FROM INTERNET REFERRAL UNITS TO INTERNATIONAL AGREEMENTS: CENSORSHIP OF THE INTERNET BY THE UK AND EU

Brian Chang*

TABLE OF CONTENTS

---

* Brian Chang is a Research Assistant with the University of Oxford's Parliaments, Rule of Law and Human Rights research project. He graduated from the University of Oxford with a B.A. (Hons.) in Jurisprudence and received an LL.M. (with Highest Honors) in International and Comparative Law from The George Washington University Law School, where he was also a Thomas Buergenthal Scholar. He would like to thank Professors Dawn Nunziato and Arturo Carrillo for their help and encouragement, Cynthia Wong and Emma Llanso for inspiring this topic, Alexandra Bornstein and the rest of the editors at Columbia for their tireless editorial assistance, and his family and partner for their never-ending support. The views presented here and any errors are the author's alone and should not be attributed to any of the persons or organizations he has had the good fortune of working with.

I. INTERNET REFERRAL UNITS

A. Setting the Scene: UK and EU Law and Policy Toward Censorship of Extremist Content Online

Imagine an Internet that is proactively monitored for "illegal," "abusive," "harmful," or "offensive" content, and such content is prevented from being uploaded, filtered, or taken down as soon as it is detected. Because of threats of litigation and liability, loss of government advertising revenue, and fear of having their services blocked, information and communications technology (ICT) companies err on the side of caution and over-censor content, removing all content that governments tell them is illegal, whether it is "unlawful terrorist content" in the UK and Tajikistan, "Gulenist" content in Turkey, "separatist" content in China, "anti-monarchy" content in Thailand and Morocco, or "fake news" in Germany, Russia, or the United States.

While this scenario may seem farfetched, it is actually becoming more of a reality with every passing day. Governments around the world began this movement by compelling ICT companies to remove illegally uploaded copyrighted material and online child sexual abuse content. The present focus of government efforts to limit online content is on terrorist or extremist content, hate speech, as well

as "fake news"[1] and online abuse. However, national governments have now begun to push ICT companies to remove all illegal, or simply offensive, content through the use of automated processes or filters. In Europe, ICT companies such as Facebook, Google, and Twitter are being increasingly pressured through threats of criminal litigation (in Germany),[2] the enactment of legislation imposing liability on ICT companies (by France, Germany, the United Kingdom, and the European Union),[3] loss of advertising revenue (in the United Kingdom), and public denouncement of ICT companies as being "shameful" and "completely irresponsible" (by UK politicians).[4] The ICT companies have responded by stepping up their efforts to remove illegal content,[5] and in 2016, they agreed on a Code of Conduct with the European Union, pledging "to have in place clear and effective processes to review notifications regarding illegal hate speech . . . [and] to review the majority of valid notifications for removal of illegal hate speech in less than 24 hours and remove or disable access to such

---

1. Guy Chazan, *Germany cracks down on social media over fake news*, FIN. TIMES (Mar. 14, 2017), https://www.ft.com/content/c10aa4f8-08a5-11e7-97d1-5e720a26771b ("The German government has presented a draft law that would impose fines of up to €50m on social networks that fail to delete hate speech or fake news . . . .").

2. *German Facebook boss to be investigated for 'ignoring racist posts,'* GUARDIAN (Nov. 10, 2015), https://www.theguardian.com/technology/2015/nov/10/german-facebook-boss-investigated-hamburg-prosecutors-hate-speech ("Hamburg prosecutors say managing director may be held responsible for social platform's alleged failure to remove hate speech . . . .").

3. Chazan, *supra* note 1; Jessica Elgot, *May and Macron plan joint crackdown on online terror*, GUARDIAN (Jun. 12, 2017) https://www.theguardian.com/politics/2017/jun/12/may-macron-online-terror-radicalisation; Arthur Beesley, *Brussels urges US social media sites to act swiftly on hate posts*, FIN. TIMES (Dec. 4, 2016), https://www.ft.com/content/b3163cca-ba32-11e6-8b45-b8b81dd5d080 (reporting that EU Justice Commissioner Vera Jourová stated, "If Facebook, YouTube, Twitter and Microsoft want to convince me and the ministers that the non-legislative approach can work, they will have to act quickly and make a strong effort in the coming months").

4. HOME AFFAIRS COMM., HATE CRIME: ABUSE, HATE AND EXTREMISM ONLINE, 2016–7, HC 609, ¶¶ 25, 36 (UK).

5. *Id.* ¶ 19; Letter from Peter Barron, Vice President of Communications and Public Affairs, Google EMEA, to Chair, Home Affairs Comm. (Mar. 30, 2017), http://data.parliament.uk/WrittenEvidence/CommitteeEvidence.svc/EvidenceDocument/Home%20Affairs/Hate%20crime%20and%20its%20violent%20consequences/written/49839.html ("We already have thousands of people working on trust and safety issues across the company and have invested hundreds of millions of pounds in tackling abuse of all kinds on our platforms.").

content."[6] Although EU Justice Commissioner Vera Jourová has said that she is not yet ready to promote EU-wide legislation similar to that being pursued in the United Kingdom, France, and Germany, she has continued to threaten legislation if ICT companies do not self-regulate to her satisfaction by May of 2018.[7] In the meantime, the EU Commission has released "guidelines and principles" demanding that online platforms increase the proactive prevention, detection, and removal of "illegal content," including not only material that constitutes incitement to terrorism, illegal hate speech, or child sexual abuse, but also material that relates to "trafficking in human beings[,] . . . violations of intellectual property rights, product safety rules, illegal commercial practices online, or online activities of a defamatory nature."[8] Establishing such a broad range of removable content demonstrates the potentially expansive reach of legislation initially targeted at hate speech and terrorist content online.

UK Prime Minister Theresa May's response to the recent terror attacks in her country has been to blame ICT companies for allowing terrorist ideology the "safe space it needs to breed" and to adopt a policy of "work[ing] with allied, democratic governments to reach international agreements that regulate cyberspace to prevent the spread of extremism and terrorist planning."[9] She had previously spearheaded an international agreement at the 2017 G7 summit in Taormina,[10] which pressured ICT companies into forming a "Global

---

6. European Commission on Code of Conduct on Countering Illegal Hate Speech Online, (May 31, 2016), http://ec.europa.eu/justice/fundamental-rights/files/hate_speech_code_of_conduct_en.pdf.

7. Daniel Boffey, *EU justice commissioner resists calls for legislation on online hate speech*, GUARDIAN (Sept. 28, 2017), https://www.theguardian.com/uk-news/2017/sep/28/eu-justice-commissioner-resists-calls-for-legislation-on-online-hate-speech.

8. *Tackling Illegal Content Online: Towards an enhanced responsibility of online platforms*, at 2, 6, COM (2017) 555 final (Sept. 28, 2017).

9. *PM statement following London terror attack*, GOV.UK (June 4, 2017), https://www.gov.uk/government/speeches/pm-statement-following-london-terror-attack-4-june-2017.

10. Elizabeth Piper, *Britain's May gets support from G7 on fight against terrorism*, REUTERS (May 25, 2017), https://www.reuters.com/article/us-g7-summit-britain/britains-may-gets-support-from-g7-on-fight-against-terrorism-idUSKBN18L2T2; Press Release, European Council, G7 Taormina Statement on the Fight Against Terrorism and Violent Extremism (May 26, 2017), http://www.consilium.europa.eu/en/press/press-releases/2017/05/26/statement-fight-against-terrorism/.

Internet Forum to Counter Terrorism"[11] to develop and share new technology and tools to automatically identify and remove content promoting incitement to violence. In their joint statement preceding a bilateral visit in June 2017, the UK and France agreed to "work together to encourage corporations to do more and abide by their social responsibility to step up their efforts to remove harmful content from their networks, including exploring the possibility of creating a new legal liability for tech companies if they fail to remove unacceptable content."[12] Since then, the UK has succeeded in obtaining supportive statements by the Five Eyes (Australia, Canada, New Zealand, the United Kingdom, and the United States), [13] in a G20 Summit declaration,[14] and at the European Council meeting of EU Member States' heads of government.[15] At the seventy-second U.N. General Assembly, the United Kingdom, France, and Italy co-hosted an event on "Preventing Terrorist Use of the Internet," at which the three nations called for tech companies "to develop solutions to remove material within 1 [to] 2 hours of upload, with the wider objective of preventing such material from being uploaded in the first place."[16]

---

11. *See Facebook, Microsoft, Twitter and YouTube Announce Formation of the Global Internet Forum to Counter Terrorism*, FACEBOOK (June 26, 2017), https://newsroom.fb.com/news/2017/06/global-internet-forum-to-counter-terrorism/.

12. Press Release, Prime Minister's Office, UK and France announce joint campaign to tackle online radicalisation (June 13, 2017), https://www.gov.uk/government/news/uk-and-france-announce-joint-campaign-to-tackle-online-radicalisation; *see* French-British Action Plan (June 13, 2017), https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/619333/french_british_action_plan_paris_13_june_2017.pdf.

13. Press Release, Home Office, Five Eye countries join Britain's call to remove terror content online (June 28, 2017), https://www.gov.uk/government/news/five-eye-countries-join-britains-call-to-remove-terror-content-online.

14. Theresa May, Prime Minister, G20 Summit July 2017: Prime Minister's press statement (July 8, 2017), https://www.gov.uk/government/speeches/g20-summit-july-2017-prime-ministers-press-statement; *see* Press Release, G20, The Hamburg G20 Leaders' Statement on Countering Terrorism (July 7, 2017), https://www.g20.org/Content/DE/_Anlagen/G7_G20/2017-g20-statement-antiterror-en.html.

15. Theresa May, Prime Minister, European Council June 2017: Prime Minister's press statement (June 23, 2017), https://www.gov.uk/government/speeches/g20-summit-july-2017-prime-ministers-press-statement; *see* Press Release, European Council, European Council Conclusions On Security and Defense, U.N. Press Release 403/17 (June 22, 2017).

16. *Statement by UK, France and Italy on the Leaders' Meeting on Preventing Terrorist Use of the Internet*, GOV.UK (Sept. 20, 2017), https://www.gov.uk/

The prospect of the United Kingdom and the European Union leading the world's liberal democracies into requiring automated censorship of "harmful" or "unacceptable" content raises a number of grave concerns. It opens the door for other countries to demand that ICT companies similarly censor content that they deem illegal, harmful, or otherwise objectionable without having to go through the normal court process. There remain no international definitions of "terrorism" or "extremism" and these concepts are frequently abused by authoritarian governments to censor their critics. Other concepts such as "fake news" and "hate speech" are similarly easy to abuse and ICT companies will be accused of hypocrisy if they resist parochial standards abroad while accepting similar standards in their home countries.

Against this background, this Article addresses the use and potential abuse of Internet Referral Units (IRUs), a novel counterterrorism response originated by the UK in 2010,[17] which has spawned copies in the EU and a growing number of countries including France,[18] Belgium,[19] and the Netherlands.[20] The UK helped to establish the EU IRU, and has been disseminating information about its Counter Terrorism IRU (CTIRU) at international counterterrorism fora [21] as part of its efforts to coordinate an

---

government/uploads/system/uploads/attachment_data/file/646510/preventing_terrorist_use_of_the_internet_statement_20_sept_2017.pdf.

17. HOME DEP'T, CONTEST: THE UNITED KINGDOM'S STRATEGY FOR COUNTERING TERRORISM, ANNUAL REPORT FOR 2015, 2016, Cm. 9310, at 24, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/539683/55469_Cm_9310_Web_Accessible_v0.11.pdf.

18. Press Release, Ministre de l'Intérieur, Rencontre avec les grands opérateurs de l'Internet (Apr. 22, 2015), https://www.interieur.gouv.fr/Archives/Archives-des-communiques-de-presse/2015-Communiques/Rencontre-avec-les-grands-operateurs-de-l-Internet; *Lutte contre la propagande terroriste: le Gouvernement mobilise les dirigeants des grands opérateurs de l'internet*, GOUVERNEMENT.FR (Apr. 23, 2015), http://www.gouvernement.fr/lutte-contre-la-propagande-terroriste-le-gouvernement-mobilise-les-dirigeants-d-internet.

19. CDC, *Clamp down on hate messages on the Internet*, FLANDERS NEWS (Jan. 4, 2017, 10:39 AM), http://deredactie.be/cm/vrtnieuws.english/News/1.2859442 ("Belgian federal police today possess a special unit, the Internet Referral Unit, screening hate messages on the internet.").

20. Door Arnout de Vries, *Internet Referral Unit Internet Police on social media*, SOCIAL MEDIA DNA (Jan. 28, 2017), http://socialmediadna.nl/internet-referral-unit-internetpolitie-op-social-media/.

21. *Baroness Shields opening speech at the Global Counter Terrorism Form*, GOV.UK (Jan. 25, 2017), https://www.gov.uk/government/speeches/baroness-shields-opening-speech-at-the-global-counter-terrorism-forum.

international response to terrorist propaganda. In February 2016, a few months before she became UK Prime Minister, then-Home Secretary Theresa May stated that:

> I would like to see the United States, Canada, New Zealand and Australia – Britain's Five Eyes Partners – taking the same approach in working with communications service providers to tackle this propaganda. We need other like-minded groups to come on board from all corners of the world to reduce the scope for terrorist groups to spew their hate online and to undermine their twisted narratives.[22]

The European Commission has become an enthusiastic convert since the establishment of the EU IRU, and is now calling on all EU Member States "to establish national Internal Referral Units,"[23] while the Global Counterterrorism Forum recently released a set of recommendations highlighting IRUs as one existing initiative from which governments should learn.[24] IRUs are a growing part of the global agenda on countering violent extremism (CVE) online, alongside counter-narratives, counter-messaging, and attempts to coerce and persuade ICT companies to do more to self-police online spaces—and they have yet to be thoroughly examined from a human rights or rule of law perspective in the sparse academic literature to date.[25]

While their precise *modus operandi* may differ, IRUs generally operate by referring content such as videos, Tweets, and posts, or user

---

22. *Home Secretary: International action needed to tackle terrorism*, GOV.UK (Feb. 16, 2016), https://www.gov.uk/government/speeches/home-secretary-international-action-needed-to-tackle-terrorism.

23. Eur. Comm'n Press Release IP/17/1789, Security Union: Commission accelerates measures to prevent radicalisation and the cyber threat (June 29, 2017).

24. GLOB. COUNTERTERRORISM FORUM, ZURICH-LONDON RECOMMENDATIONS ON PREVENTING AND COUNTERING VIOLENT EXTREMISM AND TERRORISM ONLINE 10 (2017), https://www.thegctf.org/Portals/1/Documents/Framework%20Documents/A/GCTF%20-%20Zurich-London%20Recommendations%20ENG.pdf?ver=2017-09-15-210859-467.

25. *See* Jan Ellermann, *Terror won't kill the privacy star – tackling terrorism propaganda online in a data protection compliant manner*, 17 ERA F. 555 (2016); Olivia N. Bebe, Securitising the internet: The making of an EU Internet Referral Unit at Europol (June 10, 2015) (unpublished Master's thesis, Leiden University) (on file with Leiden University Repository); Kilian Vieth, Europol Policing the Web: Internet Content & Counter-Radicalization – An Interpretive Policy Analysis Approach (Mar. 9, 2016) (unpublished Master's thesis, Freie Universität Berlin), https://cdn.netzpolitik.org/wp-upload/2017/08/MA_KilianVieth_EuropolPolicingtheWeb_finale.pdf.

accounts that produce such content, to ICT companies for review against their terms of service or community guidelines and standards, so that content or accounts found to violate these standards can be removed or suspended. These referrals may occur through regular content "flagging" tools on a website (some ICT companies, such as Google and YouTube, give IRUs a "trusted flagger" status),[26] direct contact with website owners, [27] or special dedicated reporting channels.[28] This Article will focus on these referrals, although it should be noted that IRUs may have other tasks (which will be elaborated below), and on the UK and EU IRUs.

IRUs raise a number of human rights and rule of law concerns, including their nontransparency, lack of due process safeguards and remediation mechanisms, and the possibility of discrimination in their operational focus. But the fundamental concern is that IRUs threaten freedom of expression online, including the right to receive information. IRU referrals presently focus on "terrorist" or "extremist" content, two words which are difficult to define and frequently abused by authoritarian governments to censor their critics,[29] but IRUs could easily be directed to focus on other illegal or objectionable content, such as "hate speech" or "fake news," based on the political exigencies of a democracy or the priorities of an autocratic regime. While individual referrals may take the form of voluntary requests, IRU referrals cannot be divorced from the broader context in which they are made. The threat of excessive intermediary liability (for content that is not produced or modified by the ICT companies), the potential that their service may become blocked, and other coercive pressures mean that ICT companies have found—and will continue to find—that it makes

26. HOME AFFAIRS COMM., ORAL EVIDENCE: HATE CRIME AND ITS VIOLENT CONSEQUENCES, 2016-17, HC 609, Q410 (UK) (question to Peter Barron, Vice President, Communications and Public Affairs, Google Europe, the Middle East and Africa).

27. *Scotland Yard Requests Terrorist Content Removal*, CRYPTOME (July 24, 2015), https://cryptome.org/2015/07/met-removal.htm.

28. HOME AFFAIRS COMM., *supra* note 26, Q619 (question to Nick Pickles, Senior Public Policy Manager for the United Kingdom and Israel, Twitter).

29. Courtney Radsch, *Input from the Committee to Protect Journalists to the Office of the High Commissioner for Human Rights Concerning Resolution 30/15 of the Human Rights Council on Human Rights and Preventing and Countering Violent Extremism*, COMM. TO PROTECT JOURNALISTS 3 (Mar. 18, 2016), https://cpj.org/campaigns/2016.03.18_CPJ_CVE_submission_OHCHR.pdf ("CPJ research shows that legislation related to extremism and terrorism are routinely abused by authoritarian governments to censor critical reporting and commentary.").

sense commercially to err on the side of over-censorship. This will be further discussed in Part III.E.(2). Moreover, because terms of service are vaguely worded and generally prohibit abusive or illegal content, IRU referrals may enable governments to bypass human rights law by turning ICT companies into their censors, whether the ICT companies engage in such censorship voluntarily or not.

These are not mere abstract concerns in Europe; they could easily materialize in EU or European Convention on Human Rights (ECHR) countries. Is it unthinkable that Hungary might establish an IRU that refers content for takedown if it contains references to "migrant smuggling," "extremist viewpoints against public officials," "fake news," or other "threats to national security," and threaten to impose liability on Internet intermediaries that refuse to comply with the requests of the IRU? Or that another EU Member State might institute a similar IRU after being taken over by politicians with authoritarian or fascist inclinations? Or that Turkey might establish an IRU that refers content about "Gulenists" and Kurdish "terrorists"?

None of these scenarios are far from existing precedent: the European Union added a mandate to its IRU to remove "content used by traffickers to attract migrants and refugees"[30] before the EU IRU had even been established, raising concerns about its potential unfair impact on refugees[31] and demonstrating the potential for IRUs to be abused (the new Europol Regulation contains thirty crimes that the remit of the EU IRU may be extended to cover).[32]

In 2012, a leaked copy of Facebook's guide for its contractors that moderate (and remove) content showed that they were asked to remove "all attacks on Ataturk," maps showing an independent Kurdistan as a distinct entity from Turkey, and posts supporting or depicting the Kurdistan Workers' Party (PKK) or its founder, Abdullah

---

30. EUROPOL, EU INTERNET REFERRAL UNIT, YEAR ONE REPORT, HIGHLIGHTS 3 (2016), https://www.europol.europa.eu/publications-documents/eu-internet-referral-unit-year-one-report-highlights.

31. Anna Sauerbrey, *Europol reform – but who polices the police?*, EURACTIV (Nov. 10, 2015), https://www.euractiv.com/section/justice-home-affairs/news/europol-reform-but-who-polices-the-police/ ("There is a very sensitive issue regarding who is affected when Europol deletes content that supports 'people smuggling.' Many refugees exchange views on Facebook regarding the best routes to take into Europe, as well as the obstacles they might face. Is it fair that such content could be earmarked for deletion?").

32. Regulation 2016/794 of May 11, 2016 on the European Union Agency for Law Enforcement Cooperation (Europol), Annex I, 2016 O.J. (L 135/53) [hereinafter, the Europol Regulation].

Ocalan, "unless clearly against PKK and/or Ocalan."[33] This illustrates how one government can export its censorship standards globally through ICT companies and how even the largest ICT companies may be willing to compromise their values to operate in authoritarian countries.

More recently, Facebook has censored posts and users calling attention to the ethnic cleansing of the Rohingya people in Burma/Myanmar, after acceding to a government request to ban posts "by or in support of" the Arakan Rohingya Salvation Army, a violent organization that has been characterized as a terrorist group by the Myanmar government but a freedom fighter group by its supporters.[34] The potential for censorship of "terrorist" or "extremist" content online by ICT companies is demonstrated by the overbroad censorship of many posts and the censorship of users cataloguing acts of brutal and disproportionate violence by the Myanmar military. This violence has been characterized as a "textbook example of ethnic cleansing" by the U.N. High Commissioner for Human Rights,[35] resulting in the fleeing of 500,000 civilians into Bangladesh in one month.[36]

How should the ECHR and EU systems respond to the concern that IRUs are setting a dangerous precedent of state-initiated, privately-enforced, and extra-legal censorship that could be abused to limit speech that is neither genuine incitement to violence nor terrorism? How can mature democracies resist other states' attempts to use IRUs to privatize and export their local censorship standards, given the global nature of content takedowns online? One proposal would be that states discontinue IRUs, give ICT companies strong domestic, intermediary immunity, and advocate for such immunity globally by, for example, insisting that ICT companies should only be required to take down content after receiving a court order. This would give states moral standing and enable the development of international

---

33 . *Facebook's Kurdish problem?*, AL JAZEERA (Aug. 13, 2013), http://stream.aljazeera.com/story/201308240040-0023000.

34. Julia Carrie Wong et al., *Facebook bans Rohingya group's posts as minority faces 'ethnic cleansing,'* GUARDIAN (Sep. 20, 2017), https://www.theguardian.com/technology/2017/sep/20/facebook-rohingya-muslims-myanmar.

35. *Myanmar Rohingya crackdown: 'A textbook example of ethnic cleansing,' says UN*, DEUTSCHE WELLE (Sept. 11, 2017), http://p.dw.com/p/2jhcf.

36. Albert Fox Cahn, *Facebook's Silencing of Refugees Reveals Dangers of Censorship Technologies*, JUST SECURITY (Sept. 29, 2017), https://www.justsecurity.org/45495/rohingya-censorship-demands-greater-transparency-facebook/.

human rights law norms prohibiting extrajudicial censorship of content relating to the problematic categories of "terrorism" and "extremism." In order for ICT companies to gain the right to transfer data out of their jurisdictions, states could require ICT companies to develop terms of service and mechanisms for moderating content that respect human rights, including the rights to freedom of expression and privacy, while pursuing other social goods, such as combatting terrorist propaganda and hate speech, in a collaborative process with multi-stakeholder initiatives. States could then require that ICT companies be independently audited against their human rights policies and social obligations, in order to maintain their right to transfer data across borders.

However, because states within the ECHR and EU systems regard combatting terrorism, extremism, and hate speech as pressing political priorities, they are unlikely to accept total self-regulation by ICT companies as the most effective way of addressing these problems. They will also want to continue establishing and using IRUs in the short to medium term as a means of combatting terrorist propaganda online, because many ICT companies have not been able to self-police effectively, and even the largest ICT companies have yet to develop algorithms or artificial intelligence that could moderate content without human input. States may also want to use IRU referrals as a tool to sensitize ICT companies to what they regard as terrorist or extremist content.

As long as IRUs exist, there is a need to ground IRUs within international human rights law, the rule of law, and democratic safeguards so that their potential for abuse can be minimized. This Article suggests ways to do so, by strictly limiting their use to content that incites violence; analyzing them through the paradigms of international human rights law, European human rights law, and EU law; and developing standards and safeguards for their use.

The remainder of Part I describes the UK's Counter Terrorism Internet Referral Unit (CTIRU) and the EU IRU in detail, including their *modus operandi* and future directions, and briefly presents some of the criticisms that have been made of them. These IRUs will be analyzed from the legal perspective in the subsequent sections.

Part II examines IRUs using international human rights law and the U.N. Guiding Principles on Business and Human Rights, concluding with some Model Standards for IRUs and ICT companies.

Part III looks at IRUs through the framework of the European Convention on Human Rights (ECHR). While not rejecting IRUs

entirely, it finds that the ECHR requires safeguards for due process and freedom of expression. This Part also advocates limiting the permissible restrictions to hate speech and speech inciting violence, and proposes that a narrowly tailored, positive obligation be imposed on ICT companies to ensure adequate respect for freedom of expression. Finally, this Part argues that these ECHR safeguards must govern national IRUs in the United Kingdom, Belgium, and the Netherlands, and should provide inspiration for the governance of the EU IRU.

Part IV examines IRUs through the lenses of EU law, particularly the Europol Regulation that has recently entered into force, as well as the EU Charter of Fundamental Rights (CFR), and makes suggestions about how the EU IRU may be subject to effective safeguards and oversight.

Part V concludes by making a number of recommendations to policy makers and human rights actors on how to minimize the threat of abuse of IRUs, if IRUs are not eliminated altogether.

