

## **PRIVATE SOCIAL MEDIA COMPANIES, GOVERNMENTS AND GLOBAL POPULATION: CHALLENGES AND PROSPECTS FOR THE IMPLEMENTATION OF HUMAN RIGHTS REGULATIONS**

**Malik Zia ud Din\***

Lecturer, Department of Law, The Islamia University of Bahawalpur, Punjab, Pakistan  
[malikziaudin@yahoo.com](mailto:malikziaudin@yahoo.com)

**Awais H. Gillani**

Assistant Professor, Department of Media Studies,  
The Islamia University of Bahawalpur, Punjab, Pakistan.

**Gulala Pirzada**

M. Phil. Scholar, Department of Media Studies,  
The Islamia University of Bahawalpur, Punjab, Pakistan.

### **ABSTRACT**

*Private social media corporations control speech much more stringently than any government, yet their platforms are being exploited to do serious damage to human rights. These media companies are governed mainly by themselves in top secrecy. Human rights' proponents have recommended media businesses should abide by international human rights legislation to address this. ICCPR statutes are the most well-known set of regulations for regulating speech globally. ICCPR Articles 19 and 20 might strengthen rules while allowing for more openness and supervision as a representative of billions of social media users. According to these articles, the legislation must first be construed to establish how (and if) each of its provisions act(s) suits the new aims. The law, for example, stipulates that speech may be limited to protect national security, which is one of only five legitimate justifications for speech restrictions. Governments that follow international law can make judgments based on this, but private media firms cannot because national security is not in their domain. This study analyzes and explains the main significant articles and laws related to human rights for consideration by social media corporations - ICCPR Articles 19 and 20, which cover freedom of speech – right of information to fill in some gaps.*

**Keywords:** Social media, Freedom of speech, information, media corporations, Human rights law, ICCPR, Articles 19 and 20.

### **INTRODUCTION**

Facebook, Inc. is a social media company based in California. It is the world's most secretive censoring system, which monitors more human interactions than any other government: billions of postings every day (Qin, et al., 2017). Platforms such as YouTube, other social media companies restrict speech on a massive scale according to their rules and regulations (Flew, et al., 2019).

Almost every post on the network begins as a private message, in which users exchange personal information and ideas with individuals they know offline or online. On social media, strangers also have crucial public dialogues. Many nations utilize the internet to conduct political campaigns, announce presidential decisions, offer government services, and convince populations to accept mass murder and ethnic cleansing (Tufekci, 2018).

While most of the material on the internet is benign or beneficial, social media platforms are sometimes used to transmit a range of content that hurts people. The relationship amongst media firms' conduct and the misery isn't evident, and remedy isn't as straightforward as it is in the case of items that kill people directly, such as a Ford Pinto gas tank explosion or HIV-infected clotting factors for sale to haemophiliacs (McHenry, & Khoshnood, 2014). The risk, on the other hand, is substantial.

---

\*Corresponding Author

Users might get exposed to harmful material that leads to offline damage when firms pursue income by increasing user commitment and retention (Trautman, 2019). Nonetheless, firms choose which material to govern and regulate by eliminating it, displaying warnings, verifying facts, or making it accessible to fewer individuals using algorithms millions of times per day (Bakir, & McStay, 2018). Private social media corporations decide which accounts they choose to suspend or prohibit entirely, and which ones they choose to preserve so that their owners may continue to express their right to free speech, even if it infringes on others' rights.

Companies make their own rules for all of these restrictions, often altering them proactively or in reaction to public pressure and keeping the contents of the regulations hidden from those who are affected (Jimenez, & Pulos, 2016). Employees may subsequently confess that they made a mistake by leaving or removing material. To govern more efficiently in recent years, social media organizations have employed tens of thousands of more moderators, created tools to detect hazardous material, and continuously changed their policies or made exceptions (Suzor, 2019).

Each social media platform has its own set of guidelines—such as the Facebook Community Standards—which many users are unfamiliar with. Users are less inclined to obey the rules as a result of this. Even some firm employees find them challenging to comprehend or implement. As a result, a third-party standard-setting organization is required. Users are confused and perplexed by the firms' own regulations and procedures; not only is enforcement uneven (more so in certain languages and nations than others), but it is also susceptible to government intervention. They have "produced unstable, unpredictable, and hazardous situations for users and increased government monitoring," as explained by David Kaye (Kaye, 2017).

Instead, holding to a unified set of principles, articles 19 and 20 will offer enterprises with a basis of "forceful normative reactions to unwarranted State restrictions" (Kaye, 2017) and a foundation for users to hold them responsible for adhering to fundamental customs and values. Companies will be forced to give more information about laws and enforcement if they use human rights legislation to edit material, as well as enable more effective remedies for individuals who fear companies are abusing their rights, according to Section 19—a worldwide human rights advocacy organization, and champion of freedom of expression, argues (see, ARTICLE 19, 2018).

Companies using social media require a clear, authoritative, and generally recognized source of rules and guidelines on how to govern, comprising revealing and amplifying such laws and regulations to the public. International human rights legislation is the most well recognized and legally approved source of rules for speech restriction. The "International Covenant on Civil and Political Rights (ICCPR)" is the most important legal document, with approximately 90 percent of the world's nations have signed it (Elkins et al., 2013). It should be emphasized that this does not imply that they are abiding by the treaty's provisions. The US, for example, adopts a wall of lowercase letters to restrict its compliance: five concerns, five understandings, and four declarations (Tarr, 2018).

"UN Guiding Principles on Business and Human Rights," have been created under the "Global norm of anticipated behavior have been created," hence, businesses are required to follow human rights legislation (without exceptions) (UNHCHR, 2012). Kaye noted that the "*UN Human Rights Council*" that firms like Facebook must give respect and comply with the laws because of their significant power over communications, particularly public discourse: "While the Guiding Principles are non-binding, the firms' massive impact on public life across the world strongly supports their ratification and implementation" (Kaye, 2017), ARTICLE 19 has also made a strong case for this.

Companies do not have to use the same regulations, and neither do the nations have to approve similar nationwide legislation. A varied range of individuals utilizes the majority of such platforms in several countries for various objectives that need the application of various laws. Fortunately, the ICCPR does not specify which speeches should be restricted, with just two exceptions. Instead, it establishes a procedure: a regulatory structure to ensure that the regulations are clear, geared to safeguard the relevant populations, and impose the fewest restrictions on freedom of speech possible.

