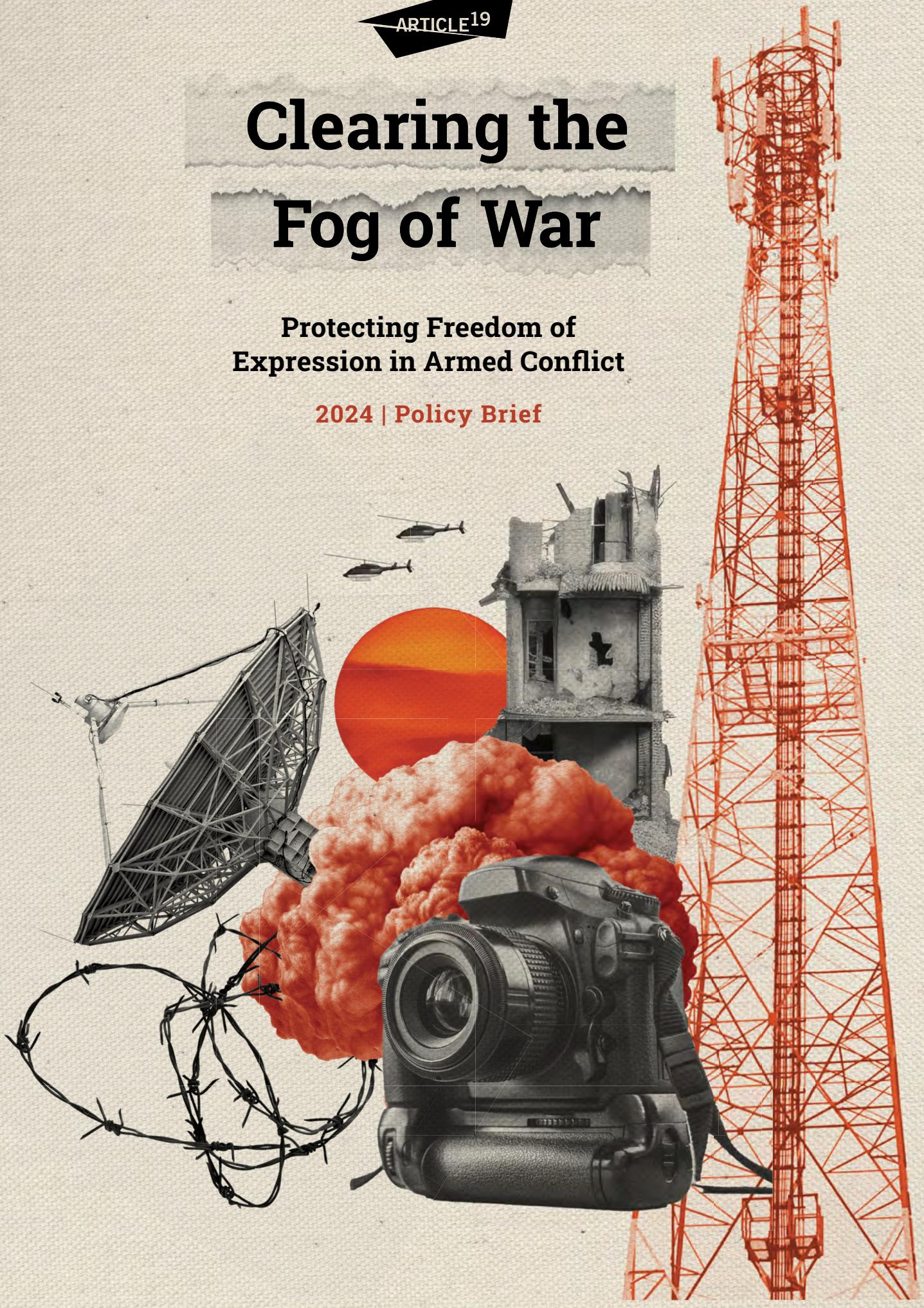
ARTICLE 19

Clearing the Fog of War

Protecting Freedom of Expression in Armed Conflict

2024 | Policy Brief



72-82 Rosebery Ave, London EC1R 4RW, UK

ARTICLE 19 is an international think–do organisation that propels the freedom of expression movement locally and globally to ensure all people realise the power of their voices.

Together with our partners, we develop cutting-edge research and legal and policy analysis to drive change worldwide, lead work on the frontlines of expression through our nine regional hubs across the globe, and propel change by sparking innovation in the global freedom of expression movement. We do this by working on five key themes: promoting media independence, increasing access to information, protecting journalists, expanding civic space, and placing human rights at the heart of developing digital spaces.

T : +44 20 7324 2500
F : +44 20 7490 0566
E : info@article19.org
W : www.article19.org
X : @article19org
Fb : facebook.com/article19org

Acknowledgements

ARTICLE 19 wishes to express its appreciation to the many experts whose thoughtful contributions during the drafting process have significantly enhanced this policy. Special thanks go to Joëlle Rizk, Samit d'Cunha, and Pierrick Devidal of the International Committee of the Red Cross; Álvaro Nistal and Patricio Grané Labat of Arnold & Porter; Aymen Zaghdoudi of Access Now; Talita Dias of Chatham House; Jiries Saadeh of Saadeh Rahman; as well as Tetiana Avdieieva and Maksym Dvorovyi of Digital Security Lab Ukraine. We also thank Arnold & Porter for their support with public international law research. While their contributions and assistance were invaluable, the views expressed in this policy are solely those of ARTICLE 19 and should not be attributed to any of the individuals, organisations, or law firms mentioned, or their clients. Additionally, we acknowledge the support of our civil society partners, academic institutions, and international organisations, whose collaboration has been vital in shaping this policy.

Contents

Abbreviations 01

Executive summary 02

Introduction 05

The standards on the protection of the right to freedom of expression during armed conflict 07

Freedom of expression standards applicable in armed conflict 08

International humanitarian law 09

The interplay between IHRL and IHL 10

Responsibilities of tech companies to respect freedom of expression and IHL during armed conflict 11

ARTICLE 19's position on interpretative gaps concerning freedom of expression in armed conflict 12

The lack of explicit recognition of freedom of expression in IHL 13

The underexplored extraterritorial application of freedom of expression obligations 15

The question of which cyber operations can be classified as an attack 17

Current responses to 'information manipulation' and 'hate speech' in armed conflict do not meet freedom of expression standards 19

Gaps in the protection of journalists and media facilities 25

The complexity of determining the legal regime applicable to internet shutdowns and their limitations under IHL 30

ARTICLE 19's position on the responsibilities of tech companies in armed conflict 34

Tech companies are becoming key conflict actors 35

Social media companies' failures during peace worsen during armed conflict 35

Flawed content moderation processes on many platforms 36

Inconsistent responses to armed conflicts 37

ARTICLE 19's proposal for a freedom of expression framework during armed conflict 43

Endnotes 47

Abbreviations

ACHR	American Convention on Human Rights
DDoS	Distributed denial-of-service
DoS	Denial-of-service
ECHR	European Convention on Human Rights
ECtHR	European Court of Human Rights
ICC	International Criminal Court
ICCPR	International Covenant on Civil and Political Rights
ICRC	International Committee of the Red Cross
ICT	Information and communications technology
IHL	International humanitarian law
IHRL	International human rights law
ISP	Internet service provider
OAS	Organization of American States
OHCHR	Office of the United Nations High Commissioner for Human Rights
OSCE	Organization for Security and Co-operation in Europe
UN	United Nations

Executive summary

Compliance with international humanitarian law (IHL) and human rights law (IHRL), as well as the associated alleviation of human suffering in times of war, relies on the free flow of information. Yet, attacks on freedom of expression and information (freedom of expression) by warring parties are on the rise. In this policy brief, ARTICLE 19 seeks to lay the groundwork for strengthening and better articulating the existing guarantees that uphold freedom of expression during armed conflict.

Conflict parties typically attempt to shape and control the information environment. Propaganda has long been used as a tactic to bolster civilian morale and influence narratives during war, at times hampering the possibility of reaching a peaceful settlement. Journalists and those who speak out against violence and atrocities have long faced silencing. However, in the digital age, governments and armed groups deploy information technologies and strategies that were inconceivable just decades ago, and civilians rely on connectivity more than ever before. Surveillance, content blocking, internet shutdowns, and the use of sophisticated 'information manipulation' and 'hate speech'[1] as tools of warfare, are the new normal in armed conflict.

In this context, it is crucial to re-examine the role of the right to freedom of expression in wartime. The rising trend of attacks on free expression must be reversed; armed conflict cannot serve as a pretext to justify censorship. Freedom of expression protects civilians. If it is impeded, compliance with IHL and IHRL suffers and atrocities proliferate.

Violations of free expression mostly occur during armed conflict not because international rules are inadequate or unknown, but because parties are unwilling to honour such rules and there are insufficient mechanisms to enforce them. In principle, free expression remains protected during armed conflicts. However, there are interpretative gaps in the existing legal frameworks. For example, IHL and IHRL do not expressly address the functional protection of media, digital threats against journalists and human rights defenders, internet shutdowns or limits to certain types of 'information manipulation', and 'hate speech' during armed conflict.

In this policy brief, ARTICLE 19 aims to fill some of these gaps. The objective is to reinforce existing rules, and promote an increased understanding of relevant legal standards among the various stakeholders, including military commanders, humanitarian and human rights organisations, private companies, and the public more generally. We believe that IHRL, particularly freedom of expression standards, can effectively complement existing protections under IHL in the face of the changing realities of modern armed conflicts and help prevent harm to those affected.

Specifically, ARTICLE 19 proposes that a framework for the protection of freedom of expression during armed conflict must be based on the following ten principles:

1 **Upholding freedom of expression during armed conflict protects civilians, as it enables the enjoyment of other human rights and fosters an environment conducive to respect for IHL.**

2 **Comprehensive protection of freedom of expression in armed conflict requires recognition of the important interplay between IHL and IHRL.**

3 **Any restriction on freedom of expression – whether it impacts individuals within or outside a state's borders – must strictly adhere to the principles of legality, legitimacy, necessity, and proportionality.**

4 **The protection of freedom of expression during armed conflict requires investment during pre- and post-conflict times.**

5 **Cyber operations must adhere to the IHL rules governing attacks if they are reasonably expected to cause – whether directly or indirectly – death, injury, physical damage, or loss of functionality.**

6 **Information operations[2] must adhere to the specific limits set by IHL and IHRL. Responses to 'information manipulation' and 'hate speech' must adopt a freedom of expression-based approach.**

7 **Internet connectivity can be a lifeline for civilians and is protected under both IHRL and IHL.**

8 **Tech companies should assume their responsibilities as key actors during armed conflict. They should take specific steps to respect IHL and uphold freedom of expression and other human rights, including by taking proactive steps to protect civilians from digital threats.**

9 **Actors operating in armed conflict, including state and non-state actors, along with humanitarian organisations and human rights actors, should abide by international freedom of expression standards. Where applicable, they should include freedom of expression considerations in their military manuals, codes of conduct, policies, and protocols. National and international courts, tribunals, and accountability mechanisms should consider freedom of expression violations as they assess potential international crimes.**

10 **Work towards greater articulation and promotion of freedom of expression standards during armed conflict must continue.**

Introduction

From Sudan to Myanmar, from Israel and Palestine to the Democratic Republic of Congo, from Ukraine to Yemen – armed conflicts affect all regions around the globe. Currently, there are more than 120 armed conflicts, which involve over 60 states and 120 non-state armed groups.[3] Not all of them capture the attention of the media and the international community, and some are quickly forgotten, providing a cloak of impunity for crimes and human rights abuses in these 'silent' wars.

Freedom of expression and information (freedom of expression) is often one of the first casualties in armed conflict. Journalists and media personnel struggle to report the news safely, internet shutdowns are on the rise, and conflict parties spread dehumanising narratives about the 'enemy' or distort information about the hostilities, undermining chances for a peaceful settlement. Private actors like social media companies can further contribute to the silencing of voices, including entire communities, through content moderation practices that fail to respect international humanitarian law (IHL) and human rights law (IHRL) rules.

As attacks on freedom of expression in armed conflict have become the norm, the need to protect this fundamental freedom has become ever more important. For individuals trapped amidst extreme violence, reliable information can be as life-saving as emergency aid. It might prevent them from accidentally entering areas of active fighting, enable them to contact their loved ones, or allow access to humanitarian relief. Journalists' ability to raise awareness about the conflict and report on crimes committed by the warring parties can permit the public and the international community to monitor events and advocate for respect of international rules. Access to internet and telecommunication networks has become as indispensable for civilians as roads and radios were at the time when IHL was in its infancy.

The impact of armed conflicts transcends the borders of the states within which they unfold. Non-belligerent states have to grapple with the complexities of responding to challenges to the information ecosystem. They are faced with decisions such as how to regulate or manage media outlets based in, or acting as mouthpieces for the belligerent state(s), as well as tech companies, including social media companies, telecommunication companies, and internet service providers (ISPs).

Governments, international bodies, private companies, humanitarian organisations, and civil society actors are increasingly focused on so-called 'information wars' and the new risks that digital technologies[4] can create for civilians. With few explicit rules in IHL, some advocate for new legal frameworks to tackle these evolving challenges.[5] At the same time, there is still a lack of clarity and awareness as to how freedom of expression standards apply in armed conflict under the existing legal framework and the contributions they can offer to addressing these issues. Furthermore, while there is a growing appreciation of the role of tech companies

in armed conflict, these discussions could benefit from a specific freedom of expression perspective.

With the present policy brief, ARTICLE 19 aims to contribute to this dialogue and fill some of these interpretative gaps. We recognise that questions around freedom of expression in armed conflict are multifaceted and of enormous complexity. We will not aim to cover all relevant aspects or questions but highlight and address several key gaps.

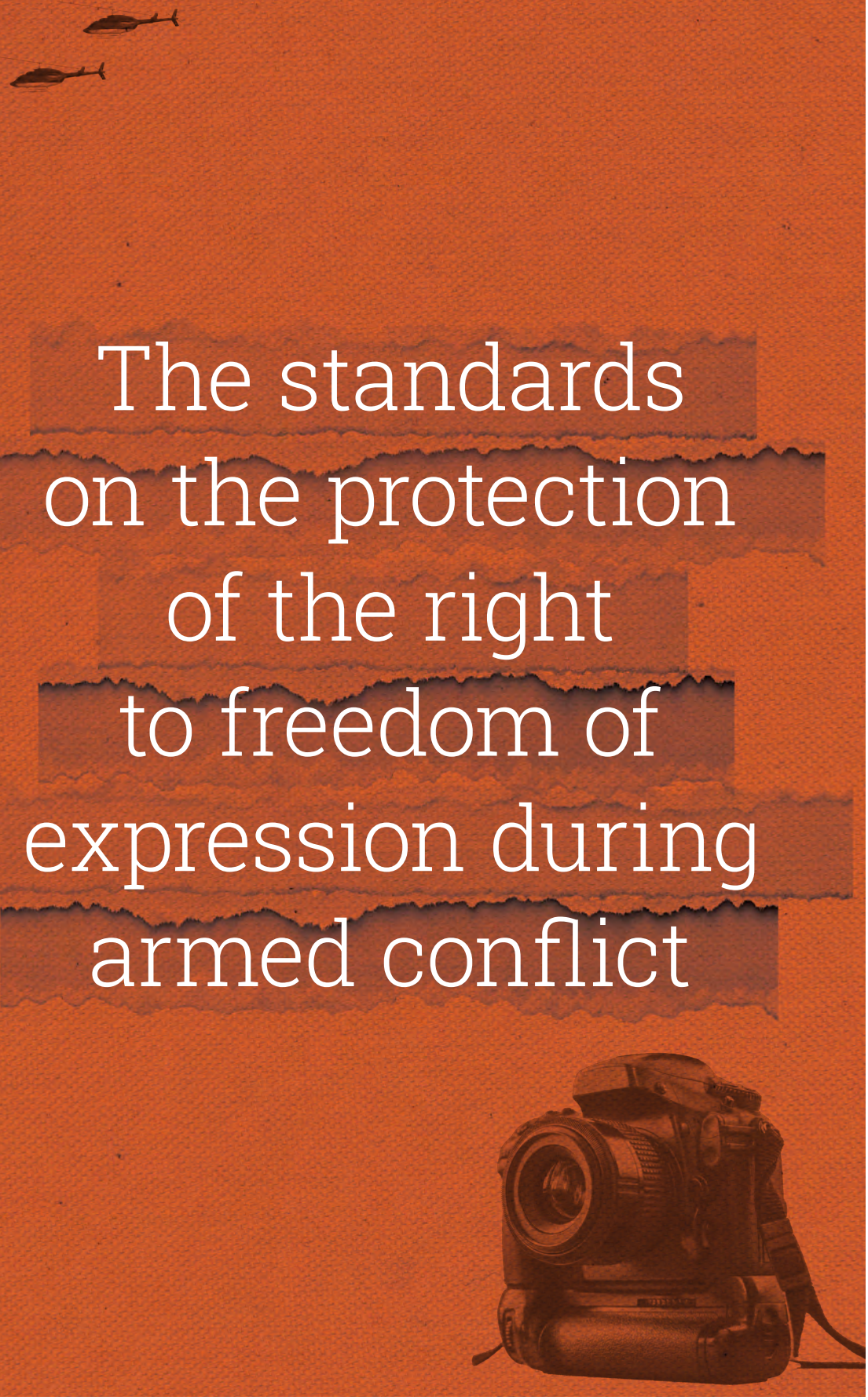
The policy brief is structured as follows:

- First, we outline standards on the protection of freedom of expression during armed conflict. This includes a summary of the applicable IHL rules, the interplay of IHL and IHRL in protecting free expression, and permissible derogations from freedom of expression obligations during armed conflict.

- Second, we address some of the key threats to freedom of expression in recent armed conflicts. We discuss the extent to which IHL provisions can protect freedom of expression and how freedom of expression obligations under human rights instruments can apply extraterritorially. Drawing from the existing IHL and IHRL obligations of state and non-state actors, we then suggest how to bridge interpretative gaps with respect to the treatment of non-kinetic military operations (operations that do not involve physical force) under IHL; the application of IHRL standards to 'information manipulation' and 'hate speech'; the protection of journalists and the media; and the legality of internet shutdowns during armed conflict.

- Third, we address the responsibilities of certain tech companies during armed conflict, namely social media companies, telecommunication companies, and ISPs. We highlight key measures that these companies should adopt to uphold IHL and IHRL, and specifically freedom of expression.

This policy brief is complemented by ARTICLE 19's upcoming policy brief on the interpretation of Article 20(1) of the International Covenant on Civil and Political Rights (ICCPR),[6] which prohibits 'propaganda for war'. ARTICLE 19 will continue to work on areas where freedom of expression intersects with armed conflicts, including the use of artificial intelligence and surveillance technologies and the relationship between freedom of expression and international criminal law.


The standards
on the protection
of the right
to freedom of
expression during
armed conflict

The right to freedom of expression during armed conflict finds protection in both IHRL and IHL. The extent to which such protections apply depends, among other things, on whether the armed conflict is international or non-international, whether the responsible party is a state actor or non-state actor, and whether the conduct by a conflict party targets the domestic population or the adversary's population outside its territory. In this section we explain those legal frameworks and standards particularly relevant for freedom of expression during armed conflict. This includes the right to derogate from freedom of expression, the interplay between IHRL and IHL, and the responsibilities of tech companies during armed conflict.

