom of expression and privacy. Vague or unclear definitions of terms such as “extremism” or “terrorism” can result in the targeting of satire, journalism, activism and community organizing, political protest, and other forms of expression, as well as undercutting existing rule of law and protections for both privacy and free expression. Additionally, public-private agreements threaten to become loopholes intended to circumvent human rights scrutiny and accountability to public institutions operating under the rule of law.

Internet companies are also exploring measures to counter what they consider "violent extremism” and “terrorism” online. In 2016, Microsoft, Facebook, YouTube, and Twitter announced a partnership[[2]](#footnote-2) to stop the spread of “online terrorist content” by committing to share digital signatures that will be used to flag images and videos. More recently, these companies, in addition to Amazon, signed the Christchurch Call[[3]](#footnote-3), an initiative aimed at "eliminat[ing] terrorist and violent extremist content online." As laudable as the intent behind these actions can be, there is a real risk of censorship as these companies create content moderation and removal practices that will adversely impact the right to freedom of expression online.

In Europe, the European Commission proposed a regulation to combat the dissemination of terrorist content online despite doubts of its necessity and effectiveness, as well as evidence showing that counter-terrorism policies have [had disproportionate impact](https://www.amnesty.org/download/Documents/EUR0153422017ENGLISH.PDF) on journalists, artists, and human rights defenders.

Having a common understanding of how human rights principles apply to digital counter-terrorism initiatives is essential for both governments and companies in order to protect freedom of expression and safeguard privacy. The Necessary & Proportionate International Principles on the Application of Human Rights to Communications Surveillance[[4]](#footnote-4), which were completed through a global consultation in which Access Now took part, provides governments and other relevant actors a framework on how existing human rights law applies to modern digital surveillance. Additionally, Access Now’s guidance on content moderation[[5]](#footnote-5) offers a set of principles for online platforms when creating content moderation policies so these protect free expression. Finally, we recognize the recommendations made by the special mandate holders in the context of the proposed EU terrorist content regulation applicable to any content moderation measure that aims to tackle terrorist content online.[[6]](#footnote-6)

Our submission addresses problematic approaches to counter-terrorism in the digital space by both governments and companies, providing recommendations for rights-respecting principles. We urge governments and companies across the world to guarantee the rights to privacy and freedom of opinion and expression, online and offline, and to always apply human rights standards when crafting policy.

1. **PROBLEMATIC APPROACHES TO COUNTER-TERRORISM**
2. **Recent Government Initiatives**

**The Christchurch Call**

On May 15, a group of 18 countries, led by New Zealand and France, along with major tech companies including Facebook and Google, published the “Christchurch Call”, a voluntary pledge committed to eliminating terrorist and violent extremist content online. The Christchurch Call was named after the city in New Zealand where tragic attacks[[7]](#footnote-7) were filmed and shared online via social media platforms. Given the increased attention on digital technology and terrorism, governments are rushing to enact policies in an effort to [clamp down](https://www.accessnow.org/australias-plans-to-regulate-social-media-bound-to-boomerang/) on violent extremist speech posted or shared on social media platforms. While the Christchurch Call certainly has its merits, just like so many other policies enacted by states aiming to counter terrorism online, it also raises important concerns.

The drafting of the text of the Call did not engage stakeholders on its actual content, only on implementation. This resulted in many pledges that put human rights at risk. Firstly, the Call uses broad, ambiguous language in addressing “terrorist and violent extremist content”. The definition of terrorist and violent extremist content can vary between countries and in some cases can be used arbitrarily to harm human rights. The Call also relies heavily on upload filters as a technical tool to prevent the dissemination of violent content. While in some cases it may be clear that content is not appropriate, in most cases the interpretation requires meaningful human involvement and evaluation of context. Lastly, the Call focuses primarily on putting the responsibility for identifying and removing violent extremist content on internet companies. Governments should never outsource the regulation of speech to private entities, as doing so removes principles of due process, government accountability and free speech-enabling intermediary liability protections.

Despite these shortcomings, unlike other recent movements in the EU and Australia[[8]](#footnote-8) aimed at achieving similar objectives, the Christchurch Call does explicitly cite human rights as both a guiding objective of government action and a limiting factor to it. It also underlines transparency obligations for companies and governments. Moving forward, to ensure human rights are prioritized in its implementation, we recommend that all signatories of the Call engage in meaningful participation with a genuine commitment to discussing issues and concepts in depth and with sufficient time for all stakeholders to provide input.