On the other hand, the International Covenant on Civil and Political Rights and other related human rights laws cannot be used to change commercial material outside of the box; it must be construed in at least two ways for this unique purpose. First, the legislation was developed and approved to be used by governments rather than private corporations. Only five grounds are allowed under the ICCPR for speech restrictions: 1) public order and health, 2) nationwide security, 3) public order, 4) public health, morality, and the moralities, and 5) characters of any others. Profiteering Companies that

emphasize profit—for example, Facebook—are unlikely to prioritize the common good, according to Evelyn Marie Black (Aswad, 2018).

It is impractical to expect businesses to make choices solely in the public interest, but regulatory restrictions and regulations may compel businesses to make decisions to safeguard customers, employees, the environment, or the general public. Companies like Facebook and Twitter should be no different. However, as shown below, several sections of human rights legislation regarding speech do not apply to a private media company.

The second way in which international human rights law must be understood for use by social media businesses is that it puts limits on speech owing to its wide and complicated language. The exclusion of publicity for combat and incitement to discrimination, hospitality, or violence in the International Covenant on Civil and Political Rights (ICCPR) should be interpreted to apply to content modification, which necessitates clear and detailed guidelines for making constant resolutions across a wide range of content to be able to publish in over 100 languages and cultures (Fick, & Dave, 2019). Other types of communication, such as "any propagation of ideas based on racial superiority or hate," are also outlawed by another international human rights convention. It seems to be much more expansive than the rules of (ICCPR) for discourse, and the two agreements must be harmonized openly.

Companies may use (International Human Rights Legislation) IHL as a blueprint to manage speaking if external norms are correctly understood; if they are not, outsiders may hold them accountable. Because the legislation mandates that speech limitations be required, legal, and established by law, this should strengthen the firms' policies. Users would have a better understanding of the regulations and would be able to hold firms accountable for mistakes.

Companies should really follow, not just respect IHL. To that aim, we provide some unique legal interpretations for private firms and outside groups that are starting to play a role in corporate speech control, for instance "Facebook's new Oversight Board", that might be a tool for bolstering Facebook's speech control. Corporations have formed what Evelyn Dweck refers to as "content cartels," (Douek, 2020), such as the Global Forum on Online Terrorism (GIFCT), which is a coalition of Internet companies entrusted with banning databases of material, as Evelyn Dweck points out.

## **1. Social Media Companies need to comply with the IHL**

### ***i. Companies that operate on social media are carrying out certain limited state functions.***

While IHL is intended to require the governments to provide human rights as enshrined in international laws and corporate social responsibility, it's a new concept. Multinational corporations have become the major violators of human rights; such violations have led the UN to recommend the corporations to comply with the human rights law in its real spirit (Arnold, 2010). It has yet to be implemented. Instead, the "UN Secretary-General nominated Professor John Ruggie," the author of the "*UN Guiding Principles on Human Rights in Business*," which, after six years of discussion, compels firms to "respect" human rights (Bellace, 2014). In 2011, the UN Human Rights Council unanimously adopted UN Guiding Principles, which have since become the most significant international instruments in the fields of business and human rights (Ruggie, 2020).

The UNGPs place a greater emphasis on process than results, allowing for a wide range of degrees of commitment. Companies should adopt policies that fulfill human rights. Companies should "avoid infringing on the human rights of others" and "address adverse human-rights repercussions with which they are associated" (see, Huijstee, & Ceresna-Chaturvedi, 2012). Social media firms have strong influence around half of the population of the world, therefore, they might influence to these people in case of seeing, reading and writing. These powers are reserved for the government when they are exercised in public. Companies employ it in a virtual realm with numerous public activities, dubbed the "electronic public domain" (Benkler, et al., 2015).

This exceptional, international power and influence set the social media companies apart from any other private firm. When platforms are utilized to exchange critical information for civic life, the platform's proprietors and personnel have an impact on whole societies' political, cultural, and economic development. The usage of social media platforms has become a forum where the public can discuss various issues of politics, livelihood, and law and order (Hanitzsch, & Vos, 2018). President Donald Trump of the United States is far from the only politician who uses social media to communicate and campaign.

Routine governance duties are increasingly being accomplished on several social media platforms. In many nations, politicians and government entities interact and deliver services through

social media. In a working paper, the "Organization for Economic Co-operation and Development (OECD)" said, "Presence and engagement on social media is no longer an issue of choice for most governments" (Mickoleit, 2014). According to the report, the executive branches of the majority of the OECD nations and numerous organizations and ministries have Facebook and Twitter profiles (Mickoleit, 2014). Social media businesses have obtained more regulatory authority than any other private company that governments normally have since establishing human rights laws. Despite the caution and aspirational nature of the UNGPs, businesses should build solid, thorough, and transparent methods for respecting and complying with international human rights principles.

Many observers have witnessed that internet businesses rule without real responsibility to the individuals they control. 'Rebecca MacKinnon' dubbed the companies "sovereigns functioning without the permission of the networked" in a seminal 2012 book (MacKinnon, 2012). In certain circumstances, business leaders have acknowledged exercising sovereign rights. "In many respects, Facebook is more like a government than a regular firm," Mark Zuckerberg said in an interview with Kate Klonick in 2018, which she titled *The New Governors*. We have a vast community of individuals, and we're defining rules, more so than other technological corporations" (Chesbrough, 2012).

Facebook and other firms, according to Klonick, are "private, self-regulating enterprises" that have built extensive regulations and methods for censoring virtual material through their "governance systems" (Medzini, 2021). "A few profoundly engaged stakeholders control enormous cultural influence, and it is being done behind closed doors, making it impossible for anybody else to examine or oppose" (see, Gillespie, 2017). "The largest danger this private system of governance presents to democratic culture is the loss of a fair chance to participate, which is reinforced by the system's lack of direct responsibility to its users," says one expert (Balkin, 2017). Fortunately, human rights legislation demands that regulations be clear so that those who are subject to them may comprehend and dispute them. The application of the law would further provide additional advantages, for example, providing businesses with a firmer foundation to resist inappropriate official pressure to stifle expression.