Freedom of expression standards applicable in armed conflict

The right to freedom of expression

The right to freedom of expression is protected by Article 19 of the Universal Declaration of Human Rights,[7] and is given legal force through Article 19 of the ICCPR and in regional human rights treaties.[8] These treaties require states to guarantee to all people the freedom to seek, receive, or impart information or ideas of any kind, regardless of frontiers, through any media of a person's choice. States may, exceptionally, limit the right to freedom of expression, provided that such limitations conform to the following strict requirements: the limitations must a) be provided by law, b) pursue one of the explicitly enumerated legitimate aims, and c) be necessary and proportionate to the aim sought (the 'three-part test' of Article 19 of the ICCPR).[9]

Article 20 of the ICCPR sets further limitations on freedom of expression and requires states to prohibit (though not necessarily criminalise) certain forms of speech, namely 'propaganda for war' and 'any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence'.[10]

As will be discussed later, additional restrictions on freedom of expression can also be found in other human rights treaties, as well as outside the human rights framework, for example in IHL or international criminal law.

Derogation from the right to freedom of expression during armed conflict

The protection offered by human rights conventions does not cease in case of armed conflict except through the effect of provisions allowing for derogation, which are contained in certain human rights treaties.[11]

Under Article 4 of the ICCPR, states may take measures derogating from certain of their obligations under that instrument – including the right to freedom of expression – 'in time of public emergency which threatens the life of the nation' to the extent 'strictly required by the exigencies of the situation, provided that such measures are not inconsistent with their other obligations under international law and do not involve discrimination solely on the ground of race, colour, sex, language, religion or social origin'. In addition, the state party must officially proclaim a state of emergency.[12]

The UN Human Rights Committee, which monitors the implementation of the ICCPR by its state parties, has recognised that armed conflict is one of the most probable scenarios that can create a 'public emergency threatening the life of the nation'.[13] However, not all armed conflicts require a derogation, and therefore the necessity of derogatory measures must be substantiated by a sufficiently detailed account of the relevant facts.[14] Derogation from a human right does not render the right entirely inapplicable.[15] According to the Human Rights Committee, states should apply derogation measures with due regard to the principles of legality, legitimacy, necessity, proportionality, and non-discrimination.[16] Although the necessity to adapt to evolving circumstances during armed conflict can make it challenging to draft precise laws, legislation should strive to be as clear as possible regarding the specific measures intended for application under the derogation regime.

Even where the specific circumstances of an armed conflict justify a derogation, derogation measures must be tailored to the nature and scope of the emergency and designed to bring to a close the threats underlying that emergency.[17] Any derogation must be temporary.[18]

Freedom of expression can be subject to derogation,[19] although the Human Rights Committee has taken the position that no circumstance can justify the derogation of the right to freedom of opinion.[20] A state party also cannot invoke a declaration of emergency as justification to engage in 'propaganda for war' and 'advocacy of national, racial, or religious hatred constituting incitement to discrimination, hostility, and violence'.[21]

International bodies have further found that, when adopting a derogation measure, a state must take account of the fundamental role played by freedom of expression in a democratic society and must demonstrate that the derogation is necessary to 'pave the way back to political freedom'.[22] Accordingly, measures that restrict public debate – even, and perhaps especially, in the case of armed conflict – are subject to scrutiny and specific limits.

International humanitarian law

Determining the scope of protection of freedom of expression during armed conflict requires an examination of IHL, a set of rules that applies to armed conflict.[23] IHL restricts the means and methods of warfare and protects persons who are not, or are no longer, participating in hostilities.[24] Among others, it seeks to strike a balance between legitimate military action and the objective of reducing human suffering.[25] The application of IHL does not depend on the reasons for the conflict and must be applied by all parties to conflict, including non-state parties.[26]

IHL draws a distinction between international armed conflicts and non-international armed conflicts. International armed conflicts arise when at least one state resorts to armed force against another state, regardless of the intensity of the hostilities or a formal declaration of war. Non-international armed conflicts arise within the territory of a state, either between governmental armed forces and non-state armed groups, or between non-state armed groups only, under the

condition that the armed group involved is sufficiently organised and the violence associated with the conflict is protracted in nature.[27]

Treaty rules concerning international armed conflicts are much more extensive than those that apply to non-international ones.[28] However, in recent years, IHL rules applicable to non-international armed conflicts have drawn closer to those applicable to international armed conflicts.[29] In particular, many of the treaty rules applicable to international armed conflicts also constitute customary international law – 'a general practice accepted as law'[30] – and are binding on all states independently of their acceptance of them. In addition, many, but not all, rules of customary international law apply to non-international armed conflicts.[31] Given that today the majority of armed conflicts are non-international, customary international law rules have acquired heightened practical relevance.[32]

IHL also applies to situations where a territory is occupied (which means it is placed under the authority of an adverse army[33]) during an international armed conflict.[34] The specific duties of the occupying power are outlined mainly in the 1907 Hague Regulations and the Fourth Geneva Convention.[35]

The interplay between IHRL and IHL

The dynamics between IHRL and IHL are key in shaping the protection of free expression in armed conflict. The prevailing view used to be that IHRL applied in times of peace and IHL applied in times of armed conflict. Modern international law, however, recognises that this distinction is inaccurate and that both regimes apply concurrently during armed conflicts. The continuous application of IHRL during an armed conflict is explicitly referenced in both IHL[36] and IHRL[37] treaties.[38] The International Court of Justice has affirmed that 'the protection offered by human rights conventions does not cease in case of armed conflict, save through the effect of provisions for derogation of the kind to be found in Article 4 of the ICCPR'.[39] The Human Rights Committee has also recognised that:

> *[t]he [ICCPR] applies also in situations of armed conflict to which the rules of [IHL] are applicable. While, in respect of certain [ICCPR] rights, more specific rules of [IHL] may be specifically relevant for the purposes of the interpretation of [ICCPR] rights, both spheres of law are complementary, not mutually exclusive.*[40]

In most cases, IHRL and IHL are mutually reinforcing and there is no conflict between the two branches of law. For example, journalists (like all individuals) enjoy the right to freedom of expression under IHRL. For its part, IHL offers journalists specific protections that complement those in IHRL instruments.[41]

In certain cases, IHL and IHRL may lead to different results or offer contradictory solutions. Pursuant to the *lex specialis* principle of interpretation, in situations where rules conflict and cannot be interpreted consistently, the more specific rule prevails. Both IHRL and IHL can constitute *lex specialis* in a given circumstance, even in the

context of armed conflict. As the Office of the United Nations High Commissioner for Human Rights (OHCHR) explained, the 'identification of which rule will have pre-eminence depends on an examination of the facts and of the particular protection included in the relevant rules'.[42]

Regional human rights bodies which, unlike the International Court of Justice, often do not have jurisdiction to directly apply IHL rules, regularly interpret the scope of the rights enshrined in human rights conventions in light of IHL standards when analysing potential human rights violations in the context of an armed conflict.[43]

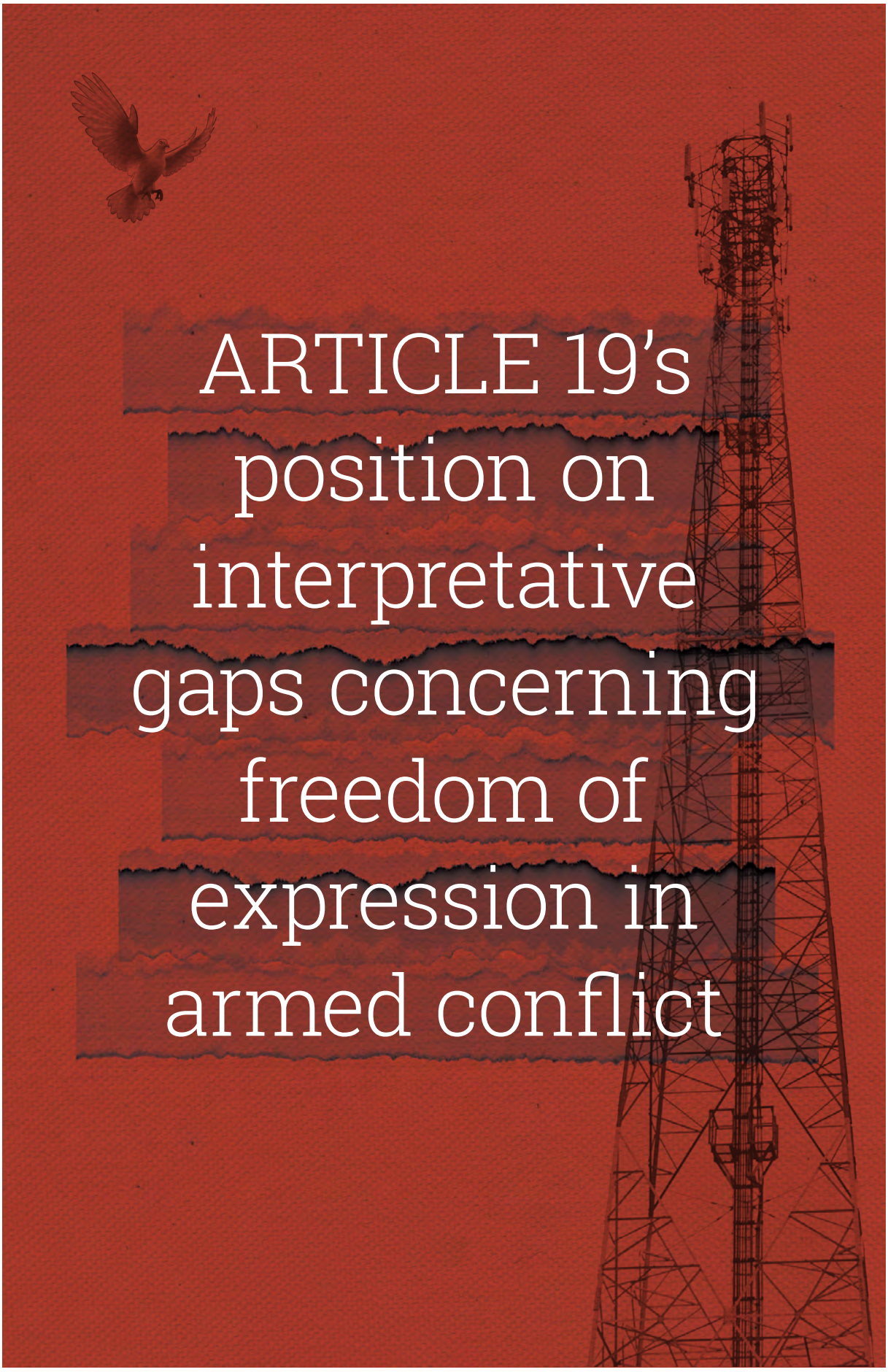
Non-state actors also have obligations under international law during armed conflict. If non-state armed groups are parties to an armed conflict, they are bound by IHL.[44] The IHRL obligations of such groups are, however, less certain. There appears to be growing recognition that non-state armed groups may, in some circumstances, have certain IHRL obligations – at least when they exert de facto governmental authority in areas under their control. However, how and to what extent those obligations exist remains an unsettled question.[45]

Comprehensive protection of free expression in conflict requires recognising the crucial interplay between IHL and IHRL.

Responsibilities of tech companies to respect freedom of expression and IHL during armed conflict

There is also growing recognition that businesses, as non-state actors, can significantly impact human rights. Although they do not have the same level of human rights obligations as states, it is increasingly acknowledged that businesses, including tech companies, have a responsibility to respect IHRL. This responsibility is articulated in the *Guiding Principles on Business and Human Rights: Implementing the United Nations 'Protect, Respect and Remedy' Framework* (the UN Guiding Principles).[46] In conflict-affected areas, the UN Guiding Principles call on businesses to not only respect major international human rights treaties, but also to 'respect the standards of [IHL]'.[47] The UN-mandated Working Group on Business and Human Rights has further emphasised that companies should adopt conflict-sensitive practices and conduct heightened due diligence to identify, prevent, mitigate, and account for how they address their adverse impacts during armed conflict.[48] These responsibilities operate independently of any state's obligations or its willingness or ability to meet its own IHRL and IHL obligations.

Furthermore, many telecommunication companies and ISPs are under government control. Consequently, their actions may be attributable to the state itself.[49]



ARTICLE 19's position on interpretative gaps concerning freedom of expression in armed conflict

Having outlined what we know about the key legal frameworks and standards that apply to freedom of expression during armed conflict, this section presents ARTICLE 19's position on how to address some of the main threats to freedom of expression in armed conflict and how to fill interpretative gaps in the protection offered by existing IHL and IHRL rules. Issues related to responsibilities of tech companies specifically and our recommendations to them are addressed in the following section.

The lack of explicit recognition of freedom of expression in IHL

ARTICLE 19 observes that IHL does not appropriately recognise the importance of freedom of expression in the context of armed conflict. Existing IHL conventions lack express provisions securing the right to freedom of expression.[50] This poses significant challenges, especially considering the dangers that arise in active conflict zones when freedom of expression is undermined.

Indeed, the digitalisation of armed conflict has increased the importance of the information space for military purposes. We are witnessing a rise in 'information manipulation' and 'hate speech',[51] internet shutdowns,[52] use of spyware,[53] and online censorship.[54] These trends can exacerbate both online and offline harm to civilians, impede broader compliance with IHL, and prolong the conflicts.

Some commentators question whether IHL is still fit to respond to these new challenges and argue that new IHL rules may be needed.[55] Indeed, IHL may have to develop further to adequately address the reality of information warfare, digital threats to civilians, and freedom of expression violations during armed conflict. However, the creation of new treaty rules could encounter substantial obstacles and risk undermining the protections currently offered by IHL and IHRL.[56] ARTICLE 19 believes that – at least initially – a proper understanding and interpretation of existing IHL rules can go a long way in protecting freedom of expression during armed conflict.

This policy brief aims to provide interpretations that can serve as a foundation for more detailed rules – developed through a collective approach that includes a variety of stakeholders – that operationalise the freedom of expression responsibilities of conflict parties and other actors.[57]

The proper interpretation of international humanitarian law can significantly help address challenges to the information ecosystem in the digital age and safeguard freedom of expression during armed conflict.

Why is there a need to further develop the understanding of IHL specifically when freedom of expression – like all human rights – continues to apply during armed conflicts? We believe that recognising that IHL can also provide protection to freedom of expression is significant for several reasons. IHL rules are tailored to the realities of conflicts, offering additional and specific protections to those found in IHRL.[58] Such rules are specifically designed

for situations of emergency and do not permit derogation.[59] Importantly, as mentioned earlier, IHL explicitly addresses the responsibilities of non-state actors, and its extraterritorial applicability is not contested.[60]

Moreover, we believe that increased recognition that violations of freedom of expression can simultaneously constitute breaches of IHL, and that certain violations of IHL – specifically, grave breaches[61] – can be prosecuted under international criminal law, can further strengthen the protection of freedom of expression during armed conflict.

ARTICLE 19's position: Compliance with international humanitarian law requires upholding freedom of expression

ARTICLE 19 acknowledges that IHL does not explicitly codify the right to freedom of expression during armed conflict. However, we argue that several IHL provisions – including the obligation to ensure respect for IHL and the fundamental principles of humanity, military necessity, distinction, and proportionality – can offer basic safeguards for freedom of expression in specific circumstances. Additionally, the protection of journalists under IHL implicitly recognises their right to freedom of expression.

Existing IHL rules were drafted without anticipating the digital revolution we are facing today, and the impact this would have on modern warfare. Like the rest of international law, IHL is a living instrument, capable of accommodating evolving interpretations.[62]

More specifically, ARTICLE 19 contends that while IHL does not explicitly codify freedom of expression, several IHL provisions can offer protection in specific circumstances and prohibit conduct that negatively impacts freedom of expression. In particular:

- IHL implicitly recognises and protects journalists' right to freedom of expression. Journalists engaged in dangerous professional missions in armed conflict enjoy the protections of civilian status.[63] The rationale for such protection is that, in order to appropriately perform their work, journalists need to be able to safely and independently exercise their right to freedom of expression.[64] This indicates that journalists have a right to freedom of expression – and the public has a right to receive information from them. It follows further that journalists should be protected and must not be targeted when reporting in dangerous missions in conflict areas.[65]

- When a belligerent state restricts freedom of expression with the intent of shielding its activities from public scrutiny and/or covering up IHL violations, that state infringes its obligation to ensure respect for IHL under Common Article 1 of the Geneva Conventions (Common Article 1) and customary IHL.[66] Common Article 1 requires the implementation

of all reasonable measures to prevent IHL violations, as well as to suppress breaches and hold accountable those responsible for violations that do take place.[67]

- IHL protects individuals' right to seek, impart, and receive information in specific circumstances. This follows from IHL's protection of civilian objects, including civilian information and communications technology (ICT) infrastructure.[68] In addition, as recognised in a resolution adopted at the 34th International Conference of the Red Cross and Red Crescent,[69] modern societies rely heavily on ICT infrastructure for communications and for the provision of essential services such as education and health care. In some circumstances, civilians also directly depend upon such infrastructure for their physical and emotional well-being. The fundamental IHL principles that can protect individuals' ability to communicate are:
  - The principles of humanity and military necessity, which apply to all military operations, whether they are kinetic or cyber in nature.[70]
  - The principle of distinction between civilians and combatants, and between civilian objects and military objectives, which prohibits indiscriminate attacks.
  - The principle of proportionality, which places a limit on the extent of incidental civilian harm that is permissible whenever military objectives are attacked, reflecting the balance that must be struck between the principles of humanity and necessity.
- Further, several specific IHL rules implicitly require the protection of civilian ICT infrastructure, the availability of the internet as a source of information, and individuals' ability to communicate. For example, humanitarian organisations and hospitals enjoy specific protection under IHL.[71] As their operations depend heavily on a functioning ICT infrastructure,[72] intentional or indiscriminate interference with such infrastructure ought to be interpreted as a violation of the IHL rules that protect these institutions. Another example is the obligation to take precautions against the effects of attacks.[73] In order to avoid the dangers resulting from military operations, civilians require the ability to access information and to communicate about such dangers.[74] The obligation to take precautions against the effects of attacks therefore implicitly includes an obligation to protect the ICT infrastructure that enables civilians to obtain information and to communicate about attacks.

The underexplored extraterritorial application of freedom of expression obligations

The question of whether states' human rights obligations apply extraterritorially is particularly pertinent in the context of international armed conflicts, and gains additional significance in the digital realm.

ARTICLE 19 notes that human rights instruments lack explicit mention of their extraterritorial application, but some contain reference to their applicability for individuals subject to or within a state's jurisdiction.[75] The specific meaning of these jurisdiction clauses is complex and marked by controversy. Jurisprudence on the matter has been inconsistent. Several possible models have been considered for the extraterritorial application of human rights treaties. For example, some human rights bodies have held that human rights obligations apply to individuals outside a state's territory when the state has 'effective control' over that territory, judged primarily by reference to the strength of the state's military presence in the area (spatial model).[76] Others look to the state's authority and control over individuals, for example where a state exercises total and exclusive control over the prisons and the individuals detained in them (personal model).[77]

The issue of extraterritoriality is particularly underexplored when it comes to freedom of expression obligations. The most prominent models for extraterritorial jurisdiction do not offer suitable solutions. This is so because many freedom of expression violations, including in armed conflict, are committed through cyber-enabled operations that are largely disconnected from traditional concepts of physical control over either territory or individuals.

