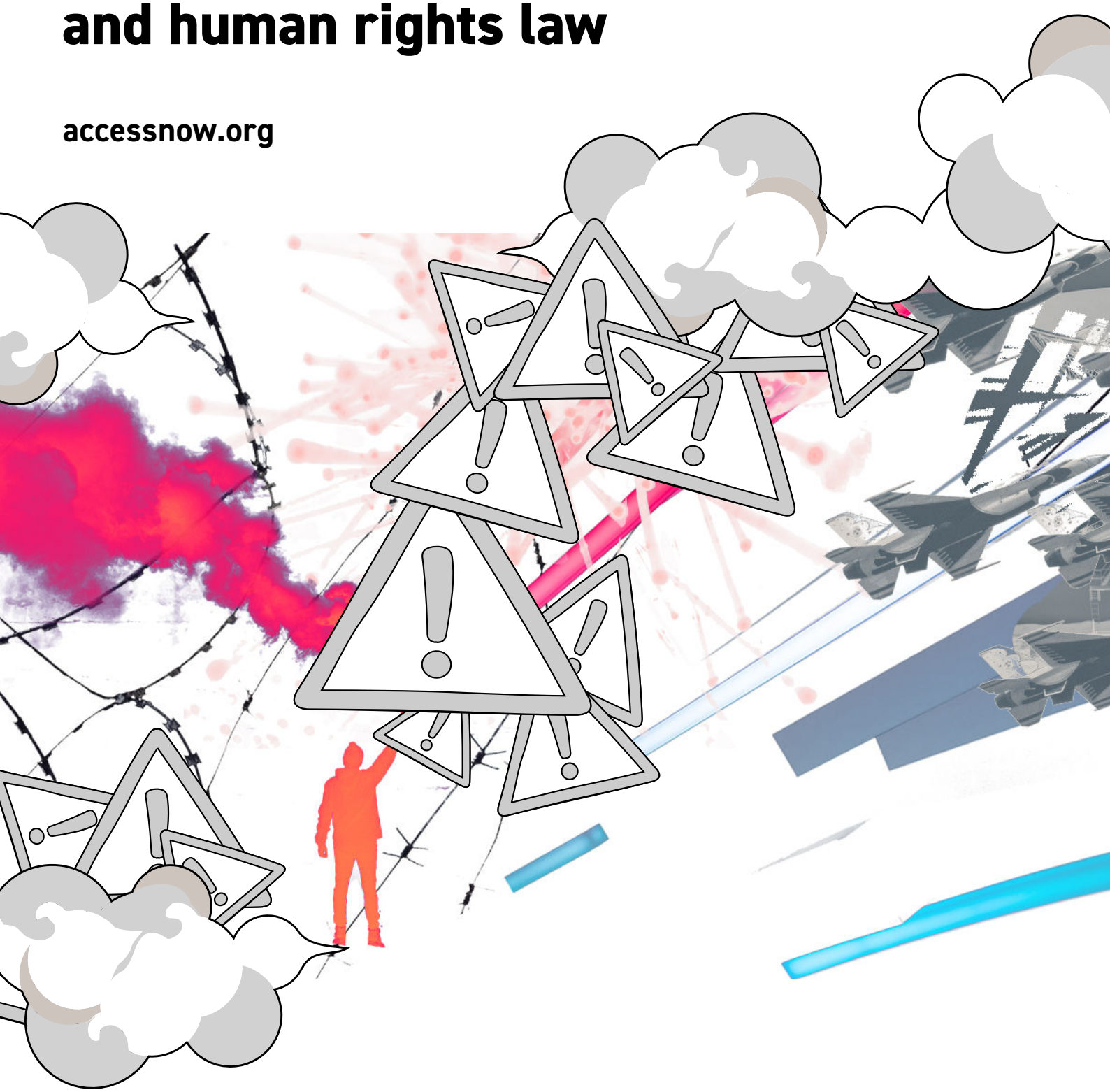
Content and platform governance in times of crisis: applying international humanitarian, criminal, and human rights law

accessnow.org



accessnow

Access Now defends and extends the digital rights of people and communities at risk. By combining direct technical support, strategic advocacy, grassroots grantmaking, and convenings such as RightsCon, we fight for human rights in the digital age.

Content and platform governance in times of crisis: applying international humanitarian, criminal, and human rights law

Acknowledgements: This policy paper was drafted by Dr Talita Dias, Research Fellow, Oxford Institute for Ethics, Law and Armed Conflict (ELAC), Blavatnik School of Government, University of Oxford, under the direction of Eliška Pírková, Senior Policy Analyst and Global Freedom of Expression Lead, and Marwa Fatafta, MENA Policy and Advocacy Director at Access Now. The authors would like to thank Giulio Coppi, Donna Wentworth, and Méabh Maguire for their support. For any questions or feedback on this paper, please contact Eliška Pírková at eliska@accessnow.org and Marwa Fatafta at marwa@accessnow.org.

December 2023

Access Now defends and extends the digital rights of people and communities at risk. By combining direct technical support, strategic advocacy, grassroots grantmaking, and convenings such as RightsCon, we fight for human rights in the digital age.

Table of contents

EXECUTIVE SUMMARY

In November 2022, Access Now, ARTICLE 19, and other NGOs signed the 'Declaration of principles for content and platform governance in times of crisis' (the 'Declaration').[1] This Declaration was borne out of a concern that, despite their significant role during armed conflict and other crises, online platforms have failed to adopt appropriate responses to the harmful effects of online content in those situations.

Online platforms have presented individuals and societies with both challenges and opportunities. On the one hand, they have been a key enabler of a range of human rights, including civil, political, economic, cultural, and social rights. Think of the Arab Spring, during which social media was key to the exercise of the individual rights to freedom of expression and assembly.[2] Online platforms have also served as rich repositories of evidence of atrocities committed in Syria, Myanmar, and Ukraine, for example. Similarly, in Ukraine, both civilians and combatants have turned to online platforms and other ICTs to inform the public about Russia's illegal invasion of Ukraine.[3]

For a long time, civil society organisations have documented platforms' inadequate and inconsistent responses to conflict, fragile governance, and crises — such as those in Ethiopia,[4] Syria, Israel/Palestine,[5] and Myanmar,[6] among others. They often failed to respect human rights or to mitigate adverse human rights impacts stemming from their systems, and have been slow or ineffective in removing or restricting hate speech and incitement to violence in real time. Rather, their responses (or lack thereof) have disproportionately impacted marginalised communities and historically oppressed groups, and have facilitated serious human rights abuses.

Furthermore, online platforms have been used by different actors — including States, non-State groups, and individuals — to cause harm or disruption. This became clear during the COVID-19 pandemic when the dissemination of false or misleading information led to fear and division in different societies.[7] The situation of the Rohingya in Myanmar is also emblematic of how online hateful rhetoric, if left to spread freely, can fuel a human catastrophe.[8] Online platforms have also been used during international and non-international conflicts by numerous parties to conflicts to

---

[1] Marwa Fatafta and Eliška Pírková, *Joint Declaration of principles for content and platform governance in times of crisis* (n.p.: *Access Now*, 2022), https://www.accessnow.org/publication/new-content-governance-in-crises-declaration/.
[2] Zeynep Tufekci, *Twitter and Tear Gas: The Power and Fragility of Networked Protest* (New Haven: Yale University Press, 2017).
[3] Ben Schreckinger, "In Ukraine war crimes go on-chain," *POLITICO,* January 17, 2023, https://www.politico.com/newsletters/digital-future-daily/2023/01/17/ukraine-war-crimes-blockchain-00078170
[4] Berhan Taye and Javier Pallero, "Open letter to Facebook on violence-inciting speech: act now to protect Ethiopians," *Access Now*, July 27, 2020 https://www.accessnow.org/open-letter-to-facebook-protect-ethiopians/.
[5] Business for Social Responsibility (BSR), *Human Rights Due Diligence of Meta's Impacts in Israel and Palestine*, (n.p.: BSR, 2022), https://www.bsr.org/en/reports/meta-human-rights-israel-palestine.
[6] Golda Benjamin, Whai Phyo Myint, and Dhevy, "Myanmar IMEI FAQ: how the junta could disconnect the resistance," *Access Now*, July 7, 2022, https://www.accessnow.org/myanmar-imei/.
[7] Nick Robins-Early, "Desperation, Misinformation: How the Ivermectin Craze Spread Across the World," *The Guardian*, September 24, 2021, https://www.theguardian.com/world/2021/sep/24/ivermectin-covid-peru-misinformation.
[8] Steve Stecklow, "Why Facebook is losing the war on hate speech in Myanmar," *Reuters,* August 15, 2018, https://www.reuters.com/investigates/special-report/myanmar-facebook-hate/.

advance their goals. The dissemination of war propaganda, disinformation, hate speech online[9] and videos depicting prisoners of war in coercive conditions have penetrated online.[10]

Against this background, the Declaration sought to advance consistent and rights-respecting principles for online platforms to adhere to in times of crises, including before, during, and after their occurrence. This report is, conceptually, a prequel to the Declaration: it complements the Declaration by outlining the legal foundations on which the latter is grounded. Those foundations are made up of three key legal frameworks that apply concurrently to online content governance in different types of crises:

- international humanitarian law;
- international criminal law; and
- international human rights law.

This report will explore the extent to which these legal frameworks apply concurrently to the use of online platforms in times of crises. We anticipate the need to develop additional guidance in the future that delves deeper into how applicable rules and principles, together with the principles devised by the Declaration, translate into concrete content governance measures before, during, and after crises.