**Proposed EU Regulation on preventing the dissemination of terrorist content online**

The aforementioned regulation[[9]](#footnote-9) proposed by the European Commission to address content online that is perceived as promoting terrorism also raises serious concerns for freedom of expression. Without evidence to demonstrate that the existing laws and measures are insufficient to address the potential propagation of terrorist content online, the proposed Regulation cannot be deemed justified and necessary. Similar to the Christchurch Call, there are several aspects of the original proposal by the European Commission that would endanger freedom of expression. Broad, ambiguous definitions are present in the text, which can lead to uncertainty and can be misread. The Regulation puts the power to remove terrorist content into the hands of undefined “competent authorities”, and places companies’ terms of service over the rule of law when judging content. The Regulation, as proposed by the Commission, would introduce serious risks of arbitrariness and have grave consequences for freedom of expression and information, as well as for civil society organizations, investigative journalism, and academic research, among other fields.

In April 2019, the European Parliament adopted[[10]](#footnote-10) its report on the proposed regulation, addressing the worst threats to freedom of expression such as the general obligation to monitor content or the system of referrals. A one-hour removal deadline for online companies, however, remained in the text.

As this law could be potentially dangerous for human rights, it must be closely monitored to assess and mitigate its negative impact.

1. **Internet Shutdowns**

Over June and July 2019, in the midst of a revolution, Sudan shut down mobile internet via its major service providers, effectively putting the country off the grid. With families unable to communicate with each other and journalists struggling to report on the organized killing and looting being committed by paramilitary forces, the internet shutdown has exacerbated human rights violations in Sudan. While the situation in Sudan appears to be the interim government’s direct attempt to impede the efforts of protesters, other state-directed shutdowns that are justified for safety reasons produce the same negative consequences. The Pakistani government has frequently[[11]](#footnote-11) implemented internet shutdowns ostensibly for security reasons, while Sri Lanka blocked access to social media platforms following the recent deadly attacks that took place on Easter this year.

Rather than stopping the spread of disinformation, these measures usually exacerbate the problem and further harm victims. Just like in Sudan, many people in Sri Lanka rely on social media platforms[[12]](#footnote-12) to communicate, meaning that the shutdown prevented individuals from confirming the safety of their loved ones. Shutdowns impede journalists from gathering information, and obstruct the spread of news and information from trustworthy sources. Furthermore, research on internet shutdowns in India showed that rather than improving the situation, shutdowns often lead to increased violence and chaos.[[13]](#footnote-13)

Despite internet shutdowns being condemned as human rights violations by the UN General Assembly and The African Commission on Human and Peoples’ Rights, certain countries such as India, where 67% of the world’s documented shutdowns in 2018 took place, have legalized them.[[14]](#footnote-14) Regardless of their purpose, internet shutdowns violate human rights by disproportionately curtailing freedom of expression, cutting access to information, and inhibiting people from assembling and associating peacefully, online and off. They must continue to be condemned and fought against on the basis of human rights.

1. **Encryption, Data Localization and Retention**

Another problematic approach to counter-terrorism lies in government-led efforts to weaken data security in the name of national security. In recent years, encryption has come under attack as lawmakers pass legislation and put pressure on companies to build tools to weaken encryption protocols. Just last year, the Australian Parliament passed a bill that allows law enforcement authorities to compel companies to provide encryption backdoors -- tools that weaken the security of a platform or device in order to gain access to private digital communications.[[15]](#footnote-15) The reports on encryption and anonymity by U.N. Special Rapporteur on freedom of expression and association, David Kaye, discuss the trend of governments neglecting their duty to protect freedom of expression and privacy by weakening encryption.[[16]](#footnote-16) States such as Turkey, Iran, and Pakistan have criminalized encryption, detaining citizens for using encrypted messaging apps, while others such as China and the United Kingdom have laws obliging companies to provide de-encryption tools to government authorities.

Weakening encryption puts users at risk. While governments could use these tools for law enforcement purposes, the existence of vulnerabilities means that they are also susceptible to use by bad actors to exploit and harm users. Proponents of these measures often argue that companies would only be compelled to provide access to particular devices; however, if such a key falls out of control of the company or government, the device will be exposed to substantial risk.[[17]](#footnote-17) Furthermore, the idea of a "secure backdoor” is deceptive. As internet services and platforms become more interconnected, government mandates to companies so they provide backdoors could lead to unanticipated, hard to detect security flaws that will further endanger users.[[18]](#footnote-18)

Governments further weaken data security and protection when they increase data retention and mandate data localization. Broad and indiscriminate data retention, particularly when companies are not transparent as to how long they retain data and who can get access to the data, places users’ privacy at risk. Meanwhile, the move towards data localization --requiring foreign companies to store data on citizens within a country's borders-- is concerning. Generally presented in terms of security and affording citizens "sovereignty” over their data, these measures are often politically motivated rather than based on technical efficiency, giving governments tools to monitor their citizens domestically. Governments, both repressive regimes and historically democratic states, have proposed data localization as a means of protecting against foreign government spying and national security risks. However, these arguments are belied by the fact that data security depends on the sophistication and architecture of infrastructure rather than the location of the data.