ARTICLE 19's position: States have extraterritorial obligations to protect freedom of expression

ARTICLE 19 asserts that if a state's actions can impact the exercise or enjoyment of the freedom of expression rights of an individual located outside its borders, freedom of expression obligations should apply extraterritorially towards that individual.

ARTICLE 19 believes that if a state's conduct affects the exercise or enjoyment of a human right by an individual outside its borders, IHRL should apply.[78] If, for example, a state interferes with the ability of individuals abroad to freely express themselves – whether through directing information operations at the population in the 'enemy' state, shutting down the internet via cyber-attacks on ICT infrastructure, or intercepting the private communications of journalists and human rights defenders in 'enemy' territory – then that state is exercising power and control over those individuals' right to freedom of expression. Accordingly, the state should be bound by freedom of expression obligations towards those individuals.[79]

Positions that are supportive of an extraterritorial scope of the right to freedom of expression – as well as the right to privacy – are gaining traction.[80] However, the issue remains unsettled. This is one of the reasons why ARTICLE 19 asserts that it is significant to recognise the protection offered to freedom of expression by IHL, which applies to conduct on 'enemy' territory in international armed conflict.

IHL does not, however, provide private enforcement mechanisms for individuals affected by IHL violations.[81] ARTICLE 19 therefore submits that it is necessary for freedom of expression obligations under IHRL to be applied extraterritorially, including during armed conflict. Failure to recognise such extraterritoriality would leave individuals whose freedom of expression rights are affected by foreign states' operations during armed conflict without full protection under freedom of expression standards and with limited means to enforce these rights.

The question of which cyber operations can be classified as an attack

There are many military operations that are non-kinetic and negatively impact freedom of expression and the information environment. For example, internet shutdowns can be implemented without the need to resort to physical damage to the communications infrastructure, through manipulation of network routing or denial-of-service (DoS) attacks.[82] Journalists and human rights defenders are also increasingly facing non-kinetic digital threats during armed conflicts, including identity fraud,[83] distributed denial-of-service (DDoS) attacks,[84] organised doxing campaigns (in which someone's personal information, including their whereabouts, is posted online with malicious intent),[85] or targeting via spyware.[86] Information operations have long been conducted against both military adversaries as well as civilian populations during armed conflicts.[87]

Which IHL rules on the conduct of hostilities (a subset of IHL rules) apply to such operations depends on the question of whether they amount to an 'attack', which is defined in IHL as an 'ac[t] of violence against the adversary, whether in offence or in defence'.[88] Indeed, while some IHL rules impose limits on any military operation, the determination of whether a military operation qualifies as an 'attack' is essential for the application of additional protections deriving from the principles of distinction, proportionality, and precaution.[89] These protections include the prohibition of attacks against civilians and civilian objects,[90] the prohibition of indiscriminate[91] and disproportionate attacks,[92] and the obligation to take all feasible precautions to avoid, or at least reduce, incidental harm to civilians and damage to civilian objects when carrying out an attack.[93]

The determination of which operations qualify as an 'attack' – and how this notion should be interpreted – has been extensively debated, especially in the context of cyberoperations. It is widely accepted that an 'attack' includes an operation that may reasonably be expected to cause death or injury (understood to include serious illness and severe mental suffering[94]) or to result in physical damage, even if the means used are non-kinetic.[95] Physical damage to communications infrastructure, whether caused by bombing or cyber mechanisms, would thus constitute an attack.

What is disputed is whether an operation that does not cause physical damage but that results in loss of functionality of the cyber infrastructure would qualify as an attack.[96] There has also been debate over whether the assessment of what constitutes the 'reasonably expected' effects for the

purposes of defining attacks should include harm due to the foreseeable direct and indirect (or reverberating) impacts of an attack.[97]

ARTICLE 19's position: The term 'attack' under international humanitarian law should not be interpreted restrictively

ARTICLE 19 opposes a restrictive interpretation of the term 'attack'. Cyber operations must adhere to IHL rules governing attacks if they are reasonably expected to cause – whether directly or indirectly – death or injury (to include serious illness and severe mental suffering), physical damage, or loss of functionality.

In ARTICLE 19's view, the term 'attack' should not be interpreted restrictively, as also advocated by the International Committee of the Red Cross (ICRC).[98]

First, we submit that both direct and indirect consequences of military operations should be considered when establishing whether they amount to an attack.[99]

Second, we believe that physical damage should not be a prerequisite for classifying a military operation as an attack. For instance, operations that render internet and telecommunication networks dysfunctional – even if temporarily – should be covered by the term 'attack'.[100]

Operations such as internet shutdowns or spyware attacks can inflict severe direct or incidental harm on civilian populations, including death and injury:

- Internet shutdowns can disrupt life-saving communication channels and hinder the operations of medical services and humanitarian actors. Their disproportionate effects on the civilian population have been consistently recognised by international human rights bodies.[101]
- Cyber operations directed against journalists and human rights defenders – who are civilians under IHL – can result in physical violence against them. For example, the use of surveillance technologies 'has been linked to arrest, intimidation and even killings of journalists and human rights defenders'.[102] They can also directly inflict severe mental harm.[103]
- Doxing can similarly expose journalists and human rights defenders to abuse and threats of physical violence.[104]

ARTICLE 19 submits that operations such as those listed above should adhere to IHL rules governing attacks if they are reasonably expected to cause – whether directly or indirectly – injury or death (to include serious illness and severe mental suffering), physical damage, or loss of functionality. Excluding them from the scope of the term 'attack' could create a protection gap. For example, while the human rights framework, including the right

to freedom of expression and the right to privacy, offers protection to civilians in non-international armed conflicts, individuals targeted by the adversary in international armed conflicts may not be equally protected due to the extraterritorial nature of such rights violations. Adopting a narrow interpretation of an 'attack' also contradicts the spirit and purpose of IHL and runs counter to the International Court of Justice's observation that IHL applies to 'all forms of warfare and to all kinds of weapons, those of the past, those of the present and those of the future'.[105]

In any case, even operations that do not qualify as an attack under IHL remain subject to limitations under IHL. These include the fundamental principles of IHL, including military necessity, humanity, the prohibition of collective punishment or the prohibition to direct operations against specifically protected objects such as medical facilities.[106] In addition, when conducting any military operation, constant care must be taken to spare the civilian population and civilian objects.[107] As the ICRC noted, 'directing disruptive cyber operations against civilian objects, including civilian data, or ignoring their incidental effects on civilian populations, would be incompatible with this rule'.[108]

Another category of non-kinetic attacks could be information operations. The causal relationships between such operations and the resulting harm can be inherently difficult to demonstrate. However, certain information operations are recognised as having the potential to lead to killings, physical harm, severe mental suffering, and fuel violence in violation of IHL.[109] Although the question of whether information operations can be classified as attacks has not received much attention, it is difficult to categorically dismiss this possibility if such operations are expected to endanger or harm civilians.[110] Further debate may be needed to explore the implications of such qualification, including for freedom of expression. Like cyber operations, information operations are subject to the principle of precaution.[111]

Current responses to 'information manipulation' and 'hate speech' in armed conflict do not meet freedom of expression standards

In recent years, the scale and spread of false, inciting, or hateful speech has drawn increased attention from governments, international organisations, private companies, and civil society. The terminology used to describe challenges to the information ecosystem varies across different actors and contexts and may include terms like 'disinformation', 'misinformation', 'information manipulation', and 'propaganda'. The ICRC primarily uses the term 'harmful information', emphasising its focus on information that can result in physical, psychological, economic, or social harm during conflict.[112] Like 'hate speech', these terms lack internationally agreed definitions.

ARTICLE 19 generally opposes efforts to define these terms, as such definitions will inevitably be ambiguous and risk overbroad, subjective interpretations that conflict with freedom of expression standards.[113] Instead, we approach the issue through an international law lens. Depending on the context, we evaluate whether restrictions on these types of expression meet the principles of legality, legitimacy, necessity, and proportionality; the limitations on states in disseminating false, misleading, or inciting speech; and the obligations of states to provide reliable information to the public and to prevent, investigate, and, where required, prosecute speech crimes, such as incitement to atrocity crimes. For this reason, in this policy brief, we do not adopt a single term or seek to define it but use terms such as 'information manipulation', 'propaganda', or 'disinformation' interchangeably.[114]

We also believe that instead of focusing on how specific instances of expression may be categorised, stakeholders should seek to address the systemic issues – such as societal tensions and biases, or the business models of tech companies – that make the information ecosystem vulnerable to manipulation, falsehoods, and radicalisation.

The nature and impact of 'information manipulation' and 'hate speech' in armed conflict

Conflict parties, whether state or non-state actors, have throughout history engaged in information operations, and continue to do so today. Conflict parties often use 'propaganda' tools to undermine proper understanding of the armed conflict, to strategically disadvantage one of the conflict parties, to maintain support among the home population and states around the globe, or to incite violence against certain groups.

ARTICLE 19 is concerned about how 'information manipulation' in armed conflict can increase people's exposure to risks and vulnerabilities. For example, if displaced people in need of humanitarian assistance are given intentionally misleading information about life-saving services and resources, they can be misdirected away from help and towards harm. 'Information manipulation' can also impact humanitarian organisations' ability to operate in certain areas. Furthermore, 'disinformation' has been widely used to incite violence against groups, intersecting with 'hate speech'. 'Hate speech' in armed conflict can contribute to the escalation of violence and the dehumanisation of groups. 'Hate speech' can also impact the conduct of combatants. For example, dehumanising the adversary can make combatants less likely to adhere to IHL or IHRL rules, thus increasing the risk of war crimes, serious human rights violations, and more aggressive treatment of civilian and detainees.[115]

In the last several years, conflict parties have frequently exploited social media platforms to spread manipulated information and 'hate speech'. These tactics – together with social media companies' own shortcomings and problematic business models – have contributed to social media becoming a driver of conflict.[116] Notwithstanding the rise of social media, reporting by legacy media continues to assume a central role during armed

conflict. Sound public interest reporting is a key antidote, but legacy media – both state-owned[117] and private[118] – can also serve as a vehicle for 'propaganda' and incitement to violence.

While 'information manipulation' and 'hate speech' are deeply concerning, so too have been the responses of some states. All too often, they have resorted to censorship, including media bans, the targeting of journalists, and communication shutdowns, while allowing impunity for serious cases of incitement particularly when originating from state agents. Ironically, governments themselves frequently act as primary sources of misleading or hateful information, both in times of peace and armed conflict, often justifying the introduction of censorship measures by pointing to the very risks that they create or exacerbate themselves.

Neither IHL nor IHRL prohibit 'disinformation' or 'hate speech' *per se*. For example, IHL only imposes non-systematic limitations on 'propaganda'.[119] Conflict parties may, for example, engage in so-called ruses of war (acts intended to mislead an adversary or induce them to act recklessly).[120] IHL also does not prohibit direct 'propaganda' operations on the civilian population of the adverse conflict party.[121]

However, certain uses of 'disinformation' do violate IHL. For example, IHL prohibits killing, injuring, or capturing an adversary by making them believe that they are entitled to protected status (perfidy).[122] Similarly, uses of 'propaganda' and 'hate speech' are prohibited if their primary purpose is to spread terror among the civilian population.[123] It is also widely recognised that the obligation to ensure respect for IHL under Common Article 1 – reflecting customary IHL – prohibits parties from encouraging, inciting, or instigating IHL violations, including by parties outside their own forces.[124] In the conduct of military operations, including in information operations, warring parties must also take constant care to spare civilians.[125]

International criminal law imposes additional limitations that may apply to situations of armed conflict. For example, the Convention on the Prevention and Punishment of the Crime of Genocide (Genocide Convention) requires states to prevent and punish 'direct and public incitement to commit genocide'.[126] In addition, the Rome Statute of the International Criminal Court (ICC) makes it a crime to order, solicit, or induce the commission of crimes within the jurisdiction of the ICC, such as genocide, crimes against humanity, and war crimes.[127]

As for IHRL, Article 20 of the ICCPR only requires states to prohibit 'propaganda for war' and 'any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence'.[128] Any restrictions on 'disinformation' and 'hate speech' whether or not such expression is covered by the prohibition of Article 20 of the ICCPR must meet standards for restrictions on freedom of expression under Article 19(3) of the ICCPR.[129] It is also well established that the mere falsity or misleading nature of certain information cannot justify restrictions, unless it affects one of the legitimate aims under Article 19(3) or Article 20 of the ICCPR.[130] At the same time, there is limited

authoritative guidance on how these standards apply and interrelate with IHL in the context of armed conflict.

ARTICLE 19 observes that the significant potential of 'disinformation' and 'hate speech' to cause severe harm to civilians during conflicts, coupled with the urgent need to address such threats during existing crises, often leads states (including non-belligerent states), companies, and other stakeholders to adopt or embrace excessively restrictive and disproportionate measures that do not comply with freedom of expression standards. Yet, responses to issues of 'hate speech' and 'disinformation' during armed conflict ought to align with principles of freedom of expression. Indeed, a censorious approach, may, in itself, contribute to harm.

ARTICLE 19's position: A non-censorious approach to 'information manipulation' and 'hate speech' should be adopted

ARTICLE 19 recalls that 'information manipulation' and 'hate speech' are not, as such, prohibited under IHL or IHRL. However, both these frameworks place specific limits on such speech and on information operations during armed conflicts. Restrictions on false, misleading, or inciting speech during armed conflict must uphold the principles of legality, legitimacy, necessity, and proportionality. As in peacetime, tackling the underlying issues and building societal resilience is often more effective than focusing on restricting expression. Open debate, access to diverse viewpoints, and preventive measures addressing the root causes of false, radical, and inciting speech require investment both before and after armed conflicts.

At the outset, ARTICLE 19 emphasises the need for clarity on the legal standards governing the use of 'information manipulation' and 'hate speech' during armed conflict. This is essential for stakeholders to formulate responses that protect and promote free expression. Here, ARTICLE 19 outlines a number of key considerations.

- IHRL imposes restrictions on state actors' information operations in armed conflict which supplement both the prohibitions under Article 20 of the ICCPR and the limits imposed under IHL. For example, ARTICLE 19 argues that the obligation to protect, respect, and fulfil the right to life (Article 6 of the ICCPR), as well as the right to be free from torture or cruel, inhuman, or degrading treatment or punishment (Article 7 of the ICCPR), should be interpreted as requiring states to refrain from communications that escalate tensions or fuel hatred and mistrust, thus increasing the likelihood of violence by either state forces or third parties.[131] We further argue that the obligation to respect, protect, and fulfil the right to freedom of expression also requires

states to refrain from encouraging or disseminating 'disinformation' during armed conflict.[132]

- These same IHRL provisions (Articles 6, 7, and 19 of the ICCPR) can be interpreted in specific circumstances as imposing a positive duty on states to disseminate information that can protect civilians affected by conflict.[133] Conflict parties that proactively issue reliable and truthful public statements and provide accurate information on events in armed conflicts can also demonstrate compliance with the obligation to respect and ensure respect for IHL under Common Article 1.[134]

- The test for state restriction on expression by individuals during armed conflict, whether categorised as 'disinformation' or 'hate speech', remains that of legality, legitimacy, necessity, and proportionality. This includes instances where international law requires the prohibition of expression, such as 'direct and public incitement to commit genocide' under the Genocide Convention.[135] The scope of the protection under Article 19 of the ICCPR will also remain the same. For example, as mentioned, the right to receive 'information and ideas of all kinds' by definition covers both accurate and false information.[136] Any restriction on false information must be closely tied to one of the legitimate aims, such as national security. The mere falsity or misleading nature of certain information is not sufficient to restrict its dissemination.[137] This is also true in wartime.

- When a state seeks to restrict information that is sponsored by a foreign government, the latter will be unable to invoke a violation of their right to free expression. A practical example might be restrictions on foreign media outlets controlled by a foreign government. Nonetheless, the restricting measure may infringe upon the right of individuals within the jurisdiction of the restricting state to receive information without interference. For example, under Article 19 of the ICCPR, individuals have the right to receive foreign 'propaganda', unless the limitations imposed by their government satisfies the three-part test.[138]

- There is often an interplay between IHL and IHRL in the context of information operations. For example, if a conflict party directs its information operations at the civilian population of the opposing state (in the case of an international armed conflict), or intends, through its information operations, to undermine the adversary's military capabilities, those operations will primarily be governed by IHL, even though they may impact the right to information of the individuals affected by them. Conversely, if state parties to an international armed conflict direct information operations at their own population – for example, to generate support – these activities would likely be regulated primarily by IHRL, although the state would still be bound by the IHL rules limiting information operations detailed earlier.[139]

- We contend that the typology that ARTICLE 19 has proposed for categorising 'hate speech', which

focuses on the severity and impact of the expression[140] remains applicable during armed conflict. This framework should inform responses to 'hate speech' during armed conflict to avoid inappropriate restrictions on the right to freedom of expression. In addition, recognising the challenges involved in drawing a distinction between incitement prohibited by Article 20(2) of the ICCPR and expression that encourages the use of violence not prohibited under IHL,[141] we consider that the six-part test from the UN Rabat Plan of Action[142] is the most appropriate tool for distinguishing these different types of expression. This test, which considers the context, speaker, intent, content, extent of the expression, and the likelihood of harm occurring, should also be applied during armed conflict.

Beyond the importance of understanding the applicable legal frameworks, ARTICLE 19 emphasises that the issue of 'disinformation' or 'hate speech' during armed conflict is – much like in times of peace – often more effectively tackled by addressing its root causes and strengthening societal resilience, rather than by focusing on restricting expression. While it might be understandable for states to respond to armed conflict in a reactive and censorious manner, such an approach might ultimately prove counterproductive and sow discontent within the population.

The importance of enabling measures that address and prevent the root of problems of 'disinformation' and 'hate speech' cannot be emphasised enough. These measures should start in peace and pre-conflict times. They should encompass both the online and offline space, including improving people's access to reliable information and the promotion of media diversity. Efforts to enhance media and digital literacy as well as citizen journalism are also vital, as they contribute to societal resilience. Weak governance structures, a stifled civic space, and the absence of independent media actors are some of the drivers of 'information manipulation' and 'hate speech', which only exacerbate during armed conflicts.

We acknowledge that governments do not typically pursue positive measures to enhance societal resilience once hostilities have already broken out, especially if they are actively involved in a conflict. This is not least given their strong interest in shaping the narrative and silencing dissent. Actors engaging with conflict parties must be aware of this and ensure that they do not encourage a censorious approach that goes beyond permissible restrictions on freedom of expression. They should ensure that they do not only remind conflict parties of the importance of preventing harm caused by 'information manipulation' and 'hate speech' but also of the harm caused by the suppression of opposing and critical viewpoints. This is why the promotion of enabling measures for freedom of expression are especially important in pre- and post-conflict periods.

While conflict parties and governments of third states have a role to play in shaping the information environment during armed conflicts, local and international media also have the power to shape public opinion and influence powerholders to work towards a peace agreement. It is essential for media

actors to respect journalistic standards, avoid contributing to narratives that can fuel tensions, and provide space for dialogue, critical voices, and diverse ideas on the conflict. They should also ensure independence, advocate for access, resist overreliance on government officials as sources of information, and maintain a critical stance towards official statements on national security issues.[143] Both international and local media should resist temptations to abandon objective coverage.[144]

ARTICLE 19 also urges civil society and humanitarian organisations to adopt an approach consistent with the principles of freedom of expression when addressing challenges to the information environment that could harm civilians. Such actors should raise awareness and encourage conflict parties and other influential actors to refrain from engaging in 'information manipulation' and 'hate speech' and, instead, to protect freedom of expression.