B. The UK Counter Terrorism Internet Referral Unit

(1) Background

The UK Terrorism Act of 2006[37] was passed in response to the July 2005 London bombings, as was U.N. Security Council Resolution 1624 (of September 14, 2005)[38] and the 2005 Council of Europe Convention on the Prevention of Terrorism,[39] which commits signatories to criminalize "public provocation to commit a terrorist offence" and "recruitment for terrorism."[40] Sections 1 and 2 of the Terrorism Act of 2006 contain criminal offenses prohibiting the intentional or reckless publication and dissemination of material that encourages, glorifies, or incites acts of terrorism. Sections 3 and 4 apply Sections 1 and 2 to "Internet activity" by creating a notice-and-takedown scheme that would allow UK police to notify a person of material in breach of Sections 1 and 2 and to hold them liable for

37. Terrorism Act 2006, c. 11 (UK), https://www.legislation.gov.uk/ukpga/2006/11/pdfs/ukpga_20060011_en.pdf?view=extent.

38. S.C. Res 1624 (Sept. 14, 2005).

39. *Council of Europe Convention on the Prevention of Terrorism*, COUNCIL EUR. (May 16, 2005), https://www.coe.int/en/web/conventions/full-list/-/conventions/rms/090000168008371c.

40. Clive Walker & Maura Conway, Online Terrorism and Online Laws 11 (Sept. 30, 2015), http://doras.dcu.ie/20841/.

endorsing the material if them failed to take it down within two working days without a reasonable excuse.

Stung by criticism that this restriction on freedom of expression "should engage a judicial officer at some stage so that the value of rights could be considered more explicitly than in the likely calculations of a commercial service provider"[41] (and likely having taken advice that the notice-and-liability provisions would be incompatible with the ECHR), the UK Government has never formally invoked the Section 3 notice power, preferring instead to remove potentially unlawful terrorist content through informal contact between the police and the Internet service provider. The CTIRU was eventually set up in 2010 for the purpose of "co-ordination and execution of voluntary and Section 3 take-down notices."[42]

The CTIRU had a slow start, with less than 2,000 pieces of content removed from the web each year in 2011 and 2012. However, it appears that the CTIRU's resourcing, and perhaps drive for results, drastically increased in 2013, with numbers rising to 17,541 that year, and 51,431 the next year.[43] This is likely to have occurred in response to the May 2013 murder of Fusilier Lee Rigby on the streets of London, which was filmed by bystanders and aired by The Sun and Independent Television News.[44] The December 2013 report of the Prime Minister's Extremism Taskforce issued in response to the Lee Rigby attack stated that "[e]xtremist propaganda is too widely available, particularly online, and has a direct impact on radicalising individuals. The poisonous messages of extremists must not be allowed to drown out the voices of the moderate majority."[45] The December 2013 report committed the Task Force to:

---

41. *Id.* at 10.

42. 717 Parl Deb HL (2010) col. WA168 (UK).

43. Metro. Police, *250,000th piece of online extremist / terrorist material to be removed*, MYNEWSDESK (Dec. 23, 2016, 14:08 GMT), http://www.mynewsdesk.com/uk/metpoliceuk/news/250000th-piece-of-online-extremist-slash-terrorist-material-to-be-removed-208698.

44. Gavriel Hollander, *Sun and ITV defend 'public interest' in showing Woolwich terror video Sky judged too 'distressing,'* PRESS GAZETTE (May 24, 2013), http://www.pressgazette.co.uk/sun-and-itv-defend-public-interest-broadcasting-woolwich-terror-video-which-sky-news-judged-too.

45. HER MAJESTY'S GOVERNMENT, TACKLING EXTREMISM IN THE UK: REPORT FROM THE PRIME MINISTER'S TASK FORCE ON TACKLING RADICALISATION AND EXTREMISM 3 (Dec. 2013), https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/263181/ETF_FINAL.pdf.

- work with Internet companies to restrict access to terrorist material online which is hosted overseas but illegal under UK law
- improve the process for public reporting of extremist content online
- work with the internet industry to help them in their continuing efforts to identify extremist content to include in family-friendly filters
- look at using existing powers to exclude from the UK those who post extremist material online who are based overseas.[46]

These objectives have largely been achieved with the aid of the CTIRU through the implementation of the CTIRU filtering list, which compiles material hosted beyond the jurisdictional reach of the CTIRU[47] and the creation of a red "STOP" button—STOP stands for "STOP Terrorists' and Extremists' Online Presence"—which enables members of the public to report online content they suspect may be of a violent, extremist, or terrorist nature directly to the CTIRU.[48]

Another expansion of the reach of the CTIRU occurred with the Counter Terrorism and Security Act of 2015, which imposed a "Prevent duty" on "specified authorit[ies]"—such as schools—to have "due regard to the need to prevent people from being drawn into terrorism."[49] The "Prevent duty guidance,"[50] issued by the Home Office, and the "Keeping children safe in education" guidance,[51] issued by the Department for Education, essentially require all schools and registered childcare providers to purchase web-filtering software that includes the CTIRU filtering list. The IRU's resourcing appears to have

---

46. *Id.*

47. *Response to FOI Request: Current status of terrorist internet filtering*, WHATDOTHEYKNOW (June 28, 2013), https://www.whatdotheyknow.com/request/160774/response/404100/attach/3/attachment.pdf.

48. Press Release, Nat'l Counter Terrorism Sec. Office, *STOP Terrorists' & Extremists' Online Presence* (Apr. 11, 2016), https://www.gov.uk/government/news/stop-terrorists-extremists-online-presence.

49. Counter-Terrorism and Security Act 2015, c. 6, § 1 (Gr. Brit.), http://www.legislation.gov.uk/ukpga/2015/6/pdfs/ukpga_20150006_en.pdf.

50. Home Office, *Prevent duty guidance*, GOV.UK (Mar. 23, 2016), https://www.gov.uk/government/publications/prevent-duty-guidance.

51. DEP'T FOR EDUC., KEEPING CHILDREN SAFE IN EDUCATION: STATUTORY GUIDANCE FOR SCHOOLS AND COLLEGES (2016), https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/550511/Keeping_children_safe_in_education.pdf.

been increased again as well, with the number of pieces of content removed increasing from 55,556 in 2015 to 121,151 in 2016.[52]

(2) Modus Operandi

The CTIRU refers content that it determines to contravene either UK terrorism legislation (alternatively termed "unlawful terrorist content" or "illegal material") [53] or company terms and conditions to ICT companies for removal.[54] The CTIRU may receive referrals from law enforcement partners or the public via its online reporting system (e.g., the red "STOP" button discussed above). However, the vast majority of referrals are generated by CTIRU officers searching the Internet for material.[55] The CTIRU also refers material to investigation teams nationally when it is identified that an offense may have been committed under the Terrorism Act of 2006 or other legislation.[56]

The police conduct periodic and regular campaigns, e.g., during every year's National Counter Terrorism Week, and engage in publicity efforts to request the public to report online material that they suspect to be extremist or terrorist by clicking on the red "STOP" button, found on each force's website.[57]

Examples of the illegal terrorist or extremist content placed on Internet sites, chat rooms, or other web-based forums that the CTIRU aims to combat include videos of violence with messages of "glorification" or praise for terrorists, as well as postings inciting people to commit acts of terrorism or violence.

The CTIRU also maintains a list of websites hosted outside of the United Kingdom, each of which is assessed for criminal liability under the provisions of the Terrorism Act of 2006 in the absence of any statutory defenses. This list is provided to web-filtering companies to block all such websites on the public estate, e.g., public libraries,[58] and in all schools and childcare facilities. The UK's major Internet service providers—BT, Virgin, Sky, and Talk Talk—have reportedly also

52. Metro. Police, *supra* note 43.
53. HOME DEP'T, *supra* note 17.
54. 772 Parl Deb HL (2016) col. 8 (UK).
55. Metro. Police, *supra* note 43.
56. *Id.*
57. Press Release, Nat'l Counter Terrorism Sec. Office, *supra* note 48.
58. *Response to FOI Request: Current status of terrorist internet filtering*, *supra* note 47.

agreed to ensure that terrorist and extremist material is captured by their filters to prevent children and young people from coming across radicalizing material, and many believe this to mean that the companies have agreed to incorporate the CTIRU list into their filters.[59]

(3) Aggregate Statistics and Lack of Transparency

Although the UK Government has regularly released aggregate statistics on the amount of content flagged for removal in response to parliamentary questions, the staffing and budget of the CTIRU are secret. The UK Government generally refuses to release this information, as well as any other operational details, stating that "for reasons of national security we do not publically disclose the detailed allocation of funding for counter terrorism by capability."[60]

The CTIRU recently published a news item, including detailed statistics and an infographic, to showcase its work after removing its 250,000th piece of content by Christmas 2016.[61] Prior to this, statistics had to be painstakingly assembled from responses to parliamentary questions.

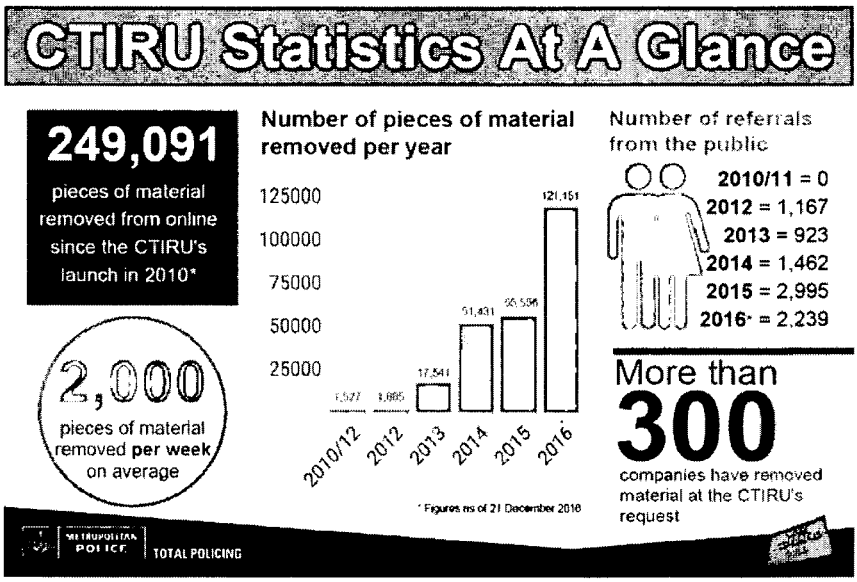
---

59. *See, e.g.*, *Counter Terrorism Internet Referral Unit*, OPEN RTS. GROUP WIKI (last visited Nov. 11, 2017), https://wiki.openrightsgroup.org/wiki/Counter_Terrorism_Internet_Referral_Unit#cite_ref-44; SWISS INST. OF COMPARATIVE LAW STUDY, COMPARATIVE STUDY ON BLOCKING, FILTERING AND TAKE-DOWN OF ILLEGAL INTERNET CONTENT, 753–78, https://rm.coe.int/1680685f10 ("In November 2014, it was announced that all major UK ISPs would be incorporating the blacklist into their adult content filters, preventing access to such websites where subscribers do not specifically opt out of such filtering.").

60. 17 Mar. 2016 Parl Deb (2016) 30893 (UK), http://www.parliament.uk/written-questions-answers-statements/written-question/commons/2016-03-14/30893.

61. Metro. Police, *supra* note 43.

Figure 1:



The number of pieces of content removed has increased drastically: from approximately sixty pieces per month in 2010,[62] to an average of over 4,500 pieces per month in 2015 and an average of 10,000 pieces per month in 2016.[63] In 2015, when the CTIRU was removing over 1,000 pieces of content a week, approximately 80% of this content was Syria- or Iraq-related and had been posted on multiple platforms.[64] The House of Lords has also recently disclosed that "[i]ndustry cooperation with CTIRU has significantly improved, leading to faster and more consistent removal of referred content, and they have established relationships with over 300 Communication Service Providers of differing sizes."[65]

62. 770 Parl Deb HL (2016) col. 6 (UK).
63. Metro. Police, *supra* note 43.
64. Nat'l Counter Terrorism Security Office, *Guidance: Online radicalisation*, GOV.UK (Nov. 26, 2015), https://www.gov.uk/government/publications/online-radicalisation/online-radicalisation.
65. 778 Parl Deb HL (2017) col. 19 (UK).

(4) Future Directions?

The UK Government's 2015 Annual Report on CONTEST, its strategy for countering terrorism, states: "[a]s set out in the Counter-Extremism Strategy, we believe that these companies should set up a body in the style of the Internet Watch Foundation to monitor and flag to industry occurrences of terrorism and extremism on their networks. We continue to work with them on this."[66] This suggests that the 2015–2017 Government saw self-regulation as the long-term solution and the IRU as a short- to medium-term solution. In this vein, Baroness Joanna Shields, who was appointed the Prime Minister's Special Representative on Internet Crime and Harms in 2016 and tasked with coordinating an international approach on Internet safety and security,[67] stressed the need for the Internet industry to match the efforts made by the Government to tackle online extremism. Baroness Shields noted the importance of companies making "this agenda [their] own" by investing in improving technological solutions that "automate the identification and removal of dangerous extremist content" and effectively combat the technological devices that support the propaganda software used by terrorists.[68]

By contrast, the Home Affairs Committee's Report on Radicalisation recommended that the CTIRU be scaled up and that "[r]epresentatives of all the relevant agencies including the Home Office, MI5 and major technology companies . . . be co-located within CTIRU."[69] The report also recommended maintaining cooperation with

66. HOME DEP'T, *supra* note 17, ¶ 2.33.

67. Home Office, *Baroness Shields appointed as the PM's Special Representative on Internet Crime and Harms and becomes solely a Home Office minister*, GOV.UK (Dec. 16, 2016), https://www.gov.uk/government/news/baroness-shields-appointed-as-the-prime-ministers-special-representative-on-internet-crime-and-harms-and-becomes-solely-a-home-office-minister.

68. Baroness Joanna Shields, *Reclaiming Technology for the Future*, HUFFINGTON POST: THE BLOG (Jan. 18, 2016), http://www.huffingtonpost.com/baroness-joanna-shields/reclaiming-technology-for_b_9008294.html (adapted from speech delivered at the DLD 'Next Next' Conference in January 2016); Baroness Joanna Shields, *How the Threat of Violent Extremism Manifests Online*, HUFFINGTON POST: THE BLOG (June 17, 2016), https://www.huffingtonpost.com/baroness-joanna-shields/how-the-threat-of-violent-extremism-manifests-online_b_10528798.html (speech delivered at the Zeitgeist Minds conference in May 2016).

69. HOME AFFAIRS COMM., RADICALISATION: THE COUNTER-NARRATIVE AND IDENTIFYING THE TIPPING POINT, 2016–17, HC 135, at 11, 33 (UK), https://www.publications.parliament.uk/pa/cm201617/cmselect/cmhaff/135/135.pdf.

the European Union after Brexit.[70] The Government did not manage to publish a written response before Parliament was unexpectedly dissolved for the early elections of June 2017. In the report, the Home Affairs Committee was critical of social media companies for "consciously failing to combat the use of their sites to promote terrorism and killings," having "teams of only a few hundred employees to monitor networks of billions of accounts," and "hiding behind their supranational legal status to pass the parcel of responsibility."[71] It recommended that the Government take action to "enforce its own measures to ensure that the large technology companies operating in this country are required to cooperate with CTIRU promptly and fully." The report also proposed a required that "the companies . . . be transparent about their actions on online extremism; instead of the piecemeal approach we currently have, they should all publish quarterly statistics showing how many sites and accounts they have taken down and for what reason."[72] More recently, members of the Home Affairs Committee have started pressuring social media companies to foot the bill for the CTIRU, although it remains to be seen whether this will materialize into a formal Government demand.[73]

C. The EU IRU

(1) Background

The EU IRU was inspired by and remains directly supported by the UK CTIRU, which has a member of staff seconded to Europol to act as a liaison between the two IRUs. On March 23, 2015, after the *Charlie Hebdo* attacks in Paris, the Justice and Home Affairs Council of the European Union agreed to establish the EU IRU by July 1, 2015, with the following core tasks:

- Coordinate and share the identification tasks (flagging) of terrorist and violent extremist online content with relevant partners;
- Carry out and support referrals quickly, efficiently and effectively, in close cooperation with the industry;

---

70. *Id.* at 11.
71. *Id.* at 34–35.
72. *Id.* at 35.
73. HOME AFFAIRS COMM., *supra* note 26, Q491–Q495 (questions to James Berry MP).

- Support competent authorities, by providing strategic and operational analysis;
- Act as a European Centre for Excellence for the above tasks.[74]

In May 2016, the European Parliament passed Regulation 2016/794 on the European Union Agency for Law Enforcement Cooperation (Europol) (Europol Regulation), which came into effect on May 1, 2017.[75] This regulation now governs the EU IRU and Europol as a whole. It contains legal limits and data protection safeguards, including a right to complain to the EU Data Protection Supervisor by data subjects, as well as provisions for parliamentary scrutiny. However, the Europol Regulation has been criticized for not doing enough to set the terms for the IRU or establish transparency or accountability.[76] The Europol Regulation and these criticisms of the regulation are examined in more detail in Part IV.

(2) Modus Operandi

The following are the EU IRU's main strategic goals through the end of 2017:

1. Effectively countering online radicalisation and recruitment efforts by terrorists, by strengthening an adaptive referral capability and mapping and influencing online terrorist propaganda networks;
2. Providing a core Internet Investigation Support Capability based on operational support and strategic analysis;
3. Striving to become a European Centre of Excellence, by strategically enhancing partnerships with cooperation partners and investing resources in Research &

74. *The European Union Internet Referral Unit at Europol*, OPEN ACCESS GOV'T (Feb. 2, 2016), https://www.openaccessgovernment.org/european-union-internet-referral-unit-europol-2/24158/.

75. Europol Regulation, *supra* note 32.

76. *See, e.g.*, Lucie Krahulcova, *Europol's Internet Referral Unit risks harming rights and feeding extremism*, ACCESS NOW (June 17, 2016, 6:11 AM), https://www.accessnow.org/europols-internet-referral-unit-risks-harming-rights-isolating-extremists/ (arguing that the IRU allows private third parties to operate outside the rule of law); Matthias Monroy, *Oversight of the new Europol regulation likely to remain superficial,* EDRI (July 12 2016), https://edri.org/oversight-new-europol-regulation-likely-remain-superficial/ (arguing that the Europol Regulation does not provide for meaningful parliamentary oversight of Europol).

> Development (R&D) Coordination as an Innovation Hub for Europol and the EU MS in the field of counter terrorism.[77]

At present, all content is flagged after a manual assessment by an EU IRU human analyst or translator.[78] YouTube has offered to give the EU IRU "trusted flagger status," so that it can upload referrals in batches without needing to fill out an online form for each individual social media profile it wishes to refer. Relevant content may be gathered by the EU IRU itself through open source collection or referred by EU Member States or third parties with their own IRUs or open-source scanning capabilities, though third parties must have an operational cooperation agreement with Europol.[79] Europol takes pains to emphasize that the final decision to remove content is a voluntary activity carried out by the concerned service providers, in accordance with their own terms and conditions.[80] But in reality, these seemingly voluntary requests occur within the broader context of coercive pressures on ICT companies to "do more," as mentioned above. There has also been discussion about automating the process to rely on algorithms to prevent re-uploading,and creating databases to which the private sector would voluntarily contribute content, which would then be flagged for review or taken down by other ICT companies.[81]

"Terrorist content" for referral is evaluated against the offenses set out in Council Framework Decision 2008/919/JHA[82] on combatting terrorism, which include "public provocation to commit a terrorist offence, recruitment for terrorism, [and] training for

---

77. EUROPOL, EU INTERNET REFERRAL UNIT: YEAR ONE REPORT HIGHLIGHTS 2, https://www.europol.europa.eu/sites/default/files/documents/eu_iru_1_year_report_highlights.pdf (last visited Nov. 15, 2017).

78. *Some Recent Trends in the Use of the Internet/ICTs for Terrorist Purposes – Part II*, VOX-POL (Nov. 9, 2016), http://www.voxpol.eu/recent-trends-use-internetict-terrorist-purposes-part-ii/.

79. Ellermann, *supra* note 25, at 563.

80. *See, e.g., id.* at 567 (noting that "final decision" rests with service provider); Press Release, Europol, Europol Coordinates EU-Wide Hit Against Online Terrorist Propaganda (May 2, 2017), https://www.europol.europa.eu/newsroom/news/europol-coordinates-eu-wide-hit-against-online-terrorist-propaganda (noting the same).

81. *See infra* section II.B.5 (describing a Franco-German proposal).

82. Council Framework Decision (EU) No. 2008/919/JHA of 28 Nov. 2008, 2008 O.J. (L300) 21–23, http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32008F0919&from=EN.

terrorism."[83] One important, self-imposed constraint is that the EU IRU uses the Consolidated U.N. Security Council Sanctions List as a basis for deciding what content to refer.[84] This means that its censorship of content should generally be limited to internationally-recognized terrorist groups,[85] although the sanctions listing process is far from perfect.[86] The statements surrounding the creation[87] and operation[88] of the EU IRU also tend to focus on the threat posed by Al Qaeda and Daesh content online; the presentation of the head of the EU IRU to the U.N. Counter-Terrorism Committee Executive Directorate (UNCTED) suggests that their operations tend to focus on Daesh propaganda.[89] Commendably, Europol's Data Protection Office has fostered internal discussion about the difference between promoting terrorism and violence and raising awareness of terrorism and violence, and it has appeared to reach the position that content that "raises awareness" should not be censored even if it includes graphics and violence and outrage or terrorism propaganda. However, one of Europol's own data protection officers, writing in a personal

---

83. Ellermann, *supra* note 25, at 564. These offenses are similar to those set out in the 2005 Council of Europe Convention.

84. CAMINO KAVANAGH ET AL., ICT4PEACE FOUND. & U.N. COUNTER-TERRORISM COMM. EXEC. DIRECTORATE, PRIVATE SECTOR ENGAGEMENT IN RESPONDING TO THE USE OF THE INTERNET AND ICT FOR TERRORIST PURPOSES: STRENGTHENING DIALOGUE AND BUILDING TRUST 7 (2016), https://www.un.org/sc/ctc/wp-content/uploads/2016/12/Private-Sector-Engagement-in-Responding-to-the-Use-of-the-Internet-and-ICT-for-Terrorist-Purposes.pdf.

85. *See* U.N. SEC. COUNCIL, *Consolidated United Nations Security Council Sanctions List*, UN.ORG, https://www.un.org/sc/suborg/en/sanctions/un-sc-consolidated-list (last visited May 4, 2017).

86. *See e.g.*, GROUP OF LIKE-MINDED STATES ON TARGETED SANCTIONS, PROPOSAL TO THE UNITED NATIONS SECURITY COUNCIL: FAIR AND CLEAR PROCEDURES FOR A MORE EFFECTIVE U.N. SANCTIONS SYSTEM 2 (2015), http://www.new-york-un.diplo.de/contentblob/4662362/Daten/6041666/151112fairclearproceduressanctions.pdf (proposing changes to bring the listing process in line with due process and international law norms).

87. Vikram Dodd, *Europol web unit to hunt extremists behind Isis social media propaganda*, GUARDIAN (June 21, 2015), https://www.theguardian.com/world/2015/jun/21/europol-internet-unit-track-down-extremists-isis-social-media-propaganda.

88. *See, e.g.*, EUROPOL, *supra* note 30 (noting that the EU IRU uses "unique linguistic capabilities" to address content issued by Al Qaeda and Daesh).

89. Stephane Duguin, Head of EU IRU, Presentation at U.N. Counter-Terrorism Committee Executive Directorate meeting on "Preventing the Exploitation of Information and Communications Technologies for Terrorist purposes, while Respecting Human Rights and Fundamental Freedoms" (Dec. 1, 2016).

capacity, acknowledged that the dividing lines may "sometimes be difficult to draw" and that "the nature of terrorist and violent extremist online content is not yet internationally agreed to the same extent as . . . child abuse material."[90] This discussion also belies the fact that, within the EU IRU, there is no internal section focused on protecting freedom of expression, like the Data Protection Office, which is tasked with ensuring respect for the value of data protection while fighting serious crime and terrorism.

There is presently little transparency regarding what kind of criteria and limits the EU IRU uses for evaluating "content used by smuggling networks to attract migrants and refugees," which makes it impossible to independently assess what content the EU IRU is referring for takedown.[91] The EU IRU claims that it is referring "[c]ontent advertising smuggling services for migrants and refugees,"[92] but it has not provided any representative samples of the content that it refers.

Because of the nature and strength of the European Union's data protection rules, before the EU IRU can refer content that contains personal data,[93] which in practice means most social media content, to ICT companies, there must be a case-by-case evaluation of whether takedown is strictly necessary, subject to restrictions stipulated by the data owners and Europol itself. While the Europol Regulation requires this evaluation to take into account the fundamental rights and freedom of the data subjects concerned, it is unclear how much weight is given to freedom of expression in this analysis and the evaluation does not take into account the public's right to receive information. This is discussed in further detail in Part IV.A *infra*.