Because social media platforms serve such a diverse range of purposes, each corporation probably may differently apply such human rights laws. Nothing is as big or influential as Facebook, and others, like dating applications or gaming sites, don't host much public debate. All private firms of any size or "operating environment" are required to comply with international human rights legislation under the UN Guidelines, however "the scale and complexity of how enterprises meet that responsibility may vary according to these factors and the severity of the enterprise's adverse human rights impacts" (UNHCHR, 2012). As a result, platform scope targeted population or intended purpose shouldn't be utilized to decide whether or not its owners should implement ICCPR content management guidelines.

When information uploaded on one or more of the company's platforms causes considerable damage, the company's duty to comply with human rights law grows proportionately. While certain applications, for example, are devoid of propaganda for war, the vast majority include information that breaches human rights. Users who have been banned from one site turn to smaller, specialist platforms such as LinkedIn, Tik-Tok, and even Pornhub to continue sharing dangerous information (see, Myers West, 2018). Platforms also change and evolve swiftly. No one could have predicted what Facebook would become when Mark Zuckerberg and others began it in their bedrooms. Although Amazon's Twitch streaming service is aimed at gamers, non-gaming content has become the most popular category (Ask, K. et al., 2019).

***ii. Private regulation has flaws that may be addressed by international human rights legislation.***

Having a unified fundamental basis for content control will aid in addressing some of the most pressing concerns with digital content management systems. In a nutshell, they are: 1) the regulations are vague and complicated, especially for the workers as they are entirely unaware of these regulations; 2) the company makes a lot of ruthless decisions that harm the people as well as their users; 3) corporate personnel's and employees impose their decisions that affect half of the people of the global population; 4) Governments may use subtle but significant pressure to impose shadow censorship.

***a. Users are unaware of or confused by a company's rules.***

Platform rules are enigmatic and mysterious to practically everyone outside of the firms who create and implement them. Thousands of tiny prints are repeated, such as the "*Facebook Community Standards*," which are divided into twenty-six parts. Most of such laws are irrelevant to the user, but users should read all six parts independently. To say the least, users are unlikely to do so. In a 2017 lab

experiment, all 543 college students claimed to have read the rules listed in Section 2.3.1 before clicking the "Join" button of a new social network (Berreby, 2017).

Even when individuals read the stated laws, they are just a schematic compared to the detailed instructions that corporate workers and contractors use to structure material (Hughes, & Ferrett, 2012). For example, although Facebook's public rules ban hatred against particular groups such as women and black people, its core regulations create exemptions for a "subset" of individuals such as women or black drivers (Wilson, & Land, 2020). Hate speech directed at specific groups is permitted on Facebook, but no one can deduce it from the community's norms. Consequently, most inquisitive and attentive users cannot comprehend the complex set of secret rules that control their online activities. Human rights legislation should help with this issue as it needs speech-restricting restrictions to be specific and understandable to individuals who are subject to them.

*b. Companies harm people by making poor decisions about whether or not to remove content.*

When it comes to setting and enforcing regulations, social media firms must traverse a panorama of real-world consequences that might arise from their activities and dormancy. The most serious of these repercussions is mass violence, with Facebook and other corporations being chastised in various countries for failing to delete posts that incite or promote it (Bhagwat, 2020). In March 2018, for example, demonstrators in Sri Lanka set fire to Muslim stores and places of worship (Abdul Razak. et al., 2021). A letter was disseminated by the "Center for Policy Alternatives" based in Colombo called consideration to Facebook remarks given through the disturbances, including the author's "demand" to murder all Muslims, including children, "because they are dogs" (Abdul Razak. et al., 2021). Facebook replied six days after receiving a message, saying it had not broken any corporate rules. Social media corporations also make mistakes by eliminating material that is significant for educational and human rights protection. Despite this, YouTube has taken down roughly 340,000 videos from its 1.7 million-strong library (Abdul Razak. et al., 2021). The regular requests procedure on YouTube hasn't worked successfully in reversing the removal, in part because only the individual who submitted the video may request that it be reinstated. This is impossible if the individual has been imprisoned or is in the process of being imprisoned.

In addition, censors often delete information that is intended to educate rather than express or incite hate. When YouTube removed a historical clip depicting the US Army burning Nazi emblems and suspended history teacher Scott Allsop's account for having a channel with archive Nazi material for his pupils, it invoked its hate speech policy (see, Kornbluh, Goodman, & Weiner, 2020). Following international human-rights legislation could reduce such errors since it stipulates that any limits on speech must be essential and reasonable, i.e., the least restrictive practicable. When mistakes happen again, consumers will have a new reason to question content moderation standards and how they are implemented. The measure should put a stop to failure to remove hazardous information by prohibiting hate speech, which includes encouragement to hostility, enmity, discrimination, violence, and war propaganda. The following sections go through these clauses in further detail.

*c. Companies impose their own set of values.*

The social media business companies make their users follow the regulations and standards on the usage of internet, as well as the ideas that support them are the result of impromptu commitment by a tiny number of US attorneys, rather than a systematic process that takes rules and diversity into consideration (Rainie, & Anderson, 2017). While businesses have made significant modifications to their policies over time, they depend on ambiguous and overwhelmingly American conceptions. The leaders of certain companies have understood that this is unjust. "What I'd really want to do is find a way to have our rules established in a manner that reflects the values of the community, so I'm not the one making those choices," Mark Zuckerberg stated in 2018. I believe there is most definitely a superior method that I have yet to discover" (Coppex, 2020). Although Facebook eventually established an oversight group, the decision-making process has essentially stayed the same.

The five principles that drive organizational policy are: 1) sound, 2) authenticity, 3) security, 4) privacy, and 5) dignity (Royakkers, et al., 2018). Freedom of expression or speech is "superior," meaning it is only limited if one of the other four criteria is violated (Baker, 2010). In certain situations, information that could typically be banned is allowed to remain on Facebook if it is deemed "newsworthy" or else in the interest of the public by Facebook (Kadri, et al., 2019). According to Facebook, it uses a "holistic and thorough" procedure that "account[s] for international human rights norms," including "Article 19 of the ICCPR", to "evaluate[s] the public interest value of the piece of

speech against the potential of damage" (Allan, 2018). While this broad requirement seems comforting, including human rights norms in a thorough evaluation is not the same as putting them at the forefront of the settlement. In the end, Facebook decides on "risk of damage" and "public interest" on a global scale and applies its conclusions to billions of individuals (Pacheco-Vega, 2019).

