

## **‘From the River to the Sea’: Protecting freedom of expression in public discourse and protest during the conflict in Israel and Palestine**

**In this briefing, ARTICLE 19 provides an overview and recommendations for states and social media companies on how international human rights standards should guide responses to the controversial slogan ‘From the River to the Sea, Palestine Will Be Free.’ The slogan, often used as a call for solidarity with Palestine, has been subject to intense public debate and restrictions in recent months. ARTICLE 19 is concerned with blunt efforts by governments, educational institutions and social media platforms to censor and even criminalise its use in a blanket manner. We remind all stakeholders that whether the use of the slogan can be restricted is not a simple ‘yes or no’ matter, as some would like to suggest. Any restrictions must be considered on a case-by-case basis and must comply with international freedom of expression standards. The same standards must be fully respected when restricting not only the use of this slogan but any expression surrounding the conflict in Israel and Palestine.**

The briefing is divided into four parts. First, we provide a brief overview of the meaning and origins of the slogan ‘From the river to the sea, Palestine will be free’ (the slogan) – a reference to the land between the Jordan River, which borders parts of the Palestinian Occupied Territory in the West Bank and eastern Israel, and the Mediterranean Sea to the west, including the Gaza Strip. Second, we offer some examples of how the use of the slogan has been restricted by states and private actors in recent months. Third, we set out relevant international freedom of expression standards that are applicable to restrictions on the use of the slogan. Fourth and finally, we offer ARTICLE 19’s position and recommendations to stakeholders on whether and under what circumstances the use of the slogan can be restricted and when it must be protected.

Our position on restricting the use of the slogan is informed by our extensive standard-setting policy work, including on issues surrounding controversial expression.\*

---

\* ARTICLE 19’s policy recommendations include [The Camden Principles on Freedom of Expression and Equality](#), [Prohibiting incitement to discrimination, hostility or violence](#) and [‘Hate Speech’ Explained: A Toolkit](#) (which provide resources on identifying and countering ‘hate speech;’ these recommendations have been incorporated into the Rabat Plan of Action, a set of guidelines on incitement to hatred, adopted by the UN Office of the High Commissioner for Human Rights in 2012); The [Johannesburg Principles on National Security, Freedom of Expression and Access to Information](#) (which outline recommendations on restrictions on freedom of expression on the basis of protecting national security); [The Right to Protest: Principles on the protection human rights in protest](#); and several policy briefs on social media companies’ regulation and on their content moderation, including [Side-stepping rights: Regulating speech by contract](#) (2018), [Watching the watchmen: Content moderation, governance, and freedom of expression](#) (2021) or the [Content moderation and freedom of expression handbook](#) (2023). We also co-developed The [Santa Clara Principles on Transparency and Accountability in Content Moderation](#) which elaborate minimum standards for transparency and accountability around social media companies’ enforcement of their own terms of service.

## Background: debate over the meaning and origins of the slogan

ARTICLE 19 observes that the meaning of the slogan has been the subject of significant discourse, including [academic scholarship and analysis](#), [legislative scrutiny](#), [political condemnations](#), [legal proceedings](#), and scores of journalistic headlines [noting 'debate'](#) over its meaning.

Prior to the brutal attack of Hamas on 7 October 2023, the slogan had a history of use as a call for equality and freedom for Palestinians; it has been understood as '[a rejoinder to the fragmentation of Palestinian land and people by Israeli occupation and discrimination](#)'. After 7 October attack and Israel's subsequent war on Gaza, the slogan has received additional attention for its use as a call for solidarity with Palestine by many people around the world. It is often chanted in protests calling for a ceasefire and condemning Israel's relentless and disproportionate attacks on Gaza.

Interpretations of the slogan, however, can have '[multiple meanings to different groups of people](#).' Some advocate that the slogan is a '[call to genocide](#)'; others find such a sentiment '[wildly mistaken](#)'. The slogan appears in the title of books as well, such as recently published [Essays for a Free Palestine: From the River to the Sea](#) by Verso Books, distributed in the United States by Penguin Random House.