Gaps in the protection of journalists and media facilities

The importance of press coverage of armed conflicts cannot be overstated. By gathering and disseminating reliable and timely information about the conduct of hostilities, journalists carry out a crucial mission. It is often thanks to journalists that serious human rights violations and war crimes are brought to light. This comes with a heavy price.

Media workers often face extreme danger in armed conflicts. They are killed, kidnapped, tortured, and subjected to various forms of systematic harassment. They are also encountering growing digital threats amid armed conflicts, such as DDoS attacks, organised doxing campaigns, and targeted spyware attacks. The intentional destruction of buildings housing media is a further common feature in armed conflict.

Journalists increasingly face digital threats during conflicts, such as DDoS and spyware attacks or doxing campaigns.

IHL provides clear protections for media professionals and media facilities. Targeting journalists – which are accorded civilian status – is clearly prohibited under IHL,[145] and attacks on them must be investigated.[146] While neither the Geneva Conventions nor the Additional Protocols define the term 'journalist', any 'citizen' journalist would be immune from attacks as a civilian as long as they do not directly participate in the hostilities. It is also clear that media facilities are civilian facilities under IHL and must not be attacked, unless they constitute appropriate military objectives (and even then, an attack on such a facility would be subject to the main protections under IHL, including the principles of proportionality and precaution).[147] When it comes to digital threats against journalists, as detailed earlier, ARTICLE 19 believes that any cyber operations that can be reasonably expected to cause – whether directly or indirectly – death, injury, including serious illness or severe mental suffering to journalists, must abide by the IHL rules governing attacks, which includes that they should not target journalists as civilians.

While most protections for journalists and media are well established, this section addresses two questions that have been the subject of some debate. First, we consider whether reporting activities can ever render media actors and facilities legitimate military targets. Second, we outline the scope of functional protection for journalists and the limits on permissible restrictions and impediments placed on media work during armed conflict.

The prohibition of targeting journalists and media facilities for reporting activities

The question of whether reporting can turn media actors and facilities into military targets has been the subject of debate. For example, while some attacks on media facilities are claimed to be accidental, many are confirmed as intentional targeting aimed at silencing the purported 'propaganda' machinery of the 'enemy', destroying its operational communications infrastructure, or silencing 'terror broadcasts'.[148] Similarly, journalists have been targeted with the justification that they worked for a media outlet affiliated with armed groups.[149] Given such alarming conduct, it is essential to reiterate the extremely limited circumstances under which media professionals, and media objects may be considered legitimate targets.

Journalists will only lose immunity from attack 'for such time as they take a direct part in hostilities'.[150] The ICRC has posited three cumulative criteria to be met for an act to amount to direct participation in hostilities: a) the act must be likely to adversely affect the military operations or military capacity of a party to an armed conflict or, alternatively, to inflict death, injury, or destruction on persons or objects protected against direct attack; b) there must be a relationship of direct causation between the act and the expected harm; and c) there must be a belligerent nexus between the act and the hostilities conducted between the parties to an armed conflict.[151]

The test to assess whether a media facility has become a military objective is different. Media objects can be considered military objectives if, by their nature, location, purpose or use, they make an effective 'contribution to military action' and if their neutralisation offers a definite military advantage.[152] This threshold is lower than the 'direct participation in hostilities' threshold governing the use of force against persons.[153]

ARTICLE 19's position: Journalists cannot be deemed legitimate targets based solely on their reporting activities

Journalists and media actors are civilians under IHL and must not be targeted during armed conflict. ARTICLE 19 submits that journalists and media actors cannot become targets based on accusations of supporting a conflict party by disseminating 'propaganda', 'hate speech', or extremist content.

Media objects do not qualify as military objectives solely for generating support for the war effort or boosting civilian morale.

Belligerents must exercise the utmost scrutiny when considering whether to target any media facility, providing clear evidence that the media object made an effective contribution to military action and that its neutralisation offered a definite military advantage. They must also show that less restrictive measures would have been insufficient to mitigate the threat posed by the media object, and that the attack was not disproportionate to the anticipated military advantage.

In ARTICLE 19's view, journalists cannot be deemed legitimate targets based solely on their reporting activities. This is because the three criteria previously mentioned – threshold of harm, direct causation, and belligerent nexus – are cumulative. An act that meets the threshold of harm but does not meet the 'direct causation' criterion does not amount to direct participation in hostilities. Journalistic reporting, by its nature, does not directly affect the military capacity of the adversary or cause harm by inflicting 'death, injury, or destruction on persons or objects'.[154] Regarding the dissemination of 'propaganda', the ICRC clarifies that while it may be considered war-sustaining, it constitutes only an *indirect* participation in hostilities, insufficient to justify attacks on media professionals.[155] The same rationale applies to other types of reporting, whether labelled as 'propaganda', extremist, or otherwise.

At the same time, we emphasise that when reporting activities constitute a crime, such as incitement to international crimes, the perpetrators must be prosecuted.

As regards media objects, it is well-accepted under IHL that objects do not constitute military objectives when they merely generate support for the war effort or boost civilian morale and nothing more.[156] Given that the threshold for using force against objects is lower than for using force against persons (lacking, for example, a direct causation criterion), it has been held that incitement to crimes may render media objects a legitimate military target.[157] This is subject to the aforementioned requirement that they make an effective 'contribution to military action' and thus their neutralisation offers a definite military advantage.

ARTICLE 19 underscores the importance of conflict parties exercising the utmost scrutiny when considering whether to target any media facility. Clear evidence must be provided to demonstrate, among others, that the media object made an effective 'contribution to military action', that less restrictive means than targeting the media object would have been insufficient to mitigate the threat it posed, and that the attack was not disproportionate to the anticipated military advantage. Moreover, any attack on media facilities should undergo independent investigation, and any violations of IHL should be prosecuted.

The functional protection of journalists and media work

The *functional* protection of journalists and media work, and the limits of permissible restrictions and impediments placed on media during armed conflict, is another

area where the law provides limited guidance. Conflict actors often practice censorship to monopolise information and control the narrative, including by hampering the ability of media actors and journalists to operate freely and independently. This censorship takes on various forms, including the following:

- Banning critical reporting on the armed conflict or the state institutions, such as the armed forces. This type of censorship typically takes place under broadly defined national security and 'disinformation' legislation, which in reality seeks to silence dissent, prosecute journalists during armed conflicts, or force foreign journalists to suspend their coverage.[158]

- Bans or restrictions on media outlets – both foreign and domestic – as well as other restrictions on media reporting are also common during conflicts, imposed both by parties to the conflict as well as third states.[159]

- Restricting media access to conflict zones to impede independent reporting.[160]

IHL treaties remain largely silent on these issues, despite their implicit recognition that journalists have a right to free expression and should be protected when reporting on conflicts.[161] For example, IHL does not regulate access of news providers to conflict zones. Article 58(a) of Protocol I requires conflict parties to move civilians away from military operations, suggesting a potential right to restrict civilians' access to conflict zones.

However, it does not allow access-denial to journalists for exclusively protective purposes.[162] As IHL does not explicitly address restrictions on news providers' access to conflict zones, this leaves the matter to be addressed by IHRL.[163] The permissibility of other restrictions on media actors is also unaddressed by IHL, creating gaps that IHRL standards should fill.

As for IHRL, these types of restrictions on media reporting are usually based on national security concerns. While those can serve as valid reasons for imposing limitations, they cannot serve as a disguised attempt to control the narrative, and must be necessary and proportionate in light of the fundamental importance of an objective coverage of armed conflicts by an independent press.

ARTICLE 19's position: Restrictions on media and journalists' work must provide evidence of necessity and proportionality

ARTICLE 19 warns that restrictions of media work during armed conflict must present evidence of invoked national security concerns, while bans on media outlets are highly unlikely to comply with freedom of expression standards. Access restrictions for journalists to conflict zones must be proportionate to the level of risks involved and should not lead to a complete denial of access for entire conflict zones.

ARTICLE 19 believes that conflict parties and third states should be guided by the following considerations:

- General bans on critical reporting on an armed conflict or the state institutions are highly unlikely to meet the requirements for restrictions on freedom of expression under IHRL (Article 19(3) of the ICCPR). In particular, in order to restrict freedom of expression on the basis of national security, states must demonstrate how the expression they want to restrict causes or concretely risks actual harm to its national security, and how the particular restriction is necessary and proportionate for the threat to be averted. Speculative national security risks to restrict freedom of expression will not meet that threshold.[164] Even in times of armed conflict, the public should have access to diverse perspectives, which can include perspectives from adversaries in situations of 'conflict and tension'.[165]

- Bans on domestic and foreign media need to demonstrate the specific threats that a certain media outlet may pose to national security or another legitimate aim. International standards stipulate that the banning of media outlets is a severe restriction of freedom of expression and is rarely justified.[166] Even restrictions that fall short of media bans need to be based on a clear law, demonstrate the legitimate aim, be transparent, and provide evidence as to the necessity and proportionality of the specific action taken.

  Generally speaking, it can be counterproductive to ban media outlets accused of disseminating 'disinformation' or inciting content as it prevents the formulation of a counter-response that challenges the harmful narrative.[167] ARTICLE 19 has thus long argued that a more effective antidote here is the promotion of a vibrant, pluralistic, professional, ethical, and viable media ecosystem, which is entirely independent of those in power.[168]

  In addition, any restrictions on broadcasters, including bans or suspension during conflicts, should respect due process and transparency. Such restrictions should be imposed by an independent media regulator, not by executive branches of the government in charge of defence, national security, or the armed forces.[169] The media regulators in question should apply the usual requirements in the relevant regulatory framework when deciding on broadcasting licence removals.[170]

- Access restrictions for journalists to conflict zones must be proportionate to the level of risks involved. Some highly dangerous sections of the front line may legitimately be categorised as off-limits for journalists. There may also be legitimate restrictions on journalistic access based on national security considerations. However, any such restrictions must be supported by evidence of their necessity and proportionality. They should be applied in a tailored manner, evaluated on a case-by-case basis, and explore less restrictive alternatives. Barring journalistic access to extensive areas or territories is highly unlikely to meet these criteria.[171]

Granting access under the condition that the journalists 'embed' with the armed forces of the conflict party also raises concerns as it hinders independent media reporting. Conditions mandating the submission of all materials and footage for review by the armed forces before publication are highly likely to constitute a disproportionate interference with a journalist's freedom of expression rights.[172]

- Many censorship measures in the digital sphere fail to meet the requirements of international freedom of expression standards. These include website blocking, surveillance, and legal and extra-legal demands of tech companies to censor online content.[173] State actors cannot use conflict situations as a *carte blanche* to justify increased censorship or surveillance. This section has focused mainly on censorship. However, we contend that untargeted or 'mass' surveillance is inherently disproportionate and a violation of human rights whether in times of peace or during armed conflict.[174]

The complexity of determining the legal regime applicable to internet shutdowns and their limitations under IHL

Various forms of internet shutdowns[175] have become an increasingly prominent feature in armed conflict and are used by governments both against their own population[176] as well as against the population of their adversaries.[177] Shutdowns in conflict zones can cause considerable harm as they might cut civilians off from life-saving information about troop movements and humanitarian corridors and impede medical facilities and humanitarian agencies from operating properly. They often occur when governments carry out armed operations, curtailing the documentation of human rights abuses and war crimes. More generally, shutdowns are often used to control the narrative.

There are different methods to restrict a population's ability to access the internet. Such methods may, for example, entail a governmental authority directing ISPs – which may be privately or government-owned – to suspend services, or the physical destruction of telecommunications infrastructure through bombing attacks.[178] Access to the internet may also be disrupted indirectly, such as when electricity is cut off.

The assessment of internet shutdowns under IHRL, including during armed conflicts, is relatively clear. However, two aspects regarding internet shutdowns during armed conflicts warrant attention. The first is determining when IHL or IHRL – or both – apply to a shutdown that is connected to a conflict.[179] Second, what limitations does IHL impose on internet shutdowns?

ARTICLE 19's position: Internet shutdowns are highly unlikely to comply with IHRL or IHL

ARTICLE 19 asserts that both IHL and IHRL impose strict limits on internet shutdowns. Because of their wide-ranging and devastating impact on civilians, shutdowns are highly unlikely to comply with either of these legal frameworks and must be avoided by conflict parties. When implemented to conceal violations of IHRL and IHL during armed conflict, they constitute a breach of the obligation to ensure respect for IHL.

ARTICLE 19 submits that the following considerations should guide conflict parties as well as telecommunication providers and ISPs faced with demands by conflict parties to shut down networks.

Shutdowns generally do not meet requirements under IHRL

Where conflict parties seek to employ measures blocking access to the internet within their own territory,[180] as well as in instances of occupation, IHRL will apply. Where such shutdowns are connected to the hostilities and there is sufficient nexus to the armed conflict, IHL rules will also apply.[181]

In addition, we believe that where shutdowns primarily aim to control information flows rather than achieve a military objective, and the main impact is on the civilian population, they should be assessed primarily under IHRL, although IHL might impose additional limits.[182]

IHRL requires governments to ensure that internet-based restrictions are provided for by law and are a necessary and proportionate response to a legitimate aim. It is highly unlikely that internet shutdowns would be permissible under Article 19 of the ICCPR.[183] More specifically, the UN Human Rights Council held that blanket shutdowns have severe consequences and can thus never be justified, and that other types of network disruptions are also likely to have indiscriminate adverse effects, rendering them disproportionate.[184] With respect to targeted shutdowns of communications services, the Human Rights Council found that they may be deemed proportionate and justifiable only in the most exceptional circumstances, as a last resort.[185] However, even in times of conflict, 'using communications "kill switches" (i.e., shutting down entire parts of communications systems) can never be justified under human rights law'.[186]

ARTICLE 19 notes that many internet shutdowns during armed conflicts are officially justified with the need of curbing 'hate speech', 'disinformation', or illegal content. Because of the arguments just laid out, responding to the circulation of these types of content with internet shutdowns, including during conflicts, will almost certainly fail to meet the three-part test. Aside from their indiscriminate and disproportionate impacts, they prove counterproductive. While they limit the accessibility and circulation of 'hate

speech' and 'disinformation' online, they also hinder fact-finding efforts, suppress access to reliable information and are likely to encourage the spread of rumours, thereby increasing the risks of division and conflict due to the uncertainty and fear that they create.[187] In addition, messages from official channels spreading 'hate speech' can carry increased weight with communities if authorities block alternative narratives, including through internet shutdowns.[188]

IHL principles also impose strict limits on internet shutdowns

IHL will come into play when the shutdown is linked to the conduct of hostilities and there is a sufficient nexus to the armed conflict. Subject to any extraterritorial human rights obligations of the party involved, in international armed conflicts IHL might exclusively apply to attacks against the communications infrastructure on the territory of an adversary state. Although the permissibility of internet shutdowns under IHL is less established than under IHRL, and IHL does not explicitly address internet shutdowns, ARTICLE 19 contends that several IHL provisions provide protections against internet shutdowns.

ARTICLE 19 points out that fundamental IHL principles, including those of military necessity and humanity, should be considered. In addition, as explained earlier, it is ARTICLE 19's position that shutdowns should be considered attacks for the purposes of IHL regardless of whether they cause physical damage. We contend that the principles of distinction, proportionality and precautions in attack impose strict limitations on internet shutdowns, not least due to their foreseeable adverse consequences for the civilian population.

We note that it is conceivable that a shutdown might be implemented for defensive purposes,[189] for example, in response to a major cyberattack, or that it can serve a legitimate military purpose, such as depriving the opposing party of communication means for orchestrating attacks. In this context, ARTICLE 19 notes that many military systems and networks rely on general infrastructure and software features that are dual-use, which means that they are used by both military and civilians.[190]

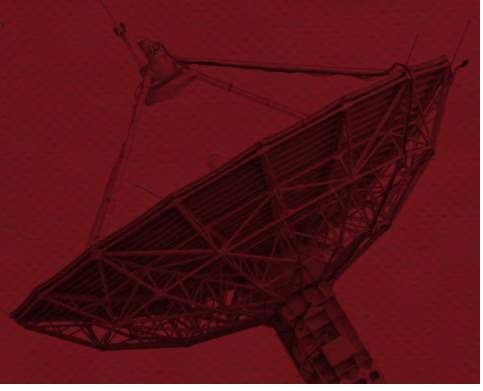
In such instances, conflict parties must, however, consider that the principle of proportionality will likely be violated when attacking communication infrastructure used by civilians. As described, shutdowns can cause significant direct and indirect harm to the civilian population. This can include mental harm, physical injury, or death. ARTICLE 19 contends that the impact on the civilian population would be so significant that it could likely be excessive in relation to any military advantage gained, and therefore disproportionate.

As mentioned earlier, shutdowns intentionally implemented to impede the documentation efforts of journalists and human rights defenders and conceal violations of human rights and humanitarian law during armed conflict may also amount to a breach of the obligation to ensure respect for IHL under Common Article 1 of the Geneva Conventions and customary international

law. Additionally, as also mentioned earlier, compliance with several specific IHL rules depends on the functioning of ICT infrastructure, such as the safeguarding of operations of humanitarian organisations and hospitals, which could be compromised by internet shutdowns.

Overall, ARTICLE 19 asserts that it is highly unlikely that measures introduced by conflict parties that completely block internet access comply with the requirements of either IHL or IHRL. Conflict parties should thus avoid shutting down access to the internet. At the very least, they should follow the recommendation by the ICRC Global Advisory Board on digital threats during armed conflicts, which states that '[i]f imperative military necessity justifies disruptions and restrictions, mitigation measures should be taken to ensure the availability of essential services and preserve the life and dignity of civilians as much as possible'.[191]

ARTICLE 19's position on the responsibilities of tech companies in armed conflict



As 'information manipulation', online censorship, internet shutdowns, and other restrictions on the free flow of information have increased during armed conflict, the role of tech companies has also expanded. There is a heightened risk that the conduct of tech companies might cause adverse human rights impacts, contribute to violations of IHL and negatively influence conflict dynamics. Even if businesses do not take a side in the conflict, '[they] are not neutral actors; their presence is not without impact'.[192]

This section focuses on the responsibility of telecommunication providers, ISPs, and the largest social media companies in the context of armed conflict. It will first outline issues, shortcomings, and complexities that are common to all these companies and then focus on those specific to the largest social media companies.

Tech companies are becoming key conflict actors

Conflict parties increasingly rely on tech companies to monitor and censor expression online, whether at the infrastructure or the content layer, and often pressure them into restricting access to services.[193] Authorities in third states may also exert pressure on companies to restrict content in alignment with their geopolitical interest, favouring one side of the conflict.

These dynamics can make it challenging for tech companies to uphold their responsibilities under international law. At the same time, despite becoming key actors in armed conflict – whether deliberately or unwillingly – many tech companies fall short of expectations. ARTICLE 19 is also conscious that some tech companies operating in countries in armed conflict, such as telecommunication providers and ISPs, are controlled by the government or have close ties to it, making them unlikely to resist government demands.

Tech companies – especially social media companies – are key actors in conflict. Yet, they often fail to live up to their responsibilities and protect civilians.