1. **Content Removal by Online Companies**

Companies are also considering measures to address what they deem violent or extremist content online, and they are subject to state pressure to take expeditious action against different forms of harmful and illegal content. This applies particularly to online platforms such as Facebook or YouTube, which face growing calls for regulation of content in the wake of tragedies like the Christchurch attack. While it’s imperative for these platforms to do their part in preventing the spread of harmful content, the measures they employ to do so must also be subject to standards and principles that respect users’ rights online.

The most important risk is that when platforms are left to decide and address for themselves what is “violent”, “extremist” or “terrorist”, they can have an outsized influence over freedom of expression online. Platforms can disproportionately shape public discourse on important issues, and can take over state functions by engaging in privatized law enforcement. Especially as major companies begin to rely on automated decision-making, without clear human rights standards for content moderation, including transparency, we are left with practices that abuse rather than respect human rights.

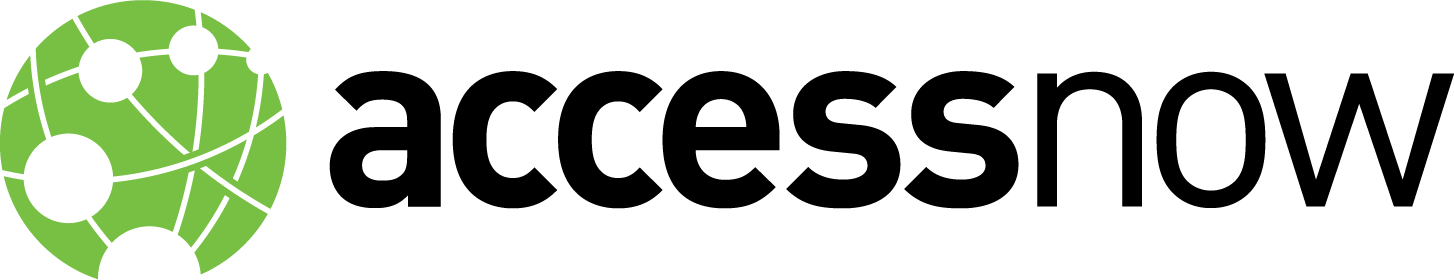
For example, the announced collaboration between tech companies to share digital signatures in an effort to combat the spread of “terrorist” content on their platforms demonstrates a danger companies run into with content moderation practices: defining what constitutes "terrorist” content. Context, therefore, is a major issue: because there is no clear, uniform definition of “terrorist” or “extremist” content, context is important in determining whether a post or video should fall under that category. Content shared by journalists for the purposes of reporting on a story or shedding light on a situation could therefore be in jeopardy of being labeled as “terrorist” or “extremist” and removed. This is a heightened concern when companies rely on automated decision-making systems which are not capable of making nuanced distinctions. And this danger is not merely theoretical: in 2017, YouTube began using machine learning protocols to sift through videos on the site, looking for objectionable content. This resulted in the shutting down, almost overnight, of the channels of around 900 individuals and groups who were documenting the civil war in Syria.[[19]](#footnote-19) Similarly, Facebook removed photos and posts documenting the ethnic cleansing of the Rohingya Muslim minority group in Myanmar after the posts had been flagged as disturbing by users.

As authorities and civil society organizations increasingly rely on social media to obtain evidence of war crimes and other abuses to build cases against perpetrators[[20]](#footnote-20), the pressure social media platforms face to remove content can lead them to impair these efforts. Companies are faced with the task of preventing the spread of harmful content online while preserving the documentation of atrocities and abuses; creating a set of international, rights-respecting standards would aid them in this task.

1. **RIGHTS-RESPECTING PRINCIPLES AND RECOMMENDATIONS**

As governments and companies consider and enact policies to combat the spread of different forms of harmful and illegal content online (including "terrorist" or "extremist" content), we urge them to keep the following principles in mind.[[21]](#footnote-21) Though by no means exhaustive, these principles and recommendations are meant to provide a basic framework for governments and companies seeking to enact policies that will respect the rights to freedom of expression and privacy.