The militant organization Hamas adopted a version of the slogan in its [2017 charter](#), leading to positions of the American Jewish Committee that it is a '[rallying cry for terrorist groups and their sympathizers](#)'. Some also claim the slogan uniformly '[carries a genocidal intent](#)', although the journal [Jewish Currents](#) and several [historians](#) disagree with this position. For instance, two academics of History and Jewish Studies (at the University of Massachusetts at Amherst) and Holocaust History (at the Hebrew University of Jerusalem) [argue](#) that the slogan is

[M]uch more complex. When interpreting it, it is important to be historically accurate and to avoid anachronism. The meaning of the slogan depends on the context in which it is used and, of course, on the intention of those who use it.

Beyond disagreement over the slogan's meaning, its history is subject to [debate](#) as well. Some scholars take [the position that the slogan emerged during the Oslo Accords](#) in the 1990s, while others date it back to the 1960s and the Palestinian nationalist movement, or simply state that '[more research needs to be done](#)'.

We also note that the political party Likud, the original platform of Benjamin Netanyahu, Israel Prime Minister, also published a version of the slogan as part of their 1977 election manifesto. The manifesto denied possibility of a two-state solution and called for only Israeli sovereignty '[between the Sea and the Jordan](#)'.

Irrespective of one's position on the meaning or origin of the slogan, the discourse surrounding it is indisputably a matter of significant public concern, impacting broad sectors of society. Such public concerns are increased given the conflict in Israel and Palestine, and during a significant year of elections worldwide.

## Restrictions on the use of the slogan

### *Restrictions imposed by states*

ARTICLE 19's research shows that since 7 October 2023, the slogan has been subject to limitations of many varieties.

Some courts and law enforcement agencies confirmed that the slogan can have different meaning (e.g. in the [Netherlands](#)) or refused to impose blanket bans on it (e.g. in [the UK](#), where the Metropolitan Police confirmed that mere chanting the slogan in protest was not a criminal offence). However, in other countries, the use of the slogan has been criminalised or otherwise sanctioned. A non-exhaustive list of some examples includes the following:

- Some governments or law enforcement authorities announced that a mere use of the slogan — in particular in protests — could be liable under criminal law for supporting terrorism, inciting hatred or violence, or denying, condoning or justifying genocide, without any additional considerations (e.g. in [Germany](#) or [the Czech Republic](#)).
- People have been prosecuted for chanting the slogan in protest under hate crime offences (e.g. in [Canada](#) or [Austria](#)), or for 'incitement to hatred' (e.g. in [Estonia](#)).
- Criminal sanctions have been applied for using the slogan on social media (e.g. in [Germany](#)).
- Administrative sanctions have also been imposed for using the slogan when organising protests on social media (see e.g. the ban of a protest in [Vienna in Austria](#) based on a social media invite).

At the same time, restrictions have been imposed on those who use the slogan in the context of committing hate crimes, such as defacing places of worship (e.g. a [synagogue](#) in Philadelphia, USA), making violent threats against Jewish people (e.g. at [Cornell University in the USA](#)), or in the wake of attacks on places of worship (e.g. [in Berlin in Germany](#)).

### *Restrictions imposed by private actors*

ARTICLE 19 also found that administrative or disciplinary sanctions have been imposed by educational institutions/university administrations or under the codes of conduct or internal rules of political parties. For instance:

- Students at various levels of education, including at public institutions, have faced suspension and bans for chanting the slogan (e.g. [Rockland Community College](#) and [Brandeis University](#) in the United States).
- Some parliaments and political parties have censored legislators or their party members for the same reasons (e.g. the [US House of Representatives](#) or [the Labour Party](#) in the UK).

**Social media companies**<sup>\*\*\*</sup> have also censored or taken a position on the use of the slogan on their platforms, while failing to provide comprehensive information explaining the volume of content removed or on what basis it is being taken down. For instance:

- In November 2023, Elon Musk, the owner of **X (formerly Twitter)**, [declared](#) that, in his view, the slogan ‘necessarily impl[ies] genocide’ and that as a result its use on the platform could ‘result in suspension’.
- In December 2023, according to an [investigation by Human Rights Watch](#), **Meta** removed hundreds of posts containing the slogan from Facebook and Instagram under ‘spam’ Community Guidelines or Standards.
- TikTok [has taken the position](#) that it does not allow using the slogan ‘in a way that threatens violence, and spreads hate’.
- On 7 May 2024, **Meta’s Oversight Board** [announced](#) that it was going to consider content decisions made by Meta in three cases involving user content containing the slogan on Facebook. In all three cases subject to current appeal, Meta declined to take down the content, as it ‘determined that, without additional context, it cannot conclude that ‘From the River to the Sea’ constitutes a call to violence or a call for exclusion of any particular group, nor that it is linked exclusively to support for Hamas’.