Although the EU IRU's most visible activities involve flagging content for takedown and assisting in operational investigations, these are not its only functions. At its launch, Europol director Rob Wainwright stated that the EU IRU "would monitor social media

---

90. Ellermann, *supra* note 25, at 564.

91. EUROPOL, *supra* note 30.

92. *Id.*

93. Article 2 of the EU Data Protection Directive defines "personal data" as "any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity." Council Directive 95/46, art. 2, 1995 O.J. (L 281) 31 (EC).

output to identify people who might be vulnerable and those preying on them."[94] Wainright also said "that the police team would be working with social media companies to identify the most important accounts operating in a range of languages that are 'underpinning what Isis is [sic] doing,'" aiming to "identify the ringleaders online," with the hope that results would be passed back to member states to take action against the individuals running the accounts.[95]

The EU IRU has also incorporated the Europol "Check-the-Web" service, an electronic reference library of terrorist online propaganda containing original statements, publications, videos, and audio produced by terrorist groups or their supporters. Competent authorities of EU Member States and third parties with operational agreements can access this content and its analysis.[96]

The EU IRU also provides Member States with operational support in their Internet investigation activities that seek to counter online radicalization and recruitment by terrorists[97] and periodically conducts intensive cooperative actions,[98] teaming with EU Member States and third parties with operational agreements to target extremist content and accounts and refer them to ICT companies. Notably, in its one of its 2017 "intensive cooperative action[s]," the EU IRU worked together with "colleagues" from the United States, which raises interesting questions about whether these collaborating U.S. agencies may have violated the First Amendment rights of account holders.[99]

The IRU is also a key player in the EU Internet Forum, engaging with online service companies to promote "self-regulation" activities by the online industry.[100] The EU IRU has led the establishment of an European Counter-Terrorism Centre (ECTC) Advisory Group on the abuse of online communication by terrorist groups for propaganda purposes and online recruitment, in order to step up the strategic cooperation and exchange of non-operational best

94. EUROPOL, *supra* note 30.
95. Dodd, *supra* note 87.
96. EUROPOL, *supra* note 30.
97. *Id.*
98. *See* Press Release, Europol, Europol Coordinates Joint Action Days to Flag Online Terrorist Content (Feb. 27, 2017), https://www.europol.europa.eu/newsroom/news/europol-coordinates-joint-action-days-to-flag-online-terrorist-content.
99. *Id.*
100. KAVANAGH ET AL., *supra* note 84, at 8.

practices with third parties. These parties have recognized expertise in the area of terrorist exploitation of online communications.[101] Given its membership in the EU Internet Forum and influence over ICT companies, the EU IRU was likely also a leading player in discussions on the ICT companies' voluntary code of conduct on hate speech[102] and the Google, Facebook, Twitter, and Microsoft partnership to prevent the uploading of extremist content online that had previously been banned through "hashes" or unique digital fingerprints,[103] similar to the proposed EU "Joint Referral Platform."[104]

(3) Mission Creep

According to Europol's Year One Report on the EU IRU, the European Council expressed concern in April 2015 regarding the waves of migration through the Mediterranean Sea. The European Council called for "the EU IRU to expand its Open Source and Internet monitoring activities, in order to contribute to the disruption of illegal immigrant smuggling networks, by detecting and requesting removal of Internet content used by traffickers to attract migrants and refugees."[105] The IRU has three full-time staff members dedicated to preventing illegal immigration[106] and the EU IRU's Year One Report says that the EU IRU has processed 122 accounts linked to illegal immigration upon request from the European Migrant Smuggling Centre (EMSC).[107]

Such expansion has been criticized for its potentially unfair impact on refugees, the magnitude of which cannot be determined without greater transparency:

---

101. EUROPOL, *supra* note 30, at 8–9.

102. European Commission, *supra* note 6.

103. *See* Rob Price & REUTERS, *Google, Facebook, Microsoft and Twitter are working together to tackle terrorist propaganda*, BUSINESS INSIDER UK (Dec. 6, 2016), http://uk.businessinsider.com/r-web-giants-to-cooperate-on-removal-of-extremist-content-2016-12?utm_source=feedburner&%3Butm_medium=referral&utm_medium=feed&utm_campaign=Feed%3A+businessinsider+%28Business+Insider%29&r=US&IR=T.

104. *See* Section I.C.5 *infra*.

105. EUROPOL, *supra* note 30.

106. *See* Europol, *Rep. to COSI on "enhancing counter terrorism capabilities at EU level: European Counter Terrorism Centre (ECTC) at Europol and counter terrorism related information sharing"* (Nov. 17, 2015), http://www.statewatch.org/news/2015/nov/eu-council-europol-ECTC-14244-15.pdf.

107. EUROPOL, *supra* note 30, at 5.

> There is a very sensitive issue regarding who is affected when Europol deletes content that supports 'people smuggling'. Many refugees exchange views on Facebook regarding the best routes to take into Europe, as well as the obstacles they might face. Is it fair that such content could be earmarked for deletion?[108]

It may also be said that this demonstrates the potential for IRUs to be abused, with political leaders initially establishing them to deal with child pornography[109] and to counter violent extremism, but gradually adding new political directives based on the exigencies of the day. Left unchecked, the remit of the EU IRU could eventually be expanded to cover the thirty crimes on which Europol is empowered to act.

(4) Statistics

In its Year One report, the EU IRU stated that it had assessed and flagged the content presented in the first two columns of Table 1, below.[110] On December 1, 2016, the head of the IRU gave a presentation to the U.N. Counter-Terrorism Committee Executive Directorate, where he presented some higher figures, contained in the below slide (Figure 2) and in the third column of Table 1, although it is unclear when these numbers were recorded.[111]

Assuming that the December 2016 numbers[112] reflect the position on or just before December 1, this would indicate a slight increase in the number of content assessments per month in the period from July to December 2016, as compared with the period from November 2015 to July 2016. There were approximately 1,250 pieces of content being assessed per month between November 2015 and July 2016 and approximately 1,340 pieces of content being assessed per month between July 1, 2016 and December 1, 2016.

---

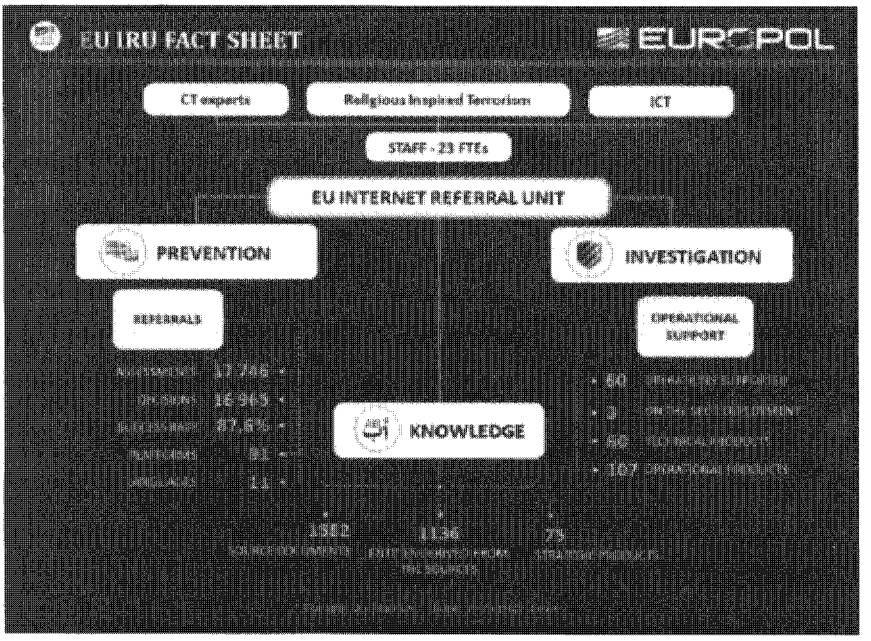
108. Sauerbrey, *supra* note 31.
109. *Cf.* INTERNET WATCH FOUND., https://www.iwf.org.uk/ (last visited May 4, 2017).
110. *Id.*
111. Duguin, *supra* note 89.
112. *Id.*

Table 1:

	11/5/2015	**7/1/2016**	**Numbers presented on 12/1/2016**
Total content assessed	1,079	11,050	17,746
Proposals for referral	690	9,787	16,695
Content removed by online service providers	511	8,949	[14,625][113]
Success rate	74%	91.40%	87.6%
Platforms identified	9	70	91
Platforms referred to	7	31	-

113. This number is calculated based on the "Success Rate" (87.6%) and number of "Decisions" (16,965) provided in the slide—87.6% of 16,965 is approximately 14,625.

Figure 2:



(5) Future Directions

The French and Germany Interior Ministers met in August 2016 and issued a joint declaration stating that:

> We want to strengthen the Internet Referral Unit at Europol by setting up an EU center to combat terrorism and radicalization on the Internet. One of the tasks of this center will include the detection of related content and their exchange as well as the prevention of a re-upload of already identified material [by use of an automated upload filter] . . . We also want to tighten the host-provider privilege in the sense of a "product liability" in the case of abuse for terrorist propaganda.[114]

This builds on European Commission proposals to establish a Joint Referral Platform together with ICT companies, which has been

114. Markus Reuter, *Summer of inner security: what the interior ministers of France and Germany really demand*, NETZPOLITIK.ORG (Aug. 24, 2016), https://netzpolitik.org/2016/sommer-der-inneren-sicherheit-was-die-innenminister-von-frankreich-und-deutschland-wirklich-fordern/.

described as a "de facto revival of the 'Clean IT' project" and which aims to prevent the unnoticed re-upload of previously removed material through mandatory monitoring of every single file that every individual in Europe uploads to the Internet, relying on content recognition by robust hashing.[115]

A proposed EU Directive on combating terrorism, which was introduced by the Commission and passed by LIBE, the European Parliament Committee on Civil Liberties, Justice and Home Affairs, on July 7, 2016, criminalized preparatory acts, including "[p]ublic incitement or praise of terrorism: public incitement to terrorism such as glorifying or justifying suicide bombers or disseminating messages or images on or off-line as a way to gather support for a terrorist cause or gain publicity for example by disseminating videos of assassinations."[116] It also imposed an obligation on EU Member States to take measures to ensure the prompt removal of illegal content hosted on their territory that constitutes public incitement to commit a terrorist offense. This may result in more, if not all, EU Member States establishing IRUs, amongst other actions.[117]

D. Criticisms of IRUs

The use of IRUs have been noted with concern by the U.N. Special Rapporteurs for Freedom of Expression and Countering Terrorism while Protecting Human Rights.[118] A number of civil liberties groups, including Access Now,[119] the American Civil Liberties

---

115. EDRI, *Algorithms – censorship à la carte?* (July 12, 2016), https://edri.org/algorithms-censorship-a-la-carte/.

116. Press Release, European Parliament, Planning terrorist attacks must be made a crime, say civil liberties MEPs (July 5, 2016), http://www.europarl.europa.eu/news/en/press-room/20160620IPR32963/planning-terrorist-attacks-must-be-made-a-crime-say-civil-liberties-meps. The text as ultimately passed in an EU Council Directive on March 15, 2017 criminalizes "the glorification and justification of terrorism or the dissemination of messages or images online and offline, including those related to the victims of terrorism as a way to gather support for terrorist causes or to seriously intimidate the population." Council Directive 2017/514, 2017 O.J. (L 88) 6, 7.

117. Michael Plachta, *Current Developments in the Counter-terrorism Efforts of the European Union*, *in* INTERNATIONAL ENFORCEMENT LAW REPORTER 320 (2016). Again, this obligation was imposed in the text that was ultimately enacted. Council Directive 2017/514, 2017 O.J. (L 88) 6, 9.

118. *See* Section III.A.1 *infra*.

119. Krahulcova, *supra* note 76.

Union (ACLU),[120] the Centre for Democracy and Technology (CDT),[121] the Committee to Protect Journalists (CPJ), [122] European Digital Rights (EDRi),[123] and German civil rights activists[124] and journalists have also criticized their use.[125] These concerns note the lack of due process, transparency, oversight and accountability, and effective remediation mechanisms, as well as the overbreadth of terms of service, the international consequences of IRUs, and the counter-productiveness of censorship. Many of these criticisms have been made of the EU IRU, though they apply generally even after the adoption of the new Europol Regulation, and are briefly set out below. The Europol Regulation will be evaluated in Part IV.

Access Now has expressed concern that the modus operandi of IRUs in referring content for voluntary removal is:

> outside the rule of law on several grounds. First, illegal content is just that — illegal. If law enforcement encounters illegal activity, be it online or off, it is expected to proceed in dealing with that in a legal, rights-respecting manner. Second, relegating dealing with this illegal content to a third private party, and leaving analysis and prosecution to their discretion, is not just lazy, but extremely dangerous. Third, illegal content, if truly illegal, needs to be dealt with that way: with a court order and subsequent removal. The IRU's blatant circumvention of the rule of law is in direct violation of international human rights standards.[126]

EDRi has criticized the EU IRU referral process for being "non-transparent and [outside of] judicial oversight." Though judicial

---

120. Hugh Handyside, *Social Media Companies Should Decline the Government's Invitation to Join the National Security State*, ACLU (Jan. 12, 2016, 2:15 PM), https://www.aclu.org/blog/speak-freely/social-media-companies-should-decline-governments-invitation-join-national.

121. Scott Craig & Emma Llanso, *Pressuring Platforms to Censor Content is Wrong Approach to Combatting Terrorism*, CDT BLOG (Nov. 5, 2015), https://cdt.org/blog/pressuring-platforms-to-censor-content-is-wrong-approach-to-combatting-terrorism/.

122. Radsch, *supra* note 29.

123. *Europol: Non-transparent cooperation with IT companies*, EDRI (May 18, 2016), https://edri.org/europol-non-transparent-cooperation-with-it-companies/.

124. Monroy, *supra* note 76.

125. Anna Sauerbrey, *Europol reform – but who polices the police?*, EURACTIV (Nov. 10, 2015), https://www.euractiv.com/section/justice-home-affairs/news/europol-reform-but-who-polices-the-police/.

126. Krahulcova, *supra* note 76.

oversight may theoretically be possible, the lack of notice about referrals to affected persons makes challenging referral decisions incredibly difficult. It has also been noted that the new Europol Regulation has no transparency requirements to inform the public about any type of information exchange between Europol and third parties and that new European Regulation does not include the European Data Protection Supervisor's recommendation that Europol provide a minimal amount of transparency by making a list of its cooperation agreements with companies publicly available.[127]

CDT has criticized the reliance on private-censorship and the corresponding issues of overbroad terms of service and lack of due process, transparency, oversight, and accountability:

> Companies' privately developed Terms of Service and content policies are typically more restrictive, and often much more restrictive, than what governments may permissibly restrict under law. Further, these programs may not be clearly articulated in law; the specific procedures and processes are often not communicated transparently with the public, and there has not been an evidence-based showing that they are necessary or effective. Finally, they are not susceptible to normal processes of democratic governance and oversight. Overzealous efforts to pursue expedited, privatized removal of content risk undermining the rule of law and fundamental values of a democratic society.[128]

The ACLU has expressed concern about the international consequences of IRUs:

> [C]ontent that companies take down through [the IRU] process is inaccessible everywhere, meaning that a single government can try to use the process to impose its more restrictive speech standards on the rest of the world . . . . Lurking beneath these kinds of content restrictions is the perennial question of what constitutes terrorism or the promotion of

---

127. EDRI, *supra* note 123.

128. CTR. FOR DEMOCRACY & TECH., COMMENTS OF THE CENTER FOR DEMOCRACY & TECHNOLOGY: TO THE SPECIAL RAPPORTEUR ON THE PROMOTION AND PROTECTION OF THE RIGHT TO FREEDOM OF OPINION AND EXPRESSION IN THE CONSULTATION ON 'FREEDOM OF EXPRESSION AND THE PRIVATE SECTOR IN THE DIGITAL AGE' 4 (2016), https://cdt.org/files/2016/02/CDT-Comments-Consultation-on-freedom-of-expression-and-the-private-sector-in-the-digital-age.pdf.

> terrorism — a question which has no clear or consistent answer in U.S. or international law, and which inevitably is subject to politics or chauvinistic impulses, and even manipulation.[129]

The Committee to Protect Journalists (CPJ) has echoed this concern, stating that IRUs pose "grave threats to freedom of expression and the right to receive information. CPJ research shows that legislation related to extremism and terrorism are routinely abused by authoritarian governments to censor critical reporting and commentary" and that "blanket restrictions on content that advocates, supports or glorifies extremism are too easily abused . . . and they are likely to prompt private Internet companies compelled to implement them to err on the side of caution."[130]

Anna Sauerbrey, a German journalist, has expressed concern about giving Europol, a police agency, the power to decide what is "good and bad internet content" and that it may be able to involve itself in hate speech on Internet forums, despite its main remit being serious crimes. She noted that Andrej Hunko, a member of the German Bundestag, Germany's legislature, as well as the Parliamentary Assembly of the Council of Europe (PACE), asked whether the German interior ministry considers hate speech, racism, and xenophobia "violent expression—and therefore candidates for deletion by Europol—and got an answer that was essentially, "yes."[131]

Finally, Access Now has noted the potential for mass content referrals to be counterproductive, as they risk silencing voices seeking to respond to or counter violent extremist narratives: "Mass take-down initiatives that take place outside of legal process frustrate corporate transparency and are not likely to deter the cultivation of 'violent extremism', and in fact may encourage it, inflaming resistance and helping 'violent extremist' recruiters discredit platforms that might otherwise support online expression and debate."[132] In this vein, Emma Llansó, Director of the Center for Democracy & Technology (CDT)'s Free Expression Project has expressed a preference for counter-speech to censorship of speech, noting that it would take a lot of resources to

---

129. Handyside, *supra* note 120.

130. Radsch, *supra* note 29, at 4.

131. Sauerbrey, *supra* note 125.

132. RAMAN JIT SINGH CHIMA, ACCESS NOW POSITION PAPER: A DIGITAL RIGHTS APPROACH TO PROPOSALS FOR PREVENTING OR COUNTERING VIOLENT EXTREMISM ONLINE 10 (2016), https://www.accessnow.org/cms/assets/uploads/2016/10/CVE-online-10.27.pdf.

determine whether accounts are promoting a counter-narrative, actual propaganda, or just discussing the topic in general. She argues that promoting a countervailing view is a much better solution in the long run.[133]

Many of these criticisms are grounded in the international human rights law framework, particularly in Article 19 of the International Covenant on Civil and Political Rights (ICCPR) and in General Comment 34 of the Human Rights Committee, which interprets Article 19. Part II turns to this framework.

II. THE INTERNATIONAL HUMAN RIGHTS LAW FRAMEWORK

A. International Human Rights Law on the Internet

International human rights law dictates that "the same rights that people have offline must also be protected online, in particular freedom of expression . . . in accordance with Article 19 of the Universal Declaration of Human Rights and of the International Covenant on Civil and Political Rights." In successive resolutions, the Human Rights Council has also emphasized the "promotion, protection, and enjoyment of human rights . . . on the Internet."[134] Successive U.N. Special Rapporteurs, as well as the Human Rights Committee in its General Comment No. 34, which is intended to constitute authoritative legal analysis of the provisions of the ICCPR,[135] have emphasized that any restrictions on freedom of expression online must pass the same test for restrictions that exists offline.[136] This means that all such

---

133. *Government Pressure To Curb Online Terrorist Speech Among CDT Priorities*, WASHINGTON INTERNET DAILY (Feb. 10, 2016) (available on file).

134. *See* Human Rights Council Res. 32/13, U.N. Doc. A/HRC/RES/32/13, at 1, 3 (July 1, 2016); Human Rights Council Res. 26/13, U.N. Doc. A/HRC/RES/26/13, at 1, 2 (June 26, 2014); Human Rights Council Res. 20/8, U.N. Doc. A/HRC/20/L.13, at 2 (July 5, 2012).

135. Michael O'Flaherty, *International Covenant on Civil and Political Rights: interpreting freedom of expression and information standards for the present and the future*, *in* THE UNITED NATIONS AND FREEDOM OF EXPRESSION AND INFORMATION: CRITICAL PERSPECTIVES 75 (Tarlach McGonagle & Yvonne Donders eds., 2015) (The author was the rapporteur who drafted General Comment No. 34, including its provisions on restrictions of rights, for the U.N. Human Rights Committee, and is at present the Director of the EU Agency for Fundamental Rights.).

136. Frank La Rue, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, ¶ 69, U.N. Doc.

restrictions must comply with paragraph 2 of Article 19 of the ICCPR.[137] One influential version of this test has been articulated by the previous U.N. Special Rapporteur on the promotion and protection of the right to freedom of expression in a 2011 report:

> When a restriction is imposed as an exceptional measure on online content, it must pass a three-part, cumulative test:
> (1) it must be provided by law, which is clear and accessible to everyone (principles of predictability and transparency);
> (2) it must pursue one of the purposes set out in article 19, paragraph 3, of the International Covenant on Civil and Political Rights, namely:
> (i) to protect the rights or reputations of others;
> (ii) to protect national security or public order, or public health or morals (principle of legitimacy); and
> (3) it must be proven as necessary and the least restrictive means required to achieve the purported aim (principles of necessity and proportionality).
> In addition, any legislation restricting the right to freedom of expression must be applied by a body which is independent of any political, commercial, or other unwarranted influences in a manner that is neither arbitrary nor discriminatory. There should also be adequate safeguards against abuse, including the possibility of challenge and remedy against its abusive application.[138]

While the protection of national security and public order is a legitimate purpose for states undertaking CVE measures, the Human Rights Committee has stated that speech offenses such as "encouragement of terrorism" and "extremist activity" as well as "praising," "glorifying," or "justifying" terrorism should be narrowly and clearly defined to ensure that they do not lead to unnecessary

---

A/HRC/17/27 (May 16, 2011); U.N. Hum. Rts. Comm., Gen. Comment No. 34, ¶ 43, U.N. Doc. CCPR/C/GC/34 (Sept. 12, 2011).

137. Article 19, paragraph 2 reads: "Everyone shall have the right to freedom of expression; this right shall include freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, in the form of art, or through any other media of his choice." International Covenant on Civil and Political Rights art. 19, *opened for signature* Dec. 16, 1966, 999 U.N.T.S. 171, 178 (entered into force Mar. 23, 1976) [hereinafter ICCPR].

138. La Rue, *supra* note 136, ¶ 69.

or disproportionate interference with freedom of expression.[139] The concern is that over-broad criminalization based on vague definitions of "terrorism" and "extremism" could result in the censorship of criticism and legitimate political expression and the silencing of non-violent groups, journalists, and political activists critical of state policy.[140] Even if these vague laws are not enforced, there may be a chilling effect on speech, because vague laws give broad discretion to authorities to determine what kind of speech is illegal, causing individuals and Internet companies to err on the side of caution by censoring content of uncertain legal status in order to avoid onerous penalties.[141]

In the IRU context, the U.N. Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism, Ben Emmerson, has stated that IRU takedown requests interfere with or restrict the right to freedom of expression. As such, IRU takedown requests must be justified and independent judicial recourse must be available.[142] He has noted that laws that allow executive authorities to block websites in the absence of any initial judicial control or *ex post facto* judicial recourse may not comply with this requirement. Finally, he reiterated that many of the efforts to combat hate speech and violent extremism by restricting speech are misguided and that strategies addressing the root causes of such viewpoints should be prioritized.[143] This reflects the longstanding orthodoxy contained within the U.N. Global Counter-Terrorism Strategy adopted by the General Assembly in 2006 [144] and re-emphasized by both former U.N. Secretary-General Ban Ki Moon in his *Plan of Action to Prevent Violent Extremism*[145] and the present

---

139. U.N. Hum. Rts. Comm., Gen. Comment No. 34, *supra* note 136, ¶ 46.

140. Ben Emmerson, *Report of the Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism*, ¶ 21, U.N. Doc. A/HRC/31/65 (Feb. 22, 2016).

141. David Kaye, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, ¶ 39, U.N. Doc. A/HRC/32/38 (May 11, 2016).

142. Emmerson, *supra* note 140, ¶ 40.

143. *Id.*

144. G.A. Res. 60/288 (Sept. 20, 2006).

145. U.N. Secretary-General, *Plan of Action to Prevent Violent Extremism*, ¶¶ 4–7, U.N. Doc. A/70/674 (Dec. 24, 2015). Former U.N. Secretary-General Ban Ki Moon also recommends to Member States that "any restrictions on freedom of expression are clearly and narrowly defined and meet the three-part test of legality, proportionality and necessity." *Id.* ¶ 50(k).

Secretary-General, António Guterres.[146] These strategies state that counterterrorism measures should focus as much on (1) "address[ing] the conditions conducive to the spread of terrorism" and (4) "facilitat[ing] the promotion and protection of human rights for all and the rule of law as the fundamental basis of the fight against terrorism," as they do on (2) "prevent[ing] and combat[ing] terrorism" and (3) "build[ing] States' capacity to prevent and combat terrorism and to strengthen the role of the United Nations system in this regard."[147] Within the online CVE sphere, counterterrorism actors adhere to this orthodoxy when they adopt measures to promote counter-messages or counter-narratives.