INTRODUCTION: SETTING THE SCENE

International humanitarian law (IHL), international criminal law (ICL), and international human rights law (IHRL) apply to different types of crises. IHL is triggered by the existence of an armed conflict, whether international or non-international in nature. ICL comes into play when international crimes (in particular genocide, crimes against humanity, war crimes, and the crime of aggression), are committed. IHRL applies to human rights violations in both armed conflict and peacetime. This report pays particular attention to gross, widespread, or systematic human rights violations.[11]

As a matter of international law, once triggered, these three legal frameworks apply together to different types of online content. They each bind States under treaties and customary international law. International custom comprises unwritten rules of international law developed through the

[9] Jack Goodman, Maria Korenyuk, Lucy Swinnen and Andrey Zakharov, "War in Ukraine: The making of a new Russian propaganda machine," *BBC News,* May 29, 2022, https://www.bbc.co.uk/news/world-europe-61441192; BBC Reality Check Team, "Ukraine crisis: Vladimir Putin address fact-checked," *BBC News*, February 22, 2022, https://www.bbc.co.uk/news/60477712; Alexey Kovalev, "Russia's Ukraine Propaganda Has Turned Fully Genocidal," *Foreign Policy*, April 9, 2022, https://foreignpolicy.com/2022/04/09/russia-putin-propaganda-ukraine-war-crimes-atrocities/; Mariia Kravchenko, "What should Russia do with Ukraine? [Translation of a propaganda article by a Russian publication]," *Medium*, April 4, 2022, https://medium.com/@kravchenko_mm/what-should-russia-do-with-ukraine-translation-of-a-propaganda-article-by-a-russian-journalist-a3e92e3cb64.

[10] Aaron Blake, "Why you should think twice before sharing that viral video of an apparent Russian POW," *The Washington Post*, March 7, 2022, https://www.washingtonpost.com/politics/2022/03/07/russian-pow-videos/.

[11] UN General Assembly, "A more secure world: our shared responsibility Report of the High-level Panel on Threats, Challenges and Change", A/59/565, para 36, 2014: https://www.un.org/peacebuilding/content/more-secure-world-our-shared-responsibility-%E2%80%93-report-high-level-panel-threats-challenges-and.

widespread and consistent practice of States, accompanied by a sense of a legal obligation (*opinio juris*). While corporations are not directly bound by international law, they have a social responsibility to respect human rights online and offline, in line with the United Nations (UN) Guiding Principles on Business and Human Rights.[12] It is a given that corporations, including online platforms, are driven by profit and other private interests. Yet it is in their own business interest to provide their users and other members of the public with a free and safe online information environment, in line with international law. International law provides companies with a tried and tested universal vocabulary for tackling harmful content in the internet's cross-boundary environment.[13] Thus, as online platforms continue to play a prominent role in the global political arena and social fabric, they should design their online content policies in line with international law, particularly IHL, ICL, and IHRL.

Given this context, this report seeks to address two overarching questions:

- What rules and principles of IHL and ICL apply to States and online platforms – including platforms themselves and their users – in times of crisis? This question is addressed in Part I below. Part I focuses on IHL, while Part II unpacks applicable rules of ICL. Relevant sub-questions include what kinds of online content violate IHL and ICL and who may be held liable for them.
- The second question tackles the interplay between these two frameworks and IHRL: How must the rules and principles of IHL and ICL be balanced against IHRL? Of particular relevance to online content governance are the rights to freedom of expression and information. This question includes the issue of which legal framework prevails when there is a norm conflict.

This report provides interpretative guidance on these and related questions, joining existing efforts to shape online content policies developed by States and platforms in times of crisis.

INTERNATIONAL HUMANITARIAN LAW (IHL)

I. Basic concept of IHL

This section provides an overview of the rules and principles of IHL that are most relevant to online content governance in times of crises. These include a) the duty to respect and ensure respect for IHL, b) obligations seeking to protect civilians from the effects of armed conflict, and c) those extending special protections to some actors. All three sets of rules are explored below.

IHL applies during armed conflict. This means that *most* of its rules and principles only come into play in the context of an armed conflict of an international or non-international character. A notable exception is the duty to respect and ensure respect for IHL, which applies even in

[12] UN Office of the High Commissioner for Human Rights, "*Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework,"* 2011*,* https://digitallibrary.un.org/record/720245?ln=en.

[13] Peggy Hicks et al, "Press briefing: Online content moderation and internet shutdowns." *UN Human Rights Office,* July 14, 2021, https://www.ohchr.org/Documents/Press/Press%20briefing_140721.pdf, 1; Evelyn Mary Aswad, "The Future of Freedom of Expression Online." *Duke Law & Technology Review* 26 (2018): 26-70, 64-67.

peacetime.[14] An international armed conflict (IAC) is one triggered by the use of military force by one State against another State — arguably irrespective of the intensity of the fighting or the status of those fighting.[15] Thus, a border incident involving the use of military force by one State in the territory of another State, including when force is used against non-State actors, is enough to trigger an IAC.[16] Conversely, a non-international armed conflict (NIAC) is a conflict between a State and a non-State actor or between non-State actors themselves. Unlike an IAC, NIACs are characterised by the crossing of a threshold of hostilities. As noted by the International Criminal Tribunal for the Former Yugoslavia (ICTY), this threshold is crossed 'whenever there is a resort to armed force between States or protracted armed violence between governmental authorities and organised armed groups or between such groups within a State'.[17] Importantly, IHL applies not only on the actual battlefield or theatre of hostilities but throughout the entire territory under effective control of the parties. [18] There, it is relevant in all situations with a sufficient nexus with the armed conflict.[19] IACs and NIACs are thus the first types of 'crises' to which this report and its interpretative guidance apply.

Broadly speaking, IHL has two aims. First, to protect civilians, prisoners of war, the sick, and other persons 'hors de combat'. Civilians are defined by exclusion: they comprise all individuals who are not combatants, i.e. not engaged in military operations. Prisoners of war are combatants taken into custody by another party to the conflict. The sick and other persons 'hors de combat' are combatants who find themselves unable to take part in military operations.

Second, IHL aims to place limits on the conduct of hostilities. Such limits come from the prohibition of using certain means and methods of warfare, such as biological, incendiary, and chemical weapons. IHL is grounded in both treaties and customary international law. The most prominent conventional rules of IHL are found in the four 1949 Geneva Conventions[20] and their 1977 Additional Protocols.[21] These rules reflect, for the most part, international custom.[22] This means that, oftentimes, conventional IHL is more protective than its customary counterpart. Conventional rules of IHL are binding on the State parties to respective treaties. They also bind non-State groups that are parties to an armed conflict and have agreed to abide by IHL treaties. While all 196 UN Member States have ratified the 1949 Geneva Conventions, their Additional

---

[14] International Committee of the Red Cross (ICRC), "Commentary on the First Geneva Convention," last modified November 8, 2023, https://ihl-databases.icrc.org/ihl/full/GCi-commentary.

[15] Dapo Akande, "Classification of Armed Conflicts: Relevant Legal Concepts," *International Law and the Classification of Conflicts,* 2012, http://dx.doi.org/10.2139/ssrn.2132573; Separate Opinion of Judge Shahabudden, Prosecutor v Duško Tadić (Tadić Case), Appeal Judgment, No. IT-94-1-A (ICTY July 15, 1999) ("Tadić Appeal Judgment 1999"), paras 7-27.

[16] Ibid.

[17] Tadić Case, Decision on the Defence Motion for Interlocutory Appeal on Jurisdiction, No. IT-94-1 (ICTY October 2, 1995) ("Tadić 1995 Judgement"), para 70.

[18] Ibid, paras 68-69.

[19] Ibid, para 69.

[20] Geneva Convention (I) for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field 1949, 75 UNTS 31; Geneva Convention (II) for the Amelioration of the Condition of Wounded, Sick and Shipwrecked Members of Armed Forces at Sea 1949, 75 UNTS 85; Geneva Convention (III) relative to the Treatment of Prisoners of War 1949, 75 UNTS 135; Geneva Convention (IV) relative to the Protection of Civilian Persons in Time of War, 75 UNTS 287.

[21] ICRC, "Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts," June 8, 1977, 1125 UNTS 3 ("Additional Protocol I"); ICRC, "Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of Non-International Armed Conflicts," June 8, 1977. 1125 UNTS 609 ("Additional Protocol II").

[22] ICRC, "Customary International Law Database," Accessed November 26, 2023, https://ihl-databases.icrc.org/en/customary-ihl/v1.

Protocols have fewer State parties. Under customary international law, IHL binds all States in the world, as well as non-State groups that are parties to an armed conflict. A non-State group becomes a party to an armed conflict when it can be characterised as an 'organised armed group'.[23] This occurs when they are under responsible military command and have the capacity to sustain military operations.[24] Individuals have rights but no direct obligations or responsibilities under IHL.

II. How does IHL apply to online platforms?

The internet and other ICTs have had a significant impact on active or potential armed conflicts. These technologies can be used to suppress internal dissent, meddle in democratic elections, intensify the spread of incitement to violence, or contribute to crimes against humanity.[25] Notably, online platforms, including social media, search engines, and messaging applications, have been an important vector for the production, hosting, and dissemination of different types of conflict-related content. But these are mere speech acts. As such, they cannot, in and of themselves, cause any of the harmful consequences usually associated with armed conflict, such as the use of military force or loss of life. After all, those targeted with the speech still need to take action in response to a piece of online content for it to have any effect in the outside world. As such, it is challenging to determine the effects of online content. Information has been disseminated at an unprecedented scale and speed online, and it is difficult to pinpoint who, when, and how people have been affected by it.

This means that the extent to which IHL applies to online content is not immediately evident. Notably, it is unlikely that an armed conflict can be triggered by operations involving the dissemination of online content *alone*.[26] IACs involve the use of military force. Thus, to trigger an IAC, digital technologies must cause effects comparable to kinetic operations.[27] For their part, NIACs involve intense fighting that lead to significant destruction and harm. Again, this outcome can hardly be caused by online content in and of itself. Nevertheless, the use or operation of online platforms by different actors in the context of an armed conflict may be covered by different rules of IHL (see Subsections 'a' to 'd' below).