*d. Censorship in the Shadows*

States may also employ company content moderation to perform quiet and unseen restrictions by pushing corporations to remove information through hidden nonpublic stations or government-run "Internet Referral Units" (IRUs). "Some States have pushed for social media firms and other hosts of user-generated information to monitor and remove content on their own initiative, rather than waiting for law-based requests from the government" (see, Kaye, 2017). Governments claim to operate under the authority of national law in certain situations but fail to follow legal mechanisms established to ensure legal procedure and sensible competition rights (Crawford, K., & Schultz, 2014). In other circumstances, they are a black-handed party-style negotiation to compel corporations to conform by pledging to disregard antitrust issues, raising the probability that intermediaries would be held accountable for user material, or overtly threatening to deny access to the platform (Klonick, 2017).

Human rights legislation might be used by businesses to fight such demands. It wouldn't work like magic, particularly on authoritarian regimes (even ones that have signed the ICCPR), but it may be helpful. Specific regimes are receptive to arguments based on international law, and are vulnerable to being disgraced if they disobey it (Goldmann, 2012). Furthermore, corporations like Facebook might utilize their immense influence to oppose unlawful government interference more vehemently, particularly in nations where political leaders depend heavily on their platforms. Blocking (or even slowing down) access to the internet in places like the Philippines or Myanmar, where Facebook is an essential social, economic, and political tool, will be impossible.

Finally, linking policy and enforcement into international human rights concepts will aid users in comprehending what these restrictions are and why they exist. It should thus be simpler to persuade businesses to amend the regulations and enforce them effectively and consistently to prevent damage.

**2. International Humanitarian Law the Businesses Companies should adhere to?**

The UNGP advises businesses to concentrate on "at least" two sources: the *"International Bill of Rights and the Declaration on Fundamental Principles and Rights at Work of the International Labor Organization."* Many treaties and declarations include international human rights legislation, but the UNGP advises businesses to concentrate "at least" on two of them: the *"International Bill of Rights and the International Labor Organization's Declaration on Fundamental Principles and Rights at Work"* (Backer, 2015). The *"Universal Declaration of Human Rights (UDHR)"*, the *"International Covenant on Civil and Political Rights (ICCPR)"*, and the *"International Covenant on Economic, Social, and Cultural Rights (ICESCR)"* are all essential weapons for limiting freedom of expression.

As stated in the Universal Declaration of Human Rights ' preamble, every person and "every institution of society...should attempt" to promote respect for and universal acceptance of human rights (UNGA, 1948). As emphasized by famous human rights expert Louis Henkin, this includes corporations. Article 19 of the UDHR establishes the rights to freedom of speech and opinion but no associated responsibilities (UNGA, 1948). The ICCPR is the most significant and precise framework for limits on freedom of speech in international law, this is the primary right that social media firms allow and limit. Article 20 of the agreement, which has been signed by 173 of the world's 195 nations, specifies that governments are only allowed to ban two categories of speech: war advocacy and advocacy that incites enmity, discrimination, or violence The diplomats who negotiated the ICCPR after WWII were all aware with the treaty, and it is still valid today, although in somewhat antiquated language. Nonetheless, they make up a small percentage of all the information that firms block under their own regulations.

Article 19 of the ICCPR contains guidelines on speech limitations that are ideally suited to regulating speech for social media firms in two ways. First, all restrictions must be "established by law," according to Article 19 (ICCPR, 1966). According to the UN Human Rights Committee, a norm may be deemed a law, which is responsible for interpreting the ICCPR. As a result, Facebook's "community standards" and the regulations of other corporations may qualify—as far as they are specific and easily understood by the general community. According to the Committee, a standard "must be stated with sufficient precision to allow a person to govern his or her behavior properly, and it must be made available to the public" (Lwin, 2020). This is no longer the case. Companies have been establishing

their internal standards with growing accuracy, and they have been releasing public versions of their rules, although these are increasingly generic and ambiguous.

The ICCPR's adaptability helps social media users. It outlines the grounds on which limits may be placed, rather than prescribing which types of material may be restricted, and it requires a clear and balanced regulating procedure. As a result, several online social media platforms may apply their requirements to a broad range of regulations.

The *"International Convention on the Elimination of All Forms of Racial Discrimination (ICERD)"* is straightly related to control virtual subject matter, and its wording contradicts the International Covenant on Civil and Political Rights, and calls for some repression, if not more, expressing concern. Article 4 of the ICERD requires states to "declare all dissemination of ideas based on racial superiority or hatred, incitement to racial discrimination, as well as all acts of violence or incitement to such acts against any race or group of persons of another color or ethnic origin as an offense punishable by law" (ICERD, 1965)

### **3. Articles 19 and 20 of the ICCPR Should Be Used by Businesses**

Articles 19 and 20 of the International Covenant on Civil and Political Rights provide minimum and maximum expressions, and they should be central to any business regulation system.

#### *a. Article 19*

Article 19 of the International Covenant on Civil and Political Rights guarantees the freedom to seek and receive any sort of information, independent of national boundaries or media outlets. 102 Section 19(3) allows speech limits if they are authorized by laws, are necessary (and reasonable), and are legal. The third of the three prongs state says that any limitation must protect at least one of the five interests listed above: 1) public order, 2) morality, 3) national security, 4) public health and 5) other people's rights (Duffy, 2014). Other than those five grounds, state parties may not prohibit speech. No limitation may also contravene any other article of the ICCPR, such as the prohibition on discrimination. The UN Human Rights Commission, UN Special Rapporteurs, and other UN agencies have clarified three parts of the Article 19(3) test. Here you'll find all of the necessary information, as well as a suggestion for a social media firm.