B. The Business and Human Rights Framework

The Business and Human Rights (BHR) framework developed by U.N. Special Representative John Ruggie rests on three pillars: (1) the "State duty to protect against human rights abuses by third parties, including business enterprises[;]" (2) the "corporate responsibility to respect human rights, meaning business enterprises should act with due diligence to avoid infringing the rights of others and to address adverse impacts with which they are involved[;]" and (3) the "need for greater access for victims to effective remedy[.]"[148] The present U.N. Special Rapporteur on freedom of expression, David Kaye, has commenced a study on freedom of expression in the digital age, adopting the BHR framework as the basis for examining how states and private businesses in the ICT sector should protect and promote freedom of expression.[149] Many governmental and multi-stakeholder organizations, including UNESCO,[150] the European

146. U.N. Sec'y-Gen., Secretary-General's remarks to the General Assembly on informal suggestion to create a new office for Counter-terrorism, U.N. (Feb. 22, 2017), https://www.un.org/sg/en/content/sg/statement/2017-02-22/secretary-generals-remarks-general-assembly-informal-suggestion.

147. *Id.*

148. John Ruggie, Special Representative of the Sec'y-Gen., *Report of the Special Representative of the Secretary General on the issue of human rights and transnational corporations and other business enterprises*, ¶ 6, U.N. Doc. A/HRC/17/31 (Mar. 21, 2011).

149. Kaye, *supra* note 141, ¶¶ 9–13.

150. REBECCA MACKINNON ET AL., FOSTERING FREEDOM ONLINE: THE ROLE OF INTERNET INTERMEDIARIES 18 (UNESCO ed., 2014), http://unesdoc.unesco.org/images/0023/002311/231162e.pdf.

Commission,[151] the Global Counterterrorism Forum,[152] and the Global Network Initiative,[153] as well as civil society initiatives, such as the Manila Principles on Intermediary Liability[154] and Ranking Digital Rights,[155] have embraced this framework.

International human rights law imposes both negative obligations on states not to violate rights online and positive obligations to ensure enjoyment of those rights. This view of positive obligations translates into a BHR Guiding Principle that states must protect against human rights abuse by third parties, including business enterprises.[156] In practice, this means that states have duties to:

> (a) Enforce laws that are aimed at, or have the effect of, requiring [ICT companies] to respect human rights, and periodically to assess the adequacy of such laws and address any gaps;
> (b) Ensure that other laws and policies governing the creation and ongoing operation of [ICT companies] . . . do not constrain but enable [ICT companies'] respect for human rights;
> (c) Provide effective guidance to [ICT companies] on how to respect human rights throughout their operations; [and]

---

151. SHIFT ET AL., ICT SECTOR GUIDE ON IMPLEMENTING THE UN GUIDING PRINCIPLES ON BUSINESS AND HUMAN RIGHTS, (2013), https://www.ihrb.org/pdf/eu-sector-guidance/EC-Guides/ICT/EC-Guide_ICT.pdf.

152. GLOB. COUNTERTERRORISM FORUM, *supra* note 24, at 11.

153. GLOB. NETWORK INITIATIVE, GNI PRINCIPLES ON FREEDOM OF EXPRESSION AND PRIVACY 1 (2008), https://globalnetworkinitiative.org/sites/default/files/GNI-Principles-on-Freedom-of-Expression-and-Privacy_0.pdf; GLOB. NETWORK INITIATIVE, IMPLEMENTATION GUIDELINES FOR THE PRINCIPLES ON FREEDOM OF EXPRESSION AND PRIVACY ¶ 2.4 (2010), https://globalnetworkinitiative.org/sites/default/files/Implementation-Guidelines-for-the-GNI-Principles_0.pdf; GLOB. NETWORK INITIATIVE, EXTREMIST CONTENT AND THE ICT SECTOR 2 (Nov. 30, 2016), https://globalnetworkinitiative.org/sites/default/files/Extremist-Content-and-ICT-Sector.pdf.

154. MANILA PRINCIPLES ON INTERMEDIARY LIABILITY, https://www.manilaprinciples.org/ (last visited Nov. 11, 2017).

155. *Corporate Accountability Index*, RANKING DIGITAL RIGHTS, https://rankingdigitalrights.org/index2017/ (last visited Nov. 15, 2017).

156. Ruggie, *supra* note 148, I.A.1; *see also*, Kaye, *supra* note 141, ¶ 8 (recognizing that individuals enjoy rights online and states have obligations to ensure those rights, including requiring "public authorities to take steps to protect individuals from the actions of private parties").

> (d) Encourage and, where appropriate require, [ICT companies] to communicate how they address their human rights impacts.[157]

It also means that states should ensure policy coherence when state actors adopt laws and policies affecting ICT companies, in order to ensure respect for human rights while pursuing different societal needs.[158] Another Guiding Principle of particular relevance when considering IRUs is that concerning access to remedy: "States must take appropriate steps to ensure, through judicial, administrative, legislative, or other appropriate means, that when [human rights] abuses occur within their territory and/or jurisdiction those affected have access to effective remedy."[159] Grievance mechanisms may be judicial, non-judicial, or non-state-based, so long as they meet the effectiveness criteria set out in the BHR Guiding Principles; that is, they must be legitimate; accessible; predictable; equitable; transparent; rights-compatible; and a source of continuous learning.[160]

C. Work of the Special Rapporteur Particularly Relevant to IRUs

In a 2016 report, the U.N. Special Rapporteur on freedom of expression, David Kaye, highlighted several challenges to freedom of expression online particularly relevant to IRUs, including: vague laws, as described above, excessive intermediary liability, extralegal restrictions, filtering, and ICT companies' internal policies and practices, including terms of service and design and engineering choices.[161] He warned that excessive intermediary liability, coupled with vague laws and extralegal requests to take down content, could encourage ICT companies to filter content excessively.[162] The UNESCO study's findings support this concern: "the stricter the intermediary liability regime in a given jurisdiction, the more likely content is to be removed either proactively by the company or upon request by authorities without challenge."[163] Kaye also noted that "[i]ntermediaries are increasingly being required to assess the validity of state requests and private complaints against general legal criteria,

---

157. Ruggie, *supra* note 148, § I.B.3.
158. *Id.* § I.B.8.
159. *Id.* § III.A.25.
160. *Id.* § III.B.31.
161. Kaye, *supra* note 141, ¶¶ 35–55.
162. *Id.* ¶ 42.
163. Mackinnon, *supra* note 150, at 11.

and remove or delink content based on such assessments."[164] Such notice-and-takedown frameworks, he explained, "have been criticized for incentivizing questionable claims and for failing to provide adequate protection for . . . intermediaries [seeking] to apply fair and human rights-sensitive standards to content regulation."[165] The Inter-American Commission on Human Rights shares Kaye's concern, having noted that private actors "lack the ability to weigh rights and to interpret the law in accordance with freedom of speech and other human rights standards,"[166] perhaps due to resource constraints, lack of oversight and accountability, or potential conflicts of interest. As a result, Kaye reasoned that if ICT companies face potential intermediary liability, they will be prone to self-censorship or over-censorship.[167]

Kaye highlighted several problems relating to terms of service (ToS), having observed that IRUs' *modus operandi* includes reporting content as a violation of sites' ToS. In particular, he expressed concern that this practice raises the prospect that states rely on private ToS to bypass human rights or domestic law norms against restricting content.[168] More generally, he noted that ToS are frequently written in such general terms that it may be difficult to predict with certainty what kinds of content would be restricted; that ToS have been inconsistently enforced;[169] and that ICT companies often fail to provide an appeals process or communicate detailed reasons for removing content or deactivating accounts.[170] He observed that private censorship is complicated by the sheer volume of content that censors have to process, that intermediaries often outsource content

164. Kaye, *supra* note 141, ¶ 43.

165. *Id.* ¶ 43.

166. *Id.* ¶ 44 (citing Inter-Am. Comm. on Hum. Rts., Freedom of Expression and the Internet, OEA/Ser.L/V/II, CIDH/RELE/INF.11/13 47–48 (Dec. 31, 2013)).

167. *Id.* ¶ 44.

168. *Id.* ¶ 53.

169. Kaye, *supra* note 141, ¶ 52 (referring to allegations that ICT companies are reluctant to address tech-related violence against women until it becomes a public relations issue and other criticisms that ICT companies have been overzealous in censoring a wide range of legitimate, but perhaps "uncomfortable" to some audiences, expression).

170. *Id.* ¶ 52.

moderation, and that they "face 'complex value judgments,' issues with cultural sensitivity, and 'difficult decisions about conflicts of law.'"[171]

Kaye noted that transparency is another issue. There is little information available about the volume and nature of government requests to restrict or remove content and wide variation in whether and how ICT companies explain reasons and processes for content-removal. He underscored that while there is often quantitative transparency, there remains a lack of qualitative transparency, for example no explanation as to why content has been taken down.[172] He also highlighted the importance of access to remedy as required by Article 2(3) of the ICCPR and the BHR Guiding Principles, but noted that there remains limited guidance on how the BHR Guiding Principles should be operationalized or assessed in the context of ICT companies and underscored the need for further research on best practices for how companies communicate ToS enforcement decisions and how they implement appeals mechanisms.[173]

Kaye concluded his report with two recommendations to states:

> (1) States bear a primary responsibility to protect and respect the right to exercise freedom of opinion and expression. In the information and communication technology context, this means that States must not require or otherwise pressure the private sector to take steps that unnecessarily or disproportionately interfere with freedom of expression, whether through laws, policies, or extralegal means. Any demands, requests and other measures to take down digital content or access customer information must be based on validly enacted law, subject to external and independent oversight, and demonstrate a necessary and proportionate means of achieving one or more aims under article 19 (3) of the International Covenant on Civil and Political Rights. Particularly in the context of regulating the private sector, State laws and policies must be transparently adopted and implemented.
>
> (2) Governments must also adopt and implement laws and policies that protect private development and the provision of technical measures, products and services

171. *Id.* ¶ 54 (quoting Emily Taylor, The Privatization of Human Rights: Illusions of Consent, Automation and Neutrality, GLOBAL COMM. ON INTERNET GOV. PAPER SERIES NO. 24 (2016)).

172. *Id.* ¶ 64.

173. *Id.* ¶¶ 65–71.

> that advance freedom of expression. They must ensure legislative, policymaking and other relevant norm-setting processes concerning rights and restrictions on the Internet in order to provide the private sector, civil society, the technical community and academia meaningful opportunities for input and participation."[174]

D. Joint Declarations by Special Rapporteurs on Freedom of Expression

The Special Rapporteurs on Freedom of Expression of a number of different human rights systems, including the United Nations, the Organization for Security and Co-operation in Europe (OSCE), the Organization of American States (OAS), and the African Commission on Human and Peoples' Rights (ACHPR), meet annually to discuss topical issues and adopt joint declarations to provide normative guidance to states. Two joint declarations of particular relevance to IRUs[175] are the 2016 Joint Declaration on Freedom of Expression and Countering Violent Extremism (Joint Declaration on CVE)[176] and the 2017 Joint Declaration on Freedom of Expression and "Fake News," Disinformation and Propaganda (Joint Declaration on Fake News).[177]

Of particular relevance to IRUs, the Joint Declaration on CVE recommends that:

> (b) All CVE/PVE programmes and initiatives should respect human rights and the rule of law, and contain specific safeguards against abuse in this regard. They should be independently reviewed on a regular basis to determine their impact on human rights, including the right to freedom of expression, and these reviews should be made public.

---

174. Kaye, *supra* note 141, ¶¶ 85–86.

175. These Joint Declarations are highly recommended reading for those interested in IRUs, but will not be reproduced in detail, as they restate much of the guidance addressed above.

176. U.N. Spec. Rapporteur on Freedom of Opinion and Expression et. al., *Joint Declaration on Freedom of Expression and Countering Violent Extremism* (May 4, 2016) [hereinafter Joint Declaration on CVE].

177. U.N. Spec. Rapporteur on Freedom of Opinion and Expression et. al., *Joint Declaration on Freedom of Expression and "Fake News," Disinformation and Propaganda* (Mar. 3, 2017) [hereinafter Joint Declaration on Fake News], http://www.osce.org/fom/302796?download=true.

> (c) The concepts of "violent extremism" and "extremism" should not be used as the basis for restricting freedom of expression unless they are defined clearly and appropriately narrowly. . . .
>
> . . . .
>
> (f) States should not subject Internet intermediaries to mandatory orders to remove or otherwise restrict content except where the content is lawfully restricted in accordance with the standards outlined above [including compliance with international human rights law, respect for the prohibition on discrimination, and the availability of independent judicial oversight]. States should refrain from pressuring, punishing or rewarding intermediaries with the aim of restricting lawful content.[178]

The Joint Declaration on Fake News elaborates further on freedom of expression, transparency and due process principles concerning intermediary liability, and recommends that:

> (d) Intermediaries should never be liable for any third party content relating to those services unless they specifically intervene in that content or refuse to obey an order adopted in accordance with due process guarantees by an independent, impartial, authoritative oversight body (such as a court) to remove it and they have the technical capacity to do that.
>
> (e) Consideration should be given to protecting individuals against liability for merely redistributing or promoting, through intermediaries, content of which they are not the author and which they have not modified.[179]

E. Sub-Conclusion: Model Standards for IRUs and ICT companies Within an International Human Rights Framework

Based on the foregoing applicable international human rights law standards and guidance, it is possible to derive a set of model standards for IRUs and ICT companies that ensure maximal protection for freedom of expression while still allowing for the pursuit of other societal values, such as the protection of civilians by suppressing

178. Joint Declaration on CVE, *supra* note 176, ¶¶ 3–4.
179. Joint Declaration on Fake News, *supra* note 177, ¶¶ 2–3.

content that incites violence.[180] These standards are elaborated in this section, and will be used later to evaluate the current IRUs in the European Union and the United Kingdom.

> (1) *Because states have a positive obligation to ensure that ICT companies respect and promote freedom of expression online, they should legislate to ensure broad intermediary immunity for third-party content that ICT companies have not modified, and to ensure that companies bear no obligation to monitor or takedown third-party content without a court order.*

This standard would ensure that ICT companies do not over-censor third-party content out of fear of future liability. Such immunity would not preclude states from imposing liability on ICT companies that refuse to take down third-party content or deactivate accounts in response to a court order issued by an independent judge, who has evaluated a content takedown request against international human rights standards, through a procedure that notifies the user who uploaded the content at issue and gives him or her a right to be heard, and to appeal.

> *(2) States should refrain from threatening ICT companies with imposing liability, blocking their services, or resorting to other coercive pressures, in order to secure the removal of third-party content that they have not modified, in the absence of a court order.*
> *(3) While content takedown requests by IRUs are not per se impermissible by international human rights standards, they constitute interferences with the right to freedom of expression, and must meet a number of requirements to adhere to international human rights standards.*
> a. There should be a clear legal framework for the IRU, adopted through the normal legislative process after extensive consultation with all relevant stakeholders. These stakeholders should include the ICT companies, civil society groups that advocate for freedom of expression, and vulnerable and minority groups that may be particularly affected by the operation of the IRU. The legislation should make clear that ICT companies will not be held liable for failing to take

180. *See also* MANILA PRINCIPLES ON INTERMEDIARY LIABILITY, *supra* note 154 (providing a more detailed set of baseline safeguards and best practice standards that are fundamentally similar to the proposed Model Standards).

down third-party content in the absence of a court order and should provide for meaningful external oversight over and judicial review of the IRU.

b. The criteria IRUs use to refer content for takedown should be accessible, clearly and narrowly defined (e.g., "advocacy of violence against civilians" or "incitement to commit a terrorist offense" but not "the promotion of extremist viewpoints"), and strictly limited to the purposes allowed by international human rights law (Article 19(3), ICCPR). These criteria should be proven to be necessary and the least restrictive means to achieve the public purpose (e.g., the availability of normal judicial process and counter-messaging systems should be considered as alternatives or as part of a broader policy framework). The inclusion of new criteria (e.g., "messages relating to migrant smuggling") in the enabling legislation should require an amendment via the normal legislative process.

c. Content takedown requests by the IRU should state clear and detailed reasons for the request; these reasons should be specified in accordance with criteria strictly limited to those provided for in the IRU's enabling legislation. Requests should make clear that the request is not a court order, and that the company will not be held liable or otherwise penalized for failing to take down the content at issue in the absence of a court order. The IRU should also request that the ICT company notify the user who uploaded the content at issue, give that user an adequate opportunity to make representations to the ICT company, and inform a user of their options to appeal or seek review of the decision if the content is taken down.

d. Content takedown requests by the IRU should only be made after a rigorous evaluation of the content at issue, to ensure that it falls squarely within the criteria for restriction, and is not otherwise protected by international human rights law (e.g., content that is journalistic in nature; content that is justified on the basis of academic freedom or contribution to robust debate in an open society; content that criticizes violence or terrorism; and content that is merely offensive, shocking, or disturbing). The IRU's algorithms, policy guidance, and training material should be continually updated to minimize the probability of wrongful takedown requests and should

> incorporate lessons from remedial mechanisms, the private sector, and civil society.
>
> e. The IRU should be transparent both quantitatively and qualitatively, publishing regular reports on the number and nature of content takedown requests and contextualizing the material that is subject to takedown requests. The IRU should publish its policy guidance at a level of detail that will enable the public to understand what types of content may be subject to takedown requests and to seek review of particular types of content that should not be targeted.
>
> f. The IRU should be independently reviewed on a regular basis to determine its impact on human rights, including the right to freedom of expression, and the results of these reviews should be made public.
>
> *(4) ICT companies should be incentivized to adopt policy commitments, terms of service, platform designs, notification and takedown procedures, and remedial mechanisms that meet the U.N. Guiding Principles on Business and Human Rights.*

One possible incentive is for states to provide more generous intermediary immunity to ICT companies certified as compliant through independent assessments by multi-stakeholder initiatives, such as the Global Network Initiative.[181] Another possible incentive is for states to restrict offshore data transfers to only those ICT companies that can prove that they are compliant with data protection laws as well as freedom of expression standards; or to ICT companies in countries that can demonstrate similarly protective laws and standards.

> a. ICT companies should develop terms of service and criteria for taking down content in accordance with human rights due diligence standards, by drawing on human rights expertise and by conducting meaningful consultations with all relevant stakeholders, including civil society groups that advocate respect for freedom of expression and vulnerable and minority groups that may be particularly affected by content takedowns.

---

181. *See, e.g.*, GLOB. NETWORK INITIATIVE, PUBLIC REPORT ON THE 2015/2016 INDEPENDENT COMPANY ASSESSMENTS (2016), http://globalnetworkinitiative.org/sites/default/files/Public-Report-2015-16-Independent-Company-Assessments.pdf (providing an independent review of member companies' compliance with the Global Network Initiative Principles and Implementation Guidelines).

b. ICT companies should adopt policies that commit to respecting human rights, including the right to freedom of expression, which meet the standards provided in the BHR Guiding Principles. This policy should include or result in operational commitments to respect and promote freedom of expression in designing and implementing terms of service, takedown procedures, and appeals mechanisms as well as through platform design and engineering choices. It should also result in a commitment to not censor or take down content in response to government requests or court orders that do not meet international human rights standards. Further, the policy should be to provide minimal compliance with such requests or court orders (e.g., by restricting content only within that jurisdiction) and to publicize these requests and orders in situations where the government threatens or exercises coercion over the ICT companies, leaving the ICT companies with no other viable options apart from ceasing operations in that jurisdiction.

c. ICT companies should develop effective notice, due process, and remedial mechanisms that are legitimate, accessible, predictable, equitable, transparent, rights-compatible, and a source for continuous learning (and therefore meet the effectiveness criteria set out in the BHR Guiding Principles).

(i) ICT companies should inform content uploaders of the reasons for removal requests, how they can challenge the removal request, and how they can appeal content removal decisions.

(ii) In terms of due process, ICT companies should ensure that content uploaders have a right to be heard before a fair and non-discriminatory adjudicatory body that provides detailed reasons for its decisions and should not take down content until the uploader has had a chance to dispute a take-down request (e.g., by requiring the uploader to fill out a form before they can continue activities on the ICT companies' platform, or by waiting a reasonable period of time before making a decision). During the time when content is awaiting a removal decision, ICT companies could indicate that the content is disputed and provide links to counter-

> messaging or alternative content from trusted content providers.[182]
> (iii) In terms of remedial mechanisms, ICT companies should have internal appeals mechanisms that enable content uploaders to be heard by a different decision-maker from the original decision-maker, as well as external appeals mechanisms that are fair, impartial, independent and transparent. The external appeals mechanisms could be run or audited by multi-stakeholder initiatives. Should a user win an appeal, the intermediary should reinstate the content. Lessons from the decisions of the initial adjudicatory and appeals mechanisms should be integrated into the decision-making processes (e.g., algorithms or policy guidance). ICT companies can limit the potential for abuse of takedown request mechanisms by limiting or banning takedown requests from particular users that appear to be abusing the reporting mechanism.
> (iv) ICT companies should be both quantitatively and qualitatively transparent, as described above, about the content takedown requests they receive, any other mechanisms they use to take down content (e.g., proactive algorithms or filters), and the decisions made by adjudicating and appeals mechanisms.
> d. ICT companies should regularly review and improve their policy commitments, terms of service, platform designs, notification and takedown procedures, and remedial mechanisms to improve them and minimize their human rights impacts.

III. THE EUROPEAN CONVENTION ON HUMAN RIGHTS AND IRUS

Although the European Court of Human Rights (ECtHR) has only a small number of cases addressing freedom of expression online and none that directly address the operation of IRUs, there are plausible arguments that IRUs violate Article 6 (the right to a fair trial in determining civil rights or criminal charges), Article 10 (the right to

182. *Contra* D. C. Nunziato, *With Great Power Comes Great Responsibility: Proposed Principles of Digital Due Process for ICT Companies*, *in* PROTECTION OF INFORMATION AND THE RIGHT TO PRIVACY – A NEW EQUILIBRIUM? (Luciano Floridi ed., 2014) (advocating an approach that is more protective of freedom of expression, which the present author believes merits consideration, but goes beyond the existing requirements of international human rights law and the business and human rights framework today).

freedom of expression) and Article 13 (the right to an effective remedy) of the European Convention on Human Rights (ECHR).[183] Should a suitable case be brought before the ECtHR, the total lack of procedural safeguards and the fundamental importance of the right to freedom of expression are likely to be persuasive factors. However, governments seeking to defend their use of IRUs may rely on formidable defenses, including (i) arguments that IRU referrals do not constitute interferences with the right to freedom of expression or determination of civil rights and obligations; (ii) arguments relying on Article 17 (the prohibition on the abuse of rights to destroy the rights of others) of the ECHR; (iii) arguments seeking to extend the logic of the ECtHR's holding in *Delfi AS v. Estonia*[184] to argue that the imposition of intermediary liability on ICT companies for failing to take measures to remove speech amounting to hate speech or incitement to violence *even without notice* is not incompatible with the ECHR; and (iv) arguments that IRUs are proportionate responses to the challenge of countering violent extremism online.

This Part explains and evaluates the different legal issues raised by IRUs, relying on the jurisprudence of the ECtHR as well as the standard-setting documents of the Council of Europe, which the ECtHR often cites in its cases concerning the Internet. This Part finds that IRU referrals constitute interferences or determinations and the ECtHR will strictly limit their use to removing hate speech and inciting terrorism. The ECtHR will not extend their use to the removal of defamatory speech. The Part concludes with the view that a well-argued case with a strong applicant—such as a Facebook user whose post criticizing violence was mistakenly censored following an IRU request—that relies on the Articles 6 and 10 jurisprudence of the ECtHR, the standard-setting documents of the Council of Europe, and the international human rights material above, is likely to obtain a judgment that limits the potential abuse of IRUs.

---

183. European Convention for the Protection of Human Rights and Fundamental Freedoms, *opened for signature* Nov. 4, 1950, art. 1, Europ. T.S. No. 5, 213 U.N.T.S. 221, arts. 6, 10, 13 (entered into force Sept. 3, 1953) [hereinafter ECHR].

184. Delfi AS v. Estonia, App. No. 64569/09, ¶ 13 (Eur. Ct. H.R. June 16, 2015) (ruling that "holding a news portal liable for clearly unlawful comments such as insults, threats and hate speech under such circumstances will in general be compatible with Article 10 of the Convention").

A. Brief Explanation of the Structure of the ECHR

In relevant part, Article 6(1) provides that: "In the determination of his civil rights and obligations . . . against him, everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal established by law."[185]

> While Article 10 provides that:
>
> 1. Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers. This Article shall not prevent States from requiring the licensing of broadcasting, television or cinema enterprises.
> 2. The exercise of these freedoms, since it carries with it duties and responsibilities, may be subject to such formalities, conditions, restrictions or penalties as are prescribed by law and are necessary in a democratic society, in the interests of national security, territorial integrity or public safety, for the prevention of disorder or crime, for the protection of health or morals, for the protection of the reputation or rights of others, for preventing the disclosure of information received in confidence, or for maintaining the authority and impartiality of the judiciary.[186]

The starting point of any analysis of potential violations therefore begins with the identification of the rights engaged and determination of whether there is such a civil right or obligation within the meaning of Article 6 or an interference with the right to freedom of expression under Article 10. This may then be followed by a determination of whether the rights have been violated. A finding of a violation may then be followed by an inquiry into whether there was an effective remedy as guaranteed by Article 13.[187]

Notably, Article 13 does not require the applicants to prove a violation of the ECHR, but requires the state to provide a preemptive remedy for individuals with "arguable claims" whose ECHR rights

---

185. ECHR, *supra* note 183, art. 6.

186. *Id.* art. 10.