Not every dissemination of online content by a State or non-State actor during an armed conflict triggers the application of IHL. The publication or sharing of digital content on those platforms must be sufficiently connected to the armed conflict for IHL to become relevant. For any conduct to

[23] Tadić 1995 Judgement (note 17 above), para 70.
[24] Article 1(1), Additional Protocol II (note 21 above); Akande (note 15 above), 32-33.
[25] See, e.g., Samantha Bradshaw, Hannah Bailey and Philip N. Howard, "Industrialised Disinformation: 2020 Global Inventory of Organized Social Media Manipulation," *Computational propaganda Research Project, Oxford Internet Institute, University of Oxford*, 2021, https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2021/02/CyberTroop-Report20-Draft9.pdf.
[26] Priya Urs, Talita Dias, Antonio Coco, and Dapo Akande, "International Law Applicable to Cyber Operations targeting the Healthcare Sector," *Oxford Institute for Ethics, Law and Armed Conflict*, April 21, 2023, https://www.elac.ox.ac.uk/new-research-report-the-international-law-protections-against-cyber-operations-targeting-the-healthcare-sector/.
[27] Michael Schmitt, ed., *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations* (2nd ed.) (Cambridge: CUP, 2017) ("Tallinn Manual 2.0"), Rule 69, para 1, at 330; Military and Paramilitary Activities in and against Nicaragua (Merits) (Nicaragua v United States of America), Judgment of 27 June 1986, ICJ Rep 14 ("Nicaragua case"), para 195.

be subject to IHL, it must be shaped by or dependent upon the armed conflict.[28] Whether or not a sufficient nexus with the armed conflict ('belligerent nexus') exists will depend on the applicable IHL rule and the factual circumstances in question. But relevant factors include: a) the status of the speaker(s) and addressee(s) (including whether they are civilians or combatants), b) the content of the information published or shared (i.e., the extent to which it is linked to the armed conflict), c) the intention behind it (in particular, whether it seeks to advance or facilitate the military aims of a party to the conflict), and d) the means of dissemination (including the particular online channels where the information is published or shared, such as a military page or group versus civilian 'online spaces').

To be sure, IHL is only formally binding on States and non-State groups that are parties to an armed conflict. It is unlikely that online platforms themselves will be qualified as parties to an armed conflict unless they organise themselves as armed groups and directly engage in military operations. However, an online platform may be formally owned by or under complete dependence or effective control of a State or non-State group that is a party to an armed conflict. In those instances, as an organ or agent of the State or non-State party to the conflict, the platform will be bound by the rules of IHL.

The more likely scenario is where online platforms are *used by the parties to an armed conflict*, including States and non-State actors. This has been prominent in the recent illegal invasion of Ukraine. Social media and other online platforms have been used by both civilian and military officials to advance their military and political goals. Notably, States have regularly made use of online war propaganda or have spread state-sponsored propaganda. Recently, the Ukrainian military has released videos of Russian prisoners of war (POW) on different online platforms, depicting their humiliation and despair. While IHL clearly states that POW must at all times be protected against acts of violence, intimidation, insults, and public curiosity, the law does not per se ban the publication of images of POWs. Thus the war in Ukraine is now providing a unique test of the relevance and application of IHL on content governance in times of war.

Online platforms have also been frequently used *by individuals or groups,* including civilians, who do not have an official or direct link to the parties to an armed conflict. Again, the war in Ukraine demonstrates how different stakeholders have turned to social media and other online platforms to document the hostilities or speak out against or in favour of the war. While those stakeholders do not have obligations under IHL, they may be entitled to *protection* by different rules of IHL, whether as civilians, prisoners of war, other persons hors de combat, or specially protected actors, including health, religious, or humanitarian personnel.

At the same time, this protection may be lost and those various actors may become lawful targets under IHL insofar as they *directly participate* in the hostilities.[29] 'Direct participation in hostilities' involves activities that aim to support one party to the conflict by *directly causing harm* to another party, either by inflicting death, injury, or destruction, or by undermining the enemy's military *operations or capacity*.[30] Clear examples include causing physical damage to military objects or

[28] Tadić Case, Appeal Judgment, No. IT-94-1-A (ICTY July 15, 1999); Antonio Cassese, "The Nexus Requirement for War Crimes," *Journal of International Criminal Justice* 10 (2015): 1395-1417.
[29] See Rule 6, ICRC Customary IHL Database (note 17 above).
[30] Niels Melzer, "Interpretative Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law," *ICRC*, June 11, 2020, https://www.icrc.org/en/doc/assets/files/other/icrc-002-0990.pdf.

personnel or disrupting military deployment, logistics, and communication. According to the ICRC, electronic interference with military computer networks and the transmission of tactical targeting intelligence for a specific attack also amount to direct participation in hostilities. These activities may well occur through online platforms. In fact, it has been reported that civilians have been using existing online platforms or purpose-built applications to directly report on the movement of Russian troops, including air strikes.[31] However, the provision of financial, administrative, and political support only amounts to indirect participation in armed conflict. Thus, it does not lead to loss of protection from targeting.[32]

In those instances, despite lacking binding legal obligations under IHL, platforms have a social responsibility to observe applicable rules, in accordance with the UN Guiding Principles on Business and Human Rights. This is true insofar as they might host or promote online content that may amount to or support a violation of IHL by a party to an armed conflict. Online platforms should also ensure that they do not become lawful military targets by directly participating in hostilities. This will happen whenever they know or should have known that their platforms are being used to directly cause significant risk to a party to the conflict, in the sense discussed earlier.[33]

Three sets of IHL rules are particularly relevant to the dissemination of online content. They may apply to and limit the use of such content by parties to an armed conflict and therefore should be followed by online platforms.

a) Obligations to respect and ensure respect for IHL

All States, whether or not parties to an armed conflict, must respect and ensure respect for IHL.[34] This obligation also applies to non-State groups that are parties to an armed conflict.[35] The obligation to ensure respect for IHL means that even before an armed conflict actually starts, States and non-State parties to an armed conflict must not encourage violations of IHL.[36] They must also do everything in their power to prevent such violations.[37] This is an obligation of conduct rather than one of result. It requires States and non-State armed groups to exercise due diligence whenever they should have known of actual or potential violations of IHL by entities whose

---

[31] Drew Harwell, "Instead of consumer software, Ukraine's tech workers build apps of war," *The Washington Post*, March 24, 2022, https://www.washingtonpost.com/technology/2022/03/24/ukraine-war-apps-russian-invasion/; Michael N. Schmitt, "Ukraine Symposium – Using cellphones to gather and transmit military information, A Postscript," *Articles of War*, November 4, 2022 ,https://lieber.westpoint.edu/civilians-using-cellphones-gather-transmit-military-information-postscript/; Yaroslav Druziuk, "A Citizen-like chatbot allows Ukrainians to report to the government when they spot Russian troops — here's how it works", *Business Insider*, April 18, 2022, https://www.businessinsider.com/ukraine-military-e-enemy-telegram-app-2022-4?r=US&IR=T.

[32] Ibid.

[33] ICRC, Interpretative Guidance (note 30 above), 60.

[34] Article 1 common to Geneva Conventions I, II, III, IV (see note 20 above); Rules 139 and 144, ICRC Customary IHL Database (note 17 above); Jean S. Pictet, *The Geneva Conventions of 12 August 1949 Commentary - IV Geneva Convention relative to the Protection of Civilian Persons in Time of Warl* (Geneva: ICRC, 1958), 16; Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory, Advisory Opinion of July 9, 2004, ICJ Reports (2004) p 136, paras 158-159.

[35] Ibid.

[36] ICRC, "*Commentary on the First Geneva Convention,*" Commentary to Article 1, paras 154 and 158-163., 2016, https://ihl-databases.icrc.org/ihl/full/GCi-commentary.

[37] Ibid, paras 127-128 and 185.

behaviour they have the ability to influence.[38] The duty to ensure respect for IHL applies with respect to parties' own armed forces or other persons acting under their effective control, or by combatants or other actors acting on behalf of other parties.[39] Thus, online content that generally or specifically supports acts or omissions that would violate IHL, before or during an armed conflict, must not be posted, shared, or amplified on online platforms. For example, content calling upon combatants to target civilians would violate this obligation. Likewise, internet shutdowns or online service bans could run contrary to this rule insofar as such actions are taken to support breaches of IHL.[40] A case in point is where Internet connection is essential to the delivery of humanitarian aid, medical, or religious services. As noted by the recent report of the UN Human Rights Commissioner on the situation in Myanmar, mobile telecommunications blockages and internet shutdowns have hampered the delivery and access to humanitarian aid by civilians.[41] When intentional, the obstruction or denial of those services may amount to war crimes (such as torture and other degrading treatment, starvation) or crimes against humanity (such as inhumane acts or persecution), provided that the other elements of crimes are met in the circumstances.[42]

b) Obligations applicable to the protection of civilians and civilian objects

Several rules of IHL — general and specific — seek to protect civilians and civilian objects during armed conflict. Some are relevant to the dissemination of digital information on online platforms by the parties to an armed conflict, as well as by civilians themselves.

The first of these rules is the basic principle of distinction between civilians and combatants, and between civilian and military objectives. It is grounded both in treaty and customary IHL and applies in both IAC and NIAC.[43] The basic principle of distinction has three key components.

- First and foremost, it requires parties to an armed conflict to *distinguish* between civilians and combatants and between civilian objects and military objectives *at all times*.[44]
- Secondly, it requires parties to *direct military operations* against *combatants* and *military* objectives.[45]
- Thirdly, it requires parties to *protect* civilians and civilian objects against the *dangers* arising from *military operations*.[46] This is an obligation of conduct requiring due diligence on the part of parties to an armed conflict.

---

[38] Antonio Coco and Talita de Souza Dias, "'Cyber Due Diligence': A Patchwork of Protective Obligations in International Law," *European Journal of International Law* 32, no 3 (2021): 771–806, 775.
[39] Nicaragua case (note 27 above), para. 220.
[40] See the UN Human Rights Council resolution (A/HRC/RES/44/20) on "The promotion and protection of human rights in the context of peaceful protests," on July 17, 2020 which stipulates that blunt measures such as blanket Internet shutdowns, sometimes for prolonged periods, contravene international law, affecting States' obligations to respect, in addition to freedom of expression and the right to peaceful assembly, a wide range of rights, including freedom of association and of movement and the rights to health and education.
[41] UN Human Rights Council, "Situation of human rights in Myanmar since 1 February 2021: Report of the United Nations High Commissioner for Human Rights," A/HRC/53/52, June 28, 2023, paras 39-41.
[42] Ibid, para 12.
[43] Articles 48, 51(2) and 52(2), Additional Protocol I (note 21 above); Article 13(2), Additional Protocol II (note 21 above); Rules 1 and 7, ICRC Customary IHL Database (note 17 above).
[44] Article 48, Additional Protocol I (note 43 above).
[45] Ibid.
[46] Article 51(1), Ibid.