Yet even those tech companies that are seemingly more independent often lack an understanding of their contribution to conflict dynamics, are generally ill-prepared for conflict, and have not invested adequately in the resources necessary to uphold IHRL and IHL during armed conflict. This is reflected in the lack of transparency regarding crisis protocols or steps taken in response to specific armed conflicts, as well as the absence of policies on heightened human rights due diligence during armed conflict.[194] Furthermore, companies often appear to lack an understanding of what it means to respect IHL.[195]

Social media companies' failures during peace worsen during armed conflict

The challenges facing the largest social media companies during armed conflict are unique and their failures repeated and notorious.

In recent years, conflict parties have often instrumentalised social media to spread 'propaganda' or incite violence during armed conflict.[196] On the other hand, social media can play a crucial and positive role in documenting armed conflicts, shaping public understanding around them, and lending a voice to those directly impacted. Particularly in contexts marked by severe repression and censorship, social media can become a primary channel for reporting events on the ground and documenting potential violations of IHL and IHRL.[197]

In many ways, the problems associated with the largest social media companies are the same in times of peace as they are in times of armed conflict, even though they are especially acute during armed conflict and more likely to lead to offline harm. These problems are primarily and fundamentally linked to a platforms' design and business model, including their data-collection practices, algorithms, and monetisation systems. For example, it has been reported that several changes instituted by X (formerly Twitter) increased the spread of 'misinformation' on X, including during the war in Gaza, Palestine. The changes included the dismantling of its verification mechanisms and the repurposing of blue checkmarks to offer algorithmic prominence for users purchasing verification. This led to a reduction in the visibility and discoverability of journalists. Most 'misinformation' observed was reported to have come from verified accounts.[198]

More generally, the business models of some of the largest social media companies – which are often based on selling access to users' attention through targeted advertising and rely on recommendation algorithms that amplify extremist, false, and violent content – have long faced criticism.[199] Concerns have also been raised about the role of monetisation systems based on content views, which incentivise the publication of attention drawing content.[200] The ability of conflict parties to use paid, targeted advertising to amplify their messages in a way that favours their narrative can further distort information about the conduct of hostilities or serve as an attempt to justify violations of IHL.[201] At times, such adverts have also been reported to breach advertising policies prohibiting violent content, in an apparent failure by social media companies to properly review advertisements against their own policies before approving them.[202]

Flawed content moderation processes on many platforms

Additional concerns arise from the flawed content moderation processes on many platforms.

These may, for example, result in the insufficient moderation of inciting content or 'disinformation' as well as the excessive moderation of protected speech during armed conflict. Concerning the latter, some social media companies have been criticised for undue removal of political expression, for not allowing graphic or violent material on their platforms, even when it is in the public interest, or for inconsistently applying any exceptions based on newsworthiness.[203]

Key problems include inadequate resources allocated to moderating content (especially in languages spoken by conflict-affected individuals), a lack of investment in comprehending the specific contexts of conflicts, or insufficient transparency regarding how social media companies respond to government demands.[204] A persistent issue has also been ineffective dialogues with local civil society actors and users in conflict settings.[205]

These shortcomings significantly complicate the tasks of moderating content at scale in conflict settings and determining which forms of expression may violate IHL or are protected under it.

Issues can also arise from content moderation policies themselves, particularly those that include lists of banned organisations and entities, such as armed groups, and prohibit any 'glorification' or 'support' of them. This approach risks reinforcing a narrative dominated by one of the parties to the conflict, while potentially limiting access to entire communities under the other party's control, placing restrictions on their communications channels and restricting journalistic coverage.[206]

Inconsistent responses to armed conflicts

Some social media companies have responded to different armed conflicts in inconsistent ways,[207] specifically with respect to their actions related to content moderation, monetisation, advertising, and other crisis response measures, and in their communication about these measures.[208] Instances have also been documented where social media companies have disproportionately suppressed content that supported one of the parties in the conflict.[209]

Whilst certain responses were specific to particular armed conflicts, some social media companies have introduced changes that, in theory, are meant to be generally applicable. For example, some social media companies have begun incorporating references to IHL and armed conflicts into their community standards.[210] While this appears to indicate a general awareness of social media companies' responsibility to respect IHL, the reference to IHL and armed conflicts is non-systematic, and the provisions in the relevant policies are, at times, questionable from an international law perspective.[211]

ARTICLE 19 also observes that, following the outbreak of an armed conflict, there can be significant public and regulatory pressure on social media companies to find rapid and effective responses to the circulation of problematic content on their platforms. For example, regulators in jurisdictions outside conflict zones might seek to influence what content on the conflict is available on platforms.[212]

Responding to the specific human rights and IHL risks arising during armed conflict and navigating regulatory pressures from conflict parties and third states is a challenging and rapidly evolving task. Despite the growing role of tech companies as actors during conflict, their responsibilities have not received the same attention as entities in the extractive industry or private military and security companies. While progress has been made

in recent years, there is comparatively little guidance on how tech companies should conduct their operations in conflict settings while upholding IHL and IHRL, particularly in light of international freedom of expression standards. ARTICLE 19's observations intend to contribute to these discussions.

ARTICLE 19's position: Tech companies must adopt a series of measures in accordance with their responsibility to respect IHL and IHRL

ARTICLE 19 urges tech companies to recognise their pivotal role in modern conflicts. Beyond merely mitigating adverse human rights impacts and ensuring respect for IHL, they should take active steps to protect civilians. At the very minimum, they should conduct heightened due diligence and incorporate conflict-specific considerations in their policies and practices. This includes protocols for staff safety, crisis communication, handling government demands, and stakeholder engagement. Tech companies must also significantly increase their transparency with respect to the steps they take in response to armed conflict.

Social media companies should revise business models that amplify problematic content and limit user exposure to diverse viewpoints, which can fuel violence and dehumanisation and affect the conduct of belligerents. They should take measures to minimise the spread of content violating IHL and reduce the risk of undue removal of public interest content and documentation of potential human rights violations and international crimes.

To address the identified issues, ARTICLE 19 recommends that tech companies adopt a series of measures to respect IHRL and IHL. These measures primarily centre on safeguarding freedom of expression during armed conflict and are not intended as an exhaustive checklist. It is incumbent upon tech companies, given their comprehensive understanding of the risks associated with their products and services, to address their impact during armed conflict and to implement effective mitigation strategies based on pertinent internal expertise and engagement with external stakeholders.

ARTICLE 19 first issues a number of recommendations directed to all tech companies covered in this policy brief, namely telecommunication providers, ISPs, and social media companies:

- Tech companies operating in armed conflicts should conduct heightened due diligence which, above all, requires them to have robust processes and protocols in place and to allocate sufficient resources to identify such conflicts, analyse risks, and implement meaningful mitigation measures (including those suggested by ARTICLE 19 in the

following paragraphs).[213] Tech companies should, in particular, be alert to changes in the information environment – increased censorship and 'propaganda', bans on media, website blocking, and increases in inflammatory speech have been described as 'red flags' pointing towards armed conflict.[214] Tech companies must also conduct an analysis of the root causes and nature of the conflict, gain an understanding of the main actors involved and the company's own relationship with said actors, and perform a specific and continuous assessment of how the company's services and operations can lead to an increase in social tensions or exacerbate conflict dynamics and undermine or breach IHRL and IHL.

- Tech companies should have a specific policy on heightened human rights due diligence during armed conflict. They should also include conflict-specific considerations throughout their policies, practices, and processes for handling human rights risks and crises. This may encompass establishing protocols for staff safety, crisis communication, and policies on handling government demands during armed conflicts.

  This requires internal expertise on conflict dynamics in general and the specific conflict zones in question. Tech companies should also acquire internal expertise in IHL,[215] complemented by specific training and, when needed, by external specialised legal counsel. Companies should also consider relevant recommendations from experts such as the ICRC, relevant UN bodies, and civil society actors.

- Tech companies should go beyond merely 'mitigating adverse human rights impacts' and take seriously their social and moral responsibility to actively promote and uphold IHRL and IHL during armed conflicts. Measures to help tech companies meet this responsibility might include offering safety features for users to protect them from surveillance, with additional safety features for groups particularly at risk during armed conflicts (such as human rights defenders and journalists), and using any leverage they may have over conflict parties or other relevant actors to promote compliance with relevant IHRL and IHL rules. Social media companies specifically should take proactive steps to promote content by humanitarian actors such as the International Red Cross/Red Crescent Movement, Médecins Sans Frontières (Doctors Without Borders), or UN agencies, or local humanitarian actors most trusted by the communities in need.

- Tech companies should adopt specific policies and practices to mitigate the harms of government demands to restrict the free flow of information or install surveillance capabilities. Companies must keep in mind that their human rights responsibilities exist irrespective of the state's willingness to comply with their own obligations under IHL and IHRL. For example, if a government derogates from certain freedom of expression obligations, this does not limit a company's own freedom of expression responsibilities.

However, when it comes to government demands, the first step should involve assessing whether these demands are in line with the *state's* obligations under IHL and IHRL. While considering the impact on staff security, companies should explore all legal avenues to challenge the implementation of requests that violate freedom of expression. Even in cases where government requests broadly comply with IHL and IHRL standards, companies should comply in a way that least affects freedom of expression. They should be transparent to users and the public about any government requests and their response.[216]

ARTICLE 19 believes that in exceptional cases, where there is a complete lack of any legal system or independent judiciary that could provide a means to challenge potential government demands, companies should refuse to comply with orders that violate human rights, if they are able to do so without risking operational and staff security (for example, if they are not located in the territory of the state in question).[217]

Tech companies should coordinate with other companies facing similar requests by the same government, as it can enhance leverage in interactions with state authorities.

- Tech companies should take proactive measures to maintain connectivity at all times in areas of armed conflict and ensure that their networks remain operational. They should also publicly disclose details about shutdown orders and explore all legal avenues to challenge them. Social media companies should invest in tools and promote circumvention technologies that are easily accessible for all users to maintain access to the platform in the event of shutdowns, and provide versions of the platforms' service that function even with significantly reduced internet speed.[218]

- Tech companies should conduct active engagement with external stakeholders to ensure their input in the design, implementation, and monitoring of the due diligence measures required under the UN Guiding Principles. At a minimum, those stakeholders should include humanitarian actors and, where feasible and subject to security considerations, local civil society actors operating in conflict settings.

- Finally, tech companies should be transparent about the measures they are adopting to respond to a conflict situation. They should also issue conflict-specific reports detailing their policies, processes, and structures for operating in conflict settings. The reports should further address the challenges encountered in previous conflict situations, how tech companies responded to them, and how past responses will impact future conflict responses. Tech companies should also issue country-specific reports for states experiencing conflicts.

ARTICLE 19 also recommends that social media companies take additional measures to account for their functions as hosting platforms and mitigate the risks arising from their content moderation

practices, recommender systems, and advertising and monetisation systems. In particular:

- Social media companies should address and revise the incentive structures underlying their recommender, advertisement, and monetisation systems, as well as their microtargeting techniques, which can influence the amplification of problematic content, suppress public interest content, and reduce the exposure of users to diverse views.

- In terms of content moderation, they should adopt comprehensive and transparent community standards that cover content moderation issues specific to armed conflict, incorporate IHL considerations, and account for the relevance of freedom of expression during armed conflicts.

  ARTICLE 19 believes that, for social media companies, respecting IHL means taking active steps to reduce the risk of disseminating content that violates IHL or incites breaches of IHL or international crimes (generally understood as genocide, crimes against humanity, war crimes, and the crime of aggression) in the context of armed conflicts. Where this requires limitations on freedom of expression, such limitations should abide by the three-part test of Article 19(3) of the ICCPR. In view of the high factual and legal complexity often arising in the assessment of individual pieces of content, this requires careful case-by-case assessments, sufficient allocation of resources, and sufficient internal and external expertise, both on the legal and factual questions involved.

- In the face of the proliferation of problematic and illegal content circulating online during conflicts, social media companies must step up their content moderation efforts, yet at the same time resist the temptation to over-enforce their content moderation policies and ensure that any restrictions abide by freedom of expression standards. This requires, among others, limiting reliance on automated tools and ensuring regular verification of their accuracy and impartiality. Social media companies should also adopt and apply policies that allow content which breaches community standards to stay online if there is an overriding public interest to publishing that content. In addition, they should know the key media actors and human rights defenders that document events on the ground.

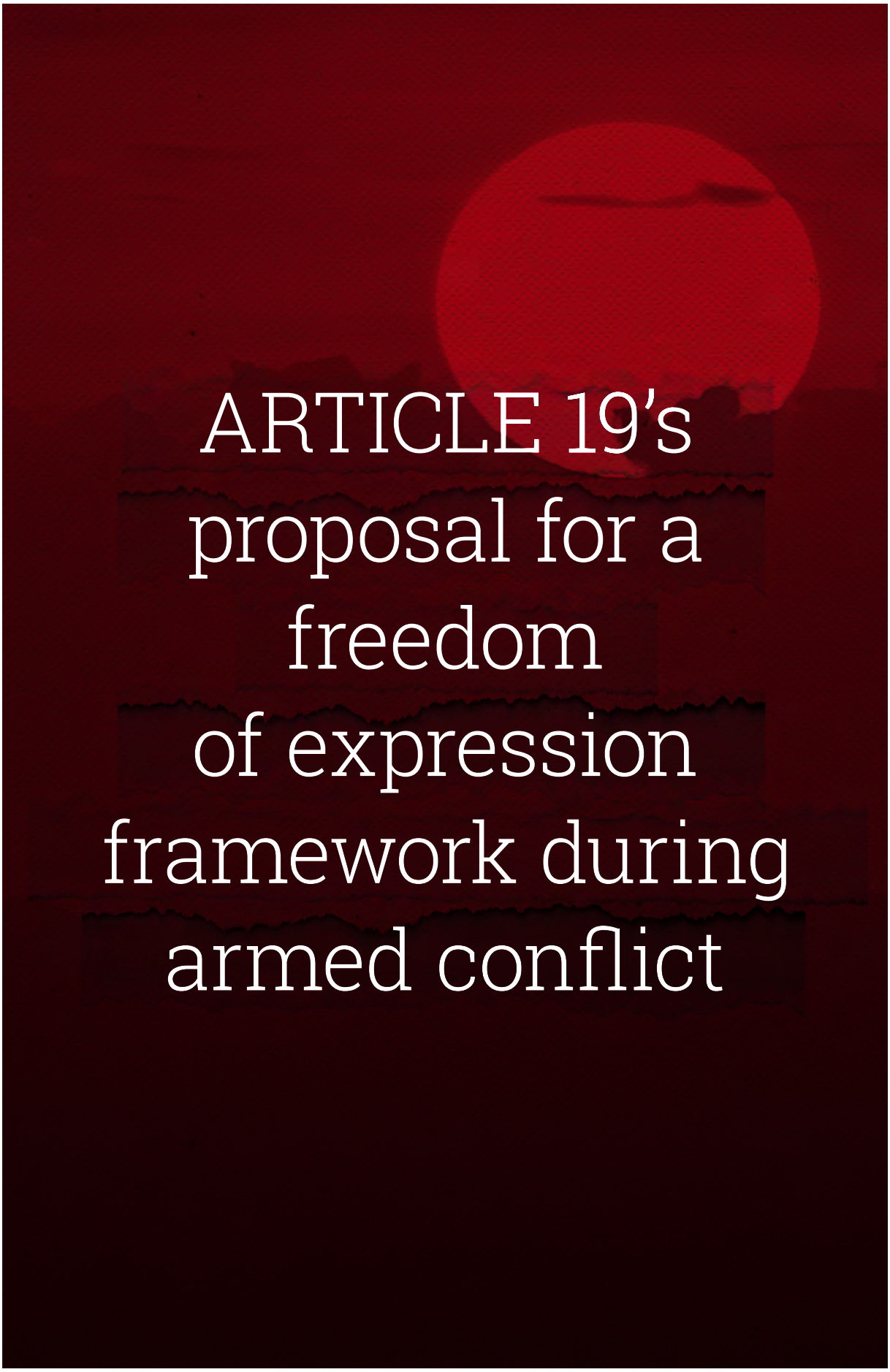
- Social media companies must exercise heightened vigilance when approving advertisements linked to an armed conflict. They should be transparent about any content policies regulating advertisements, ensure consistent application of these policies to different conflict situations, and maintain transparent and accessible advertisement repositories. Given the potential for widespread amplification of content displayed in advertisements and their increasing role in modern warfare, social media companies should in particular exercise utmost diligence when assessing whether the content they contain violates community standards, and allocate sufficient resources, in

particular human reviewers, for this purpose. We also believe that companies have a social responsibility, especially in situations of armed conflict, to ensure they do not profit from content that contributes to the vilification or even dehumanisation of one side of the conflict or the justification of violations of IHL, IHRL, or the rules on the use of force.

- Even where content is removed, social media companies should work closely with relevant accountability mechanisms and civil society organisations that focus on preserving documentation of the conduct of hostilities as well as potential evidence of human rights and humanitarian law violations.

As for regulators, ARTICLE 19 observes a tendency to focus regulatory efforts on the need to curb certain types of content, including 'disinformation', 'hate speech', or illegal content online. Content moderation processes undoubtedly need to be appropriate and robust during armed conflict and account for the potential of content resulting in harm to civilians. At the same time, regulators must be mindful that any restriction of user-generated content ought to comply with the legality, legitimacy, necessity, and proportionality requirements. We also encourage other stakeholders, including human rights and humanitarian organisations engaging with social media companies, not only to concentrate on the need to restrict content that could lead to harm, but also to advocate for upholding freedom of expression.

It remains ARTICLE 19's firm belief that, whether in times of peace or conflict, the emphasis should shift from the content itself to the systems and incentives that determine how content is generated, distributed, and amplified online.[219] This will be more effective in limiting the spread of problematic content, including content that breaches IHL.


ARTICLE 19's
proposal for a
freedom
of expression
framework during
armed conflict

Only by using both the protections offered by IHL and IHRL can the threats to the information environment in contemporary armed conflicts be addressed, including, among others, internet shutdowns, the targeting of civilians through 'information manipulation' and 'hate speech', the banning of media outlets, or attacking journalists through kinetic and non-kinetic means.

A variety of actors have a role to play in interpreting IHL in light of IHRL rules and in a way that upholds freedom of expression during armed conflict. This includes states (through their armed forces, courts, and regulators), international courts, international organisations, civil society, and academics, as well as non-state armed groups.

ARTICLE 19 encourages those and other actors working on the digital and information dimensions of armed conflicts to do so in line with international freedom of expression standards to safeguard the free flow of information and reduce civilian harm. As a foundational step, ARTICLE 19 proposes the following principles for a freedom of expression framework for states, private actors, and other stakeholders to address some of the various challenges affecting the information environment during armed conflict.

1. Upholding freedom of expression during armed conflict protects civilians, as it enables the enjoyment of other human rights and fosters an environment conducive to respect for IHL. Freedom of expression is not a luxury, but a fundamental necessity. It allows civilians to protect themselves in the midst of fighting and stay connected, it protects hospitals' and humanitarian organisations' access to the internet to function properly, and it helps avert impunity for war crimes and other atrocity crimes.

2. Comprehensive protection of freedom of expression in armed conflict requires recognition of the important interplay between IHL and IHRL. Freedom of expression enjoys complementary protections in IHL and IHRL. Despite the absence of explicit IHL provisions, freedom of expression violations can simultaneously breach several IHL rules, including the principles of humanity, distinction, and proportionality. Many of the challenges to the information environment in contemporary armed conflicts, including 'information manipulation', 'hate speech', internet shutdowns, or attacks against the media, cannot be adequately addressed solely based on IHL but require the application of the tools under IHRL in tandem with IHL rules.