#### *i. "Provided by Law"*

The *"Human Rights Committee"* has explained that there are certain conditions which need to be fulfilled, social media companies' regulations need to be taken into account "provided by law" if they "made accessible to the public" and "formulated with sufficient precision to enable an individual to regulate his or her conduct accordingly" (Gajos, 2020). However, the regulations of most platforms clearly do not fulfill this requirement. In 2018, Twitter CEO Jack Dorsey admitted before the US Congress that there is "a whole lot of uncertainty" regarding the enforcement of the company's regulations, saying, "I think you would not be able to grasp them if you went to our rules today and sat down with a cup of coffee" (Bursztynsky & Feiner, 2021). Renowned researchers explained the ambiguity of Twitter's regulations and implementation of ICCPR's principles (see, Aswad, 2018). The perplexity has real-world consequences. The vagueness of hate speech and harassment policies has triggered complaints of inconsistent policy enforcement that penalizes minorities while reinforcing the status of dominant or powerful groups, as another reputed researcher has highlighted: "The vagueness of hate speech and harassment policies has triggered complaints of inconsistent policy enforcement that penalizes minorities while reinforcing the status of dominant or powerful groups" (Kaye, 2017). Outsiders should have access to specific guidelines in order to "allow a person to govern his or her actions correctly" (Klonick, 2017).

As regulations are utilized to interpret legislation, the moderators' guides provide guidelines for applying the standards to particular instances. Corporations are traditionally refrained from disclosing their policies, claiming that doing so will enable maneuverer to "game the system"—finding methods to stay just on the wrong side of a guideline or, more broadly, ways to publish hateful or damaging material while evading removal. To begin with, most laws or regulations define a boundary between banned and permissible action, in other words, the implementation of such laws and their acceptance. In case of such a situation, the stance must be changed. Secondly, the internet subscribers are explicitly discouraged from posting harmful material to avoid being removed might infer particularly the limit has been decided by the system with a variability of postings from different identities. Thirdly, ambiguous community standards allow businesses to create exceptions without admitting them. This

would be limited if the information was made public. In any instance, corporations should disseminate guidelines with appropriate information to enable subscribers to comprehend the borders are drawn in order to comply with Article 19(3).

ii. *"Necessary"*

To constitute a requirement under Section 19, as HRC have said, limits should be of significance, reasonable or desirable (Haggart, & Keller, 2021). It must also be necessary and reasonable, implying that the limitation must be the least invasive and proportional to accomplish the stated goal. This is particularly relevant for social media corporations, which have much more possibilities for limiting material than nations. Companies may, indeed, delete items or deactivate IDs. To safeguard the liberty of speech and avoid damage, it is necessary to analyze the efficacy of other Article 19-compliant measures. Of course, the best solution is to prevent harmful material from being uploaded in the first place. The corporations started the trial by sending messages to users requesting them to confirm their identities to change the code of conduct or suspending users, which is an unavoidable technique termed "time-out" (Zargar, Joshi, & Tipper, 2013). Instead of relocating materials or accounts altogether, companies adopt less restrictive tactics such as downgrading or making content inaccessible to a restricted percentage of individuals on the site. The amount of other accounts a user may access to share information has been limited in messaging applications like WhatsApp (Church, & Oliveira, 2013). Furthermore, the United Nations Human Rights Council has advised nations to experiment with less restrictive methods of combating religious hate, such as edification and interreligious discourse.

Lastly, firms must examine if the methods they have selected really accomplish the stated goal in order to pass the "necessity" test (s), "For example," she said, "if a corporation deletes postings or bans people from its platform, it must consider if this is assisting in the creation of communities that are resilient to radicalization, knowledgeable about online falsehoods, and tolerant" (Aswad, 2018). Similarly, a corporation must assess if such safeguards lead damaging communication to flourish on smaller platforms, and what effect this has on the legitimate goal. Companies will aggressively and openly exploit their vast data, including information of breaches, to comply with this regulation, which is a critical component. They'd most likely learn a lot about how to reduce hazardous stuff on the internet. They might gain more profit by sharing information to diminish the damage since they would be required to communicate their findings to show that they have met the need prong (Mundial, 2004).

iii. *That Is, "Legitimate" Meant to Protect One or More of the Following Five Interests.*

A researcher may question that this strategy might be used by businesses whose primary purpose is to generate profits. Companies in other sectors are required to minimize pollution, cease employing minors, produce safer automobiles, and more as a result of laws and public pressure. Social media firms, in my view, may and should be urged to protect the public interest. Furthermore, whether or not their intentions are altruistic, social media businesses are now proclaiming judgments in the public good, from limiting the propagation of misinformation and extremist material to minimizing personal online abuse. Companies that examine items to fulfill user and advertising expectations may conflate profit with the public good.

Companies may not make decisions based on all five "legitimate" reasons, including content limitations based on Article 19 of the ICCPR. National security should not be used to govern a company's activities since this would be a violation of national sovereignty. Companies must comply when the government orders them to block content for legitimate national security reasons; however, if the government uses it inappropriately to push such businesses to remove such subject matter on account take national security basis; they must consider national security in their decision-making process. Authorities have often attempted to limit freedom of speech on the basis of bogus claims to safeguard national security, as documented by the Human Rights Committee (Posner, 2014). Of course, companies must recognize that national laws are being followed correctly. The government's instruction is sometimes employing as uncertain disagreement for administering policies by restriction and prohibiting Internet access. The corporation must fight back. Because public order does not stop at national boundaries, it may be controlled under Article 19's "necessity" branch to prevent specific disruptions, such as mass violence.

Facebook and other businesses already operate on this foundation in terms of the next ground, morality. For example, nudity is notably prohibited on Facebook. It claims to apply the same laws to everyone globally, even though morality has different meanings in different cultures: some cultures



approve of showing nudity in public, while others restrict them. Companies having operations in several countries may adjust their standards to match existing norms to overcome discrepancies. As far as it complies with human rights laws, such corporations may pursue assistance from indigenous NGO committees and local advisers, as I have mentioned (Coumans, 2010). Online, people's rights and reputations are more likely to be affected than offline, whether via long-term problems like defamation and slander or novel evils like human flesh-hunting and involuntary pornography (Lemley, & Volokh, 2017).