1. **For governments:**
2. Any policies or regulations aimed at combating terrorist content online must be grounded in **clearly defined legal provisions**, subject to a high level of scrutiny by an **accountable and independent legal system**.
   * Governments must not force platforms to remove content unless such content has been found unlawful under a **rights-respecting legal process** and mechanisms for **notice and redress** are in place for the accused speaker.
   * Because limits on expression should only be imposed when necessary and proportionate, governments must take care not to conflate “extremism” with the broader term “terrorist content” as the term does not have an international legal definition.
3. Any approach for countering terrorist content that constitutes surveillance — such as social media monitoring, algorithmic content reporting, or content referral programmes — must be subject to the **same normative and legal restrictions applicable to communications surveillance** in other contexts. Any practices that rely on surveillance or monitoring of online content must be provided for by law, subject to the 13 ‘Necessary and Proportionate’ principles, including:
   * Limiting such surveillance to what is strictly and demonstrably necessary to achieve a **legitimate aim**.
4. **Due process** must be provided to those whose rights have been implicated.
5. During the design and implementation of a new policy or regulation, all potential **stakeholders must be meaningfully consulted**.
6. Partnerships between government agencies and technology companies must not allow for loopholes to circumvent human rights scrutiny and accountability to rule of law institutions.
7. **For Technology Companies and Platforms:**
8. Technology companies and platforms that are used for communications or deal with user content must ensure that their **takedown and appeal policies are made clear and public.** Furthermore, users must be able to freely accept these policies.
   * Companies must notify speakers of content takedowns as quickly and transparently as possible, preferably before the takedown takes place.
   * Companies must disclose data on implementation of content policies, including terms of service enforcement and responses to third-party requests, through regular transparency reports.
9. Companies should put **appeal mechanisms** in place for content removals based on the violation of “terms of service” or internal rules. Users should be able to eventually appeal a content moderation decision before a judge when their rights are affected, and companies must not restrict or condition this right to remedy.
10. **Automated decision-making systems must be as transparent as possible**, including by publishing information about how these systems are used according to detailed criteria. Users must also be made aware about the use of automated decision-making systems and afforded the right to request human review of their case.
11. **Government-mandated removal of disputed content** should take place only when the content has been specifically adjudicated as being **illegal and a court order has been issued.**
12. Companies should perform participatory and **periodic public evaluations** to determine how content moderation decisions are impacting the fundamental rights of users



##### 

##### **Access Now (https://www.accessnow.org)** defends and extends the digital rights of users at risk around the world. By combining direct technical support, comprehensive policy engagement, global advocacy, grassroots grantmaking, and convenings such as RightsCon, we fight for human rights in the digital age.

For More Information, please contact:

**Peter Micek** | General Counsel | peter@accessnow.org | +1-888-414-0100 x709

**Wafa Ben-Hassine** | Global Policy Counsel | wafa@accessnow.org | +216 71 840 335

1. Call for submissions: Soft law and informal lawmaking in the global counter-terrorism architecture, OHCHR, <https://www.ohchr.org/EN/Issues/Terrorism/Pages/SoftLaw.aspx>. [↑](#footnote-ref-1)
2. Partnering to Help Curb Spread of Online Terrorist Content, Facebook Newsroom (5 December 2016), <https://newsroom.fb.com/news/2016/12/partnering-to-help-curb-spread-of-online-terrorist-content/>. [↑](#footnote-ref-2)
3. Christchurch Call: to Eliminate Terrorist and Violent Extremist Content Online, available at <https://www.christchurchcall.com/>. [↑](#footnote-ref-3)
4. Necessary and Proportionate: International Principles on the Application of Human Rights to Communications Surveillance, <https://necessaryandproportionate.org/principles>. [↑](#footnote-ref-4)
5. Javier Pallero, Protecting Free Expression in the Era of Online Content Moderation, Access Now (May 2019), <https://www.accessnow.org/cms/assets/uploads/2019/05/AccessNow-Preliminary-Recommendations-On-Content-Moderation-and-Facebooks-Planned-Oversight-Board.pdf>. [↑](#footnote-ref-5)
6. Mandates of the Special Rapporteur on the promotion and protection of the right to freedom of

   opinion and expression; the Special Rapporteur on the right to privacy and the Special Rapporteur

   on the promotion and protection of human rights and fundamental freedoms while countering