The Oversight Board stated that it ‘selected these cases to consider how Meta should moderate the use of the slogan given the resurgence in its use after October 7, 2023, and controversies around the slogan’s meaning’. This implies to consider additional criteria that Meta should apply when it ‘moderates’ the use of the slogan, rather than removing it automatically. At the same time, moderation of the content with the slogan on Meta’s platforms is concerning as any decision will likely be enforced by automated methods that are notoriously poor at ascertaining the context of expression (see more on this below).

Social media companies are not the only online platforms that restrict the use of the slogan. For instance, in November 2023, [it was reported](#) that Etsy, the US e-commerce company focused on selling handmade or vintage items and craft supplies, banned the slogan on sellers’ wares under [the company’s prohibited items policy](#), which includes a prohibition against ‘items that support, promote, or glorify violence’. The same reporting also pointed out that the items with the slogan ‘have been banned or heavily restricted on Amazon’ and on eBay.

---

<sup>\*\*</sup> In this briefing, the term ‘social media companies’ refers primarily to platforms with a dominant position in digital markets.

## Applicable international human rights standards

ARTICLE 19 recalls that international and regional human rights standards, including the International Covenant on Civil and Political Rights (ICCPR, Article 19), the African Charter on Human and Peoples' Rights (Article 9), the American Convention on Human Rights (Article 13), and the European Convention for the Protection of Human Rights and Fundamental Freedoms (Article 10), do not guarantee the right to freedom of expression in absolute terms. They do, however, stipulate that restrictions on the right to freedom of expression must be strictly and narrowly tailored and may not put in jeopardy the right itself. The determination whether a restriction is narrowly tailored is often articulated as a three-part test which stipulates that restrictions must be a) prescribed by law, b) pursue one or more of the explicitly enumerated legitimate aims, and c) be necessary and proportionate to the pursuit of such aim.

ARTICLE 19 notes that under these conditions, restrictions on the use of the slogan can be pursued on basis of various aims. We wish to highlight the following:

### ***Restrictions on 'hate speech'***

ARTICLE 19 has long maintained that 'hate speech,' often used to silence and intimidate minorities, is a serious human rights concern. While the term has no uniformly accepted definition, international human rights standards require restrictions on 'hate speech' to be based on the level of its severity. Namely, states must prohibit, including through criminal law:

- **Direct and public incitement to genocide**, based on the Convention on the Prevention and Punishment of the Crime of Genocide (1948) and the Rome Statute of the International Criminal Court (1998). As elaborated by *ad hoc* UN international criminal tribunals, convictions for direct and public incitement to genocide require proof of several key elements. These mandate that the expression must be a) public (in a public place, or communicated to the public), b) direct (sufficiently specific as a call to action with potential danger of an act of genocide occurring) and c) the speaker specifically intend to incite genocide and intent for genocide to occur).
- Any advocacy of hatred that constitutes **incitement to discrimination, hostility or violence**, pursuant to Article 20(2) of the ICCPR. When determining whether the expression reaches this threshold, the [UN Rabat Plan of Action](#), based on ARTICLE 19's recommendations, mandates that all incitement cases should be assessed under a uniform six-part incitement test that looks at 1) context of the expression, 2) type of speaker/proponent of the expression, 3) intent of the speaker/proponent of the expression to incite to discrimination, hostility or violence, 4) content of the expression, 5) extent and magnitude of the expression (including its public nature, its audience and means of dissemination), and 6) likelihood of the advocated action occurring, including its imminence.

States may also prohibit other types of 'hate speech,' but the restrictions must always meet the above mentioned the three-part test. Importantly, expression that is inflammatory, shocking or

offensive, but does not meet any of the thresholds described above, must be protected.

Moreover, while international human rights law requires States to prohibit the most severe forms of hate speech, this is rarely effectively achieved by censorship. States should also undertake complex legal and policy measures to tackle the underlying prejudices that 'hate speech' is symptomatic of, and address discrimination of groups at risk of it.

### ***National security and public order restrictions***

Restrictions of freedom of expression based on protecting **national security** must comply with the requirements of the three-part test. International and regional human rights bodies have consistently found that when reasons related to national security are invoked, statements of a general nature do not suffice. Instead, states 'must demonstrate in specific fashion the precise nature of the threat' to such national security caused by expression, as well as why measures taken to restrict the right to freedom of expression were necessary.