3. Any restriction on freedom of expression – whether it impacts individuals within or outside a state's borders – must strictly adhere to the principles of legality, legitimacy, necessity, and proportionality. Even in the face of national security threats, information operations and a surge in 'hate speech', any restrictions must demonstrate that they do not go beyond any legitimate interest in protecting national security, public order, or the rights of others. They should also factor in the fundamental importance of the free flow of information and the objective coverage of armed conflicts by an independent press. Even where an armed conflict justifies derogations from freedom of expression, any restrictive measures must be tailored to the nature and scope of the emergency, with due regard to the principles of legality, legitimacy, necessity, proportionality, and non-discrimination. If a state's actions can impact the exercise or enjoyment of the freedom of expression rights of an individual located outside its borders, freedom of expression obligations should apply extraterritorially towards that individual.

4. The protection of freedom of expression during armed conflict requires investment during pre- and post-conflict times. During armed conflict, freedom of expression is particularly at risk when pre-existing protections are weak. The severity of expression-related violations in conflict depends on the presence or absence of key safeguards which ought to be promoted in pre-conflict times and before the outbreak of hostilities, fostering an environment that is conducive to freedom of expression. Such measures include enhancing media independence and plurality, promoting government transparency, and societal resilience.

5. Cyber operations must adhere to the IHL rules governing attacks and military operations. Cyber operations, including among others, internet shutdowns, spyware, or DDoS attacks, can impact freedom of expression and cause harm to civilians. Such operations should adhere to the rules governing attacks if they are reasonably expected to cause – whether directly or indirectly – death, injury (including serious illness or severe mental suffering tantamount to injury), physical damage, or loss of functionality. Even when they do not qualify as attacks, they remain subject to limitations under IHL, including the principle of precaution.

6. Information operations must adhere to the specific limits set by IHL and IHRL. Responses to 'information manipulation' and 'hate speech' must adopt a freedom of expression-based approach. States and other stakeholders must resist the inclination to address false, misleading, or inciting content with measures that do not abide by freedom of expression standards as this can cause harm to civilians and prove ineffective. Measures should facilitate the promotion of counter-narratives, provide access to diverse viewpoints, and target the root causes of social division. Responses to the online dissemination of 'information manipulation' or 'hate speech' should account for the role played by social media companies' incentive structures underlying their recommender, monetisation, and advertisement systems, as well as their microtargeting techniques.

7. Internet connectivity can be a lifeline for civilians and is protected under both IHRL and IHL. Despite being prevalent in armed conflicts, internet shutdowns are highly unlikely to meet the requirements of either IHL or IHRL. Shutdowns intended to obstruct the documentation efforts of journalists and human rights defenders and conceal violations of IHRL and IHL during armed conflicts may amount to a breach of the obligation to ensure respect for IHL.

8. Tech companies should assume their responsibilities as key actors during armed conflict. They should recognise the crucial role they play in connecting people as well as acknowledge the potential harm they can cause during armed conflict. Furthermore, tech companies must take specific steps to respect IHL and uphold freedom of expression and other human rights. They should live up to their social and moral responsibility and take proactive steps to protect civilians from digital threats, going beyond merely mitigating adverse human rights impacts and ensuring respect for IHL.

9. Actors operating in armed conflict should adopt a freedom of expression-based approach to modern warfare. State and non-state actors, humanitarian organisations, tech companies, and other actors operating in conflict contexts should abide by international freedom of expression standards. This involves enhancing expertise in the digital threat landscape, engaging with digital rights and freedom of expression organisations, and integrating ARTICLE 19's interpretations and recommendations into military manuals, codes of conduct, policies, and protocols, where applicable. Free expression considerations should also be central to any conflict prevention, peacebuilding, and broader peace and security frameworks. National and international courts, tribunals, and accountability mechanisms should consider freedom of expression violations as they assess compliance with international law or potential international crimes.

10. Work towards greater articulation and promotion of freedom of expression standards during armed conflict must continue. While there is no doubt that freedom of expression applies during armed conflict, much work remains to be done to articulate what this means and to operationalise the freedom of expression responsibilities of conflict parties and other actors. This requires a collective approach that includes governments, humanitarian actors, civil society, academia, tech companies, and other relevant stakeholders. An initiative of this kind would be important in maintaining the relevance of IHL amid the current complexities of armed conflicts and their continued migration into the digital space.

Endnotes

[1] The terms 'information manipulation' and 'hate speech' are not defined in international human rights law, and international standards require different responses to different types of 'information manipulation' and 'hate speech'. For these reasons, ARTICLE 19 uses these terms in inverted commas throughout this report. Our approach to these types of expression and their definitions are detailed in the section 'Current responses to "information manipulation" and "hate speech" in armed conflict do not meet freedom of expression standards'.

[2] The International Committee of the Red Cross (ICRC) understands the term 'information operation' to mean 'the use or manipulation of information to influence or mislead the perceptions, motives, attitudes and behaviour of individuals and groups, in order to achieve political and military objectives'. See ICRC (2024) 'International humanitarian law and the challenges of contemporary armed conflicts', 26 September, footnote 170.

[3] See ICRC (2024) 'ICRC 2024 opinion paper: How is the term 'Armed Conflict' defined in international humanitarian law?', 16 April.

[4] See, for example, Rizk, J. and Coredy, S. (2023) 'What we don't understand about digital risks in armed conflict and what to do about it', Humanitarian Law & Policy [blog], 27 July.

[5] See, for example, Hutchins, T. E. (2020) 'Safeguarding civilian internet access during armed conflict: Protecting humanity's most important resource in war', *Columbia Science & Technology Law Review*, XXII, Fall: 127–80. Concluding that current safeguards under IHL are insufficient to adequately protect internet connectivity, this article 'proposes a new legal paradigm with special protections for physical internet infrastructure and the right of civilian access, while advocating the adoption of emblems (such as the Red Cross or Blue Shield) in the digital world to protect vital humanitarian communications'. See also Lahmann, H. (2020) 'Protecting the global information space in times of armed conflict', *International Review of the Red Cross*, 102, 915: 1227–48, which argues that there 'appears to be an emerging need – and room – for a broader rule against systematic and highly corrosive military information operations against civilian information spaces that is not limited to situations of armed conflict but spans the entire spectrum of peace and war' (p. 1247).

[6] UN General Assembly, *International Covenant on Civil and Political Rights*, United Nations Treaty Series, vol. 999, p. 171, 16 December 1966, Article 20(1) (cited hereafter as ICCPR).

[7] UN General Assembly, *Universal Declaration of Human Rights*, 217 A (III), 10 December 1948.

[8] The right to freedom of expression is also codified in Council of Europe, *European Convention on Human Rights* (ECHR), 4 November 1950, Article 10; Organization of American States (OAS), *American Convention on Human Rights* (ACHR), 22 November 1969, Article 13; and Organization of African Unity, *African Charter on Human and Peoples' Rights* (The African Charter), 27 June 1981, Article 9.

[9] A similar formulation can be found in Article 10(2) of the ECHR and Article 13(2) of the ACHR.

[10] It is worth highlighting that several states have made reservations to both paragraphs of Article 20 of the ICCPR. See the Office of the United Nations High Commissioner for Human Rights (OHCHR) (n.d.) 'Status of ratification interactive dashboard'.

[11] See, for example, International Court of Justice, *Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory*, Advisory Opinion, 2004 ICJ 136, 9 July 2004, para 106. See also International Court of Justice, *Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion, 1996 ICJ 226, 8 July 1996, para 25.

[12] See Article 4 of the ICCPR.

[13] See UN Human Rights Committee (2001) *General Comment No. 29, States of Emergency (Article 4)*, CCPR/C/21/Rev.1/Add.11, para 3 (cited hereafter as *General Comment No. 29*).

[14] See *General Comment No. 29*, para 3. See also *Landinelli Silva et al. v. Uruguay*, Comm. No. R.34/1978, UN Doc. CCPR/C/OP/1 at 65 (1984), para 8.3, in which it found that '[a]lthough the substantive right to take derogatory measures may not depend on a formal notification being made pursuant to article 4(3) of the Covenant, the State party concerned is duty-bound to give a sufficiently detailed account of the relevant facts when it invokes article 4(1) of the Covenant in proceedings under the Optional Protocol.'

[15] See *General Comment No. 29*, para 4.

[16] See *General Comment No. 29*, paras 3, 4, 8, and 16. See also Council of Europe (2022) 'Legal analysis of the derogation made by Ukraine under Article 15 of the European Convention of Human Rights and Article 4 of the International Covenant on Civil and Political Rights', paras 104–7, 123.

[17] *General Comment No. 29*, para 4.

[18] See Article 4(3) of the ICCPR requiring that states notify the start and end of the derogation period. See also American Association for the International Commission of Jurists (1985) 'Siracusa principles on the limitation and derogation provisions in the International Covenant on Civil and Political Rights', para 45(c).

[19] This is because freedom of expression is not contained in the list of non-derogable rights under Article 4(2) of the ICCPR.

[20] UN Human Rights Committee (2011) *General Comment No. 34, Article 19: Freedoms of Opinion and Expression*, CCPR/C/GC/34, para 5 (cited hereafter as *General Comment No. 34*).

[21] *General Comment No. 29*, para 13(e).

[22] See, for example, *Landinelli Silva v. Uruguay*, para 8.4; *General Comment No. 34*, para 13; *Mehmet Hasan Altan v. Turkey*, European Court of Human Rights (ECtHR), 20 March 2018, para 210 (in which the ECtHR warned that a public emergency 'must not serve as a pretext for limiting freedom of political debate, which is at the very core of the concept of a democratic society' and emphasised that even in times where the state faces serious threats 'one of the principal characteristics of democracy is the possibility it offers of resolving problems through public debate').

[23] IHL does not apply to situations of 'internal disturbances and tensions', such as 'riots, isolated and sporadic acts of violence and other acts of a similar nature'. See *Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of Non-International Armed Conflicts*, 1125 United Nations Treaty Series, vol. 609, 8 June 1977, Article 1(2) (cited hereafter as Protocol II).

[24] See ICRC (2022) 'What is international humanitarian law?', 5 July.

[25] See ICRC (2010) 'International law on the conduct of hostilities: Overview', 29 October.

[26] For example, a state using force in exercise of its legitimate right to self-defence under Article 51 of the United Nations Charter needs to respect the rules of IHL. See ICRC (2014) 'International humanitarian law: Answers to your questions', December, p. 9.

[27] See *The Prosecutor v. Dusko Tadic*, International Criminal Tribunal for Former Yugoslavia, Decision on the Defence Motion for Interlocutory Appeal on Jurisdiction, 2 October 1995, para 70.

[28] International armed conflicts are primarily governed by the four Geneva Conventions of 1949 – a series of treaties on the treatment of civilians, prisoner of war, and soldiers who are otherwise rendered 'hors de combat' or incapable of fighting – and Protocol I (*Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts*, 1125 United Nations Treaty Series, vol. 3, 8 June 1977). In terms of the core IHL conventions, non-international armed conflicts are covered by Article 3 common to the Geneva Conventions and Protocol II. Certain other treaty rules may apply where the state is a party to those treaties.

[29] See ICRC Casebook, 'Non-international armed conflict: Introduction'.

[30] See United Nations, *Statute of the International Court of Justice*, United States Treaty Series, vol. 993, 18 April, Article 38(1)(b).

[31] At present, the most authoritative articulation of customary IHL relating to armed conflict is the ICRC (2005) *Study on Customary International Humanitarian Law* (cited hereafter as the ICRC study). The study concludes that 136 out of 161 rules of customary international humanitarian law apply equally to non-international armed conflicts. See ICRC Casebook, 'Non-international armed conflict'; Henckaerts, J. and Doswald-Beck, L. (2005) *Customary International Humanitarian Law*, Cambridge: Cambridge University Press.

[32] ICRC (2008) 'Increasing respect for international humanitarian law in non-international armed conflicts', 1 December, p. 5.

[33] See *Hague Convention (IV) Respecting the Laws and Customs of War on Land and Its Annex: Regulations Concerning the Laws and Customs of War on Land*, International Conferences (The Hague), 18 October 1907, Article 42.

[34] According to their Common Article 2, the four Geneva Conventions of 1949 apply to any territory occupied during international hostilities.

[35] See ICRC (2004) 'Occupation and international humanitarian law: Questions and answers', 4 August.

[36] See, for example, Article 72 of Protocol I; Protocol II, preamble, para 2.

[37] See, for example, UN General Assembly, *Convention on the Rights of the Child*, United Nations Treaty Series, vol. 1577, 7 March 1990, Article 38(1); UN General Assembly, *Optional Protocol to the Convention on the Rights of the Child on the Involvement of Children in Armed Conflict*, United Nations Treaty Series, vol. 2173, 25 May 2000, preamble para 12 and Article 5; UN General Assembly, *International Convention for the Protection of All Persons from Enforced Disappearance*, United Nations Treaty Series, vol. 2716, 20 December 2006, Article 43.

[38] It is also widely acknowledged that human rights considerations may apply with greater urgency in non-international armed conflicts. As the Inter-American Commission on Human Rights has observed in the La Tablada case, it is 'during situations of internal armed conflict' that IHL and IHRL 'most converge and reinforce each other'. *Juan Carlos Abella v. Argentina*, Inter-American Commission on Human Rights (IACHR), 18 November 1997, para 160.

[39] See International Court of Justice, *Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory*, para 106.

[40] See UN Human Rights Committee (2004), *General Comment No. 31 [80]: The Nature of the General Legal Obligation Imposed on States Parties to the Covenant*, CCPR/C/21/Rev.1/Add.13, para 11 (cited hereafter as *General Comment No. 31 [80]*). See also *General Comment No. 29*, para 3; UN Human Rights Committee (2019), *General Comment No. 36, Article 6 (Right to Life)*, CCPR/C/GC/36, para 64 (cited hereafter as *General Comment No. 36*).

[41] For example, under Article 79 of Protocol I – which also applies to non-international armed conflicts as customary international law (see the ICRC study, Rule 34) – journalists engaged in dangerous professional missions in conflict areas shall be considered civilians and protected as such.

[42] See OHCHR (2011) 'International legal protection of human rights in armed conflict', p. 61.

[43] See, for example, *Hassan v. The United Kingdom*, ECtHR, Judgment, 16 September 2014, paras 33, 77, and 100–3; *Varnava et al. v. Turkey*, ECtHR, Judgment, 18 September 2009, para 185, *Ukraine v. Russia (re Crimea)*, ECtHR, Judgment, 25 June 2024, paras 912–19; *Democratic Republic of Congo v. Burundi, Rwanda and Uganda*, African Commission on Human and Peoples' Rights, Decision, 29 May 2003, para 79; *Abella v. Argentina*, paras 161 and 176–89; and *Ituango Massacres v. Colombia*, IACHR, Preliminary Objection, Merits, Reparations and Costs, Judgment, 1 July 2006, para 179.

[44] Common Article 3 of the Geneva Conventions, which is applicable in non-international armed conflicts, and the provisions contained in Protocol II are addressed to the parties to the conflict, including non-state armed groups. See also the ICRC study, Rule 139.

[45] See OHCHR (2021) 'Joint Statement by independent United Nations human rights experts on human rights responsibilities of armed non-State actors', 25 February. See also Rodenhäuser, T. (2020) 'The legal protection of persons living under the control of non-state armed groups', *International Review of the Red Cross*, 102, 915: 991–1020.

[46] See UN Human Rights Council (2011) 'Report of the Special Representative of the Secretary-General on the Issue of Human Rights and Transnational Corporations and Other Business Enterprises, John Ruggie: Guiding principles on business and human rights: Implementing the United Nations "Protect, Respect and Remedy" framework', A/HRC/17/31, 21 March, annex (cited hereafter as UN Guiding Principles); and UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (UN Special Rapporteur on freedom of expression) (2016) 'Report of the special rapporteur on the promotion and protection of the right to freedom of opinion and expression', A/HRC/32/38, 11 May, paras 9, 10. The UN Guiding Principles clarify that the responsibility to respect is a global standard of expected conduct for all business enterprises wherever they operate rather than a legal obligation. See UN Guiding Principles, Principle 11.

[47] See UN Guiding Principles, Principle 12. Before the adoption of the UN Guiding Principles, the ICRC's guide on business and international humanitarian law already stated that 'a business enterprise carrying out activities that are closely linked to an armed conflict

must also respect international humanitarian law'. Business activities may be considered 'closely linked to an armed conflict' even if they do not take place during or on the physical battlefield and even if the business did not actually intend to support a party to the hostilities. See ICRC (2006) 'Business and international humanitarian law', p. 14.

[48] See UN Development Programme (UNDP) (2022) 'Heightened human rights due diligence for business in conflict-affected contexts: A guide', 16 June.

[49] See UN Guiding Principles, Principle 4.

[50] See British Red Cross (2017) 'Media professionals and armed conflict: Protection and responsibilities under international humanitarian law, p. 28.

[51] See ICRC (2021) 'Harmful information: Misinformation, disinformation and hate speech in armed conflict and other situations of violence', July, pp. 5, 10.

[52] See Access Now (2024) 'Shrinking democracy, growing violence: Internet shutdowns in 2023', May, p. 9 ('conflicts emerged for the first time as the leading driver of internet shutdowns').

[53] See Access Now (2023) 'Spyware in warfare: Access Now documents first-time use of Pegasus tech in Azerbaijan-Armenia conflict', 25 May.

[54] See UN Special Rapporteur on freedom of expression (2022) 'Disinformation and freedom of opinion and expression during armed conflicts', UN Doc A/77/288, para 4.

[55] See, for example, Hutchins (2020) 'Safeguarding civilian internet access'; and Lahmann (2020) 'Protecting the global information space'.

[56] See Droege, C. and Giorgou, E. (2022) 'How international humanitarian law develops', *International Review of the Red Cross*, 104, 920–921: 1798–1839, pp. 1809–10.

[57] For a reflection of the role of non-binding norms and soft law in international humanitarian law, see Crawford, E. (2022) 'Non-binding norms in the law of armed conflict', *Articles of War*, 3 February.

[58] See ICRC (2014) 'Answers to your questions', pp. 39–41.

[59] See ICRC (2014) 'Answers to your questions', pp. 36–37.

[60] See ICRC (2014) 'Answers to your questions', pp. 37–38.

[61] See UN General Assembly, *Rome Statute of the International Criminal Court*, United Nations Treaty Series, vol. 2187, 17 July 1998, Article 8 (cited hereafter as Rome Statute).

[62] See Droege and Giorgou (2022) 'How international humanitarian law develops', p. 1820.

[63] See Article 79 of Protocol I. During the negotiations of Protocol I, it was considered important to provide special protection to journalists because of the crucial function they perform when reporting from conflict zones. See ICRC (1987) 'Commentary to Protocol I', Article 79.

[64] See UN Security Council (2015), 'Security Council resolution 2222 (2015) [on protection of journalists and the issue of impunity]', S/RES/2222 (2015) (cited hereafter as Security Council resolution 2222 (2015)).

[65] See Longworth, S. (2022) *Freedom of Expression in Armed Conflict: The Silence Between Spaces*, Stockholm: Stockholm University, pp. 176–77. International bodies have also recognised the connection between the physical protection of journalists as a necessary pre-requisite of their right to free expression. See Security Council resolution 2222 (2015); OHCHR (2004), 'The right to freedom of opinion and expression', E/CN.4/RES/2004/42, , para 5; Council of Europe, 'Recommendation No. R (96) 4 of the Council of Europe Committee of Ministers on the protection of journalists in situations of conflict and tension'.