187. *Id.* art. 13 ("Everyone whose rights and freedoms as set forth in [the] Convention are violated shall have an effective remedy before a national authority notwithstanding that the violation has been committed by persons acting in an official capacity.").

have been violated.[188] This means that so long as individuals can make arguable claims that their Article 6 or 10 rights have been violated, they are entitled to an effective remedy before a national authority.

In cases concerning either hate speech or speech inciting violence, the ECtHR has adopted two different approaches. First, the ECtHR has relied on Article 17[189] to find that an Article 10 claim is inadmissible, without going into a detailed analysis of whether the right is engaged or has been violated.[190] This approach covers "essentially [*only*] those rights which, if invoked, will facilitate the attempt to derive therefrom a right to engage personally in activities aimed at the destruction of any of the rights and freedoms set forth in the Convention"[191]—in practice, Articles 9, 10, and 11. In general, the ECtHR only uses this approach when the action violates the ECHR's fundamental underlying values, such as democracy, tolerance, non-violence,, and non-discrimination.[192] Examples of speech denied the protection of Article 10 by virtue of Article 17 include statements denying the Holocaust, justifying a pro-Nazi policy, linking all Muslims with a grave act of terrorism, or portraying the Jews as the source of evil in Russia.[193] This approach will not lead to such an abbreviated analysis when claimants assert violations of other ECHR rights, such as those granted in Articles 6 and 7.[194]

---

188. D. J. HARRIS ET AL., HARRIS, O'BOYLE & WARBRICK: LAW OF THE EUROPEAN CONVENTION ON HUMAN RIGHTS 767 (3d ed. 2014) [hereinafter HARRIS, O'BOYLE & WARBRICK] (citing Klass v. Germany, App. No. 5029/71, ¶ 64 (Eur. Ct. H.R. Sept. 6, 1978); Silver v. United Kingdom, App. Nos. 5947/72, 6205/73, 7107/75, 7113/75, 7136/75, ¶ 113 (Eur. Ct. H.R. Mar. 25, 1983)).

189. ECHR, *supra* note 183, art. 17 ("Nothing in this Convention may be interpreted as implying for any State, group or person any right to engage in any activity or perform any act aimed at the destruction of any of the rights and freedoms set forth herein or at their limitation to a greater extent than is provided for in the Convention.").

190. Although the court has on occasion examined the underlying speech to determine if there was incitement to violence. *See, e.g.*, Erdoğdu and İnce v. Turkey, App. Nos. 25067/94, 25068/94, ¶¶ 9, 47 (Eur. Ct. H.R. July 8, 1999) (assessing the content of an interview to see if it could be described as incitement to violence).

191. *Id.* (quoting WP v. Poland, App. No. 42264/98, ¶ 359 (Eur. Ct. H.R. Sept. 2, 2004)).

192. KAREN REID, A PRACTITIONER'S GUIDE TO THE EUROPEAN CONVENTION ON HUMAN RIGHTS, 3-010 (5th ed. 2015).

193. Delfi AS v. Estonia, App. No. 64569/09, ¶ 136 (Eur. Ct. H.R. June 16, 2015).

194. ECHR, *supra* note 183, art. 7 (the prohibition against punishment without law); HARRIS, O'BOYLE & WARBRICK, *supra* note 188, at 853 (citing Lawless v. Ireland, App. No. 332/57, ¶ 7 (Eur. Ct. H.R. July 1, 1961); Kasymakhunov and

The ECtHR's second, more common approach recognizes that Article 10 is engaged, but finds that the state's interferences with hate speech or speech inciting violence are justified under Article 10(2) of the ECHR. This is the approach that the ECtHR adopted in *Delfi AS v. Estonia* toward the state laws governing Delfi AS. In *Delfi AS*, the ECtHR found that large, professionally-managed Internet news portals that publish news articles of their own and provide, for economic purposes, a platform for user-generated comments, assume "duties and responsibilities" under Article 10(2) and therefore can be held liable for not removing without delay comments that amount to hate speech or incitement to violence.[195]

B. Addressing the Threshold Question: Article 6 Always Applies

One of the threshold questions in addressing IRU referrals would therefore appear to be whether the referred content is protected by Article 10, thereby engaging the approach the ECtHR relied upon in *Delfi AS v. Estonia*, or whether the Article 17 approach should be applied. However, this Article argues that, regardless of approach, Article 6 always applies in the context of IRUs. This means that users are always *entitled* to a determination of the lawfulness or unlawfulness of content by a fair, independent, and impartial decision-maker and an opportunity to defend the content either before a referral is made or, in exceptional cases, on appeal after a referral. Article 6 would also require that users receive clear notice when their content is referred, as discussed in Section II.F.3.iii.

In *Lawless (No. 3) v. Ireland*, the ECtHR held that Article 17 "is negative in scope and cannot be construed *a contrario* as depriving a physical person of the fundamental individual rights guaranteed by Articles 5 and 6."[196] Because applicants must exhaust domestic remedies before they can bring a case before the ECtHR, even in cases in which the ECtHR has applied the Article 17 approach, the ECtHR will generally have had determinations of the unlawfulness of the content, albeit at the domestic level. The author's review of all the cases

---

Saybatalov v. Russia, App. Nos. 26261/05, 26377/06, ¶¶ 28, 31, 37 (Eur. Ct. H.R. Mar. 14, 2013) (finding a violation of Article 7 but holding that neither applicant could rely on Articles 9, 10, or 11)).

195. Delfi AS v. Estonia, App. No. 64569/09, ¶ 115 (Eur. Ct. H.R. June 16, 2015).

196. Lawless v. Ireland, App. No. 332/57, ¶ 7 (Eur. Ct. H.R. July 1, 1961).

listed in an ECtHR factsheet on "Hate Speech"[197] discloses that all forty-one cases featured such an initial determination by a presumably fair, independent, and impartial decision-maker, regardless of which of the two approaches were taken. Hence, it seems there was at least some level of due process in the ECtHR's previous hate speech case law even if the ECtHR ultimately found the speech was not protected.

As for Article 10, the ECtHR has already read Article 6 due process guarantees into Article 10, therefore subjecting the proportionality of an interference to greater scrutiny. In borrowing Article 6 precepts into Article 10, the ECtHR's concern has been to ensure that those whose right to freedom of expression is interfered with have an effective opportunity to state their case under conditions "*in conformity with* an adversarial procedure."[198] Thus, in *Steel & Morris v. United Kingdom*, the ECtHR found that the applicants' right to freedom of expression had been violated because of the procedural unfairness and inequality of arms[199] of the defamation proceedings. The proceedings featured a large legal team representing McDonald's against the largely unrepresented applicants, who had been denied legal aid and faced major difficulties throughout lengthy proceedings in meeting their procedural burden of proving the truth of serious factual allegations against McDonald's. "Given the enormity and complexity of that undertaking" and the disproportionate size of the damages award, the ECtHR found that there was no fair balance struck between the competing interests.[200] Thus, because the claimants did not have an effective opportunity to discharge the burden of proof, the ECtHR found a violation of Article 10. The ECtHR has also found that the denial of an effective opportunity to present evidence in defamation proceedings to prove the truth of statements

197. EUR. COURT OF HUMAN RIGHTS, HATE SPEECH (2017), http://www.echr.coe.int/Documents/FS_Hate_speech_ENG.pdf.

198. Lawrence Early, *Article 10: issues of fairness, proof and evidence*, *in* FREEDOM OF EXPRESSION: ESSAYS IN HONOUR OF NICOLAS BRATZA 553, 566 (Josep Casadevall et al. eds., 2012) (emphasis added).

199. *Id.* at 553–54; *see also* Kate Gibson et al., *Regulation of the International Bar*, *in* RESEARCH HANDBOOK ON INTERNATIONAL COURTS AND TRIBUNALS 407, 408 n.2 (William A. Schabas & Shannonbrooke Murphy eds., 2017) ("The principle of equality of arms was derived from the jurisprudence of the European Court of Human Rights in the context of the right to a fair trial [Article 6]. Equality of arms requires that there be a fair balance between the opportunities afforded to the parties involved in litigation.").

200. Steel and Morris v. United Kingdom, App. No. 68416/01, ¶ 95 (Eur. Ct. H.R. Feb. 15, 2005).

may constitute a violation of Article 10.[201] While all of the above cases concern defamation proceedings, the same due process requirements may apply in cases concerning alleged hate speech or incitement to violence. As mentioned earlier, Article 17 does not deprive claimants of all their ECHR rights, including their Article 6 right to a fair trial, and it would be logical for the ECtHR to read similar due process requirements into Article 10 in cases concerning alleged hate speech or incitement to violence, because defendants need to have an effective opportunity to prove that the speech is not hate speech or incitement to violence.

In *Delfi AS v. Estonia*, while the ECtHR took the Article 10 approach with respect to the imposition of liability (as discussed above), the Court took the Article 17 approach towards the speech in question. In so doing, the ECtHR found that the majority of the impugned comments amounted to hate speech or incitement to violence and as such did not enjoy the protection of Article 10.[202] However, because the authors of the comments were anonymous and were not litigants in the domestic proceedings, in which the applicant sued Delfi for failing to take down the comments expeditiously, the ECtHR noted that the freedom of expression of the authors was not at issue.[203] Instead the case concerned whether the Estonian courts' decisions, holding Delfi liable for those comments posted by third parties, were in breach of Delfi's freedom to impart information under Article 10.[204] The case is therefore not very instructive in determining whether the users' Article 6 and 10 rights would have been engaged if they had been the ones sued.

Thus, while the Article 17 approach will presumptively apply to terrorism-related content that constitutes hate speech or incitement to violence, it will not preclude a challenge under Article 6 to determine if the impugned content is actually terrorism-related and falls outside the protection of Article 10. Having said that, it is possible and arguably preferable for courts to address IRU referrals using Article 10(2), in order to demonstrate their normative commitment to freedom

---

201. Early, *supra* note 198, at 562–63 (citing Flux (No. 4) v. Moldova, App. No. 17294/04, ¶¶ 37–38 (Eur. Ct. H.R Jun. 12, 2007); Jerusalem v. Austria, App. No. 26958/95 (Eur. Ct. H.R. Feb. 27, 2001); Folea v. Roumanie, App. No. 34434/02, ¶¶ 41–43 (Eur. Ct. H.R Oct. 14, 2008)).

202. Delfi AS v. Estonia, App. No. 64569/09, ¶ 140 (Eur. Ct. H.R. June 16, 2015).

203. *Id.*

204. *Id.*

of expression—even expression that shocks, offends, or disturbs—and to ensure that IRUs are constrained by the safeguards against unnecessary censorship that have been elaborated by the ECtHR in Article 10.

C. The Relevance of Council of Europe Standard-Setting Documents and their Content

The Committee of Ministers (CM) has passed a number of declarations and recommendations recognizing the importance of protecting freedom of expression on the Internet. These recommendations and declarations have been and will be influential in shaping Internet governance in Member States as well as ECtHR jurisprudence and in ensuring the protection of human rights online.[205] For example, in *Delfi AS v. Estonia*, the Grand Chamber of the ECtHR cited as relevant the 2003 Declaration on Freedom of Communication on the Internet adopted by the Committee of Ministers.[206] The ECtHR also cited *Recommendation CM/Rec(2011)7 of the Committee of Ministers to member states on a new notion of media* for the idea that a "differentiated and graduated approach" was favorable in regulating new media, in determining that the applicant was an "intermediary" that could be subject to intermediary liability if it did not monitor and take down content amounting to hate speech or incitement to violence.[207] The ECtHR also cited numerous declarations and recommendations of the CM in finding a right to access the Internet in *Yildirim v. Turkey*.[208] Of particular relevance to the present discussion is Recommendation CM/Rec(2012)4 on the protection of human rights with regard to social networking services, which recommends that "[s]ocial networking providers should respect human rights and the rule of law." The Recommendation further provides that:

---

205. *See generally* Lize R. Glas, *The European Court of Human Rights' Use of Non-Binding and Standard-Setting Council of Europe Documents*, 17 HUM. RTS. L. REV. 97 (2017) (analyzing the ECtHR's use of non-binding documents passed by organizations like the Council of Europe).

206. Comm. of Ministers, *Declaration on Freedom of Communication on the Internet*, COUNCIL OF EUR. (May 28, 2003), http://www.osce.org/fom/31507?download=true.

207. *Recommendation CM/Rec(2011)7 of the Committee of Ministers to member states on a new notion of media*, COUNCIL OF EUR. (Sept. 21, 2011), http://www.osce.org/odihr/101403?download=true; Delfi AS v. Estonia, App. No. 64569/09, ¶¶ 113, 125–29 (Eur. Ct. H.R. June 16, 2015).

208. Ahmet Yıldırım v. Turkey, App. No. 3111/10, ¶¶ 20–26 (Eur. Ct. H.R. Dec. 18, 2012).

> A number of self- and co-regulatory mechanisms have already been set up in some Council of Europe member States in connection with standards for the use of social networking. It is important that procedural safeguards are respected by these mechanisms, in line with the right to be heard and to review or appeal against decisions, including in appropriate cases the right to a fair trial, within a reasonable time, and starting with the presumption of innocence.[209]

The latest CM recommendation recommends that "[a]ny measure taken by State authorities or private-sector actors to block, filter or remove Internet content, or any request by State authorities to carry out such actions complies with the conditions of Article 10 of the Convention regarding the legality, legitimacy and proportionality of restrictions" and that "Internet users or other interested parties have access to a court in compliance with Article 6 of the Convention with regard to any action taken to restrict their access to the Internet or their ability to receive and impart content or information."[210]

D. Article 6 and IRU Referrals

While Article 6 is most often invoked in the criminal context, it also has a civil limb. Article 6(1)[211] guarantees a fair trial in the determination of civil rights and obligations, especially entitling everyone whose (i) *civil rights* are being (ii) *determined* "to a fair and public hearing within a reasonable time by an independent and impartial tribunal established by law."[212] This Section argues that Article 6(1) applies to alleged violations of terms of service as breach of

---

209. *Recommendation CM/Rec (2012)4 of the Committee of Ministers to member States on the protection of human rights with regard to social networking services*, COUNCIL OF EUR. (Apr. 4, 2012), https://wcd.coe.int/ViewDoc.jsp?p=&id=1929453&Site=CM&direct=true.

210. *Recommendation CM/Rec (2016)5 of the Committee of Ministers to member States on Internet freedom*, COUNCIL OF EUR. (Apr. 13, 2016), https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=09000016806415fa. As this article was finalized for publication, the Committee of Ministers adopted a new Recommendation on the roles and responsibilities of internet intermediaries, which is particularly relevant to content restrictions and generally accords with the findings and recommendations of this article. *See Recommendation CM/Rec(2018)2 of the Comm. of Ministers to Member States on the Roles and Responsibilities of Internet Intermediaries*, COUNCIL OF EUR. (Mar. 7, 2018), https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=0900001680790e14.

211. ECHR, *supra* note 183, art. 6(1).

212. *Id.* art. 6.

contract obligations therefore obliging IRUs and ICT companies to provide notice of a pending determination and of the right to challenge that determination before a fair, independent, and impartial tribunal and requiring states to safeguard the Article 6(1) rights discussed below in sub-section (3).

(1) Terms of Service Violations Involving Civil Rights (Contract Rights)

The ECtHR has held that the concept of "civil rights and obligations" has an autonomous meaning, such that a state's classification of the right is not decisive.[213] The ECtHR generally bases the concept on a distinction between public and private law, holding that private law rights are always "civil rights and obligations." The uniform position in European national law is that the rights and obligations of private persons in their relations are "civil rights and obligations," including contract rights and obligations, and torts.[214] Terms of service (ToS) are binding contractual obligations to which users must agree[215] in order to use Internet companies' services and generally contain clauses that require users not to post content that is illegal, contains hate speech, or incites violence and terrorism. Therefore posting such speech would be a violation of these terms, which would in turn constitute a breach of a contract obligation. In such situations, Internet companies may suspend or terminate the contract.[216]

---

213. HARRIS, O'BOYLE & WARBRICK, *supra* note 188, at 379 (citing Konig v. Germany, App. No. 6232/73 (Eur. Ct. H.R. Sept. 6, 1978)).

214. *Id.* at 380; *see also* WILLIAM A. SCHABAS, THE EUROPEAN CONVENTION ON HUMAN RIGHTS: A COMMENTARY 273 (2015) ("[O]bviously, article 6 will apply to private law disputes concerning tort or civil responsibility and contractual matters.").

215. *See e.g.*, *Twitter Terms of Service*, TWITTER (Sept. 30, 2016), https://twitter.com/tos?lang=en ("1. Who May Use the Services? You may use the Services only if you agree to form a binding contract with Twitter").

216. *See, e.g.*, *id.* ("You may use the Services only in compliance with these Terms [including the Twitter Rules] and all applicable laws, rules and regulations."); *Google Terms of Service*, GOOGLE (Apr. 14, 2014), https://www.google.com/policies/terms/ ("Don't misuse our Services. . . . You may use our Services only as permitted by law . . . . We may suspend or stop providing our Services to you if you do not comply with our terms or policies or if we are investigating suspected misconduct."); *Statement of Rights and Responsibilities*, FACEBOOK (Jan. 30, 2015), https://www.facebook.com/terms ("You will not post content that: is hate speech, threatening, or pornographic; incites violence . . . . You will not post content or take any action on Facebook that infringes or violates

(2) An IRU Referral is Part of a "Determination" of a Terms of Service Violation

While it is obvious that content removal by Internet companies is the result of a determination of a ToS violation, the more difficult question is whether an IRU referral is also a determination. The ECtHR has held that for proceedings to constitute a "determination," they must be "directly decisive," and a "tenuous connection or remote consequences do not suffice."[217] On its face, an IRU referral does not appear to qualify as a "determination." However, because IRU referrals initiate and result in "determinations" of terms of service violations, they arguably form part of the process of "determination." Therefore, Article 6(1) would be engaged in from the moment a referral is made.

(3) Article 6(1) Rights–Implications for IRU Referrals/Content Takedown Determinations

This Section concludes by arguing that, because states are obliged by Article 6(1) to provide rights of effective access to a court, access to a fair hearing, equality of arms, and an independent and impartial tribunal established by law, states must fashion a regulatory regime that ensures that individuals whose content is taken down are notified before the "determination" that their content is in violation of the ToS and given an effective means of challenging or appealing a determination. Although states have a margin of appreciation or degree of discretion in determining the precise details of their regulatory regimes, the regimes have to comply with the principle of proportionality and may not impair or destroy the very essence of these rights.

Article 6(1) guarantees access to a court, though this is not an absolute right.[218] Although not expressly mentioned in the text of the ECHR, the ECtHR has held that the right to access a court can be inferred from the text of the ECHR as it is a key feature of the rule of law and is the only interpretation that would prevent states from

---

someone else's rights or otherwise violates the law. . . . If you violate the letter or spirit of this Statement, or otherwise create risk or possible legal exposure for us, we can stop providing all or part of Facebook to you.").

217. Le Compte v. Belgium, App. Nos. 6878/75, 7238/75, ¶ 47 (Eur. Ct. H.R. June 23, 1981).

218. Golder v. United Kingdom, App. No. 4451/70, ¶ 24 (Eur. Ct. H.R. Feb. 21, 1975).

avoiding their ECHR obligations by closing their courts.[219] This right is one of effective access to the courts and may entail the state provision of legal assistance,[220] depending on the facts of the case, as seen in *Steel and Morris v. UK*.[221] While there is no right as such to civil legal aid, "Article 6(1) may sometimes compel the state to provide for the assistance of a lawyer when such assistance proves indispensable for an effective access to court"[222] or is needed to ensure procedural fairness or equality of arms, depending "upon the importance of what is at stake for the applicant in the proceedings, the complexity of the relevant law and procedure and the applicant's capacity to represent him or herself effectively."[223] Legal aid is not required when there is no arguable case on the facts.[224]

Although the right of access to a court is not absolute and states may restrict or regulate the right of access, they may not impose such restrictions such that "the very essence of the right is impaired."[225] The restrictions must also "have a 'legitimate aim' and comply with the principle of proportionality[.]"[226] This means that although states have some discretion, their regulations may not destroy the right of access and must be justified on the basis of human rights principles. In practice, because the ToS of U.S. Internet companies, such as Facebook, Google, and Twitter, require all disputes to be brought in U.S. courts, ECHR states must either provide adequate legal aid for those individuals whose content is not clearly unlawful so that they may bring a case to the U.S. courts or provide an alternative forum where users can effectively challenge or appeal determinations that their content is in violation of ToS.

Article 6(1) also entitles everyone to a "fair and public hearing within a reasonable time by an independent and impartial tribunal established by law,"[227] which includes a right to participate effectively,

---

219. HARRIS, O'BOYLE & WARBRICK, *supra* note 188, at 388–89.

220. Airey v. Ireland, App. No. 6289/73, ¶ 26 (Eur. Ct. H.R. Oct. 9, 1979).

221. Steel and Morris v. United Kingdom, App. No. 68416/01, ¶¶ 68–69, 72 (Eur. Ct. H.R. Feb. 15, 2005).

222. HARRIS, O'BOYLE & WARBRICK, *supra* note 188, at 400 (quoting Airey v. Ireland, App. No. 6289/73, ¶ 26 (Eur. Ct. H.R. Oct. 9, 1979)).

223. Steel and Morris v. United Kingdom, App. No. 68416/01, ¶ 61 (Eur. Ct. H.R. Feb. 15, 2005).

224. HARRIS, O'BOYLE & WARBRICK, *supra* note 188, at 400.

225. Ashingdane v. United Kingdom, App. No. 8225/78, ¶ 57 (Eur. Ct. H.R. May 28, 1985).

226. *Id.*

227. ECHR, *supra* note 183, art. 6.

a right to an adversarial trial, and the principle of equality of arms. In the IRU context, this means that users are entitled to an initial determination of the lawfulness of content by a fair, independent, and impartial decision-maker, which allows a user an effective opportunity to defend the content either before a referral is made, or in exceptional cases, on appeal after a referral decision. In practice, although many users may not exercise this right, it is important that IRUs and ICT companies notify users of this right in order for it to be effective. Similarly, states and ICT companies must provide adequate alternatives if they wish to minimize the number of people using civil proceedings while ensuring the expeditiousness of content takedowns. This would be in line with Article 13 and the recommendation of the Committee of Ministers to member states that "[s]ocial networking providers should respect human rights and the rule of law[.]"[228]

E. Article 10 ECHR and IRU Referrals

As mentioned earlier, while IRUs may be justified on the basis of Article 17, evaluating them against Article 10 would better demonstrate society's normative commitment to the fundamental value of freedom of expression and better ensure that IRUs are constrained by the safeguards against unnecessary censorship that have been elaborated in Article 10(2).

In general, when determining whether there has been a violation of Article 10, the ECtHR relies upon the following analysis: (1) identify the rights engaged; (2) identify the interference by a public authority with the rights engaged; (3) determine whether the interference is prescribed by law; (4) determine what objectives the interference aims to protect; and (5) decide "whether the interference is 'necessary in a democratic society,' *i.e.* whether the state gives, and gives evidence for, relevant and sufficient reasons for the interference" and whether, allowing the state a margin of appreciation, "those reasons are proportionate to the limitation of the applicant's enjoyment of his right."[229]

---

228. *Recommendation CM/Rec(2012)4 of the Committee of Ministers to member States on the protection of human rights with regard to social networking services*, COUNCIL OF EUR. (Apr. 4, 2016), https://search.coe.int/cm/Pages/result_details.aspx?Reference=CM/Rec(2012)4.

229. HARRIS, O'BOYLE & WARBRICK, *supra* note 188, at 521.

(1) Identification of the Rights Engaged: Negative and Positive Obligations?

While Article 10 is classically seen as imposing negative obligations on the state not to interfere with freedom of expression, the Grand Chamber of the ECtHR has recently confirmed that Article 10 may also impose positive obligations on the state to protect freedom of expression against interference by private persons.[230]

It may therefore be argued that IRUs (1) interfere with the rights to freedom of expression and to impart information and ideas of the persons whose speech is referred to ICT companies, (2) interfere with the public's right to receive information,[231] and (3) interfere with the ICT companies' right to publish information and ideas; but also that (4) states have a positive obligation to ensure that ICT companies respect users' and content providers' freedom of expression.

In determining whether a positive obligation exists in a particular situation, the ECtHR attempts to strike a "fair balance" between the general interest of the community and the interests of the individual. The ECtHR has stated that:

> The scope of [the positive] obligation will inevitably vary, having regard to the diversity of situations obtaining in Contracting States and the choices which must be made in terms of priorities and resources. Nor must such an obligation be interpreted in such a way as to impose an impossible or disproportionate burden on the authorities.[232]

The ECtHR has considered the following factors to be relevant: the nature of the expression rights at stake, e.g., political, artistic, commercial, hate speech; the capacity of the expression to contribute to debate on a topic of public interest or about the exercise of public powers; the nature and scope of the restrictions on expression rights; the availability of alternative venues for expression; and the weight of

---

230. Palomo Sanchez v. Spain, App. Nos. 28955/06, 28957/06, 28959/06, 28964/06, ¶ 59 (Eur. Ct. H.R. Sept. 12, 2011) (citing Dink v. Turkey, App. Nos. 2668/07, 6102/08, 30079/08, 7072/09, 7124/09, ¶ 106 (Eur. Ct. H.R. Sept. 14, 2010); Özgür Gündem v. Turkey, App. No. 23144/93, ¶¶ 42–46 (Eur. Ct. H.R. Mar. 16, 2000); Fuentes Bobo v. Spain, App. No. 39293/98, ¶ 38 (Eur. Ct. H.R. May 5, 2000)).