Additionally, in the Johannesburg Principles, ARTICLE 19 recommended that restrictions on freedom of expression in the name of national security should be prohibited unless: a) the expression is intended to incite imminent violence; b) it is likely to incite such violence; and c) there is a direct and immediate connection between the expression and the likelihood or occurrence of such violence.

ARTICLE 19 also notes that many restrictions on national security rely on vaguely defined 'terrorism' provisions, including incitement to terrorism. While there is no universally agreed definition of terrorism under [international law](#), [international](#) and [regional](#) human rights bodies have called on states to ensure that counter-terrorism measures comply with international human rights standards. They highlighted that offences – such as 'encouragement of terrorism' or 'extremist activity' – and offences of 'praising', 'glorifying', or 'justifying' terrorism – should be clearly defined to ensure that they do not lead to unnecessary or disproportionate interference with freedom of expression.

Similarly, the threshold for prohibiting expression on the basis of protection of **public order** or public safety is high and must be evidence based, rather than premised on speculations. In the context of protests, the potential for public order disturbances should not be the basis for denying the right to protest. [Any prior restraint of a blanket nature](#), especially when based on the content of expression, are almost always illegitimate.

### ***Responsibilities of private actors, including social media companies***

While states primarily hold duties under international law, the [Guiding Principles on Business and Human Rights: Implementing the UN 'Protect, Respect and Remedy' Framework](#) (the Guiding Principles) recognise that private companies have a responsibility to respect human rights, independent of state obligations or the implementation of those obligations. These

responsibilities include the obligation to conduct due diligence and human rights impact assessments to identify, prevent, and mitigate any potential negative human rights impacts of their operations. They should also incorporate human rights safeguards by design to mitigate adverse impacts, and act collectively to increase their leverage to address such impacts vis-à-vis government authorities.

When it comes to social media companies, the [UN Special Rapporteur on Freedom of Expression recommended](#) that they should, among other things:

- Establish clear and unambiguous terms of service/community guidelines in line with international human rights norms and principles and remain consistent in their application;
- Produce transparency reports about government requests to moderate content and the actions taken in that regard, as well as transparency reports regarding the scope of moderation and decisions made under their own terms of service/community guidelines;
- Ensure that sanctions for non-compliance with their terms of service are proportionate, and
- Provide effective remedies for affected users in case of violations.

These responsibilities are particularly important as social media companies often span multiple jurisdictions, some of which may feature fewer protections for freedom of expression online.

### **ARTICLE 19's position on the restrictions on the use of the slogan**

At the outset, ARTICLE 19 wishes to stress out that we do not take a position as an organisation on what the slogan 'From the River to the Sea' means or what its origin is. As we have already observed, there is significant academic and other discourse on this topic, and we do not think it is our role to engage in a substantive debate about it with subject matter experts, including historians and other commentators.

Regardless, we highlight that it ultimately does not matter what we as an organization believe to be the meaning of the slogan. Under international human rights standards, to impose restrictions on a particular usage of the slogan, the restrictions must be analysed according to a variety of criteria on a case-by-case basis. Such analysis is not subject or dependent on the views of ARTICLE 19 or others on the meaning or the origin of the slogan.

In the previous sections of this briefing, we already showed that a variety of claims are being made, arguing that the slogan constitutes hate speech or hate crime, promotes or incites genocide, or supports 'terrorism'. In the following analysis, we outline the legal criteria that would need to be met in each of these cases to reach the necessary threshold for restrictions and what is an appropriate balance between the right to freedom of expression and other aims on the basis of which the restrictions are imposed.

### ***Can ‘From the River to the Sea’ constitute incitement to genocide?***

**The use of the slogan should not be considered incitement to genocide unless very specific circumstances, including specific intent to incite genocide, are met while the severity threshold is very high.**

Some commentators claim that the slogan is a [general call to genocide](#) (see e.g. the [scrutiny](#) in the Czech Republic, or [declarations](#) of Elon Musk). ARTICLE 19 finds such declarations to be serious, and accordingly invite a serious analysis of the scope of when the international legal prohibition on incitement to genocide would apply.