[66] The ICRC study, Rules 139 and 144; Longworth (2022) *Freedom of Expression in Armed Conflict,* p. 388.

[67] See ICRC (2016) 'Commentary on the First Geneva Convention', 2nd edition, Common Article 1, paras 145–46.

[68] See 'Protecting civilians and other protected persons and objects against the potential human cost of ICT activities during armed conflict', resolution passed at the 34th International Conference of the Red Cross and Red Crescent, 28–34 October 2024, preamble, para 7; UN Security Council (2006) 'Security Council Resolution 1738 (2006) [Protection of civilians in armed conflict]', S/RES/1738, para 3 (confirming that 'media equipment and installations constitute

civilian objects, and in this respect shall not be the object of attack or of reprisals, unless they are military objectives'); Security Council resolution 2222 (2015) para 10.

[69] See ICRC (2024), 'Protecting civilians and other protected persons and objects against the potential human cost of ICT activities during armed conflict', preamble, para 2.

[70] See ICRC (2019), 'International humanitarian law and cyber operations during armed conflicts', 28 November, p. 4 ('For the ICRC, there is no question that IHL applies to, and therefore limits, cyber operations during armed conflict – just as it regulates the use of any other weapon, means and methods of warfare in an armed conflict, whether new or old').

[71] See, for example, the ICRC study, Rule 55, expressing the customary IHL obligation to allow and facilitate rapid and unimpeded passage of humanitarian relief for civilians in need. See also obligation in the ICRC study, Rule 110 to allow for the wounded, sick, and shipwrecked to receive, to the fullest extent practicable and with the least possible delay, the medical care and attention required by their condition. For the protection of hospitals, see, for example, Article 19 of the First Geneva Convention (Diplomatic Conference of Geneva of 1949, *Convention (I) for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field*, 12 August 1949); Article 18 of the Fourth Geneva Convention (Diplomatic Conference of Geneva of 1949, 'Convention (IV) Relative to the Protection of Civilian Persons in Time of War', 12 August 1949); Article 12 of Protocol I; Article 11 of Protocol II; the ICRC study, Rules 25, 28, and 29.

[72] See UN Human Rights Council (2022) 'Internet shutdowns: Trends, causes, legal implications and impacts on a range of human rights', Report of the Office of the High Commissioner for Human Rights, A/HRC/50/55, 13 May, paras 37–39.

[73] See Article 58(c) of Protocol I; the ICRC study, Rule 22.

[74] See Longworth (2022) *Freedom of Expression in Armed Conflict*, p. 385.

[75] See Article 2(1) of the ICCPR; Article 1 of the ECHR; and Article 1 of the ACHR. The African Charter does not contain any reference to jurisdiction.

[76] See *General Comment No. 31 [80]*, para 10. For the interpretation of the European Convention of Human Rights in this context, see *Al-Skeini and Others v. The United Kingdom*, ECtHR, 7 July 2011, paras 130–40.

[77] See, for example, *Ocalan v. Turkey*, ECtHR, 12 May 2005, para 91.

[78] Depending on the specific circumstances, it is also conceivable that a state's jurisdiction may extend extraterritorially through the activities of entities, such as companies, that are based in its territory or subject to its jurisdiction, if those companies' activities have a direct and reasonably foreseeable impact on the right to freedom of expression of individuals outside the state's borders. See van Benthem, T., Dias, T., and Hollis, D. B. (2022) 'Information operations under international law', *Vanderbilt Journal of Transnational Law*, 55, 5: 1217–86, pp. 1253–54, referring to the Human Rights Committee's argument in *General Comment No. 36*, para 22.

[79] This does not mean that a state with the power to violate freedom of expression obligations abroad would necessarily be bound by the entire catalogue of human rights guaranteed in the ICCPR or regional instruments.

[80] See, for example, UN Special Rapporteur on freedom of expression (2022) 'Disinformation and Disinformation and freedom of expression during armed conflicts', paras 50–52 ('the power of effective control should be considered not only over the person or the territory where they are located but over their human rights'). See also the position of the minority in International Group of Experts at the Invitation of the NATO Cooperative Cyber Defense Center of Excellence (2017), *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations* (2nd edn), Cambridge: Cambridge University Press, chapter 6, Rule 34, para 10 (cited hereafter as *Tallinn Manual 2.0*); Milanovic, M. (2011) *Extraterritorial Application of Human Rights Treaties: Law, Principles, and Policy*, Oxford: Oxford University Press, pp. 209–22 (arguing that a state's negative obligation to respect human rights should be territorially unbound). Of note is the judgment in *Wieder and Guarnieri v. The United Kingdom*, ECtHR, 12 September 2023, where the Court found that the interception, storing, or processing of an individual's data, which affects their right to privacy, falls within the jurisdiction of the ECHR if the surveillance occurs within the state's own territory, even if the individual concerned is located outside that territory (albeit framing the case as not being about extraterritorial application at all). For commentary, see Milanovic, M.

(2024) 'Wieder and Guarnieri v. UK: A justifiably expansive approach to the extraterritorial application of the right to privacy in surveillance cases', EJIL:Talk! [blog], *European Journal of International Law*, 21 March.

[81] Instead, IHL is mainly enforced through international criminal law, which focuses on the prosecution of individuals for war crimes.

[82] See Access Now (2022) 'Taxonomy of internet shutdowns: How internet shutdowns are implemented', 1 June, pp. 6–30.

[83] In Syria, journalists reported identity fraud, with Facebook pages and X (formerly Twitter) accounts being opened under their names, or accusations of crimes appearing on websites. See International Press Institute (IPI) (2013) 'Latest threat to journalists covering Syria: Identity fraud', IFEX, 17 April.

[84] The CyberPeace Institute has recorded DDoS attacks against Ukraine since January 2022, including against media outlets; CyberPeace Institute (n.d.) 'Timeline: How have cyberattacks and operations evolved over time since the military invasion of Ukraine?'. See, for example, IPI (2022) 'Ukrainian news site NikVesti offline following DDoS attack', 18 May.

[85] In Ukraine in 2022, there was an organised doxing campaign against members of the Ukrainian military and journalists. See, for example, Institute for Strategic Dialogue (2022) 'Project Nemesis, doxxing and the new frontier of informational warfare', 23 June; Roth, A. (2016) 'Hackers have doxed all the reporters covering east Ukraine's war. Twice', *The Washington Post*, 27 May.

[86] In the context of the Nagorno-Karabakh conflict in Azerbaijan, at least 12 Armenian public figures and officials, including journalists and human rights defenders, were targeted with NSO Group's Pegasus spyware between October 2020 and December 2022. See Amnesty International (2023) 'Armenia/Azerbaijan: Pegasus spyware targeted Armenian public figures amid conflict', 25 May.

[87] See Rodenhäuser, T. and D'Cunha, S. (2023) 'IHL and information operations during armed conflict', Articles of War [blog], Lieber Institute West Point, 18 October.

[88] See Article 49(1) of Protocol I. Such attacks under the IHL framework need to be distinguished from the type of attack under governed by the UN Charter that would give rise to the right to self-defence.

[89] See ICRC (2019) 'International humanitarian law and cyber operations during armed conflicts', p. 7; *Tallinn Manual 2.0*, chapter 17, Rule 92, para 2. See also NATO Cooperative Cyber Defence Centre of Excellence (CCDCOE) (n.d.), 'Attack (international humanitarian law)'. Note that the principal rules of IHL governing targeting in international armed conflicts – including the principles of distinction and proportionality and the prohibition of indiscriminate attacks – also apply to non-international armed conflicts as customary international law. See the ICRC study, Part I.

[90] See Articles 51–52 of Protocol I; the ICRC study, Rules 1–10.

[91] See Article 51(4) of Protocol I; the ICRC study, Rules 11–13.

[92] See Article 51(5)(b) of Protocol I; the ICRC study, Rule 14.

[93] See Article 57 of Protocol I; the ICRC study, Rules 15–21.

[94] *Tallinn Manual 2.0*, chapter 17, Rule 92, para 8.

[95] *Tallinn Manual 2.0*, chapter 17, Rule 92, para 3; ICRC (2019) 'International humanitarian law and cyber operations during armed conflicts', p. 7.

[96] See *Tallinn Manual 2.0*, chapter 17, Rule 92, paras 10–13. See also CCDCOE (n.d.), 'Attack (international humanitarian law)' (including an overview of state positions on the interpretation of what constitutes 'damage' for assessing whether an operations amounts to an 'attack'). Related to this question is also the unsettled question whether data – for example, civil registries, insurance data, medical data – benefit from IHL protections. See Gisel, L. and Rodenhäuser, T. (2019) 'Cyber operations and international humanitarian law: Five key points', Humanitarian Law & Policy [blog], 28 November.

[97] Gisel, L., Rodenhäuser, T., and Dörmann, K. (2020) 'Twenty years on: International humanitarian law and the effects of cyberoperations during armed conflicts', *International Review of the Red Cross,* 102, 913: 287–334, pp. 312–13.

[98] See ICRC (2019) 'International humanitarian law and cyber operations during armed conflicts', pp. 7–8.

[99] See ICRC (2019) 'International humanitarian law and cyber operations during armed conflicts', pp. 7–8.

[100] See ICRC (2019) 'International humanitarian law and cyber operations during armed conflicts', pp. 7–8.

[101] See UN Human Rights Council (2022) 'Internet shutdowns', paras 33–39.

[102] See OHCHR (2021) 'Use of spyware to surveil journalists and human rights defenders: Statement by UN High Commissioner for Human Rights Michelle Bachelet', 19 July.

[103] See, for example, Bowcott, O. (2020) 'UN warns of rise of "cybertorture" to bypass physical ban', *The Guardian*, 21 February (in which former UN Special Rapporteur on torture and other cruel, inhuman or degrading treatment or punishment commented that '[c]ybertechnology can also be used to inflict, or contribute to, severe mental suffering while avoiding the conduit of the physical body, most notably through intimidation, harassment, surveillance, public shaming and defamation, as well as appropriation, deletion or manipulation of information'.)

[104] See International Center for Journalists (2022) 'The chilling: A global study of online violence against women journalists', 2 November, p. 43–44.

[105] International Court of Justice, *Legality of the Threat or Use of Nuclear Weapons*, para 86.

[106] ICRC (2023) 'The principles of humanity and necessity', March, pp. 1–2; *Tallinn Manual 2.0*, chapter 17, para 13.

[107] See *Tallinn Manual 2.0*, chapter 17, Rule 114.

[108] See ICRC (2024) 'International humanitarian law and the challenges of contemporary armed conflicts', p. 58.

[109] See ICRC (2024) 'International humanitarian law and the challenges of contemporary armed conflicts', p. 58.

[110] See also Rodenhäuser, T. (2023) 'The Legal Boundaries of (Digital) Information or Psychological Operations Under International Humanitarian Law', *International Law Studies*, 100: 541–73, p. 566 arguing that 'in some circumstances, information or psychological operations may amount to attacks as defined in IHL and therefore be subject to the IHL principles and rules on the conduct of hostilities'.

[111] See ICRC (2024) 'International humanitarian law and the challenges of contemporary armed conflicts', p. 59.

[112] See Rizk, J. (2024) 'Why is the ICRC concerned by "harmful information" in war', Humanitarian Law & Policy [blog], 10 September.

[113] See ARTICLE 19 (2021) 'Response to the consultations of the UN Special Rapporteur on freedom of expression and 'disinformation'.

[114] Importantly, this policy does not address matters concerning the interpretation of Article 20(1) of the ICCPR, which prohibits 'propaganda for war'. These issues will be addressed in a separate upcoming ARTICLE 19 policy.

[115] See McCarthy, A. H. (2018) 'Erosion of the rule of law as a basis for command responsibility under international humanitarian law,' *Chicago Journal of International Law*, 18, 2: 553–93. See also Rizk (2024) 'Why is the ICRC concerned by "harmful information" in war'.

[116] See, for example, UN Human Rights Council (2018) 'Report of the independent international fact-finding mission for Myanmar', A/HRC/39/64, 12 September, para 74; Jackson, J. et al. (2022) 'Facebook accused by survivors of letting activists incite ethnic massacres with hate and misinformation in Ethiopia', Bureau of Investigative Journalism, 20 February; Madung, O. (2022) 'From dance app to political mercenary: How disinformation on TikTok gaslights political tensions in Kenya', Mozilla, 7 June.

[117] For example, in Myanmar, the government-controlled media has been described as 'just propaganda outlets'. See Reporters Without Borders (n.d.) 'Myanmar'.

[118] For instance, the role of Radio Télévision Libre des Mille Collines has been regarded as crucial to creating the racial hostility that allowed the Rwandan genocide to occur; see ARTICLE 19 (2005) 'War of words: Conflict and freedom of expression in South Asia', May, p. 13. More recently, mass media in Ethiopia have been reported as having contributed to fuelling ethnic violence in Ethiopia. See Shifa, M. and Pabón, F. A. D. (2022) 'The interaction of mass media and social media in fuelling ethnic violence in Ethiopia', Accord, 15 March.

[119] Lahmann (2020) 'Protecting the global information space', pp. 1233–34.

[120] See Article 37(2) of Protocol I; the ICRC study, Rule 57.

[121] See *Tallinn Manual 2.0*, chapter 17, Rule 93, para 5.

[122] See Article 37(1) of Protocol I; the ICRC study, Rule 65.

[123] See Article 51(2) of Protocol I; the ICRC study, Rule 2.

[124] See Rodenhäuser (2023) 'The legal boundaries of (digital) information or psychological operations under international humanitarian law', pp. 552–53.

[125] See ICRC (2024), 'International humanitarian law and the challenges of contemporary armed conflicts', p. 59.

[126] UN General Assembly, *Convention on the Prevention and Punishment of the Crime of Genocide*, United Nations Treaty Series, vol. 78, p. 277, 9 December 1948, Article III(c).

[127] See Article 25(3)(b) of the Rome Statute. For the distinction between 'direct and public incitement to commit genocide' and the accessory modes of liability, see Timmermann, W. K. (2006) 'Incitement in international criminal law', *International Review of the Red Cross*, 88, 864: 823–52.

[128] As mentioned above, several states have made reservations to both paragraphs of Article 20 of the ICCPR. See OHCHR (n.d.) 'Status of ratification interactive dashboard'.

[129] See *General Comment No. 34*, paras 49, 50; OHCHR (2013) 'Report of the United Nations High Commissioner for Human Rights on the expert workshops on the prohibition of incitement to national, racial or religious hatred', A/HRC/22/17/Add.4, 11 January, para 18 (known as the Rabat Plan of Action).

[130] See UN Special Rapporteur on freedom of expression (2021) '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, Irene Khan', A/HRC/47/25, 13 April, para 40; *General Comment No. 34*, paras 47, 49.

[131] For similar reasoning, see UN Human Rights Committee (2001) *General Comment No. 36*, para 59 ('Failure to comply with [the] obligations under article 20 [of the ICCPR] may also constitute a failure to take the necessary measures to protect the right to life under article 6').

[132] See UN Special Rapporteur on freedom of expression(2022) 'Disinformation and freedom of expression during armed conflicts', para 110; see also the 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 African Commission on Human and Peoples' Rights Special Rapporteur on Freedom of Expression and Access to Information (2017) 'Joint declaration on freedom of expression and "fake news", disinformation and propaganda', 3 March, para 2(c) (cited hereafter as 'Joint declaration on freedom of expression and "fake news"').

[133] See *Case of Perozo et al. v. Venezuela*, IACtHR, 2009, para 151. See also 'Joint declaration on freedom of expression and "fake news"', para 1(d); UN Special Rapporteur on freedom of expression (2021) 'Disinformation and freedom of opinion and expression', para 38 (arguing that the right to freedom of expression places a positive obligation on states to proactively put information of public interest in the public domain).

[134] See Longworth (2022) *Freedom of Expression in Armed Conflict*, p. 132.

[135] See ARTICLE 19 (2019) '"Hate speech" explained: A toolkit', 23 December, p. 67.

[136] See *Salov v. Ukraine*, ECtHR, Judgment, 6 September 2005, para 113 ('Article 10 of the [European] Convention [on Human Rights, on freedom of expression] as such does not prohibit discussion or dissemination of information received even if it is strongly suspected that this information might not be truthful.'). See also *General Comment No. 34*, paras 47, 49.

[137] UN Special Rapporteur on freedom of expression (2021) 'Disinformation and freedom of opinion and expression', para 40.

[138] 'Joint declaration on freedom of expression and "fake news"', paras 1(c) and 1(h).

[139] For a discussion of the interplay between IHRL and IHL, ee Longworth (2022) *Freedom of Expression in Armed Conflict*, p. 414–16.

[140] The typology contains three main categories of 'hate speech': first, hate speech that must be prohibited under international law, including incitement to genocide under the Genocide Convention or advocacy of discriminatory hatred constituting incitement to hostility, discrimination, or violence under Article 20(2) of the ICCPR. Second, hate speech that may be restricted under Article 19(3) of the ICCPR. Third, lawful 'hate speech' raising concerns in terms of tolerance but that must be protected under Article 19(3) of the ICCPR. See ARTICLE 19 (2019) '"Hate speech" explained: A toolkit'.

[141] For example, it can be difficult to distinguish between hate speech prohibited under Article 20(2) of the ICCPR and expression that reflects the realities of conflicts, encourages use of violence that is not prohibited by IHL (for example, violence that is aimed at combatants) and is not motivated by hatred or aimed at a group due to their protected characteristic. In non-international armed conflicts where conflicting groups may define themselves along ethnic or religious lines, drawing this distinction can be particularly challenging.

[142] The UN Rabat Plan of Action provides authoritative guidance to states on implementing their obligations under Article 20(2) of the ICCPR to prohibit 'any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence'.

[143] See also du Toit, P. (2014) 'Reporting atrocities: A toolbox for journalists covering violent conflict and atrocities', Internews, 13 November.

[144] In the war in Gaza, Palestine, following 7 October 2023, both local and international media were accused of biased coverage. See Johnson, A. and Ali, O. (2024) 'Coverage of Gaza war in the New York Times and other major newspapers heavily favored Israel, analysis shows', *The Intercept*, 9 January; Graham-Harrison, E. and Kierszenbaum, Q. (2024) 'Journalists see their role as helping to win': How Israeli TV is covering Gaza war', *The Guardian*, 6 January; İnceoğlu, Y. G. (2024) '"Dead" versus "killed": A closer look at the media bias in reporting Israel–Palestine conflict', *The Wire*, 1 November (finding that 'depending upon the side they favour[ed]' in the Israel–Palestine conflict after October 2023, 'media outlets have been tactfully employing language, which is resulting in hate speech and warmongering').

[145] More specifically, under Article 79 of Protocol I – which , as mentioned, also applies to non-international armed conflicts as customary international law – journalists engaged in dangerous professional missions in conflict areas shall be considered civilians and protected as such.

[146] The ICRC study, Rule 158; see also the Council of Europe (1996) 'Recommendation No. R (96) 4 of the Committee of Ministers of the Council of Europe on the protection of journalists in situations of conflict and tension', 3 May.