   Terrorism (7 December 2018), <https://spcommreports.ohchr.org/TMResultsBase/DownLoadPublicCommunicationFile?gId=24234>. [↑](#footnote-ref-6)
7. Charlotte Graham-McLay et al, Christchurch Mosque Shootings Were Partly Streamed on Facebook, The New York Times (March 14, 2019), <https://www.nytimes.com/2019/03/14/world/asia/christchurch-shooting-new-zealand.html?searchResultPosition=58>. [↑](#footnote-ref-7)
8. Lucie Krahulcova & Brett Solomon, Australia’s plans for internet regulation: aimed at terrorism, but harming human rights, Access Now (26 March 2019), <https://www.accessnow.org/australias-plans-to-regulate-social-media-bound-to-boomerang/>. [↑](#footnote-ref-8)
9. A Proposal for Regulation on Preventing the Dissemination of Terrorist Content Online (2018), <https://ec.europa.eu/commission/sites/beta-political/files/soteu2018-preventing-terrorist-content-online-regulation-640_en.pdf>. [↑](#footnote-ref-9)
10. European Parliament legislative resolution of 17 April 2019 on the proposal for a regulation of the European Parliament and of the Council on preventing the dissemination of terrorist content online (2019), <https://www.europarl.europa.eu/doceo/document/TA-8-2019-0421_EN.html> [↑](#footnote-ref-10)
11. Berhan Taye, Pakistan shuts down the internet three times in one week, Access Now (6 November 2018), <https://www.accessnow.org/pakistan-shuts-down-the-internet-three-times-in-one-week/>. [↑](#footnote-ref-11)
12. Berhan Taye, Sri Lanka: shutting down social media to fight rumors hurts victims, Access Now (22 April 2019), <https://www.accessnow.org/sri-lanka-shutting-down-social-media-to-fight-rumors-hurts-victims/>. [↑](#footnote-ref-12)
13. Jan Rydzak, Of Blackouts and Bandhs: The Strategy and Structure of Disconnected Protest in India, Stanford Global Digital Policy Incubator (7 February 2019), available at <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3330413>. [↑](#footnote-ref-13)
14. Berhan Taye et al. THE STATE OF INTERNET SHUTDOWNS AROUND THE WORLD THE 2018 #KEEPITON REPORT, <https://www.accessnow.org/cms/assets/uploads/2019/06/KIO-Report-final.pdf>. [↑](#footnote-ref-14)
15. Australia joins Russia and China in undermining users’ security and threatening human rights, Access Now Press Release (6 December 2018), <https://www.accessnow.org/australia-joins-russia-and-china-in-undermining-users-security-and-threatening-human-rights/>. [↑](#footnote-ref-15)
16. David Kaye, Encryption and Anonymity Follow Up Report, U.N. Human Rights Special Procedures (June 2018), available at <https://www.ohchr.org/Documents/Issues/Opinion/EncryptionAnonymityFollowUpReport.pdf>. [↑](#footnote-ref-16)
17. Jon Haggerty & Arthur Rizer, Australia's Encryption Law Deals a Serious Blow to Privacy and Security, The National Interest (19 December 2018), <https://nationalinterest.org/feature/australia%E2%80%99s-encryption-law-deals-serious-blow-privacy-and-security-39212>. [↑](#footnote-ref-17)
18. Harold Abelson et al., Keys under doormats: mandating insecurity by requiring government access to all data and communications, Journal of Cybersecurity (17 November 2015), <https://doi.org/10.1093/cybsec/tyv009>. [↑](#footnote-ref-18)
19. Avi Asher-Schapiro, YouTube and Facebook are Removing Evidence of Atrocities, Jeopardizing Cases Against War Criminals, The Intercept (2 November 2017), <https://theintercept.com/2017/11/02/war-crimes-youtube-facebook-syria-rohingya/>. See also, Malachy Browne, YouTube Removes Videos Showing Atrocities in Syria, New York Times (22 August 2017), <https://www.nytimes.com/2017/08/22/world/middleeast/syria-youtube-videos-isis.html>. [↑](#footnote-ref-19)
20. “These are the Crimes we are Fleeing”: Justice for Syria in Swedish and German Courts, Human Rights Watch Report (October 2017), <https://www.hrw.org/report/2017/10/03/these-are-crimes-we-are-fleeing/justice-syria-swedish-and-german-courts>. [↑](#footnote-ref-20)
21. For guidance specifically on countering violent extremism, *see* Raman Jit Singh Chima, Access Now Position Paper: A Digital Rights Approach to Proposals for Preventing or Countering Violent Extremism Online, Access Now (November 2016), <https://www.accessnow.org/cms/assets/uploads/2016/10/CVE-online-10.27.pdf>. [↑](#footnote-ref-21)