Though the first component of the principle applies continuously during armed conflict — even *before* parties engage in specific military operations — the second and third components only apply in the course of military operations. Thus, parties to an armed conflict and online platforms should *always* distinguish between civilian and military users and the online content that they produce or share. However, whether or not the other two components of the principle of distinction — the rule of targeting and the rule of protection — are relevant to online platforms will depend on whether or not the dissemination of online content amounts to a *military operation*.

The concept of 'military operation' is broader than that of an attack, which triggers different rules applicable to the conduct of hostilities (see section 'b' below). The former refers to military operations during which violence is used, that is, 'all movements and acts related to hostilities that are undertaken by armed forces' (as opposed to ideological, political, or religious campaigns).[47] 'Attacks', on the other hand, only cover 'combat action',[48] i.e., specific 'acts of violence against the adversary, whether in offence or in defence'.[49] Violence, in this context, has been understood to mean 'acts *reasonably expected* to cause *injury or death* to persons or *physical damage or destruction* to objects by means or effects'.[50] Thus, it is difficult to see how the dissemination of online content may amount to an attack, unless it can be reasonably expected to cause physical effects. This may occur when the content is exceptionally grave, such as direct incitement to imminent and likely violence. However, when operationally connected to or directly seeking to further battlefield action, the dissemination of online content may be part of a military operation. This may occur, for instance, when parties to an armed conflict gather or share intelligence via online platforms, or intercept online military communications on those platforms. Electronic surveillance of civilians may also amount to a military operation, insofar as the civilian data gathered advances a military goal. For example, journalists, activists, and researchers might be in possession of sensitive information that might have strategic military value for a party to the conflict.

In those instances, parties to an armed conflict must not only distinguish between civilians and combatants but also refrain from directing any military operation, including those involving online content, against civilians and civilian objects. As noted by the ICRC, the better view is that civilian data also constitutes a civilian object that must be protected from the effects of military operations and attacks.[51] Parties must also take steps to protect civilians and civilian objects from the dangers arising from military operations. In the digital age, this means that online content that is either part of a military operation against civilians or civilian objects (including their data) is probably contrary to IHL and should therefore be removed or otherwise limited. Likewise, when operating in contexts

---

[47] ICRC, "Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), 8 June 1977 - Commentary of 1987" ("1987 Commentary to Protocol I"), Commentary to Article 48, https://ihl-databases.icrc.org/en/ihl-treaties/api-1977/article-48/commentary/1987?activeTab=undefined.

[48] Ibid, "Commentary to Article 49", https://ihl-databases.icrc.org/en/ihl-treaties/api-1977/article-49/commentary/1987?activeTab=undefined.

[49] Article 49(1), Additional Protocol I (note 21 above).

[50] Tallinn Manual 2.0 (note 27 above), Rule 92, para 2 (emphasis added).

[51] Kubo Maćak and Tilman Rodenhäuser, "Towards common understandings: the application of established IHL principles to cyber operations," *ICRC*, March 7, 2023: https://blogs.icrc.org/law-and-policy/2023/03/07/towards-common-understandings-the-application-of-established-ihl-principles-to-cyber-operations/.

where an armed conflict is taking place, platforms should exercise due diligence in trying to protect civilian users, civilian content, and other civilian data such as personal information, from the dangers arising from military operations carried out online or offline.

A related rule of IHL is the principle of precaution, which requires parties to an armed conflict to take constant care to spare the civilian population, civilians, and civilian objects in the conduct of *military operations*.[52] This principle entails that feasible precautions must be taken to avoid, and in any event to minimise, incidental loss of civilian life, injury to civilians, and damage to civilian objects.[53] This is an obligation of conduct requiring active steps to prevent or mitigate harm to civilians and civilian objects that is incidental to the targeting of combatants or military objectives. Like the principle of distinction, this rule applies to military operations, not just attacks.

In the context of the principle of precaution, a 'military operation' has been understood to mean 'any movements, manoeuvres and other activities whatsoever carried out by the armed forces with a view to combat'.[54] Thus, whether the dissemination of online content amounts to a military operation triggering the applicability of the principle of precaution depends on its direct link to military combat operations. Examples include tactical or operational data-gathering or sharing, including through online platforms. The spread of false or misleading information may also be directly connected to a military manoeuvre if it seeks to secure a military advantage, such as the targeting of a certain person or object. Insofar as a certain piece of online content can be said to directly advance military operations, then it is subject to the principle of precaution. For example, online disinformation spread as part of a false flag operation to deceive the adversary must not lead to violence or harm against civilians. In the online environment, this principle requires parties to an armed conflict to exercise due diligence in preventing incidental harm to civilians or civilian objects, including civilian content and data. Online platforms should observe the same rule and exercise the same degree of diligence whenever operating in conflict zones.

The principle of proportionality in attacks follows on from the basic rule of distinction and is closely connected to the principle of precaution. In IHL, proportionality prohibits parties to an armed conflict from launching an *attack* that may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated.[55] Thus, the applicability of this principle is limited to situations involving an attack, that is, acts of violence that are reasonably expected to cause injury or death to persons, or damage or destruction to objects by means or effects, whether in offence or in defence. As noted earlier, only on rare occasions will the dissemination of online content rise to the level of an attack, i.e., be reasonably expected to cause physical harm to civilians or civilian objects. This will depend on the seriousness of the content at hand, the intention of the speaker and the objective causal link between the information and any physical harm, i.e. whether it is reasonably expected to cause such an effect.

---

[52] Rule 15, "ICRC Customary IHL Database" (note 17 above); Article 57(1), Additional Protocol I (note 21 above); Article 13(1) Additional Protocol II (note 21 above).
[53] Ibid.
[54] ICRC, "1987 Commentary to Protocol I" (note 47 above), "Commentary to Article 57", para 2191, https://ihl-databases.icrc.org/en/ihl-treaties/api-1977/article-57/commentary/1987?activeTab=undefined.
[55] Rule 14, "ICRC Customary IHL Database" (note 17 above); Articles 51(5)(b) and 57, Additional Protocol I (note 21 above).

As noted earlier, the dissemination of particularly serious pieces of online content may amount to an attack for the purposes of triggering the applicability of the principle of precaution in IHL.

Examples may include direct incitement to or threats of violence that are reasonably expected to cause harm to civilians or civilian objects. Thus, parties to an armed conflict must refrain from publishing or disseminating such content, insofar as the civilian harm reasonably expected is excessive to the military advantage anticipated from the attack. Online platforms should seek to remove or limit the dissemination of those pieces of content. However, deference should be given to the parties to an armed conflict when making this assessment.

Another rule of IHL seeking to protect civilians that might be relevant in the online context is the prohibition of 'acts or threats of violence the primary purpose of which is to spread terror among the civilian population'.[56] This rule applies not only to acts of actual violence but also threats thereof.[57] These may well be speech acts disseminated online.[58] Moreover, the prohibition of terrorising civilians does not require civilians to be actually terrorised — an intention or purpose to do so suffices.[59] However, in the context of IHL (as opposed to ICL), such purpose is to be objectively assessed based on the nature of the act in question. Relevant factors include the content and context of the online information in question. Examples of online content that might amount to prohibited terrorisation of civilians include incitement to or propaganda for violence or hostility or direct threats of violence, such as the threat to use nuclear weapons to annihilate a civilian population or part thereof.[60] The online dissemination of false information, in and of itself, is unlikely to amount to the terrorisation of civilians, except when it also involves threats of violence.

Online content may also violate the related rule prohibiting parties to a NIAC from ordering the displacement of the civilian population, in whole or in part, for reasons related to the conflict, unless the security of civilians or imperative military reasons so demand.[61] As others have noted, the meaning of 'ordering' should be understood broadly to include not just direct orders but also any deliberate action by a party to a NIAC.[62] Thus, any online content that conveys an order or any deliberate action seeking to displace civilians in a NIAC, such as false or misleading information seeking to coerce civilians to flee their homes, is contrary to IHL.

c) Obligations applicable to specially protected actors

Beyond civilians, IHL specifically protects certain actors that are particularly vulnerable in armed conflict. The first set of those actors comprises medical, religious, and humanitarian personnel.[63]

---

[56] Article 51(2), Additional Protocol I (note 21 above); Article 13(2), Additional Protocol II (note 21 above); Rule 2, "ICRC Customary IHL Database" (note 17 above).
[57] ICRC, "1987 Commentary to Protocol I" (note 47 above), "Commentary to Article 51", para. 1940, https://ihl-databases.icrc.org/en/ihl-treaties/api-1977/article-51/commentary/1987?activeTab=undefined.
[58] Tallinn Manual 2.0 (note 27 above), Rule 98, para 6.
[59] Prosecutor v Galić (Galić Case), Appeals Judgement, No. IT-98-29-A (ICTY November 30, 2006), paras 103-104.
[60] Tallinn Manual 2.0 (note 27 above), Rule 98; Tilman Rodenhäuser, "International Humanitarian Law and The Limits of Information or Psychological Operations during Armed Conflicts". In *The Oxford Process on International Law Protections in Cyberspace: A Compendium*, April 2021, https://www.elac.ox.ac.uk/wp-content/uploads/2022/10/Oxford-Process-Compendium-Digital.pdf, 434-449.
[61] Rule 129, ICRC Customary IHL Database (note 17 above); Article 17, Additional Protocol II (note 21 above).
[62] Mélanie Jacques, *Armed Conflict and Displacement: The Protection of Refugees and Displaced Persons under International Humanitarian Law* (Cambridge: CUP, 2012), 62; Rodenhäuser (note 60 above), 443-444.
[63] Rules 25, 26, 31, 32, ICRC Customary IHL Database (note 17 above).