[147] See Article 52 of Protocol I. As established by Article 52(3) of Protocol I, in cases of doubt, objects normally dedicated to civilian purposes are to be presumed not to be used to make an effective contribution to military action. Indiscriminate attacks on media facilities are prohibited. See also UN Security Council (2006) 'Security Council Resolution 1738', para 3. For a broad analysis of the protections of journalists and news media personnel, see Balguy-Gallois, A. (2004) 'Protection des journalistes et des médias en période de conflit armé', *International Review of the Red Cross*, 86, 853: 37–67 (English translation).

[148] See Burri, N. (2015) *Bravery or Bravado? The Protection of News Providers in Armed Conflict*, Leiden: Brill Nijhoff, p. 287.

[149] See, for example, Davies, H., et al. (2024) '"The grey zone": How IDF views some journalists in Gaza as legitimate targets', *The Guardian*, 25 June.

[150] See Article 51(3) of Protocol I, which reflects customary international law. See the ICRC study, Rule 6.

[151] Melzer, N. (2009) 'Interpretive guidance on the notion of direct participation in hostilities under international humanitarian law', ICRC, May, p. 46.

[152] The definition of military objectives is found in Article 52(2) of Protocol I.

[153] See Burri (2015) *Bravery or Bravado?*, p. 291.

[154] See Burri (2015) *Bravery or Bravado?*, p.186.

[155] Melzer, N. (2009) 'Interpretive guidance', p. 51. It also reflects the view of the majority of the International Group of Experts in *Tallinn Manual 2.0*, chapter 18, Rule 139, para 9.

[156] See, for example, Balguy-Gallois, A. (2004) 'Protection des journalistes et des médias en période

de conflit armé', p. 52 (p. 11 of the English translation).

[157] This was the view in International Criminal Court for the Former Yugoslavia (2000) 'Final report to the prosecutor by the committee established to review the NATO bombing campaign against the Federal Republic of Yugoslavia', 13 June, para 47 ('If the media is used to incite crimes, as in Rwanda, then it is a legitimate target. If it is merely disseminating propaganda to generate support for the war effort, it is not a legitimate target'). The report followed the NATO bombing of the Serbian State radio and television (RTS) building in 1999, justified by NATO as neutralising a propaganda tool.

[158] Examples include Myanmar (Committee to Protect Journalists (2021) 'Bitter reversal: Myanmar military coup wipes out press freedom gains', 28 July); Ethiopia (Anna, C. (2021) 'Ethiopia seeks to restrict media reporting on yearlong war', *AP News*, 26 November); and Russia (Milanovic, M. (2022) 'The legal death of free speech in Russia', EJIL:Talk! [blog], *European Journal of International Law*, 8 March).

[159] See, for example, *Reuters* (2022) 'Russia blocks access to BBC and Voice of America websites', 4 March; as well as the EU ban on the Russian state-controlled media (first RT and Sputnik, and subsequently also Rossiya RTR/RTR Planeta, Rossiya 24/Russia 24, and TV Centre International) following the invasion of Ukraine, adopted with the justification that 'RT and Sputnik are essential and instrumental in bringing forward and supporting Russia's aggression against Ukraine'. See European Commission (2022) 'Ukraine: Sanctions on Kremlin-backed outlets Russia Today and Sputnik', 2 March.

[160] For example, since the Gaza blockade imposed by Israel in 2007, journalists are not allowed to enter the territory without authorisation from Israel. See ARTICLE 19 (2023) 'Israel and Palestine: Stop the assault on free speech and protect civilians', 13 November. For Sudan, see *Nuba Reports* (2017) 'Sudan's silent conflicts: State censorship in the war zones', 3 May; for Ukraine, see Scott, L. (2023) 'New rules limit media's ability to cover Ukraine war', *Voice of America*, 31 March.

[161] Burri (2015) *Bravery or Bravado?*, pp. 255–56; Longworth (2022) *Freedom of Expression in Armed Conflict*, p. 176.

[162] See Geiss, R. (2008) 'The protection of journalists in armed conflicts', *German Yearbook of International Law*, 51: 289–320, pp. 302–3.

[163] See Geiss (2008) 'The protection of journalists in armed conflicts', pp. 303–4.

[164] *General Comment No. 34*, para 30.

[165] *Sürek and Ôzdemir v. Turkey*, ECtHR, 8 July 1999, paras 61–64.

[166] UN Special Rapporteur on freedom of expression (2022) 'Disinformation and freedom of expression during armed conflicts', para 64.

[167] UN Special Rapporteur on freedom of expression (2022) 'Disinformation and freedom of expression during armed conflicts', para 103.

[168] See ARTICLE 19 (2021) 'Response to the consultations of the UN Special Rapporteur on freedom of expression and 'disinformation', Recommendations; "Hate speech" explained: A toolkit', p. 52.

[169] Examples in contravention of these standards include the EU Council's suspension of *Russia Today*, *Sputnik*, and other media outlets in 2022 (see European Commission (2022) 'Ukraine: Sanctions on Kremlin-backed outlets Russia Today and Sputnik') as well as Israel's shutdown of Al Jazeera in 2024 (see ARTICLE 19 (2024) 'Israel: Al Jazeera ban is an attack on media freedom and war reporting', 8 May).

[170] It should also be possible to appeal against the application of administrative measures to an independent court or other adjudicatory body. See UN Special Rapporteur on Freedom of Opinion and Expression, the OSCE Representative on Freedom of the Media, the OAS Special Rapporteur on Freedom of Expression, and the African Commission on Human and Peoples' Rights Special Rapporteur on Freedom of Expression and Access to Information (2015) 'Joint declaration on freedom of expression and responses to conflict situations', 4 May, para 4(a)(b) (cited hereafter as 'Joint declaration on freedom of expression and responses to conflict situations').

[171] See also *General Comment No. 34*, para 45.

[172] To be able to cover the war in Gaza, Palestine, following 7 October 2023, some international journalists embedded with the Israel Defense Forces. The forces have imposed strict conditions, including the requirement for reports to be submitted for

review before publication and the prohibition of interaction with Palestinians. See ARTICLE 19 (2024) 'Israel and Palestine: Allow international media access to Gaza', 25 January.

[173] For example, on 4 March 2022, the Russian authorities blocked access to online media – both in Russian territory and Crimea – and pressured social media companies to remove content. See Longworth (2022) *Freedom of Expression in Armed Conflict*, p. 384; OHCHR (2022) 'Update on the human rights situation in Ukraine', 28 March, para 58.

[174] See 'Joint declaration on freedom of expression and responses to conflict situations', para 8(a).

[175] Internet shutdowns are usually described as the 'intentional disruption of internet or electronic communications, rendering them inaccessible or effectively unusable, for a specific population or within a location, often to exert control over the flow of information'. See, for example, Access Now (2022) 'Taxonomy of internet shutdowns', p. 5.

[176] For instance, since the outbreak of hostilities in Ethiopia between the Ethiopian Defense Forces and the regional forces in Tigray, there were a series of internet shutdowns reportedly imposed by the government of Ethiopia. See NetBlocks (2020) 'Internet disrupted in Ethiopia as conflict breaks out in Tigray region', 4 November. In February and March 2022, the Russian government blocked social media platforms including X (formerly Twitter) as well as Meta-owned Facebook and Instagram. See Freedom House (n.d.), 'Freedom of the Net 2022 – Russia'.

[177] By the end of 2022, Ukraine suffered at least 22 shutdowns as the Russian military used missile strikes to attack communications infrastructure, while also reportedly launching cyberattacks against Ukrainian ISPs. See Access Now (2023), 'Weapons of control, shields of impunity', February, pp. 4, 10.

[178] For an overview of the various technical mechanisms for implementing internet shutdowns, see Access Now (2022) 'Taxonomy of internet shutdowns'.

[179] Telecommunications law is also relevant in assessing the legality of internet shutdowns but will not be covered in more detail in this policy. However, it is important to note that Articles 34 and 35 of the International Telecommunications Union Constitution have been invoked by some states as granting legal authority to block communications, including to implement internet shutdowns. As the UN Human Rights Council found, these provisions must, however, be applied together with and subject to the additional obligations that states have assumed under international human rights law to respect the right to freedom of expression and other applicable human rights. UN Human Rights Council (2022) 'Internet shutdowns', para 18.

[180] This was the case in Myanmar and Ethiopia. See Human Rights Watch (2020) 'Myanmar: End world's longest internet shutdown', 19 June; NetBlocks (2020) 'Internet disrupted in Ethiopia'.

[181] Longworth (2022) *Freedom of Expression in Armed Conflict*, pp. 379–81, 389.

[182] Longworth (2022) *Freedom of Expression in Armed Conflict*, pp. 383–84.

[183] See UN Human Rights Council (2022) 'Internet shutdowns', para 13 (stipulating that 'Internet shutdowns […] generally do not meet those requirements. Given their indiscriminate and widespread impacts, internet shutdowns very rarely meet the proportionality test.')

[184] UN Human Rights Council (2022) 'Internet shutdowns', para 13.

[185] UN Human Rights Council (2022) 'Internet shutdowns', para 13.

[186] See 'Joint declaration on freedom of expression and responses to conflict situations', para 4(c).

[187] See also UN Special Rapporteur on freedom of expression(2022) 'DDisinformation and freedom of expression during armed conflicts', para 70 ('Shutting or slowing down [the internet] aggravates rather than combats disinformation, propaganda or incitement.').

[188] See Deffenbaugh, N. (2024) 'Dehumanization: Practicing humanity', Humanitarian Law & Policy [blog], 27 June.

[189] Longworth (2022) *Freedom of Expression in Armed Conflict*, p. 402.

[190] Longworth (2022) *Freedom of Expression in Armed Conflict*, p. 393.

[191] See ICRC (2023), 'Global Advisory Board on digital threats during armed conflicts', final report, 9 October, Recommendation 4.

[192] UN Human Rights Council Working Group on Business and Human Rights (2020) 'Report on business, human right and conflict-affected regions: Towards heightened action', A/75/212, 21 July, para 43.

[193] Myanmar has provided particularly concerning examples of these practices with the government – both before and after the February 2021 coup – ordering Telenor and ISPs to block websites containing 'fake news'; to install intercept spyware; to shut down connectivity; and to provide sensitive consumer data. See PROTECT Consortium (2021) 'Rebuilding an architecture of oppression'; Telenor (2022) 'Updates on Telenor in Myanmar'. In Sudan, a day after the conflict between government forces and the Rapid Support Forces broke out in April 2023, MTN Sudan blocked internet services at the request of Sudan's telecommunications regulator for a few hours. See Reuters (2023) 'Sudanese telecoms provider MTN restores internet service – MTN official', 16 April. Tech companies had to respond to demands from both Russia and Ukraine following the full-scale invasion in February 2021 with both sides requesting to block access or restrict online content. See, for example, Satariano, A. and Frenkel, S. (2022) 'Ukraine war tests the power of tech giants', *The New York Times*, 28 February. Israel's takedown requests to social media companies have increased tenfold after the start of the war in Gaza, Palestine, following 7 October 2023. See Brewster, T. (2023) 'Israel has asked Meta and TikTok to remove 8,000 posts related to Hamas war', *Forbes*, 14 November.

[194] See UN Special Rapporteur on freedom of expression (2022) 'Disinformation and freedom of expression during armed conflicts', paras 74–99.

[195] See surveys by the Business & Human Rights Resource Centre among tech companies of steps undertaken in the context of the armed conflicts in Ukraine and Gaza, Palestine, for example: Business & Human Rights Resource Centre (2022) 'Russian invasion of Ukraine: Analysis of companies' human rights due diligence', 24 May (including a survey conducted among the technology sector); Business & Human Rights Resource Centre (2024) 'Switched off: Tech company opacity & Israel's war on Gaza', 18 April.

[196] In 2018, for example, following the massacre of Rohingya Muslims by the military in Myanmar, the independent international fact-finding mission on Myanmar concluded that Facebook had been 'a useful instrument for those seeking to spread hate in a context where, for most users, Facebook is the internet'. See , UN Human Rights Council (2018) 'Report of the independent international fact-finding mission for Myanmar', para 74. In Ethiopia, by way of further example, the Bureau of Investigative Journalism reported that Facebook posts that were 'inciting violence or making false claims designed to encourage hate between ethnic groups in Ethiopia have been allowed to circulate freely' (see Jackson et al. (2022) 'Facebook accused by survivors').

[197] See the work of Mnemonic to preserve digital information documenting human rights violations, including on social media platforms.

[198] Other relevant changes increasing the amount of 'misinformation' on X included the introduction of a 'pay-per-view' monetisation model for premium users, the elimination of headline previews for links shared on the platform, and restrictions to X's application programming interface (API). See Brooking, E. T., Mashkoor, L., and Malaret, J. (2023), 'Distortion by design: How social media platforms shaped our initial understanding of the Israel-Hamas conflict', Atlantic Council's Digital Forensic Research Lab, 21 December.

[199] See Zuboff, S. (2018) *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: PublicAffairs; Ranking Digital Rights (2020) 'It's the business model: How big tech's profit machine is distorting the public sphere and threatening democracy'.

[200] Hao, K. (2021) 'How Facebook and Google fund global misinformation', *MIT Technology Review*, 20 November.

[201] For an analysis of the online advertising ecosystem following the 2022 full-scale invasion of Ukraine, see Yeung, C., et al. (2023) 'Online advertising in Ukraine and Russia during the 2022 Russian invasion', in *Proceedings of the ACM Web Conference 2023*, New York: Association for Computing Machinery, pp. 2787–96.

[202] See, for example, Martin, L., Goujard, C., and Fuchs, H. (2023) 'Israel floods social media to shape opinion around the war', *Politico*, 17 October.

[203] See, for example, Human Rights Watch (2023) 'Meta's broken promises: Systemic censorship of Palestine content on Instagram and Facebook'.

[204] See UN Special Rapporteur on freedom of expression (2022) 'Disinformation and freedom of expression during armed conflicts', paras 90, 94.

[205] See ARTICLE 19 (2022) 'Content moderation and freedom of expression: Bridging the gap between social media and local civil society'; Internews (2023) 'Safety at stake: How to save Meta's trusted partner program', 2 August.

[206] This issue has been particularly problematic in the context of Meta's Dangerous Organizations and Individuals policy. See Oversight Board (2024) 'Sudan's Rapid Support Forces Video Captive', in which the board decided in reference to banning of the armed group Rapid Support Forces in Sudan that '[g]iven the situation in Sudan, where the RSF has de facto influence or control over parts of the country, civilians who rely on Facebook, including the RSF's communications channels, for critical security and humanitarian information, could be at greater risk through the restrictions placed on those communications channels'. Regarding the banning of Hamas, Business for Social Responsibility found in September 2022 that Palestinians were more likely to violate Meta's Dangerous Organizations and Individuals policy because of Hamas' presence as a governing entity in Gaza and political candidates' affiliations with designated organisations. See Business for Social Responsibility (2022) 'Human rights due diligence of Meta's impacts in Israel and Palestine in May 2021', September, p. 8. See also Human Rights Watch (2023) 'Meta's broken promises', pp. 29–33. See also ARTICLE 19 (2024) 'Content moderation and local stakeholders in Colombia', p. 37 (detailing the disappearance from social media of the voices of the Revolutionary Armed Forces of Colombia (FARC) guerillas despite being a party to the peace process in Colombia).

[207] For instance, following Russia's invasion of Ukraine, Meta and X (then Twitter) were reported to have employed significant resources to adjust their content moderation processes to the situation in Ukraine while such measures had not been adopted previously in other conflicts outside of Europe. See, for example, Bidle, S. (2022) 'Facebook's Ukraine-Russia moderation rules prompt cries of double standard', *The Intercept*, 13 April.

[208] For example, YouTube's monetisation policies (as applicable in November 2024), stated that 'in light of the war in Ukraine, we are pausing YouTube's monetization of Russian Federation state-funded media channels'. Its advertiser-friendly content guidelines (as applicable in November 2024) also provide that '[d]ue to the war in Ukraine, content that exploits, dismisses, or condones the war is ineligible for monetization until further notice'. The respective policies and guidelines do not reference any other armed conflicts. Meta has also reportedly prohibited advertising from all companies controlled by the Myanmar military after the 1 February 2021 coup (see *The Diplomat* (2021) 'Facebook bans all Myanmar military-linked accounts and ads', 25 February) and from Russian state media after the full-scale Ukraine invasion in 2022, but it does not appear to have taken comparable measures in response to other armed conflicts, including most recently in Palestine. For a direct comparison of steps adopted, see Meta (2022) 'Meta's ongoing efforts regarding Russia's invasion of Ukraine', 26 February; and Meta (2023) 'Meta's ongoing efforts regarding the Israel-Hamas war', 13 October. With respect to other armed conflicts active around the same time, Meta has not published equally detailed communications about steps taken, although some were mentioned in its annual Human Rights report. See Meta (2023) 'Human rights report: Insights and actions'.

[209] For example, the UN Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, Irene Khan, found that during the war in Gaza, 'censorship of content on Palestinian rights and views has increased significantly across platforms, including Meta, X, Google and Telegram'. See UN Special Rapporteur on freedom of expression (2024) 'Global threats to freedom of expression arising from the conflict in Gaza', A/79/319, 23 August, para 52.

[210] For example, Meta's 'Coordinating harm and promoting crime' policy (as applicable in November 2024) prohibits exposing the identity of prisoners of war and putting them at risk of harm.

[211] For example, TikTok's 'Violent and hateful organizations and individuals' policy (as applicable in November 2024) bans accounts from 'violent political organisations' which it defines as 'non-state actors that commit violent acts primarily against state actors (such as national military) rather than civilians, as part of ongoing political disputes (such as territorial claims)'. While references to targeting state actors rather than civilians seem inspired by IHL standards (a distinction is made with violent extremists, understood by TikTok as non-state groups 'that threaten or use violence against civilians for

political, religious, ethnic, or ideological reasons'), they do not properly align with the terminology and specific IHL rules.

[212] A few days after the 7 October 2023 attacks on Israel by Hamas, EU Commissioner Breton wrote to various social media companies demanding responses regarding the systems in place to prevent the spread of disinformation and illegal content on their respective platforms. See ARTICLE 19 (2023) 'Europe: Tackling content about Gaza and Israel must respect rule of law', 18 October; ARTICLE 19 (2024) 'EU: Call for precise interpretation of the Digital Services Act', 18 January.

[213] See UNDP (2022) 'Heightened human rights due diligence'.

[214] See UNDP (2022) 'Heightened human rights due diligence', p. 21.

[215] Depending on the size and resources of the company, internal expertise in IHL might not be feasible, in which case seeking specialist external advice will be key.

[216] See UN Special Rapporteur on freedom of expression (2018) '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, paras 50–52.

[217] See ARTICLE 19 (2024) 'Myanmar: Crackdown on freedom of expression with 24-hour monitoring', 1 April.

[218] See Rigot, A. (2022) 'Design from the margins: Centering the most marginalized and impacted in design processes – from ideation to production', Harvard Kennedy School Belfer Center for Science and International Affairs, pp. 44–46.

[219] On how to regulate content moderation while protecting freedom of expression (for example, through measures limiting the risks of overly personalised content on social media platforms), see ARTICLE 19 (2021) 'Watching the watchmen: Content moderation, governance and freedom of expression'. On how to tackle the excessive market power of social media giants, see ARTICLE 19 (2021) 'Taming big tech: A pro-competitive solution to protect free expression'.

@ info@article19.org

www.article19.org

@article19org

@article19

www.linkedin.com/company/article19/

facebook.com/article19org