The Human Rights Council

The role and responsibility of multinational companies in moderating online behaviour

Research Report

Leiden Model United Nations 2021

The power of the Internet

Forum: *The Human Rights Council (HRC)*

Issue: *The role and responsibility of multinational companies in moderating online behaviour*

Position: *President*

Introduction

In this digital world, moderating online behaviour is important. Multinational companies rely on online content and behaviour to ensure the success of their businesses, especially in their social media networks, websites, online advertising etc. Such companies have offices in different countries.

A multinational company generally has offices and/or factories in different countries and a centralized head office where they coordinate global management. Some of these companies, also known as international, stateless, or transnational corporate organizations, may have budgets that exceed those of some small countries. ("Multinational Corporation (MNC)")

As in any online behaviour, there are risks which multinational companies should address and moderate to ensure users' safety and lawful usage of any online facility as well as gaining advantage for their businesses, depending on the elimination of hateful and dangerous illegal content while accompanying the allowance of the full right of expression of the users.

Content moderation by these companies is done through numerous ways including employees who check the content (human moderators) and the use of AI which uses "Machine Learning ". Moderation is usually based on the company's policies, rules and regulations to guarantee proper online behaviour and security of the company and its customers. A wrong post or misinformation could affect the credibility of the company and it might lose its reputation. Moderators review all posts and comments and delete ones that include offensive language, misinformation, disinformation, and hateful content which clashes with the policy regulations of that company.

Definition of key terms

Machine learning

Machine learning for moderating content refers to the process of assigning and inputting a group of information such as keywords, phrases, words, sample images or videos and posting rules to the AI moderator. The collection of information and references is also referred to as the base model. (“WHAT IS CONTENT MODERATION: OUR COMPLETE GUIDE”)

Content moderators

Content moderators are responsible for ensuring businesses and end-users are fully protected from harmful, disturbing posts and deceitful offers made by online trolls and scammers. They uphold the guidelines and objectives specified by a particular brand. In other words, they are the people behind the processes involved in screening user-generated posts, along with giving the green or red light for content uploaded by end-users. Content moderators also have the power to remove or ban members who violate in-house rules or threaten fellow users. (“WHAT IS CONTENT MODERATION: OUR COMPLETE GUIDE”)

Objectionable content

Is content that most users would find to be offensive, upsetting, or inappropriate. Also includes content that can infringe upon the rights of any third party company such as copyright, trademark, privacy, publicity and fraud. (TERMS & CONDITIONS FOR OBJECTIONABLE CONTENT)

Background Information

Although content moderation is not a new phenomenon, it faces a lot of backlashes in today’s society as the content online is becoming more abundant and diverse. Therefore, effective content filtering and procedural clarity are required to avoid overregulation, customer backlash, potential reputational harm, and income loss.

There's no denying that content moderation has a vital purpose. However, multinational companies are often criticized for the way they regulate or moderate content especially social media companies which need to ensure freedom of speech along with ensuring abiding by the international law as well as defining clearly what is considered as offensive and harmful. Countries have laws and legislations regarding online behaviour but more action needs to be done to enforce these laws; however, some countries use the definition of what constitutes lawful or not to restrict freedom of speech. For example, the decision of the Nigerian government to suspend Twitter because they deleted a message by President Buhari’s because they considered it as a violation of their policy.

According to UN Human Rights, “shutdowns like this matter because they restrict people’s ability to access information, also affecting other rights including work, health and education. They also have massive economic costs and undermine development” (“Moderating online content: fighting harm or silencing dissent?”).

Websites provide venues for a worldwide community of contributors to submit user-generated material. The user base is as diverse as the contributed material, which implies a large portion of the content published is undesirable to the majority of users and, by extension, the platforms. Child pornography, gratuitous violence, and unpleasant, hate-filled remarks are all frequently uploaded (or attempted to be uploaded) by users.

Possible Solutions

1) Improve content moderation processes

2) Transparent rules and restrictions around content moderation

3) Involve stakeholders including civil society in content moderation evaluation

Work Cited

Chen, James. “Multinational Corporation (MNC).” *Investopedia*, Investopedia, 27 Sept. 2021, https://www.investopedia.com/terms/m/multinationalcorporation.asp

“What Is Content Moderation & Types of Content Moderation.” *New Media Services,* 21 Dec. 2020, https://newmediaservices.com.au/fundamental-basics-of-content-

“Moderating online content: fighting harm or silencing dissent?” *United Nations Human Rights,* 23 July 2021, https://www.ohchr.org/EN/NewsEvents/Pages/Online-content-regulation.aspx

“What is content moderation?”, *Besedo,* 20 November 2020 https://besedo.com/resources/blog/what-is-content-moderation/

“ Guiding principles on business and human rights”, *United Nations Human Rights,* 16 June 2011, https://www.ohchr.org/documents/publications/guidingprinciplesbusinesshr_en.pdf

Mathew Ingram, “The challenges of global content moderation”, *Columbia Journalism Review,* 10 June 2021, https://www.cjr.org/the_media_today/the-challenges-of-global-content-moderation.php

“Expanding the debate about content moderation: scholarly research agendas for the coming policy debates”, *Internet Policy Review*, 21 Oct 2020,https://policyreview.info/articles/analysis/expanding-debate-about-content-moderation-scholarly-research-agendas-coming-policy

Richard Wingfield,” Risks and responsibilities of content moderation in online platforms” , *Open Global Rights*, 1 August 2018, https://www.openglobalrights.org/risks-and-responsibilities-of-content-moderation-in-online-platforms/

“The Quest for truth: content moderation in action ” , https://www.pwc.com/us/en/industries/tmt/library/content-moderation-quest-for-truth-and-trust.html