Parties to an armed conflict must not only respect but also actively protect those actors. This means that they must exercise due diligence to prevent, stop, and redress harm or disruption to their activities. This rule applies at all times, irrespective of the type of operation; that is, whether it amounts to an attack, a military operation, or a mere disruption to the delivery of medical, religious, and humanitarian services.

Thus, parties to an armed conflict must refrain from disseminating online content that is detrimental to the activities of medical, religious, and humanitarian personnel, such as hate speech, or mis- or disinformation directed against their staff.[64] They must also exercise due diligence to prevent or mitigate the effects of those types of online content, such as by adopting a basic legal framework requiring platforms to moderate such content in line with IHL. Irrespective of the adoption of such national legal frameworks, online platforms should seek to delete or otherwise limit the dissemination of these types of content on their websites or applications.

Numerous rules of IHL also protect prisoners of war. Notably, some of these rules prohibit or otherwise limit the disclosure of private information relating to prisoners of war. Examples include Articles 13(2) and 14 of the Third Geneva Convention, which protect prisoners of war from insults and public curiosity and entitle them to respect for their persons and honour.[65] According to the ICRC, this means that the dissemination of videos or images depicting prisoners of war, such as the ones seen in the context of the war in Ukraine,[66] is generally unlawful, even if it serves as proof of life.[67] This is so unless the identity of prisoners can be protected, or there is a compelling public or individual interest in revealing their identity, such as when prisoners are missing or accused of war crimes.[68]

INTERNATIONAL CRIMINAL LAW (ICL)

I. Basic concepts of ICL

Unlike IHL, ICL only imposes criminal responsibility on individuals, i.e., natural persons. This applies under customary international law and under existing treaties defining international crimes. States have considered the concepts of State and corporate criminal liability for international crimes, particularly in the context of the UN International Law Commission (ILC)'s Draft Code of Crimes Against the Peace and Security of Mankind.[69] But both were ultimately

[64] See ICRC, "Harmful Information – Misinformation, disinformation and hate speech in armed conflict and other situations of violence: ICRC initial findings and perspectives on adapting protection approaches," July 9, 2021, https://www.icrc.org/en/publication/4556-harmful-information-misinformation-disinformation-and-hate-speech-armed-conflict.

[65] See ICRC, "Convention (III) relative to the Treatment of Prisoners of War," August 12, 1949; and "Commentary of 2020, Articles 13 and 14", https://ihl-databases.icrc.org/en/ihl-treaties/gciii-1949/article-13/commentary/2020?activeTab=undefined and https://ihl-databases.icrc.org/en/ihl-treaties/gciii-1949/article-14?activeTab=undefined.

[66] Blake (note 10 above).

[67] ICRC, "1987 Commentary to Protocol I" (note 47 above), paras 1624-1625, https://ihl-databases.icrc.org/en/ihl-treaties/api-1977/article-13/commentary/1987?activeTab=undefined.

[68] Ibid, paras 1626-1627.

[69] International Law Commission (ILC), "Summaries of the Work of the International Law Commission Draft code of crimes against the peace and security of mankind* (Part II) — including the draft Statute for an international criminal court," last updated on June 23, 2023, https://legal.un.org/ilc/summaries/7_4.shtml.

rejected given divergences across domestic legal frameworks. This means that neither States nor corporations, including online platforms, may, at present, be held criminally responsible for the commission of crimes under international law.

However, that is not to say that States are precluded from providing, in their domestic laws, for State or corporate criminal liability for the commission of or participation in international crimes. In fact, many States already do so. Notably, under Lebanese domestic law, corporations may be held responsible for committing or participating in different crimes. This has meant that the Special Tribunal for Lebanon heard two cases involving the responsibility of media companies for contempt of court offences.[70] This is an international tribunal established by the UN Security Council with jurisdiction over domestic crimes committed in the context of the terrorist attack of 14 February 2005.[71] The attack killed 22 people, including former Lebanese prime minister of Lebanon, Rafik Hariri. Likewise, it is open to States to conclude specific treaties or develop rules of customary international law that provide for State or corporate criminal responsibility for international crimes, including existing or new ones.

Nevertheless, States do have certain obligations under treaty and customary international law to prevent or punish certain international crimes, including genocide, war crimes, and arguably crimes against humanity.[72] The same is true of certain transnational offences, such as the financing of terrorism and drug trafficking.[73] For the most part, these are obligations of conduct, requiring States to exercise due diligence in seeking to prevent and punish those offences.[74] In particular, under certain treaties, such as the Genocide Convention[75] as well as the Geneva Conventions and their Additional Protocol I,[76] States are required to enact specific domestic legislation criminalising certain war crimes, as well as to investigate them as far as possible and prosecute perpetrators as appropriate.[77] Likewise, if the commission of international crimes can be attributed to a State, the latter will be held internationally responsible for a breach of international law, in accordance with the rules of State responsibility.[78] Insofar as some of these obligations are also grounded in IHL and

[70] Al Jadeed S.A.L. & Ms Khayat case, Appeals Judgment, No. STL-14-05 (Special Tribunal for Lebanon - STL, March 8, 2016); Akhbar Beirut S.A.L. & Mr Al Amin case, Appeals Judgment, STL-14-06 (STL, July 15, 2016).

[71] UN Security Council, "Resolution 1757, S/RES/1757(2007"), May 30, 2007. See also Stéphane Dujarric, "Statement attributable to the Spokesperson for the Secretary-General on the Special Tribunal for Lebanon," January 12, 2023, https://www.un.org/sg/en/content/sg/statement/2023-01-12/statement-attributable-the-spokesperson-for-the-secretary-general-%E2%80%93-the-special-tribunal-for-lebanon.

[72] Federica D'Alessandra and Shannon Raj Singh, "Operationalizing Obligations to Prevent Mass Atrocities: Proposing Atrocity Impact Assessments as Due Diligence Best Practice," *Journal of Human Rights Practice* 14, no 3 (2022) 769–793, 777-780.

[73] See Article 9, UN General Assembly, International Convention for the Suppression of the Financing of Terrorism, December 9, 1999, No. 38349; Article 4, UN Economic and Social Council (ECOSOC), United Nations Convention Against Illicit Traffic in Narcotic Drugs and Psychotropic Substances, December 19, 1988, 95 UNTS 1582.

[74] Case Concerning Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosnia-Herzegovina v Yugoslavia), Judgment of February 26 2007 (Merits), ICJ Rep 2007 ("Bosnian Genocide case"), para 430.

[75] UN General Assembly, "Convention on the Prevention and Punishment of the Crime of Genocide", December 9, 1948, 78 UNTS 277.

[76] Article 50 Geneva Convention I; Article 51, Geneva Convention II; Article 130, Geneva Convention III; Article 147, Geneva Convention IV (see note 20 above); Articles 11 and 85, Additional Protocol I (note 21 above).

[77] Rule 158, ICRC Customary IHL Database (note 17 above).

[78] See Articles 1-11, UN General Assembly, "Responsibility of States for internationally wrongful acts," A/RES/62/61, January 8, 2008; Bosnian Genocide case (note 74 above), para 431 (with respect to genocide).

IHRL, corporations, including online platforms, have a responsibility to refrain from and prevent, as far as possible, the commission of international crimes.[79]

There are four so-called 'core' or 'atrocity' crimes defined and criminalised under international law, including treaties and customary international law. These are genocide, war crimes, crimes against humanity and the crime of aggression. Their existence makes up the second type of crisis that this report addresses. Unlike IHL, ICL applies in both peacetime and armed conflict. The exception is war crimes, which constitute serious violations of IHL and thus can only occur in armed conflict.

Genocide can be defined as the commission of certain serious acts, such as killing and causing serious bodily harm to individuals, committed with an intent to destroy, in whole or in part, a national, ethnical, racial, or religious group.[80] Genocide is to a group what murder is to an individual — a crime seeking the physical or biological destruction of entire groups or a significant part thereof. Neither the destructive intent nor conduct need to be successful, meaning that a serious act, accompanied by a discriminatory intent to destroy the group, suffices.

In contrast, crimes against humanity are, in broad terms, grave violations of individual human rights, committed in a systematic way. They comprise the commission of one or more 'inhuman' acts, such as murder, extermination, or enslavement, as part of a widespread or systematic attack directed against any civilian population, with knowledge of the attack.[81]

For their part, war crimes are serious violations of IHL criminalised under treaty or customary international law. These include the grave breaches of the Geneva Conventions and their Additional Protocol I as well as other serious violations of the laws and customs of war applicable in both IAC and NIAC.[82] Examples of war crimes in both types of armed conflict are murder or wilful killing, torture or other forms of other inhuman treatment against protected persons, and intentional attacks against civilians.[83]

The crime of aggression is the most controversial of the four core crimes, given that it criminalises the use of force by one State against another, in violation of Article 2(4) of the UN Charter and its customary counterpart. In the Rome Statute of the International Criminal Court (ICC), the crime of aggression is defined as 'the planning, preparation, initiation or execution, by a person in a position effectively to exercise control over or to direct the political or military action of a State, of an act of aggression which, by its character, gravity, and scale, constitutes a manifest violation of the Charter of the United Nations.'[84] Conversely, under customary international law, the crime of aggression includes 'wars of aggression' and 'invasions', i.e., only particularly serious illegal uses of force.[85]

---

[79] UN Guiding Principles on Business and Human Rights (note 12 above).
[80] Article I, Genocide Convention (note 75 above).
[81] Article 7, UN General Assembly, Rome Statute of the International Criminal Court (last amended 2010), 17 July 1998 ("Rome Statute"); Article 2, ILC, Draft articles on Prevention and Punishment of Crimes Against Humanity, *Yearbook of the International Law Commission* vol. II, Part Two , A/74/10 (2019).
[82] Article 8, Rome Statute (note 80 above); Article 3, UN Security Council, Statute of the International Criminal Tribunal for the Former Yugoslavia (as amended on 17 May 2002), May 25, 1993 ("ICTY Statute"); Tadić 1995 Judgement (note 17 above), paras 71-137.
[83] Ibid.
[84] Article 8 bis, Rome Statute (note 81 above).
[85] Marko Milanovic, "Aggression and Legality: Custom in Kampala," *Journal of International Criminal Justice* 10, no. 1 (2012): 165–187, 184; Talita de Souza Dias, "The Activation of the Crime of Aggression before the International Criminal