231. Observer & Guardian v. UK, App. No. 13585/88, ¶ 59 (Eur. Ct. H.R. Nov. 26, 1991).

232. Appleby v. United Kingdom, App. No. 44306/98, ¶ 40 (Eur. Ct. H.R. May 6, 2003).

countervailing rights of others or the public.[233] Thus, the ECtHR has held that there are positive obligations to protect journalists who have been threatened for their political speech[234] and, in some cases, where employees have been dismissed after publicly criticizing their employers,[235] but not in a case where the applicant claimed that the state has a positive obligation to force a private publisher to publish his commercial advertisement.[236]

In *Appleby v. United Kingdom*, the ECtHR rejected applicants' claim that Article 10 imposes a positive obligation on the state to secure a "freedom of forum" for them to access a privately-owned shopping center in order to set up a stand to campaign against a local council decision.[237] The ECtHR was swayed by the limited nature of the shopping center's restrictions, which only prevented stands in passageways and the entrance area, and the presence of alternative

---

233. REID, *supra* note 192, at 617 (citing *id.* ¶¶ 42–43, 47–49); *Frăsilă and Ciocîrlan v. Romania*, App. No. 25329/03, ¶ 55 (Eur. Ct. H.R. May 10, 2012); *Remuszko v. Poland*, App. No. 1562/10, ¶ 65 (Eur. Ct. H.R. July 16, 2013).

234. Dink v. Turkey, App. Nos. 2668/07, 6102/08, 30079/08, 7072/09, 7124/09, ¶ 106 (Eur. Ct. H.R. Sept. 14, 2010) (finding that Turkey had a positive obligation under Article 10 to protect an Armenian journalist against attack by members of an extreme nationalist group); Özgür Gündem v. Turkey, App. No. 23144/93, ¶¶ 42–46 (Eur. Ct. H.R. Mar. 16, 2000) (finding that Turkey had a positive obligation under Article 10 to take investigative and protective measures where the journalists and staff of a pro-PKK newspaper had been the victims of a campaign of violence and intimidation).

235. Fuentes Bobo v. Spain, App. No. 39293/98, ¶ 38 (Eur. Ct. H.R. May 5, 2000) (finding that Spain had a positive obligation to protect the freedom of expression of a TV producer who was fired after he made statements criticizing his employer in two radio broadcasts); *but see* Palomo Sanchez v. Spain [GC], App. Nos. 28955/06, 28957/06, 28959/06, 28964/06, ¶¶ 76–79 (Eur. Ct. H.R. Sept. 12, 2011) (finding that Spain had no positive obligation to annul dismissal of employees by private company after they published insulting material about their employers in a union newsletter).

236. Remuszko v. Poland, App. No. 1562/10, ¶ 65 (Eur. Ct. H.R. July 16, 2013).

237. Appleby v. United Kingdom, App. No. 44306/98, ¶¶ 13–47 (Eur. Ct. H.R. May 6, 2003); *but see* Pruneyard Shopping Center v. Robbins, 447 U.S. 74, 81–85 (1980) (allowing state constitutional provisions to adopt more expansive liberties than the Federal Constitution, permitting individuals reasonably to exercise free speech and petition rights on the property of a privately-owned shopping center to which the public was invited; this did not violate the property rights of the shopping-center owner so long as any restriction did not amount to taking without compensation or contravene any other federal constitutional provisions). The applicants in *Appleby v. United Kingdom* sought to rely on *Pruneyard Shopping Centre v. Robbins* and extensive state law jurisprudence.

means, such as getting permission to set up stands within individual stores, distributing leaflets on public paths, or door-to-door calling.[238] However, the ECtHR explicitly left open the possibility that "a positive obligation could arise for the State to protect the enjoyment of the Convention rights [such as Article 10] by regulating property rights," citing a corporate town where the entire municipality is controlled by a private body as a possible example.[239]

The ECtHR has allowed states to impose positive obligations on Internet companies and has also imposed positive obligations to protect rights online. The ECtHR has already demonstrated a willingness to allow states to impose liability—and thus positive obligations—on online news portals to remove hate speech and incitements to violence.[240] It has also permitted states to convict individuals for online copyright infringements,[241] based on state laws imposing a positive obligation not to infringe copyrights, and required—or imposed a positive obligation on—states to have legal frameworks that would require Internet service providers to identify users who advertise child sexual abuse online.[242] In the last of these cases, *K.U. v. Finland*, the ECtHR reasoned that "users of telecommunications and Internet services must have a guarantee that their own privacy and freedom of expression will be respected," but such protection "cannot be absolute and must yield on occasion to other legitimate imperatives, such as the prevention of disorder or crime or the protection of the rights and freedoms of others."[243] The ECtHR went on to explain that the legislature had to strike a balance between these competing social goods, but that the outcome cannot not require

238. REID, *supra* note 192, at 619 (citing Appleby v. United Kingdom, App. No. 44306/98, ¶ 48 (Eur. Ct. H.R. May 6, 2003)).

239. Appleby v. United Kingdom, App. No. 44306/98, ¶ 47 (Eur. Ct. H.R. May 6, 2003) (citing *Marsh v. Alabama*, 326 U.S. 501 (1946), which held that a privately-owned corporate company town with all the characteristics of other municipalities was subject to the First Amendment rights of free speech and peaceable assembly).

240. Delfi AS v. Estonia, App. No. 64569/09, ¶ 6 (Eur. Ct. H.R. June 6, 2015).

241. *See* Ashby Donald v. France, App. No. 36769/08, ¶¶ 44, 45 (Eur. Ct. H.R. Jan. 10, 2013); Neij v. Sweden, App. No. 40397/12, 11 (Eur. Ct. H.R. Feb. 19, 2013) (affirming Sweden's conviction of defendants for torrenting copyrighted data).

242. K. U. v. Finland, App. No. 2872/02, ¶ 49 (Eur. Ct. H.R. Feb. 3, 2009).

243. *Id.*; *see also* Case C-131/12, Google Spain SL v. Agencia Española de Protección de Datos, 2014 E.C.R. 317, ¶¶ 92–94 (where the Court of Justice of the European Union demonstrated a willingness to impose positive obligations on search engines to index search results that are "inadequate, irrelevant or no longer relevant, or excessive in relation to [the purposes for which the data was collected and processed] and in the light of the time that has elapsed").

user confidentiality at the expense of the positive obligation to effectively deter sexual crimes against minors.[244]

Should states have a positive obligation to ensure that ICT companies respect users' and content providers' freedom of expression? Freedom of expression, which may include the freedom to use speech that offends, shocks, or disturbs, is fundamentally important in a democracy.[245] The capacity of freedom of expression on the Internet to contribute to public debates,[246] combined with the dominance of a few Internet companies in providing platforms where vigorous public debate happens and the fact that these Internet companies already moderate user-generated expression for violations of their ToS, supports an ECtHR requirement for states to ensure that these companies give sufficient regard to the protection of freedom of expression when formulating and implementing ToS. That said, states should bear in mind the need to take a "differentiated and graduated approach," as defined by the Council of Europe, in deciding what "duties and responsibilities" under Article 10(2) to delegate to internet companies.[247] Such a positive obligation does not have to privilege freedom of expression over the competing social goods mentioned above, but should ensure that freedom of expression is appropriately valued when the companies are designing and implementing ToS. No such obligation need be imposed on intermediaries that do not modify third-party content or companies that do not moderate user-generated content.

On the other hand, it may be argued that Internet companies that moderate user-generated content are private individuals with property rights and economic interests that may be infringed by the coercive imposition of a positive obligation to respect users' freedom of expression. It may also be argued that there are both private and public interests in allowing these companies to adopt their own content moderation policies, including the preservation of "safe" communities where offensive content and online abuse against women and other

---

244. K. U. v. Finland, App. No. 2872/02, ¶ 49 (Eur. Ct. H.R. Feb. 3, 2009).

245. Handyside v. United Kingdom, App. No. 5493/72, ¶ 49 (Eur. Ct. H.R. Dec. 7, 1976).

246. Ahmet Yildirim v. Turkey, App. No. 3111/10, ¶¶ 48, 54 (Eur. Ct. H.R. Dec. 18, 2012); Times Newspapers Ltd (Nos. 1 and 2) v. United Kingdom, App. Nos. 3002/03, 236676/03, ¶ 27 (Eur. Ct. H.R. Mar. 10, 2009); Delfi AS v. Estonia, App. No. 64569/09, ¶ 110 (Eur. Ct. H.R. June 16, 2015).

247. *Recommendation CM/Rec(2011)7*, *supra* note 207, ¶ 7.

vulnerable groups are taken down,[248] even if such content would be protected against state censorship. These are compelling arguments, which the ECtHR is unlikely to take lightly. The ECtHR is likely to conduct a proportionality analysis in weighing the competing rights, interests, and public goods to "strike a fair balance" between them and may be persuaded to adopt a narrowly-tailored, positive obligation on Internet companies to respect freedom of expression when moderating content. This result would be in line with the BHR Guiding Principles, which find that states have the primary duty to respect human rights, but also an obligation to ensure that businesses respect human rights. This would also be in line with the recommendation of the Committee of Ministers to member states that "social networking providers should respect human rights and the rule of law."[249]

The beneficial result of having such a narrowly tailored positive obligation is that ICT companies will not over-censor content for fear of facing liability. However, even if the ECtHR does not find in favor of such a positive obligation, it may still find that IRUs constitute unjustified interferences with Article 10.

(2) Whether IRU Referrals Constitute an Interference with Article 10

The question of whether an IRU referral of content constitutes an "interference by public authority" is complicated by the purportedly voluntary nature of the referral and the fact that the decision to remove content is made by the Internet company to whom the referral is made. States may rely on these points to argue that there is no interference with Article 10. On the other hand, Ben Emmerson QC, one of the United Kingdom's most renowned ECtHR practitioners[250] and the U.N. Special Rapporteur on Counter-Terrorism and Human Rights,[251] has stated his view that "any measure taken to prevent or remove messages communicated through the Internet or other forms of technology constitute an interference with the right to freedom of expression and must be justified. . . . Independent judicial recourse

248. *See, e.g.*, *Community Standards*, FACEBOOK, https://www.facebook.com/communitystandards (last visited Nov. 4, 2017) ("We want people to feel safe when using Facebook.").

249. *Recommendation CM/Rec(2012)4*, *supra* note 228.

250. *See Ben Emmerson QC: Practice CV*, MATRIX CHAMBERS (2016), https://www.matrixlaw.co.uk/wp-content/uploads/2016/03/Ben-Emmerson.pdf.

251. *See* Section II.A *supra*.

must be available."[252] There is almost no ECtHR case law on these novel questions, aside from the case of *Saliyev v. Russia*,[253] discussed below. That said, analogies may be drawn to cases involving the use of injunctions to restrain publication[254] or the failure to investigate and protect members of the press from violence and threats.[255] On balance, if it can be shown that referral requests are accompanied by coercive pressures from the state, it is probable that the ECtHR will extend its jurisprudence to find that there is an interference with Article 10, even though Internet companies such as Facebook or Twitter are arguably not public authorities.

The case of *Saliyev v. Russia* involved the withdrawal from circulation of an edition of a privately-owned newspaper containing an article alleging that government corruption was accepted.[256] While some of the principal facts were disputed,[257] the ECtHR found that the editor-in-chief had withdrawn the newspaper because of the applicant's article, out of fear of possible sanctions related to the content of the article.[258] The ECtHR reasoned that while there is no general right of access to the media, save in instances where the media is controlled by a monopoly, this case did not concern a right of access to the media, but rather an "interference" with the applicant's rights under Article 10. The Court explained that this was an interference despite the fact that the withdrawal was made by the editor-in-chief, a private citizen, of a privately-owned newspaper because the article was already in the public domain and the newspapers were withdrawn because of the content of the applicant's article, out of fear of possible sanction.[259] The implication of this case is that the ECtHR held that a state-coerced withdrawal of a publication by a privately-owned

252. Emmerson, *supra* note 140, ¶ 40.

253. Saliyev v. Russia, App. No. 35016/03 (Eur. Ct. H.R. Oct. 21, 2010).

254. *See, e.g.*, Sunday Times v. United Kingdom, App. No. 6538/74, ¶¶ 42–68 (Eur. Ct. H.R. Apr. 26, 1979) (holding injunction restraining publication of news on basis that publication would constitute contempt of court is a violation of Article 10).

255. Özgür Gündem v. Turkey, App. No. 23144/93, ¶¶ 41–42 (Eur. Ct. H.R. Mar. 16, 2000).

256. Saliyev v. Russia, App. No. 35016/03, ¶¶ 7–9 (Eur. Ct. H.R. Oct. 21, 2010).

257. This includes whether the editor-in-chief had withdrawn the newspaper after a private conversation that led him to believe the politicians accused of corruption were "untouchable."

258. *Id.* ¶¶ 12, 52–61.

259. *Id.* ¶ 56.

newspaper could be characterized as an "interference" with the applicant's rights under Article 10. Applicants bringing challenges to IRUs may be able to analogize the editor-in-chief's actions here to the ICT companies "voluntarily" removing content after a government request, if there is coercive pressure, such as the threat of litigation or liability, in the background.

The ECtHR then went on to examine the question of whether the interference was done by a public authority, relying on a test it had previously set out in *Radio France v. France*:

> In order to determine whether any given legal person other than a territorial authority falls within the category [of 'governmental organisations'], account must be taken of its legal status and, where appropriate, the rights that status gives it, the nature of the activity it carries out and the context in which it is carried out, and the degree of its independence from the political authorities. [260]

On the facts of the case, the ECtHR found that the newspaper was a state authority.[261] However, this test should not be taken to be definitive of what constitutes an "interference by public authority."[262] The *Radio France* test was originally developed to determine what constitutes "governmental organisations" as opposed to "non-governmental organisations" within the meaning of Article 34. Its purpose was to determine whether Radio France could qualify as a "non-governmental organisation" in order to bring a case against France.[263] In the IRU context, skillful human rights lawyers should be able to argue that what is determinative of an "interference by public authority" is whether a third party acts or does not act as a result of coercive pressure from the state, therefore persuading the ECtHR to extend its jurisprudence in this area.

In a case concerning the UK CTIRU, the ECtHR is likely to find that CTIRU referrals are not truly voluntary as they are accompanied by coercive pressures. Sections 1 and 2 of the Terrorism Act of 2006[264] prohibit the intentional or reckless publication and

---

260. *Id.* ¶ 64 (citing Radio France v. France, App. No. 53984/00, Decision [Extracts], ¶ 26 (Eur. Ct. H.R. Sept. 23, 2003)).

261. Saliyev v. Russia, App. No. 35016/03, ¶ 69 (Eur. Ct. H.R. Oct. 21, 2010).

262. ECHR, *supra* note 183, art. 10.

263. Radio France v. France, App. No. 53984/00, Decision [Extracts], ¶¶ 24–26 (Eur. Ct. H.R. Sept. 23, 2003).

264. Terrorism Act 2006, c.11 §§ 1–2 (UK).

dissemination of material that encourages, glorifies, or incites acts of terrorism. Although these sections have never been used to prosecute ICT companies, the Solicitor-General stated in recent testimony before the Home Affairs Committee that social media companies could be charged for reckless dissemination.[265] The Home Affairs Committee has also expressed its support for penalizing social media companies for failing to remove illegal content, including violations of the Terrorism Act of 2006, recommending that the UK "[g]overnment consult on a system of escalating sanctions to include meaningful fines for social media companies which fail to remove illegal content within a strict timeframe."[266] The Home Affairs Committee has also criticized social media companies for profiting from advertising next to content produced by ISIS supporters, stating that "it is shameful that they have failed to use the same ingenuity to protect public safety and abide by the law as they have to protect their own income."[267] The UK Government has also withdrawn all advertising, worth £3,878,600 in 2016, from YouTube after reports that ads appeared alongside YouTube videos created by supporters of terrorist groups such as ISIS. The Government stated that its advertising "will not be reactivated until such time as Google can give definitive assurance that government messages will be delivered in a safe and appropriate way."[268] Given coercive pressure, including the threat of criminal sanction and the withdrawal of all government advertising, it will be difficult to convince a court that CTIRU referrals are purely voluntary requests, especially in a media environment where both government and opposition politicians often aim to score positive publicity by criticizing ICT companies.[269]

---

265. HOME AFFAIRS COMM., *supra* note 26, Q673, Q688, Q745.

266. HOME AFFAIRS COMM., HATE CRIME AND ITS VIOLENT CONSEQUENCES, 2016-17, HC 609, ¶ 37 (UK).

267. *Id.* ¶ 36.

268. *Id.* ¶ 23.

269. Madhumita Murgia, *Tech firms pledge to improve response to terror propaganda*, FIN. TIMES (Mar. 31, 2017), https://www.ft.com/content/32c861c0-156c-11e7-80f4-13e067d5072c (reporting a statement by Yvette Cooper, the Labour MP who chairs the Home Affairs Committee, regarding Google's "lack of effort and social responsibility"); Press Release, Amber Rudd, Home Sec'y, Home Secretary statement: meeting with Communication Service Providers (Mar. 30, 2017), https://www.gov.uk/government/news/home-secretary-statement-meeting-with-communication-service-providers (stating "I'd like to see the industry to go further and faster in not only removing online terrorist content but stopping it going up in the first place").

(3) Whether IRU Interferences Are Prescribed By Law

For an interference to be "prescribed by law," it must pass the tests of *accessibility* and *foreseeability* and include sufficient *safeguards against abuse.*[270] *Accessibility* means that the citizen must be "able to have an indication that is adequate in the circumstances of the legal rules applicable to a given case."[271] *Foreseeability* means that the law must be "formulated with sufficient precision to enable the citizen—if need be, with appropriate advice—to foresee, to a degree that is reasonable in the circumstances, the consequences which a given action may entail."[272] The requirement of sufficient safeguards against arbitrary interference means that any available procedure of judicial review must be effective in constraining state authorities' abuse.[273] Thus, in *Ahmet Yıldırım v. Turkey*, the ECtHR held that blocks on websites are not necessarily incompatible with the ECHR *per se*, but that:

> [A] legal framework is required, ensuring both tight control over the scope of bans and effective judicial review to prevent any abuse of power . . . . In that regard, the judicial review of such a measure, based on a weighing-up of the competing interests at stake and designed to strike a balance between them, is inconceivable without a framework establishing precise and specific rules regarding the application of preventive restrictions on freedom of expression.[274]

The ECtHR found that "the measure in question produced arbitrary effects and could not be said to have been aimed solely at blocking access to the offending website, since it consisted in the wholesale blocking of all the sites hosted by Google Sites," and concluded by finding that:

> [T]he judicial review procedures concerning the blocking of Internet sites are insufficient to meet the criteria for avoiding abuse, as domestic law does not provide for any safeguards to ensure that a blocking

---

270. HARRIS, O'BOYLE & WARBRICK, *supra* note 188, at 649.

271. *Id.* (citing Sunday Times v. United Kingdom, App. No. 6538/74, ¶ 49 (Eur. Ct. H.R. Apr. 26, 1979)).

272. *Id.* (citing Muller v. Switzerland, App. No. 10737/84, ¶ 29 (Eur. Ct. H.R. May 24, 1988)).

273. Ahmet Yıldırım v. Turkey, App. No. 3111/10, ¶ 68 (Eur. Ct. H.R. Dec. 18, 2012).

274. *Id.* ¶¶ 57–68.

> order in respect of a specific site is not used as a means of blocking access in general.[275]

Although the UK Government may attempt to argue that CTIRU referrals are prescribed by law by pointing to sections 1 and 2 of the Terrorism Act of 2006 to indicate to citizens what types of material will be referred to Internet companies, the CTIRU's operational framework is deficient for a number of reasons. Firstly, because there is no legal framework constraining the operation of the CTIRU, the choice of these laws as the basis for referral is entirely a matter of executive discretion: the UK Government could make an executive decision tomorrow to refer other types of content, including copyrighted content, fake news, hate speech, or "non-violent extremist" content. Judicial review of such decisions under the UK's Human Rights Act[276] is an ineffective safeguard as individuals will find it difficult to prove that they are victims of a violation of their ECtHR rights since they will not know if their content was removed as the result of a IRU referral. Classical judicial review in administrative law is also unlikely to prove to be a sufficient safeguard, as long as the Government confines itself to referring different types of illegal content and argues that this is an executive decision for which it should be granted a large measure of judicial deference.

Secondly, because the IRU operates on the basis of policy decisions instead of a legal framework and most of its operational details are shrouded behind the UK Government's claim of "national security,"[277] citizens will find it difficult to foresee what types of terrorism-related content will be referred. For example, will pictures of beheadings in a news report be referred? Will counter-narratives that contain extremist content be referred?).

Thirdly, because the IRU does not presently notify persons whose content is referred for removal, they will find it difficult to prove that their content was removed as a result of a governmental act or to claim victim status for the purposes of the Human Rights Act, resulting in insufficient judicial safeguards against arbitrary interference.

Finally, as a general criticism, the foreseeability test requires "quality" laws that are not excessively vague, which is always an issue

---

275. *Id.* ¶ 68.

276. Human Rights Act 1998, c.42 (UK).

277. 17 Mar. 2016 Parl Deb (2016) col. 30893 (UK), http://www.parliament.uk/written-questions-answers-statements/written-question/commons/2016-03-14/30893.

with offenses involving "incitement" or "glorification" of "terrorism" or "extremism." The Section 1 "glorification of terrorism" offense in the Terrorism Act of 2006 was heavily criticized during its legislative debate for being overbroad and for potentially encompassing opposition to totalitarian regimes worldwide. Labour MP Jeremy Corbyn argued that "[w]hat some would call a freedom fight going on in another country others might term a terrorist offence. Nelson Mandela was branded a terrorist by Margaret Thatcher; he was later branded a freedom fighter."[278]

This analysis demonstrates the value of using Article 10 rules to analyze IRUs rather than the Article 17. The Article 10 approach will ensure that the law is accessible and sufficiently precise to enable citizens to regulate their conduct and that it includes sufficient safeguards against arbitrary interference with the right to freedom of expression.

(4) Whether IRU Interferences Are Justified

(i) The Precedential Scope of Delfi AS v. Estonia

Whether the interference is justified—or "necessary in a democratic society"—is a question that will have to be addressed on its own merits. While there may be a temptation to generalize principles from *Delfi AS v. Estonia*[279] as it is a Grand Chamber judgment, this temptation should be resisted. Grand Chamber judgments are ordinarily given more weight than Chamber judgments, since they are decided by seventeen judges rather than seven. However, one of the judges who decided the case, Judge Robert Spano, has warned that "the case is to some extent unique and unsuitable for broad interpretive conclusions over and above the facts presented by the case."[280] The ECtHR itself emphasized that the holding was limited to "large professionally managed Internet news portals run on a commercial basis, which published news articles of its own and invited its readers to comment on them." The ECtHR further provided that the holding did not concern:

278. Matthew Tempest, *Terrorism Act comes into force*, GUARDIAN (Apr. 13, 2006), https://www.theguardian.com/politics/2006/apr/13/uksecurity.terrorism.

279. Delfi AS v. Estonia, App. No. 64569/09, ¶ 115 (Eur. Ct. H.R. June 16, 2015).

280. Robert Spano, *Intermediary Liability for Online User Comments under the European Convention on Human Rights*, 17 HUM. RTS. L. REV. 1, 15 (2017).

> other fora on the Internet where third-party comments can be disseminated, for example an Internet discussion forum or a bulletin board where users can freely set out their ideas on any topics without the discussion being channeled by any input from the forum's manager; or a social media platform where the platform provider does not offer any content and where the content provider may be a private person running the website or a blog as a hobby.[281]

Properly seen in its context, the precedential scope of *Delfi AS v. Estonia* may be limited to the following: States *may* impose liability on "*active* intermediaries," like *Delfi*, for not removing user-generated content that amounts to hate speech and incitement to violence against a person where the user comments are clearly unlawful and have been posted anonymously or under a pseudonym and there are no domestic mechanisms in place to afford the injured party a real and effective opportunity to pursue the actual authors.[282] This holding is reinforced by two subsequent cases finding that the rules concerning defamation online are different.[283] In *Magyar T.E. and Index.hu Zrt v. Hungary* and *Pihl v. Sweden*, the ECtHR took the position that online speech that is merely defamatory or otherwise listed in Article 10(2) should be dealt with using the Article 10(2) framework—(1) interference prescribed by law; (2) legitimate aim and necessity in a democratic society. This arguably creates a safeguard against abuse of IRUs to initiate removal of content without judicial oversight other than speech that violates Article 17.