We recall that international law sets a *very high bar* for establishing incitement to genocide. Convictions for direct and public incitement to genocide necessarily require examining the context and intent of the speaker. Therefore, it is not enough (and would not satisfy the required elements justifying or requiring criminalisation or sanctions) to assess a set of words alone. People have intent; words do not.

For this reason, there is a particular danger in calls such as the blanket assertion by the US House of Representatives to [censure](#) a Palestinian-American lawmaker on grounds that the slogan is ‘clearly entailing Israel’s destruction and denial of its fundamental right to exist’; such a conclusion, again, incorrectly attributes intent to mere words. The claim made by the Czech Ministry of Interior that speech should be restricted where it is ‘denying, questioning, condoning or justifying genocide’ also does not, on its face, describe ‘direct and public incitement to genocide’.

To the extent that the comments reference ‘denying’ of genocide, [ARTICLE 19 reiterates](#) that ‘memory laws’ prohibiting denial of historical events, or similar prohibitions, are not necessary in a democratic society and are typically counterproductive.

Thus, on a case-by-case basis, a communication of the slogan ‘From the River to the Sea’ would only meet this high standard if it can be proven that specific intent to incite genocide and for it to occur exists, in addition to awareness of a ‘substantial likelihood’ for it occurring as a result of the speech. This high bar for incitement to genocide would be difficult if not nearly impossible to establish where individual protestors, for instance, are merely chanting the slogan while attending an event or in protest.

### ***Can ‘From the River to the Sea’ constitute incitement to discrimination, hostility, or violence?***

**The use of the slogan should not be considered incitement to discrimination, hostility or violence unless very specific circumstances are met while all restrictions must be assessed under the six-part test of the Rabat Plan of Action.**

When imposing criminal sanctions for using the slogan, some states have insinuated that it constitutes or incites ‘hate’ (e.g. [Canada](#) – ‘hate motivation’, [Estonia](#) – ‘incitement to hatred’, or [Austria](#) – ‘hatred crime against Israeli regime’).

ARTICLE 19 recalls that although states can restrict freedom of expression that amounts to incitement to discrimination, hostility, or violence (as per Article 20 para 2 of the ICCPR), all incitement cases should be assessed under a uniform six-part incitement test highlighted earlier ([the six-part test of the Rabat Plan of Action](#)).

As we pointed out in our analysis on the scope of incitement to genocide (see above), this test cannot be satisfied by looking at words alone, but necessarily requires an analysis of several other factors. Again, without being able to weigh factors – including the context of the particular expression, the intent of the speakers, or the likelihood of violence, discrimination or hostility against the targeted group occurring – there is no way to possibly ‘measure’ incitement.

For example, if law enforcement authorities want to sanction the chanting of the slogan in protest, they would have to assess whether the chants were likely to lead to violence against Jewish people in a particular situation and the likelihood of such violence was imminent. The subjective feelings of Jewish or other persons offended or insulted by the chants should not be pertinent in the assessment. The focus of the restrictions should not be to protect the feelings; instead, restrictions should aim at prevent specific likelihood violence that may be incited by the expression.

On the other hand, the incitement standard would be met if there was credible proof of intent to incite violence, discrimination, or hostility in a particular context, and likelihood of such prohibited action occurring (e.g. if the slogan was shouted directly at a specific group of people and caused an immediate and credible risk of violence targeted at them). This case would be clearly distinguished from the use of slogan as a general call for solidarity with Palestine, which would not qualify as incitement.

Last but not least, even where the high threshold for incitement is reached, international standards mandate that criminal sanctions should only be used as a last resort measure, considering the principles of necessity and proportionality.

### ***Can ‘From the River to the Sea’ constitute a ‘hate crime’?***

**Under very specific circumstances, the use of the slogan could be considered as evidence of a bias motivation of the base criminal offence.**

ARTICLE 19 notes that under international and regional standards, states have positive obligations to investigate, prosecute and remedy the potential crimes motivated by racial, ethnic, religious or other hatred. However, we also observe that ‘hate speech’ is often confused with ‘hate crime.’

[The term ‘hate crime’](#) refers to the commission of a criminal offence (‘base’ criminal offence) in which victims are targeted on the basis of a protected characteristic (‘bias motivation’). This means that the particular expression is not the ‘base criminal offence’ and basis for the restriction; it can be an indicator of ‘bias motivation’ of such offence.