As a general rule, all core crimes may be committed by direct perpetrators acting alone, jointly with others, and/or through one or more persons. Individuals may also be held criminally liable for participating in the commission of those offences as accomplices or accessories. Under customary international law and the ICC Statute, individuals may be criminally liable as accomplices for planning, instigating, ordering, or otherwise aiding and abetting a crime. The exception is the crime of aggression as defined in the Rome Statute, which limits liability for this crime to persons in a position to effectively exercise control over or direct the political or military action of a State.[86] However, under international custom, the crime of aggression may be committed not only by State leaders themselves but also accomplices, including corporate leaders and ordinary individuals.[87] Individuals may also be liable for international crimes as military commanders or civilian superiors when they fail to prevent or punish the crimes committed by their subordinates. All core crimes may be punishable if completed or attempted.[88] Furthermore, the inchoate crime of direct and public incitement to commit genocide is also punishable under international law.[89] Unlike instigation, incitement to commit genocide does not require any actual acts of genocide to occur as a result.[90]

II. How does ICL apply to online platforms?

This means that speech acts, including online, are, in and of themselves, unlikely to amount to the *commission* of international crimes. However, they may well constitute different forms of *complicity* in those crimes.[91] This may occur insofar as the perpetrator has the relevant intention to plan, instigate, order, or aid and abet the crime in question.[92] In the case of aiding and abetting — the widest category of complicity — the speech act(s) must 'substantially contribute' to the commission of the principal crime.[93] Substantial contribution implies an objective, direct, and more than *de minimis* effect[94] on the offence carried out by the principal perpetrator in the 'offline' world. But this is not a very high threshold: the conduct need not be criminal. What matters is that it has some factual effect on the principal crime.[95] The mental element required for aiding and abetting differs between customary international law and the ICC Statute. The former is satisfied by an intention to aid and abet plus knowledge that the conduct will or is substantially likely to

Court: Some Overlooked Implications Arising for States Parties and Non-States Parties to the Rome Statute," *Journal of Conflict and Security Law* 24 (2019): 567–591, at 585.

[86] Article 25(3 bis), Rome Statute (note 81 above).

[87] Nikola Hajdin, "Responsibility of Private Individuals for Complicity in a War of Aggression," *American Journal of International Law* 116, no 4 (2022): 788-797, 788; Claus Kreß, "On the Activation of ICC Jurisdiction over the Crime of Aggression," *Journal of International Criminal Justice* 16, no 1 (2018): 1–17, 6; Milanović (note 85 above), 183–184.

[88] E.g., Article 25(3)(f) Rome Statute (note 81 above); Robert Cryer, Friman Hakan, Darryl Robinson and Elizabeth Wilmshurst, *An Introduction to International Criminal Law and Procedure* (2nd ed) (Cambridge: CUP, 2010), 380.

[89] Article III(c), Genocide Convention (note 75 above); Article 25(3)(e), Rome Statute (note 81 above).

[90] Prosecutor v Nahimana et al, Appeal Judgement, No. ICTR-99-52 (ICTR November 28, 2007) ("Nahimana et al case"), paras 480, 677-678.

[91] See, e.g., UN Human Rights Council, "Report of the Independent International Fact-Finding Mission in Myanmar," A/HRC/39/64, September 17, 2018, paras 83-89.

[92] Tadić Appeal Judgment 1999 (note 15 above), para 229; Prosecutor v Orić, Appeals Judgment, No. IT-03-68-A (ICTY July 3, 2008), para 288; Prosecutor v Halilović, Trial Judgment, No IT-01-48-T (ICTY, November 16, 2005), para 286; Prosecutor v Blaškić, Appeal Judgment, No. IT-95-14-A (ICTY July 29, 2004), para 46.

[93] See Nahimana et al case (note 90 above), paras 480, 482.

[94] Cryer, Hakan, Robinson, Wilmshurst (note 88 above), 317.

[95] Ibid, 371.

assist a specific crime.[96] Conversely, Article 30 of the ICC Statute requires both knowledge and intention with respect to the principal's crime for aiding and abetting liability.

Instigation, which includes prompting, urging, or encouraging a crime, is another category of complicity that could potentially cover a wide range of online content. Yet, a relatively strong causal link between the instigation itself and the principal crime is required. While the instigation need not be a necessary condition for the principal offence, it must be a clear contributing factor to its occurrence.[97] Under customary international law, the instigator must also intend to instigate the principal offence with substantial knowledge of its likelihood and acceptance thereof.[98] The ICC Statute also recognises soliciting and inducing as a mode of liability.[99] However, it requires both knowledge and intention with respect to the principal crime.[100]

This all means that the dissemination of a range of online content — from false or misleading information, to data leaks, doxing, hate speech, and war propaganda — could *potentially* amount to one or more forms of complicity in the commission of international crimes. However, proving the causal link between the speech act and the actual result, as well as the necessary mental element under customary international law or the ICC Statute, will be decisive for anyone to be held liable as an accomplice to an international crime. This is true insofar as platforms' conduct of allowing content to be shared, hosted, or amplified on their platforms aids and abets in or otherwise contributes to the commission of an international crime, provided that the requisite mental element is present. Likewise State leaders or platform directors could be liable under the theory of superior liability insofar as their actions or omissions constitute a failure to prevent or punish the criminal actions of their subordinates of which they should have been aware.

As noted earlier, unlike aiding and abetting instigation, direct and public incitement to commit genocide is an inchoate crime in that it does not require any result or causal link between the latter and the incitement.[101] However, the mental element required for this offence is very high: inciters must themselves have a genocidal intent, i.e., the intent to destroy, in whole or in part, a national, ethnical, racial, or religious group.[102] Moreover, the incitement must be public and direct. This means that online content must be disseminated publicly or to a sufficiently wide audience to amount to this offence.[103] While directness does not require explicit expressions of incitement, these must be sufficiently clear to the particular audience in question.[104] Thus, context is key in determining whether seemingly neutral expressions are in fact coded language directly inciting

---

[96] Ibid, 373-374; Prosecutor v Taylor, Appeal Judgment, No. SCSL-03-01-A (Special Court for Sierra Leone, September 26, 2013), paras 413-438.

[97] Prosecutor v Blaškić, Judgment, No. IT-95-14-T (ICTY March 3, 2000), para 270; Prosecutor v Orić. Judgment, No IT-03-68-T (ICTY June 30, 2006), para 274; Sylvestre Gacumbitsi v The Prosecutor. Appeal Judgement, No ICTR-2001-64-A (International Criminal Tribunal for Rwanda - ICTR July 7, 2006), para 129.

[98] The Abbaye Ardenne Case, Trial of S.S. Brigadefuhrer Kurt Meyer (Case No. 22), Aurich, Germany (Canadian Military Court, December 10-18, 1945), available at https://www.legal-tools.org/doc/e9d181/pdf, 97-98; Cryer, Hakan, Robinson and Wilmshurst (note 88 above), 377.

[99] Article 25(3)(b), Rome Statute (note 81 above).

[100] Article 30, Rome Statute (note 81 above).

[101] Nahimana et al case (note 90 above); Akayesu Case, Trial Judgement, No. ICTR-96-4-A (ICTR September 2, 1998) ("Akayesu case"), para 562.

[102] Nahimana et al case (note 90 above), para 1012.

[103] Ibid, para 1011.

[104] Akayesu case (note 101 above), para 557.

genocide. As in the case of Rwandan cartoons and radio broadcasts,[105] referring to individuals or groups as animals that ought to be killed, such as snakes, rats, cockroaches, or other parasites, against a background of inter-communal violence, may amount to direct and public incitement to genocide.[106] The same is true in the digital environment, as the situation in Myanmar illustrates. There, the context of systematic oppression and violence against the Rohingya was key to the finding that Facebook posts referring to them as "Muslim dogs" and other hateful expressions could amount to genocide, including direct and public incitement thereto.[107] Online hate speech below this threshold is unlikely to amount to direct or public incitement to commit genocide or the *commission* of other international crimes, such as crimes against humanity.[108]

III. Online content and international prosecution

All four core crimes have been subject to the jurisdiction of various international criminal tribunals over time.[109] At present, the ICC is the only permanent international criminal tribunal with jurisdiction over genocide, crimes against humanity, and war crimes perpetrated in the territory or by a national of one of its States parties or a State that has accepted its jurisdiction for particular situations,[110] such as Ukraine.[111] The ICC's jurisdiction over the crime of aggression is more restrictive and requires not only the ratification of the Rome Statute but also specific acceptance of the definition and conditions for the prosecution of this crime by *both* the State party where the crime took place and the State of nationality of the accused.[112] Only 45 (out of the 123)[113] State parties to the Statute have accepted those provisions.[114] The UN Security Council may nonetheless refer to the Court cases involving any UN member State, irrespective of their status as parties to the ICC Statute, including for the crime of aggression.[115] Thus, to be prosecuted before the ICC, online speech acts must not only amount to one of the four core crimes but also fulfil those jurisdictional requirements as well as certain admissibility conditions, including sufficient gravity and the inability or unwillingness of domestic courts to prosecute the offence.[116]

The jurisdiction of the ICC and other international tribunals is ultimately grounded in States' own jurisdiction to prescribe and adjudicate these crimes domestically, including in cases where they