(ii) The Tests for Necessity and Proportionality

The ECtHR has summarized the fundamental principles underlying whether an interference with freedom of expression is "necessary in a democratic society" as follows:

> (i) Freedom of expression constitutes one of the essential foundations of a democratic society and one of the basic conditions for its progress and for each individual's self-fulfilment. Subject to paragraph 2 of Article 10, it is applicable not only to 'information' or

281. *Id.*

282. Spano, *supra* note 280, at 15.

283. Magyar T.E. and Index.hu Zrt v. Hungary, App. No. 22947/13, ¶ 69 (Eur. Ct. H.R. Feb. 2, 2016); Pihl v. Sweden, App. No. 74742/14, ¶ 28 (Eur. Ct. H.R. Mar. 9, 2017).

> 'ideas' that are favourably received or regarded as inoffensive or as a matter of indifference, but also to those that offend, shock or disturb. Such are the demands of pluralism, tolerance and broadmindedness without which there is no 'democratic society'. As set forth in Article 10, this freedom is subject to exceptions, which ... must, however, be construed strictly, and the need for any restrictions must be established convincingly . . . .
> (ii) The adjective 'necessary', within the meaning of Article 10 § 2, implies the existence of a 'pressing social need'. The Contracting States have a certain margin of appreciation in assessing whether such a need exists, but it goes hand in hand with European supervision, embracing both the legislation and the decisions applying it, even those given by an independent court. The Court is therefore empowered to give the final ruling on whether a 'restriction' is reconcilable with freedom of expression as protected by Article 10.
> (iii) The Court's task, in exercising its supervisory jurisdiction, is not to take the place of the competent national authorities but rather to review under Article 10 the decisions they delivered pursuant to their power of appreciation. This does not mean that the supervision is limited to ascertaining whether the respondent State exercised its discretion reasonably, carefully and in good faith; what the Court has to do is to look at the interference complained of in the light of the case as a whole and determine whether it was 'proportionate to the legitimate aim pursued' and whether the reasons adduced by the national authorities to justify it are 'relevant and sufficient' . . . In doing so, the Court has to satisfy itself that the national authorities applied standards which were in conformity with the principles embodied in Article 10 and, moreover, that they relied on an acceptable assessment of the relevant facts. . . .[284]

While the ECtHR has not developed a consistent methodological approach toward proportionality in its Article 10 jurisprudence, the UK Supreme Court has distilled the ECtHR's approach towards proportionality into the following four-step analysis:

---

284. Delfi AS v. Estonia, App. No. 64569/09, ¶ 131 (Eur. Ct. H.R. June 16, 2015).

> [I]t is necessary to determine:
> (1) whether the objective of the measure is sufficiently important to justify the limitation of a protected right,
> (2) whether the measure is rationally connected to the objective,
> (3) whether a less intrusive measure could have been used without unacceptably compromising the achievement of the objective, and
> (4) whether, balancing the severity of the measure's effects on the rights of the persons to whom it applies against the importance of the objective, to the extent that the measure will contribute to its achievement, the former outweighs the latter. . . . In essence, the question at step four is whether the impact of the rights infringement is disproportionate to the likely benefit of the impugned measure.[285]

These tests will be applied in the evaluation of IRUs in the next section.

(iii) Evaluating IRUs: Justified in Limited Circumstances

While Article 10(2) allows for the limitation of the right to freedom of expression for a number of reasons, it is arguable that the objectives of protecting national security and public safety and the prevention of public disorder from terrorism are the most important. Variants of arguments that "it is the first responsibility of government in a democratic society to protect and safeguard the lives of its citizens" are heard in courthouses[286] and parliaments[287] around the world. This could therefore be an important limiting device for the use of IRUs, which may only be justified for the purpose of preventing terrorism, while normal judicial processes and other less serious interferences are required for less socially important objectives, such as the protection of copyrights or reputations, in view of the fundamental importance of the right to freedom of expression.

---

285. Bank Mellat v. Her Majesty's Treasury [2013] UKSC 39, [2014] 1 AC 700, [¶ 74] (appeal taken from Eng. & Wales).

286. A and Others v. Sec'y of State for the Home Dep't [2005] UKHL 71, [2006] 2 AC (HL) 221, [¶ 99] (appeal taken from Eng. & Wales).

287. 147 CONG. REC. 21,012 (2001) (statement of Rep. Inslee) ("The first duty of government is to protect the physical security and safety of its citizens. That is the first duty of government.").

The challenge will be in addressing the expansion of IRUs to refer content relating to "serious crimes," such as human trafficking. Here, traditional protections of speech and safeguards against prosecution for merely speaking about criminal acts run up against the desire to prevent the Internet from being used as a medium to facilitate crime and perception that the unprecedented nature of the Internet in facilitating the exchange of information and flow of data can enable crime to escape law enforcement. This issue has been raised by the decision of the European Council to ask the EU IRU to detect and request removal of Internet content "used by traffickers to attract migrants and refugees."[288] The answer to whether such IRU referrals are justified is likely to turn on whether (1) "a less intrusive measure could have been used"[289] that will achieve the objective, (2) the balancing—or proportionality *stricto sensu*—analysis, and (3) the tendency of judges to insist that normal prosecutorial or judicial processes be followed in securing the removal of content or accounts of "criminals" given the extrajudicial and extra-prosecutorial process of IRU referrals.

While IRU referrals may be defended as being less restrictive than wholesale blocks on websites,[290] criminal prosecutions,[291] or the imposition of intermediary liability on Internet companies, they are still more restrictive than a court order to take down the material or a request to add references to counter-messages to the impugned content.[292] States may choose their desired level of protection and are

---

288. Press Release, EUR. COUNCIL, Special meeting of the European Council, 23 April 2015 - statement, http://www.consilium.europa.eu/en/press/press-releases/2015/04/23-special-euco-statement/ (last visited Nov. 20, 2017).

289. Bank Mellat v. Her Majesty's Treasury [2013] UKSC 39, [2014] 1 AC 700 [¶ 74] (appeal taken from Eng. & Wales).

290. *See, e.g.*, Ahmet Yıldırım v. Turkey, App. No. 3111/10, ¶ 68 (Eur. Ct. H.R. Dec. 18, 2012) (holding that it was a violation of freedom of expression to entirely block access to Google Sites given that only one user's content violated Turkish laws); Cengiz v. Turkey, App. Nos. 48226/10 and 14027/11, ¶ 64 (Eur. Ct. H.R. Dec. 1, 2015) (finding a violation of freedom of expression where YouTube was entirely blocked because of ten offending videos). *But see* Akdeniz v. Turkey, App. No. 20877/10, ¶ 25 (Eur. Ct. H.R. Mar. 11, 2014) (declining to find violation of freedom of expression despite wholesale blocking of music sites because sites were primarily used for commercial means and music was available elsewhere).

291. *But see* Ashby Donald v. France, App. No. 36769/08, ¶ 39 (Eur. Ct. H.R. Jan. 10, 2013) (declining to find a violation of freedom of expression where criminal sanctions imposed on defendants illegally selling photographs online).

292. *See* Wegrzynowski and Smolczewski v. Poland, App. No. 33846/07, ¶ 40 (Eur. Ct. H.R. July 16, 2013) (noting that Polish courts could allow a court order

given a margin of appreciation in choosing the means of securing removal of content, but this needs to be balanced against the impact on individuals' right to freedom of expression—it cannot impair or destroy the very essence of the right. Here, the existence of due process safeguards, as mentioned above in Sections III.A and III.C.iii, are necessary to ensure adequate protection of freedom of expression and the standard-setting documents of the Council of Europe may be relied upon to support the desirability of such safeguards.

F. Sub-Conclusion: The ECHR Provides Guidance on the Use and Abuse of IRUs

To conclude, while the ECHR does not prohibit the use of IRUs that target content inciting violence or terrorism for referral to Internet companies, it provides a useful set of disciplines to ensure that they are not abused, are strictly limited in purpose, and have adequate safeguards.

These include, as a minimum requirement under Article 6(1), that the state provides effective opportunities to challenge ICT companies' determinations, which may be in the form of a "fair and public hearing within a reasonable time by an independent and impartial tribunal established by law,"[293] resort to which may be minimized by the provision of adequate alternatives that enable notice, challenge, and appeal of determinations. While Article 17 may be used to justify procedures meeting the minimum standards for content that incites violence or hate speech, it may not justify the use of IRUs to take down content that is not directed against the fundamental values of the ECHR. Such content restrictions instead have to be addressed within the framework of Article 10(2). The Article 10(2) framework is preferable in general, as its use demonstrates the state's normative commitment to freedom of expression and provides more robust safeguards against abuse.

The Article 10(2) framework will require that IRUs are grounded in a legal framework that sets out the types of content that can be referred and meets the tests of accessibility, foreseeability, and sufficient safeguards against arbitrary abuse. It will also require that the decisions over what types of content are referred as well as referral decisions are subject to proportionality analysis, particularly whether

---

requesting a reference to favorable libel judgments be added to a news article online, which may be applied mutatis mutandis).

293. ECHR, *supra* note 183, art. 6.

there are less restrictive means, whether the referral has a disproportionate impact on the rights of the individuals affected, and whether there are adequate due process safeguards to ensure respect for the rights of the individuals.

Finally, as discussed in Sexction III.E.(1), the use of Article 10 in a suitable case—not necessarily limited to cases involving IRUs—presents an opportunity for courts to read into the ECHR a narrowly-tailored positive obligation on the state to require social media companies to provide adequate protection for freedom of expression, in line with both the ECHR's jurisprudence on positive obligations and the U.N. Guiding Principles on Business and Human Rights.

IV. THE EU IRU, EUROPOL REGULATION, AND THE EU CHARTER OF FUNDAMENTAL RIGHTS

A. The Europol Regulation: Sufficient Oversight and Safeguards?

Regulation 2016/794 (the Europol Regulation)[294] came into effect on May 1, 2017. This Regulation governs Europol as a whole, including the IRU, and contains legal limits, robust data protection safeguards, including a right to complain to the EU Data Protection Supervisor, and provisions for parliamentary scrutiny. This Part examines and evaluates the Europol Regulation provisions for the IRU and its safeguards and oversight. This Part reaches the conclusion that, while the new Europol Regulation grounds the IRU within a legal framework, it does not address the significant due process and transparency concerns about the IRU. It does create political oversight in the form of a Joint Parliamentary Scrutiny Group, but this oversight will be hindered by the fact that the European Parliament does not have any more privileged access to documents than an ordinary EU citizen. While it contains data protection safeguards, these are presently insufficient to address the freedom of expression and due process concerns raised by the EU IRU.

(1) Operational Provisions and Their Interpretation

Article 3 of the Europol Regulation sets out Europol's objectives: to support and strengthen action by the competent authorities of the Member States, boost mutual cooperation in preventing and combating serious crime affecting two or more

294. Europol Regulation, *supra* note 32.

Member States, and preventing terrorism and forms of crime that affect a common interest covered by a Union policy as listed in Annex I.[295] The Regulation also specifies that Europol's objectives shall cover criminal offences in order to (a) "procure the means of perpetuating" crimes within its competence, (b) "facilitate or perpetrate" crimes within its competence, or (c) "ensure the impunity" of those committing crimes within its competence.[296]

Article 4 of the Europol Regulation defines Europol's tasks and Article 4(1)(m) specifically sets out the task of the EU IRU, which is to:

> support Member States' actions in preventing and combating forms of crime listed in Annex I which are facilitated, promoted or committed using the Internet, including, in cooperation with Member States, the making of referrals of Internet content, by which such forms of crime are facilitated, promoted or committed, to the online service providers concerned for their voluntary consideration of the compatibility of the referred Internet content with their own terms and conditions.[297]

As can be seen, the EU IRU is empowered to support Member States' actions to prevent and combat *all* thirty forms of crime listed in Annex I.[298] This raises the concern that the EU IRU could be used to secure the extrajudicial removal of all content relating to Annex I crimes, bypassing the normal judicial process and need for court orders. The main limiting factor preventing this from occurring is political: the requirement that the EU IRU must support and act in cooperation with Member States means that Member States have to request the EU IRU's support for it to act on a particular type of crime listed in Annex I. A Council proposal that the EU IRU be empowered to act autonomously in processing data and making referrals was rejected in the final text of the Europol Regulation, as discussed below.

Recital (76) of the Europol Regulation contains a perambulatory statement likely to be relied on by European courts and authorities in interpreting the Regulation and which could be used to emphasize the need for Europol to respect rights granted by the EU Charter of Fundamental Rights, which will be discussed in Section IV.B, including the right to freedom of expression under Article 11, due

---

295. *Id.* annex I.
296. *Id.* recital (6).
297. *Id.* art. 4(1)(m).
298. *Id.* annex I.

process rights such as the right to an effective remedy and to a fair trial under Article 47, and the right to good administration under Article 41:

> This Regulation respects the fundamental rights and observes the principles recognized in particular by the Charter of Fundamental Rights of the European Union, in particular the right to the protection of personal data and the right to privacy as protected by Articles 8 and 7 of the Charter, as well as by Article 16 TFEU.[299]

In interpreting the ability of the EU IRU to refer content relating to all crimes listed in Annex I Regulation, European courts and authorities should also take note of the fact that the Council Presidency proposed a Recital that was rejected in the final text, which would have explained that the task of the EU IRU was to secure the removal of content relating to *all* forms of criminal activities under Europol's competence.[300]

The final text does not include any Recital explaining the task of the IRU. The final text of the EU IRU's task set out in Article 4(1)(m) also rejected the Presidency's proposal that Europol should take the lead in processing information and making referrals, but emphasizes that the IRU's task is to support Member States in preventing and

---

299. *Id.* ¶ 76.

300. Presidency of the Council of the Eur. Union, Meeting Document: The Functioning of the Internet Referral Unit (EU IRU) Based on the Future Europol Regulation, at 5 (Sept. 17, 2015), https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fdrakulablogdotcom3.files.wordpress.com%2F2015%2F10%2Fds01497-en151.docx. The proposal stated that "Article 4(1)(m) would be accompanied by a new recital (9a) explaining the new task set out in Art. 4(1)(m):

> (9a) Whilst the Internet provides a common global infrastructure for the exchange of ideas, services and goods, it can also be used for cross-border criminal activities. Europol should therefore be able to process information, including personal data, also from the Internet and other publicly available sources to support the Member States in preventing and combating forms of crime that fall under Europol's competence when criminal acts are facilitated and committed using the Internet. As criminal activity on the Internet has increased in recent years, [such as the amount of online material facilitating or promoting terrorism, illegal migrant smuggling], Europol should, in close cooperation with the Member States, identify such content, analyse it, and take appropriate action to secure its removal in voluntary cooperation with online service providers." *Id.*

combatting the crimes listed in Annex I and to make referrals "in cooperation with Member States."[301]

It has been reported, but not confirmed, that the EU IRU informs the designated contact-point of a Member State simultaneously when referring content to an ICT company, giving EU Member State a deadline to object.[302] However, if Member States do not take action to evaluate content referrals, but simply allow the deadline to expire, this does not constitute adequate oversight. Such a system also raises questions about how content that cannot be easily referred to any one Member State is addressed: is content hosted outside the European Union that is not attributable to an EU citizen referred to all EU Member States or none?

The Europol Regulation includes robust data protection safeguards, which will apply to the work of the EU IRU insofar as it deals with "personal data,"[303] which in practice should cover most

---

301. *Id.*; Europol Regulation, *supra* note 32, art. 4(1)(m). The Council Presidency's proposal, with amendments to final text reflected in **bold** and ~~strikethroughs~~:

> "(m) ~~To process information in order~~ to support Member States in preventing and combating forms of crime listed in Annex 1 which are facilitated**, promoted** or committed using the internet, ~~and to process personal data in so far as it is necessary for such a purpose. This includes receiving reports, collecting and analysing information from publicly available sources, notably the internet, identifying content which facilitates or promotes such forms of crime, and taking action to secure its removal in voluntary cooperation with online service providers~~ **including, in cooperation with Member States, the making of referrals of internet content, by which such forms of crime are facilitated, promoted or committed, to the online service providers concerned for their voluntary consideration of the compatibility of the referred internet content with their own terms and conditions.**"

302. Sauerbrey, *supra* note 31.

303. Europol Regulation, Article 2(h) defines "personal data" to mean "any information relating to a data subject," while Article 2(i) defines "data subject" to mean:

> an identified or identifiable natural person, an identifiable person being a person who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data or an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that person.

Europol Regulation, *supra* note 32, art. 2.

social media posts and accounts. Yet, the task of the EU IRU involves an exception to the normal data protection rules on exchange of personal data with private parties. Specifically, Article 26(5) of the Europol Regulation "stipulates that Europol may not transfer personal data to private parties except where, on a case-by-case basis strictly necessary and subject to any possible restrictions stipulated by the data owners or Europol itself."[304] The case-by-case evaluation has to take into account whether "the transfer is strictly necessary for the performance of the task [of the EU IRU] and the following conditions are met:

> (i) the transfer concerns an individual and specific case; and
> (ii) no fundamental rights and freedoms of the data subjects concerned override the public interest necessitating the transfer in the case at hand.[305]

This provision appears to ensure regard for fundamental rights by requiring a case-by-case evaluation that takes into account the fundamental rights and freedoms of the data subjects concerned. Still, this provision may be criticized for reversing the presumption in favor of rights, generally allowing the public interest, e.g., in securing the removal or terrorist propaganda or "material that attracts refugees to Europe," to prevail unless the right of the data subject to freedom of expression is sufficiently threatened. It is also unclear how much weight will be given to freedom of expression. In practice, a referral of individual social media posts could be regarded as minimally intrusive on a data subject's right to freedom of expression and the public's right to receive information, as well as the ICT companies' right to impart information and freedom to conduct a business without unreasonable interferences. As a result, it is unlikely that the data subject's right to freedom of expression will ever override the public interest in suppressing terrorist propaganda, material used to attract migrants or refugees to Europe, and other criminal activity. The evaluation may be different when a referral of an entire social media account is considered, but a restrictive view of the data subject's right to freedom of expression could be taken. In this view, the data subject can always create another account, so the interference with the data subject's right to freedom of expression is minimal and insufficient to override the

304. Ellermann, *supra* note 25, at 570.
305. Europol Regulation, *supra* note 32, art. 26(5)(c).

public interest. Similar arguments may apply to other rights, such as the data subject's right to an effective remedy.

One way to remedy these deficiencies and ensure that sufficient weight is given to fundamental rights and freedoms is for the Europol Data Protection Officer, the European Data Protection Supervisor, national supervisory authorities, the Fundamental Rights Agency, or the Court of Justice of the European Union (CJEU) to adopt an interpretation of Article 26(5)(c),[306] which, read together with Recital (76)[307] and Article 51 of the EU Charter of Fundamental Rights, [308] would recognize the public's interest in the EU IRU respecting fundamental rights. These rights would include the public's rights to receive information, to freedom of expression, to a fair trial and an effective remedy, and to good administration, as well as the ICT companies' rights to impart information and to conduct a business.

(2) Data Protection Safeguards

The Europol Regulation has robust data protection safeguards, which could help protect freedom of expression and other public interests. The Europol Regulation requires the appointment of a functionally-independent Data Protection Officer to assist in monitoring compliance with the Regulation. [309] It further requires Europol to be supervised by the European Data Protection Supervisor, who "shall be responsible for monitoring and ensuring the application of the provisions of this Regulation relating to the protection of fundamental rights and freedoms of natural persons with regard to the processing of personal data by Europol, and for advising Europol and data subjects on all matters concerning the processing of personal data." The supervision also has a number of duties, including:

> (a) hearing and investigating complaints, and informing the data subject of the outcome within a reasonable period;
> (b) conducting inquiries either on his or her own initiative or on the basis of a complaint, and informing the data subject of the outcome within a reasonable period;

---

306. *Id.* art. 26(5)(c).
307. *Id.* recital (76).
308. Charter of Fundamental Rights of the European Union, art. 51, 2012 O.J. (C 326) 395 [hereinafter CFR].
309. Europol Regulation, *supra* note 32, recital (49), art. 11(1)(l), art. 41.

> (c) monitoring and ensuring the application of [the Europol] Regulation and any other Union act relating to the protection of natural persons with regard to the processing of personal data by Europol;
> (d) advising Europol, either on his or her own initiative or in response to a consultation, on all matters concerning the processing of personal data, in particular before it draws up internal rules relating to the protection of fundamental rights and freedoms with regard to the processing of personal data;
> (e) keeping a register of new types of processing operations notified to him or her by virtue of Article 39(1) and registered in accordance with Article 39(4);
> (f) carrying out a prior consultation on processing notified to him or her.

In addition, the European Data Protection Supervisor (EDPS) is empowered to:

> (a) give advice to data subjects on the exercise of their rights;
> (b) refer a matter to Europol in the event of an alleged breach of the provisions governing the processing of personal data, and, where appropriate, make proposals for remedying that breach and for improving the protection of the data subjects;
> (c) order that requests to exercise certain rights in relation to data be complied with where such requests have been refused in breach of Articles 36 and 37;
> (d) warn or admonish Europol;
> (e) order Europol to carry out the rectification, restriction, erasure or destruction of personal data which have been processed in breach of the provisions governing the processing of personal data and to notify such actions to third parties to whom such data have been disclosed;
> (f) impose a temporary or definitive ban on processing operations by Europol which are in breach of the provisions governing the processing of personal data;
> (g) refer a matter to Europol and, if necessary, to the European Parliament, the Council and the Commission;
> (h) refer a matter to the Court of Justice of the European Union under the conditions provided for in the TFEU;

> (i) intervene in actions brought before the Court of Justice of the European Union.

The Europol Data Protection Supervisor thus has the mandate and power to ensure that the EU IRU operates in compliance with all of its fundamental rights obligations, including the obligation to respect and promote the rights to freedom of expression, a fair trial and effective remedy, and good administration. Unfortunately, this potential has yet to be realized.

The Europol Regulation also provides data subjects with a right to obtain information about whether Europol processes any personal data relating to them;[310] and a right to request rectification or erasure of data.[311] The right to request erasure of data is subject to refusal or restriction for reasonable grounds, such as the protection of security and public order or the prevention of crime.[312] However, such refusals can be challenged through a complaint to the EDPS.[313] While these rights alone are unlikely to ensure greater protection of the fundamental rights of data subjects, because there is no procedure to notify data subjects that their content was referred for takedown, these rights may allow data subjects whose content was taken down by ICT companies to discover if the EU IRU was involved. This presents an opportunity for strategic litigation, such as the *Schrems* case,[314] by data subjects whose content is unfairly referred by the EU IRU.

Article 47 creates a right to lodge a complaint with the EDPS if a data subject believes Europol's processing of his or her personal data does not comply with the Europol Regulation.[315] Read together with Article 18, which sets out the limited purposes for which Europol is allowed to process personal data, and Annex II, which sets out the limited categories of personal data and categories of data subjects whose data may be collected and processed, it may be possible for a person who falls outside of these categories, which largely revolve around criminal activity, to lodge a complaint and succeed as long as that person is not connected to or suspected of a crime.

---

310. *Id.* art. 36.
311. *Id.* art. 37.
312. Europol Regulation, *supra* note 32, recital.
313. *Id.* art. 37(9).
314. Case C-362/14, Maximillian Schrems v. Data Prot. Comm'r, 2015 EUR-Lex CELEX LEXIS 627 (Sept. 23, 2015) (involving Maximillian Schrems, a Facebook user, challenging the transfer of his data to the US by Facebook).
315. Europol Regulation, *supra* note 32, art. 47.

(3) Transparency

The Europol Regulation has been criticized for having weak transparency provisions.[316] While Article 65(1) on Transparency[317] states that Regulation 1049/2001, on public access to documents,[318] shall apply to documents held by Europol, this is an extremely limited right of access. Unlike some Freedom of Information laws, Regulation 1049/2001 does not require EU officials to create a new document in response to a request, but only to furnish documents that already exist, such as minutes of meetings, email correspondence, policy papers, and briefing papers.[319] Article 67 provides for Europol to create rules to protect sensitive information and, in the past, even benign requests have been refused.[320] As an example, Cornelia Ernst, Member of European Parliament, asked the Commission which national authorities were in Europol's "Internet analysis groups," but three months later was told that it was "confidential information."[321] Having said that, Article 65(3) states that "decisions taken by Europol under Article 8 of Regulation 1049/2001, to refuse or partially refuse access to documents, may be the subject of a complaint to the European Ombudsman or of an action before the Court of Justice of the European Union, in accordance with Articles 228 and 263 TFEU respectively."[322] Article 65(4) provides for the publication of summaries of the outcome of the meetings of the Europol Management Board, but this is a very limited disclosure obligation.

As mentioned above, data subjects do have the right to access personal data relating to them and to complain about refusals to provide them with such data.