Though national laws specify definite solutions, they are ineffective in the internet setting; social media company regulation may be far quicker and more efficient and critical. Rather than limiting the damage in real-time, legal remedies are often geared to hold a few culprits accountable after the harm has happened. When persons cross national boundaries, they run into jurisdictional issues. In many situations, they aren't meant to tackle the new problems at all. Nevertheless, if "rights and reputation" is mainly construed, it might further be utilized to cover any content moderation concern, necessitating a definition of its scope. For example, a person may be worried about the impact of online insults on their reputation, but interpreting Section 19 to prohibit all insults would be an overreach.

Eventually, business companies related to social media are required to secure and supervise public health by disseminating valuable information while also limiting the spread of harmful misinformation (Ghenai, & Mejova, 2018). As the COVID-19 epidemic has brutally shown, preserving public health while promoting misinformation meant to harm, it is difficult for multinational corporations to make them ideal candidates for social media regulation. Several businesses currently govern in this manner. When the Pinterest team learned their network was used to promote anti-vaccine misinformation, they decided to cease giving search results for relevant phrases; now, such searches only return material from public health organizations (Eckert, 2020). On Facebook, misinformation, prominent health claims, and adware with health claims are rated lower (Soomro, & Hussain, 2019). YouTube videos that advertise dangerous medicines and cures have been banned by Google. Companies look into disinformation and give instructional material from health experts (Gharib, 2021). To summarize, the public health foundation for speech control provides businesses with a huge chance to defend the community's interest.

*b. Article 20*

Article 20 outlines a couple of types of information needed to avoid at all costs while always adhering to the conditions of Article 19. I provide some clarifications below since war mongers encourage bigotry, hate, or viciousness are not obvious enough for corporate governance. When the International Convention on Civil and Political Rights (ICCPR) was being negotiated in the late 1940s, propaganda for war was any of the couple categories of expression that the treaty effectively prohibited. Since then, it has received little consideration in the UN proceedings and IHL, possibly because governments are the source of much pro-war propaganda (Munro, 2014).

To begin with, the word "war" applies only to international law-enforced aggression conflicts; it does not relate to self-determination or the right to self-determination and independence, and it does not refer to civil wars (Chadwick, 2020). The phrase "propaganda" rather than "incitement" shows that the ICCPR's drafters intended to ban a broader category of material than simply advocating for war or encouraging a populace to support it (Ronen, 2010). They realized it takes time and repeated messaging to convert one set of people fiercely against another. They "thought that a clause that was restricted to outlawing incitement to war would have little chance of creating a permanent peace and averting future hostilities," according to Kearney, after much deliberation (Kearney, & Kearney, 2007). Instead, they opted to outlaw "the persistent and relentless presentation of an opinion with the intent of establishing a climate of hostility and lack of understanding between the peoples of two or more nations to lead them to armed confrontation" (Farris, & Coleman, 2020). A social media corporation may need bravery to limit or re-move information based on Article 20(1), but the potential reward should be considered for the good act on the part of the company. Companies have greater leverage in interacting with certain governments than they have utilized in the past, as indicated in Section II.B (iv).

The Article 20(2) of the ICCPR specifies, "any promotion of national, racial, or religious hatred that constitutes incitement to discrimination, hostility, or violence" must be prohibited by legislation. This unique construction, which includes the phrases "hate" and "hostility," as well as "advocacy" & "incitement," has sparked significant controversy and misunderstanding. Section 20(2) is also a little out of date since it only relates to the three foundations of identification—nationality, ethnicity, and

religion—rather than additional regulations such as age, gender, and gender identity that are incorporated in various national legislation and social media business policies. Navanethem Pillay, the then-United Nations High Commissioner for Human Rights, started a project in 2012 to clarify Article 20(2), which resulted in the Rabat Plan of Action (Esmaili, Marboe, & Rehman, 2017).

"Hatred' and' hostility' refer to intense and irrational emotions of opprobrium, enmity, and detestation toward the target group; the term 'advocacy' is to be understood as requiring an intention to promote hatred publicly towards the target group; and the word 'incitement' refers to statements about national, racial, or religious groups that create an imminent risk of discrimination, hostility, or violence against persons belonging to those groups," it says (Egan, 2013). This poses several issues, comprising the question of operations of social media corporations or anybody who can decide whether a danger is imminent. Context, speaker, purpose, speech content, form, speech act degree, likelihood, and urgency are elements of the Rabat Project's six-part threshold test for speech criminality. On the other hand, Urgency is a lowly criterion for regulating subject matter available virtually. For instance, it is often too late to intervene if a group waits until mass violence is impending before responding to dangerous chemicals.

According to the Rabat strategy, there must be an intention to provoke (Esmaili, Marboe, & Rehman, 2017). Intent is difficult to determine, particularly online, and is usually variable: the individual who creates incendiary information frequently means to inspire violence. The Rabat strategy has potential, but it will need to be tweaked in order to restrict internet material. The analytical framework for "hazardous speech," described as any human contact that enhances the possibility that the public would condone or engage in violence against another group, is another relevant test (Sherman, 2017). Furthermore, the Rabat test is most successful when combined with consequences instead of deletion of subject matter, for example, prohibitions on specific users or administrations. Lastly, the ban in section 20 is susceptible to legality, necessity, and legality in section 19, just like other speech limitations.

## CONCLUSION

We are optimistic that the following suggestions will inspire fresh debates, and a better grasp of IHL by the firms will be chosen explicitly for the social media forum. There is still additional work to be completed, since implementing international laws in the absence of international law might make corporate content change more difficult and ineffectual. In light of this, we want to raise a few of the numerous unresolved issues concerning how such businesses utilize and comply with the law.

- How should end-to-end encrypted services like WhatsApp, which restrict firms from accessing the material, be addressed under international human rights legislation? Is there a need for additional regulatory requirements in this case?
- In the context of platform rules, what do "least restrictive methods" entail, considering features like encryption and goals similar to user secrecy?
- How can Article 19 compliance be determined when businesses are regulated for a variety of reasons, some of which are covered by the treaty and others that aren't?
- As part of the ratification process, states can make reservations about a treaty. Should firms be subjected to a similar procedure if they are otherwise reluctant to comply with the law?