---

[105] See Al Jazeera World, "Rwanda: From hatred to reconciliation," September 29, 2015, https://www.aljazeera.com/program/al-jazeera-world/2015/9/29/rwanda-from-hatred-to-reconciliation.
[106] See Nahimana et al case (note 90 above), paras 477-672. See also UN Human Rights Council, Report of the detailed findings of the Independent International Fact-Finding Mission on Myanmar, A/HRC/39/CRP.2, September 17, 2018, paras 1316-1318.
[107] Amnesty International, "The Social Atrocity: Meta and the Right to Remedy for the Rohingya," September 29, 2022, https://www.amnesty.org/en/documents/asa16/5933/2022/en/, 6-7, 16, 24; UN Human Rights Council, A/HRC/39/64 (note 91 above) paras 71-74 and 84-87.
[108] See Nahimana et al case (note 90 above), Partly Dissenting Opinion of Judge Theodor Meron, paras 5-8.
[109] Antonio Cassese, *International Criminal Law* (2nd edition) (Oxford: OUP, 2008), 11-12.
[110] See Articles 5 and 12, Rome Statute (note 81 above).
[111] ICC, "Situation in Ukraine," March 2, 2022, https://www.icc-cpi.int/situations/ukraine.
[112] See Articles 8bis and 15bis, Rome Statute (note 81 above).
[113] ICC, "The States Parties to the Rome Statute,"https://asp.icc-cpi.int/states-parties.
[114] UN Treaty Collection, "Amendments on the crime of aggression to the Rome Statute of the International Criminal Court," June 11, 2010, https://treaties.un.org/Pages/ViewDetails.aspx?src=IND&mtdsg_no=XVIII-10-b&chapter=18&clang=_en.
[115] Article 13(b), Rome Statute (note 81 above).
[116] Article 17, Rome Statute (note 81 above).

take place in the territory as well as extraterritorially. Extraterritorial jurisdiction may be based on different grounds or principles depending on the offence in question. Accepted grounds for extraterritorial jurisdiction under customary international law are active nationality (based on the nationality of the accused), passive personality (based on the victim's nationality), protective principle (based on the security interests of the State) and universality (based on the offence's nature as one affecting the interests of the international community as a whole).[117]

It is widely accepted that genocide, crimes against humanity, and war crimes are subject to universal jurisdiction so that they may be prosecuted and punished by any State in the world, irrespective of any link with the crime or individual perpetrator. This includes the prosecution of online speech acts that amount to one of those three crimes. Whether or not aggression is subject to universal jurisdiction remains contested.[118] Nevertheless, it is uncontroversial that the victim State has the power to prosecute this crime domestically.[119] Growing acceptance among States for the creation of a special international tribunal for the crime of aggression committed by Russian nationals against Ukraine[120] may suggest that the victim State may be able to delegate its domestic jurisdiction over the crime of aggression to one or more States acting collectively.[121] Thus, prosecution of online speech acts in the context of the war in Ukraine amounting to the crime of aggression may well fall within the jurisdiction of this prospective tribunal.

---

[117] Cryer, Hakan, Robinson and Wilmshurst (note 88 above), 52-57.

[118] Beth Van Schaack, "Par in Parem Imperium Non Habet: Complementarity and the Crime of Aggression," *Journal of International Criminal Justice* 10, no. 1 (2012): 133, 134, 137–45; Roger S. Clark, Roger, "Complementarity and the Crime of Aggression," *The International Criminal Court and Complementarity: From Theory to Practice, Volume II*, edited by Carsten Stahn and Mohamed M. El Zeidy. Cambridge: CUP, 2011, 731; Dapo Akande, "Prosecuting Aggression: The Consent Problem and the Role of the Security Council," *Oxford Legal Studies Research Paper,* no 10/2011. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1762806.

[119] Akande (note 118 above), 32-35.

[120] Irina Paliashvili, "Calls mount for Russia to face tribunal for aggression against Ukraine," *Atlantic Council*, February 28, 2023. https://www.atlanticcouncil.org/blogs/ukrainealert/calls-mount-for-russia-to-face-tribunal-for-aggression-against-ukraine/; "Ukraine war: MEPs push for special tribunal to punish Russian crimes," *European Parliament News*, January 19, 2023, https://www.europarl.europa.eu/news/en/press-room/20230113IPR66653/ukraine-war-meps-push-for-special-tribunal-to-punish-russian-crimes.

[121] Oliver Corten and Vaios Koutroulis, "Tribunal for the crime of aggression against Ukraine - a legal assessment." *Eurpean Parliament,* December 2022, https://www.europarl.europa.eu/RegData/etudes/IDAN/2022/702574/EXPO_IDA(2022)702574_EN.pdf; Chatham House, "A criminal tribunal for aggression in Ukraine," March 4, 2022, video, 54:52, https://www.youtube.com/watch?v=XdHGf50fCCk&t=156s; EJIL: The Podcast!, "Episode 14 – "From Russia With War," March 7, 2022, podcast, 56:00, https://www.ejiltalk.org/ejilthe-podcast-episode-14-from-russia-with-war/.

INTERPLAY BETWEEN IHL, ICL, AND INTERNATIONAL HUMAN RIGHTS LAW (IHRL)

There is no question that IHL applies alongside IHRL in situations of armed conflict.[122] When the two sets of rules clash, the better view is that the norm conflict cannot be resolved in the abstract. IHL does not necessarily provide for the most specific, protective, or otherwise appropriate legal framework in armed conflict. For instance, even though Article 3 common to the Geneva Conventions and its customary counterpart provide that only a 'regularly constituted court' may pass judgement on an accused person, IHRL provides for additional guarantees of judicial independence and impartiality that can never be dispensed with, even during armed conflict.[123] Likewise, while IHL's provisions on direct participation of civilians in hostilities might be appropriate for battlefield scenarios, they may not provide adequate protection for civilians caught up in situations akin to law enforcement pursuits in urban settings.[124] Thus, it is not accurate to say that IHL automatically displaces IHRL in situations of armed conflict — international or non-international. Which exact rules or principles apply will depend on the specific circumstances at hand and can only be determined on a case-by-case basis.[125] Arguably, the more removed a situation is from the battlefield, the more likely it will be that IHRL is the more specific and appropriate legal framework.

It is also uncontroversial that ICL and IHRL apply together. On the one hand, ICL criminalises certain acts that constitute serious human rights violations. On the other, IHRL limits the discretion of States to criminalise and prosecute individuals in domestic and international settings. For instance, the principle of legality[126] and fair trial guarantees[127] protect individuals from arbitrary criminal prosecutions. The ICC Statute specifically stipulates that the interpretation and application of its provisions 'must be consistent with internationally recognized human rights'.[128]

---

[122] International Court of Justice, "Legality of the Threat or Use of Nuclear Weapons (Advisory Opinion of July 8, 1996)," ICJ Rep 1996, para 25; "Case Concerning Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory (Advisory Opinion of July 9, 2004)," ICJ Rep 136, paras 105-106; "Case concerning Armed Activities on the Territory of the Congo (The Democratic Republic of the Congo v Uganda"), Judgement of December 19, 2005 (Merits), ICJ Rep 168, para 216, https://www.icj-cij.org/case/116#:~:text=The%20Court%20also%20found%20that,the%20principle%20of%20non%2Dintervention.

[123] Rule 100, "ICRC Customary IHL Database" (note 17 above); UN Human Rights Committee, "General Comment No. 29 (Article 4 of the International Covenant on Civil and Political Rights): Derogations during a State of Emergency," CCPR/C/21/Rev.1/Add.11, August 31, 2001, para 2998; Inter-American Commission on Human Rights, "Report on Terrorism and Human Rights," OEA/Ser.L/V/II.116, Doc. 5 rev. 1 corr., para 3019, October 22, 2002.

[124] Supreme Court of Israel, "Public Committee against Torture in Israel v Government of Israel," Case No. HCJ 769/02, paras 18, 22, 40, December 13, 2006, http://elyon1.court.gov.il/files_eng/02/690/007/A34/02007690.a34.pdf.

[125] Cordula Droege, "Elective affinities? Human rights and humanitarian law," *International Review of the Red Cross* 90 (2008): 501-548.

[126] E.g., Prosecutor v Ali Muhammad Ali Abd-Alrahman, Judgment on the appeal of Mr Abd-Al-Rahman against the Pre-Trial Chamber II's 'Decision on the Defence "Exception d'incompétence"', No. ICC-02/05-01/20-302 (ICC, November 1, 2021), paras 83-92; Talita Dias, *Beyond Imperfect Justice: Legality and Fair Labelling in International Criminal Law* (Leiden, Boston: Brill, 2022), 35-84.

[127] E.g., Prosecutor v Lubanga, Judgment on the Appeal of Mr. Thomas Lubanga Dyilo against the Decision on the Defence Challenge to the Jurisdiction of the Court pursuant to article 19 (2) (a) of the Statute of 3 October 2006, No. ICC-01/04-01/06-772 (ICC, December 2006, 14), para 37.

[128] Article 21(3), Rome Statute (note 81 above).

This means that if more than one meaning can be assigned to a criminal provision, the interpretation that is more in line with IHRL must prevail.

The rights to freedom of expression and information, recognised under different human rights treaties[129] and customary international law,[130] are particularly relevant to online content governance. Individuals have the right 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 their choice, including online.[131] This includes content 'that may shock, offend, or disturb' and applies 'irrespective of the truth or falsehood' of the information.[132] Not only speakers are entitled to the rights to freedom of expression and information. Audiences as well as platforms (including their owners) also have the right to seek, receive, and impart online content.[133]

The rights to freedom of expression and information give rise to both negative and positive State obligations. The negative duty to respect the freedoms of expression and information require States to refrain from interfering with those rights in an arbitrary manner. This means that content must not be censored or otherwise limited, except in accordance with the law, for a legitimate purpose and in a necessary and proportionate manner.[134] Likewise, States themselves must not spread information that they know or should have known is false or misleading.[135] Conversely, the positive duty to protect the right to seek, receive, and impart information requires States to ensure a plural, independent, and robust media and information environment, favourable to public debate and critique.[136]

Those rights and their corresponding obligations continue to apply in times of crises, including during armed conflict and in the face of international crimes. They also apply in other situations

---

[129] Anne Lowe, "Customary International Law and International Human Rights Law: A Proposal for the Expansion of the Alien Tort Statute," *Indiana International and Comparative Law Review* 23, no 3 (2013) 523-552, 535, 537.