For example, in cases of racially or religiously motivated murder or physical assault of Jewish or Muslim individuals, murder or physical assault are 'base' criminal offense; while 'bias motivation' would exist if the perpetrator chose the target on the basis of real or perceived racial or religious characteristics. If the perpetrator uttered hateful comments immediately prior to, during or after committing the crime, this can be used as evidence of their motivation. The speech aspect of the crime is then treated as an aggravated circumstance of the crime and attracts a higher penalty for the 'base' criminal offence.

In the case of the slogan, ARTICLE 19 highlights that its mere utterance cannot be automatically considered a hate crime. Only when it is used in *conjunction* with a crime targeting an individual based on a protected characteristic can it be considered evidence of a bias motivation for the underlying offence. Examples of this might include instances where the use of the slogan is accompanied by direct physical violence against individuals, or specifically causes damage to places of religious worship.

### ***Can 'From the River to the Sea' constitute 'support for terrorism'?***

**The use of the slogan should not be restricted on basis of protecting national security and public order unless restrictions meet the requirements of international freedom of expression standards.**

ARTICLE 19 is aware that there have been efforts to portray the use of the slogan as a general expression of 'support or promoting terrorism' (e.g. in the [Czech Republic](#)).

As a general matter, we recall that under international freedom of expression standards, any restrictions must satisfy the tripartite test of legality and necessity and proportionality to the legitimate aim. We find that the examples of restrictions on the use of slogan on the grounds of 'support for terrorism' mentioned earlier would not meet these requirements. They especially do not meet the requirements of necessity and proportionality, as it is unclear why banning words is either necessary to achieving a legitimate aim, or a proportional method for doing so.

Even where some speakers have claimed that the slogan is aligned with militant organizations, such as Hamas, these interpretations are disputed in public discourse, including by historians, academics, and others. It is therefore unclear whether the mere utterance of the slogan is intended to express support for militant organizations (or has anything to do with them at all).

As a result, the criminalisation of usage of the slogan under terrorism offences without considering all additional factors outline above will fail to be justified as a legitimate restriction under international law.

### ***Can 'From the River to the Sea' be restricted in protest?***

**The use of the slogan should not be restricted in protest unless restrictions meet the requirements of international freedom of expression standards**

ARTICLE 19 notes that considerable attention has been paid to the use of the slogan in the context of protests surrounding the conflict in Israel and Palestine, with the slogan used as a common rallying cry.

We recall that under international and regional human rights standards, States have positive obligations to protect the rights engaged in protest, including the right to freedom of expression. All limitations on expression in protest must meet the three-part test and other requirements set forth above. They should also protect other rights engaged in protest and ensure that law enforcement authorities follow recommendations of international and regional bodies and [best practices](#) in this area, including those regarding surveillance of protesters (i.e., use of the slogan by itself should not be grounds for engaging in surveillance).

As a result, chanting the slogan alone at a rally or its depictions on protest banners would not meet the threshold for legitimate restriction, and for prosecuting or otherwise sanctioning protesters. The instances where the restrictions on the use of the slogan could be justified would need to involve a specific situation where the slogan is connected to actions beyond speech, such as situations where it would be used in the context of incitement to violence or violent threats against counter-protesters.

### ***How should states respond to the use of the slogan 'From the River to the Sea'?***

In summary, ARTICLE 19 recommends that States should:

- Promote an environment where opinions and ideas, even controversial ones, can be shared and debated, particularly where they involve matters of significant public concern.
- Instead of resorting to censorship, adopt a range of positive measures to promote equality, intercultural dialogue, engagement, understanding, and tolerance. They should engage in broad efforts to combat negative stereotypes of, and discrimination against, Jewish, Muslim and other individuals and groups.
- Refrain from implementing blanket bans on the slogan 'From the River to the Sea' or any other slogans as such, since blanket prohibitions do not meet requirements for restrictions on freedom of expression under international human rights law. They are also typically counter-productive to the aim of promoting equality in the society and fail to address the underlying social roots of prejudice and intolerance.

- When deciding whether to limit use of the slogan in specific circumstances, make sure that all restrictions meet the three-part test. This means they must have a basis in law, which is publicly available and accessible, and be formulated with sufficient precision to enable individuals to regulate their conduct accordingly. They must pursue one of the legitimate aims explicitly enumerated in international human rights treaties and be necessary and proportionate to the pursuit of such aim. Assessments must be made on a case-by-case basis that seriously considers all the relevant factors.
- Ensure that hate crimes against Muslim and Jewish populations are subject to independent, speedy, and effective investigations and prosecutions.