## REFERENCES

- Abdul Razak, M. I., & Mohamed Saleem, A. (2021). COVID-19: The Crossroads for Sinhala–Muslim Relations in Sri Lanka. *Journal of Asian and African Studies*, 00219096211021873.
- Allan, R. (2018). Hard Questions: Where Do We Draw the Line on Free Expression?. *Facebook Newsroom*, August, 9.
- Araújo, T., & da Rocha, G. C. (2017). Journalism in the Age of Personalization: A Psychophysiological Approach to the Attention Economy.
- Arnold, D. G. (2010). Transnational corporations and the duty to respect basic human rights. *Business Ethics Quarterly*, 20(3), 371-399.
- Ask, K., Spilker, H. S., & Hansen, M. (2019). The politics of user-platform relationships: Co-scripting live-streaming on Twitch. tv.
- Aswad, E. M. (2018). The future of freedom of expression online. *Duke L. & Tech. Rev.*, 17, 26.

- Backer, L. C. (2015). Moving forward the UN guiding principles for business and human rights: Between enterprise social norm, state domestic legal orders, and the treaty law that might bind them all. *Fordham Int'l LJ*, 38, 457.
- Baker, A. (2010). Constitutional Law-Freedom of Expression-Statute Criminalizing Creation, Sale, and Possession of Depictions of Animal Cruelty Was Substantially Overbroad and Facially Invalid as a Violation of First Amendment Protection of Speech. *Supra*, 80, 21.
- Bakir, V., & McStay, A. (2018). Fake news and the economy of emotions: Problems, causes, solutions. *Digital journalism*, 6(2), 154-175.
- Bellace, J. R. (2014). Hoisted on their own petard? Business and human rights. *Journal of Industrial Relations*, 56(3), 442-457.
- Benkler, Y., Roberts, H., Faris, R., Solow-Niederman, A., & Etling, B. (2015). Social mobilization and the networked public sphere: Mapping the SOPA-PIPA debate. *Political communication*, 32(4), 594-624.
- Berreby, D. (2017). Click to agree with what? No one reads terms of service, studies confirm. *The Guardian*.
- Bhagwat, A. A. (2020). The Law of Facebook. *UC Davis Law Review*, Forthcoming.
- Bursztynsky, J., Feiner, L. (2021). Twitter CTO Parag Agrawal will replace JackDorsey as CEO. CNBC-TECH DRIVERS, Available at <https://www.cnbc.com/2021/11/29/twitter-ceo-jack-dorsey-is-expected-to-step-down-sources-say.html> (accessed on 30-12-2021)
- Chadwick, E. (2020). Terrorism and Self-determination. In *Research Handbook on International Law and Terrorism*. Edward Elgar Publishing.
- Chesbrough, H. (2012). Open innovation: Where we've been and where we're going. *Research-Technology Management*, 55(4), 20-27.
- Church, K., & De Oliveira, R. (2013, August). What's up with WhatsApp? Comparing mobile instant messaging behaviors with traditional SMS. In *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services* (pp. 352-361).
- Coppex, J. (2020). *Cambridge Analytica: une crise plurielle. Analyse du scandale: Cambridge Analytica-Facebook* (Doctoral dissertation, University of Geneva).
- Coumans, C. (2010). Alternative accountability mechanisms and mining: the problems of effective impunity, human rights, and agency. *Canadian Journal of Development Studies/Revue canadienne d'études du développement*, 30(1-2), 27-48.
- Crawford, K., & Schultz, J. (2014). Big data and due process: Toward a framework to redress predictive privacy harms. *BCL Rev.*, 55, 93.
- Deutch, J. (2020). Image Activism After the Arab Uprisings| Challenges in Codifying Events Within Large and Diverse Data Sets of Human Rights Documentation: Memory, Intent, and Bias. *International Journal of Communication*, 14, 17.
- Douek, E. (2020). The rise of content cartels. *Knight First Amendment Institute at Columbia*.
- Duffy, M. J. (2014). Arab media regulations: Identifying restraints on freedom of the press in the laws of six Arabian Peninsula countries. *Berkeley J. Middle E. & Islamic L.*, 6, 1.
- Eckert, E. K. (2020). *The Information Needs and Information Behaviors of the Fat Acceptance Community* (Doctoral dissertation, Long Island University, CW Post Center).
- Egan, S. (2013). Strengthening the United Nations human rights treaty body system. *Human Rights Law Review*, 13(2), 209-243.
- Elkins, Z., Ginsburg, T., & Simmons, B. (2013). Getting to rights: Treaty ratification, constitutional convergence, and human rights practice. *Harv. Int'l LJ*, 54, 61.
- Esmaili, H., Marboe, I., & Rehman, J. (2017). *The Rule of Law, Freedom of Expression and Islamic Law*. Bloomsbury Publishing.
- Farris, M. P., & Coleman, P. (2020). *First Principles on Human Rights: Freedom of Speech* (No. 232). Heritage Foundation Special Report.
- Fick, M., & Dave, P. (2019). Facebook's flood of languages leave it struggling to monitor content. *Reuters*, April, 23.
- Flew, T., Martin, F., & Suzor, N. (2019). Internet regulation as media policy: Rethinking the question of digital communication platform governance. *Journal of Digital Media & Policy*, 10(1), 33-50.