(4) Parliamentary Oversight

Article 51 of the Europol Regulation provides for scrutiny and oversight by a Joint Parliamentary Scrutiny Group (JPSG), which has a mandate to monitor Europol's activities in fulfilling its mission.[323] This includes the impact of those activities on the fundamental rights

316. *See* Monroy, *supra* note 76.
317. Europol Regulation, *supra* note 32, art. 65(1).
318. Council Regulation 1049/2001, 2001 O.J. (L 143) 43 (EC).
319. *Id.* art. 10.
320. Europol Regulation, *supra* note 32, art. 67.
321. Sauerbrey, *supra* note 31.
322. Europol Regulation, *supra* note 32, art. 65.
323. *Id.* art. 51.

and freedoms of natural persons. This JPSG will be set up by the national Parliaments together with the Civil Liberties, Justice and Home Affairs (LIBE) Committee of the European Parliament.[324] It may subpoena the Chairperson of the Europol Management Board, its Executive Director, or their Deputies, to appear before the JPSG at its request to discuss matters relating to Europol activities, taking into account their obligations of discretion and confidentiality. The JPSG may also require the EDPS to appear before it at least once a year to discuss matters relating to the protection of fundamental rights and freedoms of natural persons. The JPSG is also entitled to receive a modest number of documents and to exercise the right of public access to request other documents.[325] The documents to which the JPSG is entitled include threat assessments, strategic analyses, and general situation reports relating to Europol's objective, as well as the results of studies and evaluations commissioned by Europol, annual and multiannual work programs, and annual reports. They do not include policy documents or briefing papers.[326]

Thus, the provision for parliamentary oversight has been subject to the following criticism:

> Despite all of this activity, the new opportunities for parliamentary oversight and access to information provided for by the regulation are likely to remain superficial. The explicit intention is not to scrutinise Europol's day-to-day work. The idea is merely to "politically monitor Europol's activities." This includes examining their impacts on "the fundamental rights and freedoms of natural persons." If the Members of Parliament do indeed identify problems, they may draw up "summary conclusions" and submit them to the Parliaments. . . .
>
> The JPSG also does not have any wider rights to gain information. Europol is supposed to transmit "relevant documents" including "threat assessments, strategic analyses and general situation reports," as well as the results of studies and evaluations commissioned by Europol. However, this only applies to non-classified

---

324. *See id.* (stating that "this shall constitute a specialised Joint Parliamentary Scrutiny Group (JPSG) established together by the national parliaments and the competent committee of the European Parliament").

325. Ellermann, *supra* note 25, at 577–78.

326. Europol Regulation, *supra* note 32, art. 51.

> documents and thus continues a previous practice denying MEPs access to important information.[327]

Finally, Recital 38 requires the Commission to evaluate the activities that Europol should undertake on the basis of the Europol Regulation, in particular the corresponding practice of direct exchanges of personal data with private parties by May 1, 2019.[328] The JPSG will be entitled to receive this evaluation.

(5) Sub-Conclusion: "The Most Controlled Police Agency" or Insufficient Oversight and Safeguards?

Europol has been called "the most controlled police agency" because of the number of accountability mechanisms it has, particularly for data protection, including the Data Protection Officer, the EDPS, national supervisory authorities, and joint parliamentary scrutiny.[329] Nevertheless, it does not have sufficient oversight and safeguards to ensure the protection of other fundamental rights, such as freedom of expression and the right to an effective remedy and a fair trial. However, there are particular opportunities for the data protection safeguards to ensure that there is meaningful protection of other fundamental rights in their work.

B. EU Charter of Fundamental Rights

(1) Brief Introduction of the EU Charter of the Fundamental Rights

The EU Charter of Fundamental Rights (CFR)[330] should not be confused with the European Convention of Human Rights (ECHR), which is much older and was opened for signature in 1950. The CFR was proclaimed in 2000 but did not have binding effect until 2009. The CFR was originally intended to form the Bill of Rights for the European Constitution, but was rejected in 2005 by Dutch and French referendums, and thus did not come into effect until it was ratified as part of the Lisbon Treaty.[331] Although the CFR contains the rights

---

327. Monroy, *supra* note 76.
328. Europol Regulation, *supra* note 32, recital (38).
329. Ellermann, *supra* note 25, at 576–77.
330. CFR, *supra* note 308.
331. Francesca Ferraro & Jesús Carmona, *Fundamental Rights in the EU: The Role of the Charter after the Lisbon Treaty*, EUROPEAN PARLIAMENTARY

contained in the ECHR, it goes beyond the ECHR, covering a wealth of new rights, such as the right to protection of personal data (Article 8),[332] the freedom to conduct a business (Article 16),[333] and the right to good administration (Article 41).[334] Unlike the ECHR which has specific limitations clauses for some Articles, e.g., Article 10(2), which provides for restrictions on the right to freedom of expression, the CFR has a general limitation clause, Article 52(1),[335] which is provided below.

Article 52(3) of the CFR provides that the Charter should follow the definition and scope of ECHR rights where they correspond, but that this should not prevent EU law from providing more extensive protection of human rights.[336] Judge Allan Rosas of the Court of Justice of the European Union (CJEU) has also explained that the Explanations relating to the CFR,[337] which according to Article 6(1) of the Treaty on European Union[338] shall be taken into account in the interpretation of the CFR, state that the definition and scope of the corresponding ECHR rights should be determined not only by the ECHR text, but also by ECtHR and CJEU case law.[339] The CJEU also "draws inspiration from the constitutional traditions common to the Member States and from the guidelines supplied by international treaties for the protection of human rights on which the Member States have collaborated or to which they are signatories" and has said that the ECtHR has "special significance" in that respect.[340]

RESEARCH SERVICE 9–10 (2015), http://www.europarl.europa.eu/RegData/etudes/IDAN/2015/554168/EPRS_IDA(2015)554168_EN.pdf.

332. CFR, *supra* note 308, art. 8.

333. *Id.* art. 16.

334. *Id.* art. 41.

335. *Id.* art. 52(1).

336. *Id.* art. 52(3).

337. Explanations Relating to the Charter of Fundamental Rights, 2007 O.J. (C 303) 17 [hereinafter Explanations].

338. Consolidated Version of the Treaty on European Union art. 6(1), Oct. 26, 2012, 2012 O.J. (C 326).

339. Allan Rosas, *Five Years of Charter Case Law: Some Observations*, *in* THE EU CHARTER OF FUNDAMENTAL RIGHTS AS A BINDING INSTRUMENT: FIVE YEARS OLD AND GROWING 11, 14 (Sybe de Vries, Ulf Bernitz, & Stephen Weatherill eds., 2015).

340. Case C-274/99, P Connolly v. Comm'n, 2001 E.C.R. I-1611, ¶ 37; *see also* Koen Lenaerts, *Exploring the Limits of the EU Charter of Fundamental Rights*, 8 EUR. CONST. L.R. 375, 395 (2012); Joined Cases 46/87 and 227/88, Hoechst v. Comm'n, 1989 E.C.R. 2919, ¶ 13.

The CFR is binding on all EU institutions, bodies, offices, and agencies, including Europol and the EU IRU, and on Member States only when they are implementing EU law. Article 51 obliges them to respect the rights, observe the principles, and promote the application thereof while respecting the limits of the powers of the Union.

In our discussion of the CFR and the EU IRU, at least the following rights are relevant.

> Article 7: Respect for private and family life
> Everyone has the right to respect for his or her private and family life, home and communications.
>
> Article 8: Protection of personal data
> 1. Everyone has the right to the protection of personal data concerning him or her.
> 2. Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified.
> 3. Compliance with these rules shall be subject to control by an independent authority.
>
> Article 11: Freedom of expression and information
> 1. Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers.
> 2. The freedom and pluralism of the media shall be respected.
>
> Article 16: Freedom to conduct a business
> The freedom to conduct a business in accordance with Union law and national laws and practices is recognised.
>
> Article 41: Right to good administration
> 1. Every person has the right to have his or her affairs handled impartially, fairly and within a reasonable time by the institutions, bodies, offices and agencies of the Union.
> 2. This right includes:

(a) the right of every person to be heard, before any individual measure which would affect him or her adversely is taken;
(b) the right of every person to have access to his or her file, while respecting the legitimate interests of confidentiality and of professional and business secrecy;
(c) the obligation of the administration to give reasons for its decisions.

Article 47: Right to effective remedy and a fair trial
Everyone whose rights and freedoms guaranteed by the law of the Union are violated has the right to an effective remedy before a tribunal in compliance with the conditions laid down in this Article.
Everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal previously established by law. Everyone shall have the possibility of being advised, defended and represented.

Article 52: Scope and interpretation of rights and principles
(1) Any limitation on the exercise of the rights and freedoms recognised by this Charter must be provided for by law and respect the essence of those rights and freedoms. Subject to the principle of proportionality, limitations may be made only if they are necessary and genuinely meet objectives of general interest recognised by the Union or the need to protect the rights and freedoms of others.

Article 54: Prohibition of abuse of rights
Nothing in this Charter shall be interpreted as implying any right to engage in any activity or to perform any act aimed at the destruction of any of the rights and freedoms recognised in this Charter or at their limitation to a greater extent than is provided for herein.[341]

341. CFR, *supra* note 308, arts. 7, 8, 11, 16, 41(1), 41(2), 47, 52(1), 54.

(2) Application of the CFR to the EU IRU

Article 52(3) of the CFR provides that the Charter should follow the definition and scope of ECHR rights where they correspond and that the CJEU draws inspiration from the case law of the ECtHR.[342] As such, the analysis of the ECHR and case law and the conclusions in Part III above should apply to the EU IRU, *mutatis mutandis*, since Articles 7, 11, 47, and 54 of the CFR correspond to ECHR Articles 8, 10, 6(1), 13, and 17, respectively.[343]

i. Preliminary Conclusions, Based on ECHR Principles and Case Law

Thus, applying the principles and case law of the ECHR, Article 47—the right to effective remedy and a fair trial—is engaged when ICT companies decide to take down material as a violation of their ToS. When ICT companies take down material as a result of EU IRU referrals, Article 47 imposes an obligation on the EU to provide a fair and public hearing within a reasonable time by an independent and impartial tribunal and provide legal aid to those who lack sufficient resources.[344] The EU should take into account the Fundamental Rights Agency's recent opinion on "Improving access to remedy in the area of business and human rights at the EU level."[345] The opinion recommends that "[t]he EU could provide stronger incentives for the creation of remedy mechanisms at company level (operational-level grievance mechanisms), including multi-stakeholder initiatives with several businesses joining forces with other actors" and that "[t]he EU should make available information on existing judicial and non-judicial mechanisms for the benefit of the general public, legal practitioners and victims."

While Article 54, prohibiting the abuse of rights,[346] may be used to justify procedures meeting the above minimum standards for content that incites violence or hate speech, it may not justify the use of IRUs to take down content that is not directed against the fundamental values of the European Union, since Article 54 CFR

342. *Id.* art. 52(3).

343. *See* Explanations, *supra* note 339, at 20, 21, 29–30, 35.

344. CFR, *supra* note 308, art. 47.

345. EUR. UNION AGENCY FOR FUNDAMENTAL RIGHTS, IMPROVING ACCESS TO REMEDY IN THE AREA OF BUSINESS AND HUMAN RIGHTS AT THE EU LEVEL 14 (2017), http://fra.europa.eu/en/opinion/2017/business-human-rights.

346. CFR, *supra* note 308, art. 54.

corresponds to Article 17 ECHR.[347] Article 11, Freedom of expression and information,[348] will therefore be engaged when content other than incitement to violence or hate speech is involved. IRU referrals will have to be addressed within the Article 52, entitled "Scope and interpretation of rights and principles,"[349] which provides a framework of limitations on rights. Moreover, because Article 11 of the CFR corresponds to Article 10 of the ECHR, the limitations on such content may therefore not exceed those provided for in Article 10(2).[350]

Turning to the Article 52 framework, the EU IRU meets the requirement of being grounded in an adequate legal framework that sets out the types of content that can be referred and is likely to meet the tests of accessibility, foreseeability, and sufficient safeguards against arbitrary abuse. The EU IRU also conducts proportionality analysis on the impacted fundamental rights on a case-by-case basis, but this analysis lacks adequate due process safeguards per Articles 41 and 47. Particularly, this proportionality analysis fails to adequately protect the Article 41(2) right to be heard and to be given reasons for a body's decisions.[351] What is more, it is systematically deficient, because it prioritizes the public interest in suppressing "illegal" content over the rights of data subjects, the public, and ICT companies. Unless the system of data protection safeguards is interpreted in a way that ensures sufficient protection for other fundamental rights, there are insufficient independent safeguards for freedom of expression and other fundamental rights and interests.

Finally, as discussed in Section III.E.(1), the use of Article 11 presents an opportunity for courts to read into the CFR a narrowly-tailored positive obligation on the EU to require social media companies to provide adequate protection for freedom of expression, in line with the ECHR's jurisprudence on positive obligations and the U.N. Guiding Principles on Business and Human Rights.

ii. Application of CJEU Case Law & EU Law to the EU IRU

Examining the CJEU's case law on the right to freedom of expression and the Internet thus far, it seems unlikely that the CJEU

347. ECHR, *supra* note 183, art. 17.
348. CFR, *supra* note 308, art. 11.
349. *Id.* art. 52.
350. Explanations, *supra* note 339, at 21.
351. CFR, *supra* note 308, art. 41(2).

would approve any proposal for the EU IRU to maintain a system of general monitoring enforced by the threat of liability on ICT companies without actual notice.

The CJEU has ruled against a system of "general monitoring" in which information service providers are compelled to proactively monitor and filter content from the Internet in multiple cases. Article 15 of the EU's e-Commerce Directive states that information service providers cannot be compelled to undertake general monitoring, while Article 12 provides for a limited form of intermediary immunity for mere conduits, "caches," and "hosts" in the absence of notice.[352] In *Scarlet Extended*,[353] the CJEU found that the right to freedom of expression was engaged by a general filtering system on broadband networks for the purpose of copyright enforcement, because it may not be able to accurately distinguish between lawful and unlawful content.[354] The right to privacy was also engaged, since it would have to systematically examine all content and identify the IP addresses of the individual users.[355] The CJEU held that the e-Commerce Directive also prohibits the imposition of a general obligation to monitor for content and to require "notice and stay down" of content.[356] Scholar Monica Horten has explained that the ruling in *Scarlet Extended* effectively

> means that anything involving continuous monitoring of all content for an unlimited period of time would comprise a general obligation to monitor and would be illegal under EU law. This does not preclude filtering measures being ordered, but there are strict legal criteria that should be met. The ECJ has stated that filtering measures must be necessary and proportionate, they should be targeted, and the determination of the filtering criteria or the content to be filtered should be ordered by a court or a body independent of political influence and should be subject to judicial oversight. In addition, the ECJ claims such

352. Council Directive 2000/31, arts. 12, 15, 2000 O.J. (L 178) 1.

353. Case C-70/10, Scarlet Extended v. Société Belge de auteurs, compositeurs, et éditeurs SCRL (SABAM), 2011 E.C.R. I-12006.

354. *Id.* ¶ 52.

355. *Id.* ¶ 51.

356. MONICA HORTEN, CTR. FOR DEMOCRACY & TECH., CONTENT RESPONSIBILITY: THE LOOMING CLOUD OF UNCERTAINTY FOR INTERNET INTERMEDIARIES (2016), https://cdt.org/files/2016/09/2016-09-02-Content-Responsibility-FN1-w-pgenbs.pdf.

> measures should not impose excessive costs on the broadband providers.[357]

Similarly, in *Google France v. Louis Vuitton* and *L'Oreal v. eBay*, the CJEU held that search engines and auction websites cannot be held responsible for third party content if they have no knowledge of it, and they do not have a general obligation to monitor.[358] Similarly, under the net neutrality provisions adopted by the European Union in 2015, blocking by network providers is not permitted, unless the intermediary has received a court order.[359]

iii. Application of New Rights in the CFR to the EU IRU

Turning to the application of new rights in the CFR to the EU IRU, it may be argued that the freedom to conduct a business gives ICT companies additional protection, albeit presently weak,[360] against extra-legal threats or coercive pressures, while the right to good administration requires that EU citizens whose right to freedom of expression is affected by the EU IRU have a right to be heard,[361] and to have reasons for EU IRU decisions.

V. CONCLUSION AND RECOMMENDATIONS

This Part presents some recommendations based on the parts above that are tailored to ensure that IRUs operate within the rule of

---

357. Monica Horten, *Freedom of Expression, Human Rights Standards, and Private Online Censorship*, *in* CYBERSECURITY AND HUMAN RIGHTS IN THE AGE OF CYBERVEILLANCE 94, 94 (Joanna Kulesza & Roy Balleste eds., 2016).

358. FRANCESCO BUFFA, FREEDOM OF EXPRESSION IN THE INTERNET SOCIETY 36 (2016).

359. Council Regulation 2015/2120, art. 11, 2015 O.J. (L 310) 1.

360. *See* Xavier Groussot, Gunnar Thor Pétursson & Justin Pierce, *Weak right, strong Court – the freedom to conduct business and the EU Charter of Fundamental Rights*, *in* RESEARCH HANDBOOK ON EU LAW AND HUMAN RIGHTS 326 (Sionaidh Douglas-Scott & Nicholas Hatzis eds., 2017).

361. *See* Alexander Türk, *Administrative law and fundamental rights*, *in* RESEARCH HANDBOOK ON EU LAW AND HUMAN RIGHTS 120, 130 (Sionaidh Douglas-Scott and Nicholas Hatzis eds., 2017) ("The Union courts' core formula postulates that the right to be heard applies 'in all proceedings initiated against a person which are liable to culminate in a measure adversely affecting that person'. While it has been applied with great flexibility across the Union's policy sectors and the Union courts have deviated from it on occasion, this formula encapsulates as guiding considerations the requirements of *individualisation* and *adverse effect*.").

law, in accordance with human rights standards, and with democratic safeguards, so that their potential for abuse can be minimized. It should be noted that IRUs are merely one response—albeit a growing and potentially troubling one—to the legitimate problem of extremist content online. There is a trend towards automated content filters being developed by ICT companies. Nevertheless, many of the recommendations below are relevant to automated content filters, such as the provision of broad intermediary immunity, notice and takedown procedures, remedial mechanisms, and the drafting of content removal standards that respect human rights.

More broadly, it should be emphasized that censorship of extremist content online alone is an inadequate response to win the fight against violent extremism; combating violent extremism requires comprehensive responses that include effective counter-messaging, offline interventions, and the development of inclusive societies that respect human rights.[362] Governments, policy-makers, and policy-shapers are presently focusing immense attention on compelling social media companies to do more to censor extremist content, even though they are already acting cooperatively.[363] At the same time, governments are failing to adequately address the use of encrypted messaging apps like Telegram by Daesh and other terrorist groups, despite indications that Telegram is being used both to spread propaganda and as a means of communication by terrorist groups, because Telegram does not cooperate with requests from law enforcement.[364] The United Kingdom's independent reviewer of counterterrorism laws has stated that he sees no or very little need for

362. *See* U.N. Secretary-General, *supra* note 145, ¶¶ 44, 49–55; U.N. Security Council Counter-Terrorism Comm., *Comprehensive international framework to counter terrorist narratives*, U.N. Doc. S/2017/375 (Apr. 28, 2017); Global Counterterrorism Forum, *supra* note 24, at 3.

363. Alan Travis, *Tech firms could do more to tackle extremism – but so could politicians*, GUARDIAN (Sept. 19, 2017), https://www.theguardian.com/uk-news/2017/sep/19/tech-firms-could-do-more-to-tackle-extremism-but-so-could-politicians (quoting UK Independent Reviewer of Terrorism Legislation Max Hill: "In Germany there was a proposal to levy heavy fines on tech companies whenever they failed to take down extremist content. I am not sure that is absolutely necessary . . . I have sat with the relevant police unit that identifies the extremist material. I have seen them communicate with the tech companies and the co-operation that flows from that. It is a question of the bulk of the material rather than a lack of cooperation in dealing with it.").

364. MARTYN FRAMPTON ET AL., POLICY EXCHANGE, THE NEW NETWAR: COUNTERING EXTREMISM ONLINE (2017), https://policyexchange.org.uk/wp-content/uploads/2017/09/The-New-Netwar-2.pdf.

new legislation on the use of social media by terrorists beyond the Terrorism Acts of 2000 and 2006. Instead, he stresses instead the importance of investigation and criminal prosecution of existing laws and notes that driving extremist content underground would be counter-productive if would-be terrorists could still access it on the dark web, which would make it harder for law enforcement to detect and much harder for good people to argue against.[365] Ultimately, the battle for hearts and minds will not be won by censoring content, but by persuading potential terrorists that there are better alternatives to violence.

To the United Kingdom:

Adopt legislation to put the CTIRU on a legislative footing in line with Model Standard (1), presented in Section II.E.

Encourage politicians to refrain from threatening or coercing ICT companies into over-censoring content online, in order not to undermine freedom of expression online. They should be encouraged to adopt positions similar to that of Baroness Joanna Shields, in persuading ICT companies to do more to combat terrorist propaganda voluntarily, while ensuring respect for democratic values and freedom of expression.

Repeal sections 3 and 4 of the Terrorism Act of 2006 or adopt policy guidance making it clear that IRU referrals are not section 3 notices and that sections 3 and 4 will not be used in the absence of a court order.

Provide adequate notice and remediation mechanisms for users whose content is referred by the CTIRU to be able to effectively challenge referral decisions.

To EU Member States, including the United Kingdom:

EU Member States should legislate to ensure broad intermediary immunity in the absence of modification of third party content or notice and cease extra-legal threats and coercive pressure on ICT companies, so that the work of IRUs is truly voluntary.

---

365. Max Hill, *Responding to Terrorists' Use of Social Media: Legislation, Investigation and Prosecution*, INDEP. REVIEWER OF TERRORISM LEGIS. (Sept. 3, 2017), https://terrorismlegislationreviewer.independent.gov.uk/responding-to-terrorists-use-of-social-media-legislation-investigation-and-prosecution/.

EU Member States should not impose a general obligation on companies to monitor and takedown third-party content without a court order, though they may facilitate the development of multi-stakeholder initiatives that are voluntary.

National supervisory authorities should request that the EU IRU adopts measures to adequately protect Article 11 (freedom of expression) and Article 47 (the right to effective remedy and a fair trial) of the CFR, in line with Model Standard (3).

EU Member States should take measures to incentivize ICT companies to provide remedial mechanisms meeting Model Standard (4), presented in Section II.E, such as by restricting offshore data transfers to ICT companies that can demonstrate they are compliant with freedom of expression standards as well as data protection law.

EU Member States should cease the practice of IRU referrals for "material attracting migrants and refugees to Europe" and should not request the IRU to make extra-judicial referrals for any content that does not meet a strict definition of terrorism or violent extremism.

To the Europol Data Protection Officer and European Data Protection Supervisor (EDPS):

Adopt an interpretation of Articles 4(1)(m) and 26(5) of the Europol Regulation, in line with Recital (76) and Article 51 of the CFR, that ensures sufficient weight is given to the right to freedom of expression of the data subject, the public, and the ICT company and that the data subject's Article 47 rights are respected through the provision of notice, a right to be heard, and an effective mechanism to challenge a decision to refer content.

Ensure that this interpretation results in the development and implementation of effective procedures for safeguarding Article 11 and Article 47 in the operation of the EU IRU.

To the Joint Parliamentary Scrutiny Group:

The JPSG should request that the heads of Europol and the EU IRU demonstrate how they are ensuring the protection and promotion of the right to freedom of expression and the right to an effective remedy in the operation of the EU IRU.

The JPSG should request that the EDPS demonstrate how he is ensuring the protection and promotion of the right to freedom of expression and the right to an effective remedy in the operation of the EU IRU.

The JPSG should consider requesting that the Fundamental Rights Agency provide an Opinion on how to respect and promote fundamental freedoms, especially those found in Articles 11 and 47 CFR, in the operation of the EU IRU.

To civil society groups:

Consider strategic litigation, through finding an EU citizen whose freedom of speech has been violated by an ICT company, potentially an artist or political satirist, or someone who modified Daesh propaganda to make a statement against terrorism.

The EU citizen could make a complaint to the EU Data Protection Supervisor, arguing that the EU IRU improperly processed his or her data. The goal is to establish a precedent and practice of data protection safeguards being used to protect freedom of expression, thus working towards changing the practices of the EU IRU.

The EU citizen could bring a case against the ICT company in order to establish that his Article 11 and 47 CFR rights, and Article 6, 10, and 13 ECHR rights require that the ICT company provide notice.

The EU citizen should seek the opportunity to bring litigation before the CJEU.

Inform EU citizens that if their right to freedom of expression is improperly restricted by an ICT company or the EU IRU, they have a right to an effective remedy against both.

Ask the EDPS, pursuant to Article 43(3) of the Europol Regulation, to advise EU citizens on how they can ensure their right to freedom of expression is protected when they exercise it.

To ICT companies:

When content is removed pursuant to a referral from an IRU, give notice to the affected person that the content was referred to them by an IRU, so that they may challenge the decision if they believe that it was improper.

ICT companies should generally adopt policy commitments, terms of service, platform designs, notification and takedown procedures, and remedial mechanisms that meet the U.N. Guiding Principles, in line with the Model Standard (4), presented in Section II.E.

ICT companies should ensure that content removal guidelines are drafted in consultation with multi-stakeholder groups that are

committed to the protection of freedom of expression and other human rights online, and that those guidelines are publicly available.

To the European Court of Human Rights:

Consider elaborating a system that only permits IRUs to make content referrals for hate speech and a narrowly defined category of "terrorist" or "violent extremist" speech. Such a system should have effective notification, due process, and appeals safeguards.

Consider imposing a narrowly-tailored positive obligation on states to require Internet companies to respect freedom of expression when moderating content, in line with the U.N. Guiding Principles on Business and Human Rights.