### ***Should social media companies ban or restrict ‘From the River to the Sea’?***

Before setting out recommendations, ARTICLE 19 would like to highlight the following issues raised by the response to the slogan by social media companies:

- **Failure to consider freedom of expression standards in content moderation:** Community standards of social media platforms typically have lower standards for restrictions on freedom of expression than those permitted under international human rights law. We acknowledge that social media companies are in principle free to restrict content on the basis of freedom of contract, but they should still respect human rights, including the rights to freedom of expression, in line with the Guiding Principles. Blanket banning of the slogan ‘From the River to the Sea’ would clearly fall short of the responsibility to respect human rights.
- **Use of automation in content moderation:** We are aware that social media companies rely on automated systems/algorithms in their content removal processes or for fast-tracking reports of violations of their community standards. As we emphasised in the preceding sections on the obligations of state actors, the analysis of when expression falls under certain category that can be restricted is highly context- and case-specific. For example, concluding that expression should be prohibited as incitement to violence depends on numerous factors including intent, nature of the speaker, and likelihood of actions occurring. These may also be [dependent](#) on linguistic, political, social, cultural, and economic dimensions. We note that these factors are precisely the aspects that automated systems *struggle the most with*. Relying only on automated systems to remove the slogan, without highly context-dependent analysis, will inevitably lead to over-censorship of content on matters of public concern and ongoing armed conflict.
- **Lack of understanding of context, language and culture:** We have previously warned that social media companies should take steps to address and eliminate biases toward certain languages and cultures. The lack of such understanding presents particular risk in the case of the slogan (that originates in Arabic and from the MENA region, and thus heavily implicates non-English speakers). Lack of moderators who understand different contexts and languages

and are equipped with sensitivities to the communities and demographics most impacted by the slogan, might lead to over-removals.

- **Lack of transparency and internal remedies:** Users of social media may often not even know the reason for removal of content, including the slogan 'From the River to the Sea', from the platform. There may also be a lack of clarity about how community guidelines are being applied.

ARTICLE 19 reiterates that social media platforms have responsibilities to respect human rights. These responsibilities are amplified in instances where social media companies serve as forums for matters of significant public concern. Their content moderation policies and practices applicable to use of the slogan on their platforms must be based on multiple complex considerations from a rights-centred approach. In line with existing ARTICLE 19's policies, we recommend in particular the following:

- As a general matter, social media companies should reflect and closely align their internal policies (community guidelines and terms of service) with international human rights standards on freedom of expression. The internal policies should be articulated with sufficient precision. They should provide specific examples addressing the way in which their standards are applied in practice, such as through case studies. This should be accompanied by guidance as to the factors that are taken into account in deciding whether or not content should be restricted.
- Social media companies should refrain from imposing blanket bans on the slogan on their platforms as these will be generally inconsistent with their human rights responsibilities. Removal of content containing the slogan should only happen when it violates a specific aspect of community guidelines/policy (e.g. on harassment or hate speech).
- When considering whether a particular piece of content with the slogan violates some part of the community guidelines, the context assessment should be done specifically to incorporate input from members of impacted communities and with cultural sensitivity. In line with the [Santa Clara Principles](#), moderation decisions should be made by those with sufficient awareness of any relevant context and familiarity with the relevant languages (including Arabic and Hebrew) and regions where the slogan is being used.
- Although we are aware that social media companies will rely on automated tools in their content moderation decision, including for the use of the slogan, they should at the very least ensure human review of content flagged through these mechanisms.
- Social media companies should also conduct human rights impact assessments of their automated content management systems, and in particular the extent to which they are likely to lead to over-removal of (legitimate) content containing the slogan.

- Social media companies should put in place effective complaints mechanisms to remedy the wrongful removal of content, including the content containing the slogan, or other disproportionate restrictions on the exercise of their users' right to freedom of expression (such as account blocking). In doing so, they should respect basic due process rights.
- Social media companies should improve their transparency reporting on content moderation practices. They should always specify whenever they remove content – including the content containing the slogan – on the basis of their Terms of Service/community guidelines at the request of governments or 'trusted' third parties, such as local NGOs or associations. They should also provide information about the number of complaints they receive about alleged wrongful removals of content, including the content that contains the slogan, and the outcome of such complaints (i.e. whether content was restored or not).