- Gajos, M. (2020). Facebook in Myanmar: the challenges and promises of applying the United Nations Guiding Principles on Business and Human Rights to a social media company. *Problemy Współczesnego Prawa Międzynarodowego, Europejskiego i Porównawczego*, 18.
- Gharib, M. (2021). WHO Is fighting false COVID Info on social media. How's that going?. *NPR*, February, 9, 2021.
- Ghenai, A., & Mejova, Y. (2018). Fake cures: user-centric modeling of health misinformation in social media. *Proceedings of the ACM on human-computer interaction*, 2(CSCW), 1-20.
- Gillespie, T. (2017). Platforms are not intermediaries. *Geo. L. Tech. Rev.*, 2, 198.
- Goldmann, M. (2012). We need to cut off the head of the king: Past, present, and future approaches to international soft law. *Leiden Journal of International Law*, 25(2), 335-368.
- Haggart, B., & Keller, C. I. (2021). Democratic legitimacy in global platform governance. *Telecommunications Policy*, 45(6), 102152.
- Hanitzsch, T., & Vos, T. P. (2018). Journalism beyond democracy: A new look into journalistic roles in political and everyday life. *Journalism*, 19(2), 146-164.
- Hughes, P., & Ferrett, E. D. (2012). *Introduction to health and safety in construction*. Routledge.
- Huijstee, V. R., & Ceresna-Chaturvedi, L. (2012). How to Use the UN Guiding Principles on Business and Human Rights in Company Research and Advocacy. *SOMO*. [somo.nl/publicationen/Publication\\_3899/at\\_download/fullfile](http://somo.nl/publicationen/Publication_3899/at_download/fullfile).
- International Covenant on Civil and Political Rights (ICCPR) (Dec. 16, 1966)
- Jimenez, G. C., & Pulos, E. (2016). 1. Corporations and their Social Responsibility. *Good Corporation, Bad Corporation: Corporate Social Responsibility in the Global Economy*.
- Kadri, T. E., & Klonick, K. (2019). Facebook v. Sullivan: Public figures and newsworthiness in online speech. *S. Cal. L. Rev.*, 93, 37.
- Kaye, D. (2017). Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression on his Mission to Japan. *United Nations Human Rights Council A/HRC/35/22/Add, 1*, 29.
- Kearney, M. G., & Kearney, M. (2007). *The prohibition of propaganda for war in international law*. Oxford University Press on Demand.
- Klonick, K. (2017). The new governors: The people, rules, and processes governing online speech. *Harv. L. Rev.*, 131, 1598.
- Kornbluh, K., Goodman, E. P., & Weiner, E. (2020). Safeguarding Digital Democracy. *German Marshall Fund of the United States*.
- Lemley, M. A., & Volokh, E. (2017). Law, virtual reality, and augmented reality. *U. Pa. L. Rev.*, 166, 1051.
- Lwin, M. (2020). Applying International Human Rights Law for Use by Facebook. *JREG Bulletin*, 38, 53.
- MacKinnon, R. (2012). Consent of the networked: The worldwide struggle for Internet freedom. *Politique étrangère*, 50(2), 432-463.
- McHenry, L., & Khoshnood, M. (2014). Blood Money: Bayer's Inventory of HIV-Contaminated Blood Products and Third World Hemophiliacs. *Accountability in research*, 21(6), 389-400.
- Medzini, R. (2021). Enhanced self-regulation: The case of Facebook's content governance. *New Media & Society*, 1461444821989352.
- Mickoleit, A. (2014). Social media use by governments: A policy primer to discuss trends, identify policy opportunities and guide decision makers.
- Mundial, B. (2004). Reducing Poverty, Sustaining Growth. Scaling Up Poverty Reduction.
- Munro, V. (2014). *Hate Crime in the Media: A History*. ABC-CLIO.
- Myers West, S. (2018). Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms. *New Media & Society*, 20(11), 4366-4383.
- Nationen, V. (1985). *International Convention on the Elimination of All Forms of Racial Discrimination: Adopted and Opened for Signature and Ratification by General Assembly Resolution 2106 A (XX) of 21 December 1965; Entry Into Force: 4 January 1969, in Accordance with Article 19; Canada Signed the Convention on August 24, 1966 and Ratified it on October 14, 1970*. Minister of Supply and Services Canada.

- Pacheco-Vega, R. (2019). (Re) theorizing the politics of bottled water: Water insecurity in the context of weak regulatory regimes. *Water*, 11(4), 658.
- Posner, E. A. (2014). *The twilight of human rights law*. Oxford University Press, USA.
- Qin, B., Strömberg, D., & Wu, Y. (2017). Why does China allow freer social media? Protests versus surveillance and propaganda. *Journal of Economic Perspectives*, 31(1), 117-40.
- Rainie, L., & Anderson, J. (2017). *The Future of Jobs and Jobs Training*. Pew Research Center.
- Ronen, Y. (2010). Incitement to terrorist acts and international law. *Leiden Journal of International Law*, 23(3), 645-674.
- Royakkers, L., Timmer, J., Kool, L., & van Est, R. (2018). Societal and ethical issues of digitization. *Ethics and Information Technology*, 20(2), 127-142.
- Ruggie, J. G. (2020). The social construction of the UN Guiding Principles on Business and Human Rights. In *Research Handbook on Human Rights and Business*. Edward Elgar Publishing.
- Sherman, M. J. (2017). *Brandenburg v. Twitter*. *Geo. Mason UCRLJ*, 28, 127.
- Sinpeng, A., Gueorguiev, D., & Arugay, A. A. (2020). Strong fans, weak campaigns: Social media and Duterte in the 2016 Philippine election. *Journal of East Asian Studies*, 20(3), 353-374.
- Soomro, T. R., & Hussain, M. (2019). Social Media-Related Cybercrimes and Techniques for Their Prevention. *Appl. Comput. Syst.*, 24(1), 9-17.
- Suzor, N. P. (2019). *Lawless: The secret rules that govern our digital lives*. Cambridge University Press.
- Tarr, G. A. (2018). *Understanding state constitutions*. Princeton University Press.
- Trautman, L. J. (2019). Governance of the Facebook Privacy Crisis. *Pitt. J. Tech. L. & Pol'y*, 20, 43.
- Tufekci, Z. (2018). How social media took us from Tahrir Square to Donald Trump. *MIT Technology Review*, 14, 18.
- UN Committee on the Elimination of Racial Discrimination (UNCERD), *General Recommendation No. 35: Combating Racist Hate Speech*, U.N. Doc. CERD/C/GC/35, at 12 (Sept. 26, 2013), <https://undocs.org/CERD/C/GC/35> [<https://perma.cc/A24H-3NYD>].
- UN General Assembly, *Universal Declaration of Human Rights*, Dec. 10, 1948, 217 A (III), at 1-2, <https://www.un.org/en/universal-declaration-human-rights/> [hereinafter UDHR].
- United Nations. Office of the High Commissioner for Human Rights. (2012). *The Corporate Responsibility to Respect Human Rights: An Interpretive Guide*. United Nations Publications.
- Wilson, R. A., & Land, M. K. (2020). Hate Speech on Social Media: Content Moderation in Context. *Conn. L. Rev.*, 52, 1029.
- Zargar, S. T., Joshi, J., & Tipper, D. (2013). A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks. *IEEE communications surveys & tutorials*, 15(4), 2046-2069.