[130] See Article 19, UN General Assembly, "International Covenant on Civil and Political Rights", December 16, 1966, 999 UNTS 171 ("ICCPR"); Art 10, Council of Europe, "European Convention for the Protection of Human Rights and Fundamental Freedoms, as amended by Protocols Nos. 11 and 14", November 4, 1950, ETS 5 ("ECHR"); Article 13, Article 13, Organization of American States (OAS), American Convention on Human Rights, "Pact of San Jose", Costa Rica, November 22, 1969 ("ACHR"); Article 9, Organization of African Unity (OAU), African Charter on Human and Peoples' Rights ("Banjul Charter"), June 27, 1981, CAB/LEG/67/3 rev. 5, 21 I.L.M. 58 (1982).

[131] UN Human Rights Council, "Disinformation and freedom of opinion and expression - Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression," A/HRC/47/25, April 13, 2021, para 37.

[132] Ibid, para 38, citing UN Human Rights Committee, "CCPR General Comment 34, Article 19: Freedoms of opinion and expression," CCPR/C/GC/34, September 12, 2011 ("General Comment 34"), paras 11, 47, 49; Handyside v the United Kingdom, App no 5493/72 (European Court of Human Rights - ECtHR, December, 7 1976), para 49; Salov v Ukraine, App no 65518/01 (ECtHR, September 6, 2005), para 113 (noting that "Article 10 of the [ECHR] does not prohibit discussion or dissemination of information received even if it is strongly suspected that this information might not be truthful"). See also UN Special Rapporteur on Freedom of Opinion and Expression, the Organization for Security and Co-operation in Europe (OSCE) Representative on Freedom of the Media, the OAS Special Rapporteur on Freedom of Expression and the ACHPR Special Rapporteur on Freedom of Expression and Access to Information, "Joint Declaration on Freedom of Expression And 'Fake News', Disinformation and Propaganda," March 3, 2017 ("Joint Declaration") para 1(c), https://www.osce.org/files/f/documents/6/8/302796.pdf, preambular para 7.

[133] Manfred Nowak, *U.N. Covenant on Civil and Political Rights: CCPR Commentary* (Kehl, Germany: N.P. Engel, 2005), at 463; The Sunday Times v United Kingdom, App no 6538/74 (ECtHR, 26 April 1979), para 66; NIT S.R.L v Moldova, App no 28470/12 (ECtHR, 5 April 2022), paras 174, 192.

[134] General Comment 34 (note 132 above), paras 11 and 21.

[135] Joint Declaration (note 132 above), para 2(c); UN Human Rights Council, A/HRC/47/25 (note 131 above), para 88.

[136] Joint Declaration (note 132 above), preambular para 9; General Comment 34 (note 132 above), paras 14, 40; *Dink v Turkey*, App nos 2668/07, 6102/08, 30079/08, 7072/09 and 7124/09, (ECtHR, September 14, 2010), para 137.

and contexts marked by gross, widespread, or systematic human rights violations, such as civil unrest and authoritarian regimes.

Balancing between ICL, IHL, and IHRL in the online information space requires, first and foremost, respect for the freedoms of expression and information. These rights may only be limited if certain conditions are met. These are found, for example, in Article 19(3) of the ICCPR, which stipulates that:

> The exercise of the rights [to the freedom of expression and information] carries with it special duties and responsibilities. It may therefore be *subject to certain restrictions*, but these shall only be such as are *provided by law* and are *necessary*:
>
> (a) For respect of the rights or reputations of others;
>
> (b) For the protection of national security or of public order (ordre public), or of public health or morals.[137]

Similar provisions are found in Article 10(2) of the ECHR and Article 13(2) of the ACHR. In essence, they all recognise that the rights to freedom of expression and information may be limited where the public interest so demands, provided that the 'tripartite test' of legality, legitimacy, and necessity and proportionality is met.[138]

Legality stands for the requirement that any limitation on the freedoms of expression and information must be grounded in law. Any laws limiting speech, whether written or unwritten,[139] must be adopted by an independent legislative body.[140] They must also be accessible and foreseeable to the ordinary person.[141] Legitimacy refers to the different public interest grounds that may justify a restriction on freedom of expression or information.[142] Though the list of legitimate grounds is exhaustive in the ICCPR, ECHR, and ACHR, each ground is flexible enough to accommodate a variety of public interests in different States. But their interpretation must be consistent with other human rights, particularly non-discrimination.[143] Finally, necessity refers to the least restrictive means to achieve the legitimate aim in question and includes an assessment of the proportionality between the means chosen, the aim(s) sought, and the seriousness of the content restricted.[144]

Specific online content governance measures that the joint application of IHL, ICL, and IHRL calls for in different types of crises will be the subject of a subsequent report. But three overarching points can be made.

---

[137] Emphasis added.
[138] See Evelyn Mary Aswad, "To Protect Freedom of Expression, Why Not Steal Victory from the Jaws of Defeat?" *Washington and Lee Law Review* 77, no 2 (2020): 609-659, 609, 618 and 622.
[139] Karl Josef Partsch, "Freedom of conscience and expression, and political freedoms," *The International Bill of Rights: The Covenant on Civil and Political* Rights, (New York: Columbia University Press, 1981), 220.
[140] General Comment 34 (note 132 above), para 24.
[141] Ibid, para 25.
[142] Aswad (note 138 above), 625.
[143] General Comment 34 (note 132 above), paras 26, 32.
[144] Ibid, paras 33–36.

First, online speech acts amounting to violations of IHL and international crimes already satisfy the requirements of legality and legitimacy. This is true insofar as these unlawful acts are provided by law (IHL and ICL, respectively) and constitute legitimate means to ensure respect for the rights of victims. However, limitations to online content that breaches IHL or ICL must still be necessary and proportionate in the circumstances. This will usually mean that only the most serious speech acts should be considered as violations of IHL and ICL and punished as such. In accordance with the principle of legality, States must consider the accessibility and foreseeability of relevant criminal laws when assessing whether the individual speech act constitutes an offence. This means that an ordinary person must have understood, in general terms, that their speech was criminal and punishable, without reference to any particular legal provision.[145] Examples of speech acts that are clearly criminal under international law include direct and public incitement to commit genocide, and serious instances of hate speech making a substantial contribution to the commission of international crimes. More restrictive moderation measures like content takedowns and accounts suspension should be reserved for those serious speech acts, in line with the requirements of necessity and proportionality. For less serious online content that does not breach IHL or ICL but could infringe IHRL, less restrictive measures, such as de-amplification or labelling, may be more appropriate.

Secondly, and relatedly, Article 20 of the ICCPR already lists certain particularly serious speech acts whose prohibition by States is presumably legitimate. These are propaganda for aggressive war and advocacy of national, racial, or religious hatred that constitutes incitement to hostility, discrimination, or violence. For States that are bound by Article 20 of the ICCPR, Article 19(3)'s tripartite test must still be observed.[146] This means that when prohibiting war propaganda or incitement, States must adopt accessible and foreseeable laws. Likewise, any limitations on those speech acts must be necessary and proportionate to their seriousness and the aim of the limitation. Consequently, criminal sanctions should be reserved for the most serious forms of war propaganda and incitement, such as where there is an intent to cause harm and a serious and imminent risk thereof.[147] Likewise, online platforms should ensure that deleting such pieces of content or banning their authors is indeed necessary and proportionate in the circumstances, in light of other available content moderation measures.

Thirdly, all other types of online content falling below this threshold, i.e., content that cannot be limited via the tripartite test, must be protected. To be sure, there is no definitive category of 'protected speech', since all speech acts are in principle subject to limitations. However, some types of speech should receive heightened protection given the public interest in their

[145] Dias (note 126), 49.

[146] UN Human Rights Committee, "CCPR General Comment No. 11: Article 20 Prohibition of Propaganda for War and Inciting National, Racial or Religious Hatred," 29 July 1983, para 2; General Comment 34 (note 132 above), para 52; UN General Assembly, "Annual report of the United Nations High Commissioner for Human Rights," A/HRC/22/17/Add.4, January 11, 2013, paras 18 and 22; UN General Assembly, "Promotion and protection of the right to freedom of opinion and expression," A/67/357, September 7, 2012, para 41; UN General Assembly, "Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression," A/HRC/38/35, 6 April 2018, para 8; UN General Assembly, "Promotion and protection of the right to freedom of opinion and expression," A/74/486, October 9, 2019, para 13; Aswad (note 138 above), 629.

[147] UN General Assembly, A/HRC/22/17/Add.4 (note 146 above), para 29; UN General Assembly, A/67/357 (note 146 above) paras 46–47, 79; UN General Assembly, A/74/486 (note 146 above), paras 14–15.

dissemination, not the least to ensure individuals' right to receive information. This includes political speech[148] and statements whose dissemination is in the public interest, such as impartial journalistic reporting on public affairs.[149]

CONCLUSION

IHL, ICL, and IHRL provide different yet interlocking legal protections that apply concurrently to the governance of online content by States and platforms in three types of crises: armed conflict, international crimes, and human rights violations, respectively. Speech acts — online or offline — are not the quintessential way in which violations of IHL and international crimes are committed. Likewise, these acts are subject to the rights to freedom of expression and information, which apply online as they do offline. Thus, whether any particular piece of online content is to be considered a violation of IHL or ICL and limited on this basis can only be determined by working out the interplay between these legal frameworks and IHRL.

While this can only be assessed on a case-by-case basis, it is possible to conclude that only very serious types of online content amount to violations of IHL or international crimes, such that their limitation, in whichever form, is consistent with the safeguards imposed by the rights to freedom of expression and information. Examples of serious types of online content include incitement to genocide, violence, and other serious acts, as well as war propaganda. Limitations to these speech acts, even when lawful and legitimate, must still be calibrated to the seriousness of the act in question and the purpose of the limitation, taking into account all available content governance measures. It is for States and online platforms to make this assessment in concrete cases. This report has sought to guide them through this process, by outlining how IHL, ICL, and IHRL should be balanced in different types of crises.

[148] Case of Mouvement Raëlien Suisse v Switzerland, App no 16354/06 (ECtHR, July 13, 2012), para 61.
[149] Council of Europe, "Recommendation CM/Rec(2022)4 of the Committee of Ministers to member States on promoting a favourable environment for quality journalism in the digital age," March 17, 2022, https://search.coe.int/cm/pages/result_details.aspx?objectid=0900001680a5ddd0.