SOCIAL MEDIA CONTENT MODERATION CHALLENGES FOR VULNERABLE GROUPS: A CASE STUDY ON TIKTOK INDONESIA

Rizky Amalia Wismashanti
Faculty of Social and Political Science, Universitas Indonesia, Indonesia
Email: rizky.amalia22@ui.ac.id

ABSTRACT
As a popular social media platform, TikTok relies heavily on User-Generated Content (UGC), which is also one of its main attractions. The rapid production of UGC has implications for TikTok’s ability to moderate a large amount of content in a short time. Even though algorithmic technology and automatic filtering are used to assist this process, a moderator team that is responsive and efficient in carrying out moderation actions is needed. This need has implications for the growth of the Commercial Content Moderation (CCM) industry, where Tiktok employs a team of moderators to check and moderate content uploaded by users. The research aims to explore the complexities and unique aspects of Indonesian Tiktok content moderation and its implications for the development of UGC. This research uses an in-depth interview method approach with Tiktok Moderator Team informants as primary data, as well as related secondary uses as supporting data. The research focuses on how moderation practices are carried out by the Indonesian Tiktok Moderator Team in handling content related to vulnerable groups in society such as women, children, and the elderly, who in the last seven years have colored the dynamics of Tiktok in Indonesia. By highlighting the challenges faced by the Moderator Team in finding a balance between giving users freedom of expression and maintaining the security of the Tiktok platform, this research offers recommendations for improving content moderation practices on the platform, especially for the protection of these vulnerable groups.

KEYWORDS Tiktok; user-generated content; commercial content moderation; vulnerable

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International
INTRODUCTION

The existence of social media has changed the way individuals, communities and organizations produce, share and consume information (Kietzmann et al., 2018). There have been many studies highlighting the “bright side of social media” such as how social media opens engagement between companies and consumers in a more democratic way, or how companies improve public relations, customer service, product development or personnel decision making (Wagner, 2017; Kumar et al., 2016; Sabate et al., 2014). Apart from the many opportunities offered by social media, research has also found the “dark side of social media” which holds great risks for individuals, communities, organizations and even the whole society. More and more attention is paid to concerns such as online harassment (Jhaver et al., 2018), cyberbullying (O’Keeffe et al., 2011), trolling (Buckels et al., 2014), privacy (Pai & Arnott, 2013), fake news or hoaxes (Allcott & Gentzkow, 2017). The prevention of anti-social behavior has caused many platforms to implement moderation mechanisms, including trolling and harassment (Jhaver et al., 2018). Thus, it is important to find a balance between giving users freedom of expression and maintaining the integrity and security of social media platforms, one of which is through setting content moderation.

Commercial Content Moderation (CCM) As an Industry

Content moderation is the practice of screening user generated content posted on internet sites, social media and other online platforms to determine its suitability for a particular site, region or jurisdiction (Roberts, 2017). In its development, the practice of content moderation has shifted from open and voluntary moderation based on social communities such as Usenet groups, Wikipedia, etc. become a contemporary large-scale moderation practice using professional moderators and basic algorithms (Roberts, 2017). This happens as the rapid production of content by users or User-Generated Content (UGC) has implications for the platform to check and moderate a very large amount of content in a short time.

Even though the use of algorithmic technology and automatic filtering is used to assist this process, a moderator team is needed that is responsive and efficient in carrying out this moderation action. This need has implications for the growth of the Commercial Content Moderation (CCM) industry where online platforms or companies employ moderator teams or use technology to monitor and moderate content uploaded by users (Roberts, 2016).

In this regard, Facebook's Head of Global Policy Management, Monika Bickert, has been very candid about the challenges facing her company's platform. This was stated at a conference held at Santa Clara University in February 2018, as Alexis Madrigal reported at The Atlantic that the challenge of content moderation is distinct from many competitive battles and platform shifts companies have proven capable of surmounting. According to him, this is not a major technical challenge that can be solved by deploying many engineers to solve the problem. Bickert acknowledges that many of the issues facing Facebook's content team are unlikely to be resolved quickly or easily any time soon. And as for Artificial Intelligence (AI), Madrigal reports an emphatic response from Bickert. “That is a question we are often asked: When will AI save us all? We are far from it” (Roberts,
This means that the global human labor market for social media moderation as both a professional and part-timer will increase. As happened at Google and Facebook, the workforce will increase across all sectors where commercial content moderation takes place, from boutiques in the cloud to digital micro-chunk tasks in Amazon Mechanical Turk (Roberts, 2018).

The CCM moderator team usually consists of internal company staff or can also use the services of third parties who are experts in content moderation. The moderator's job is to view, review, and classify user-uploaded content according to the specified platform policies. Actions taken may include removing content, hiding infringing content, warning users, or taking other appropriate action. CCM may also involve using technology such as algorithms and automated filters to detect prohibited content. While these technologies can assist in the moderation process, the final decision is still made by a team of human moderators who have deeper judgment and understanding of the policy context and platform. The importance of CCM lies in maintaining the security, integrity, and online platform reputation. By effectively moderating content, platforms can create a more secure, positive and compliant environment for their policies, thereby enhancing user experience and building user trust in the platform.

On Twitter, moderation mechanisms include using a small number of human moderators who manually remove abusive posts, moderating through a voting mechanism where registered users vote positively or negatively on each post, and allowing users to flag abusive content. Another mechanism many platforms rely on is giving users the ability to mute, block, or report the offending user (Jhaver et al., 2018). On most platforms, and particularly on Twitter, blocking or muting an account allows a user to stop receiving notifications from that account, and the account's posts to not appear in the user's timeline or newsfeed. The difference between blocking and muting is as follows: blocking an account prevents that account from seeing the blocker's posts or sending direct messages to the blocker. Conversely, muted accounts will still be able to see the user's posts, "favorites," and reply to them. Muting an account is more socially complicated than blocking it: the muted user is not notified that he is muted, and he can continue to post to the muted user without knowing it (Jhaver et al., 2018). Recipients can't see their posts, whereas blocked users immediately notice that they are blocked if they try to post to the blocker. Even in camp with this feature Community Notes contributors can add information related to images and context will be located at the bottom of the tweet. This context will also appear when other users upload the same image in the future.

While Facebook and Twitter's moderation platforms may be similar in terms of their use of algorithms and the involvement of a team of human moderators, both platforms use a combination of technology and a team of human moderators, the proportion and engagement of the two can vary. Thus, it is interesting to understand the moderation practices of various social media platforms, from the perspective of human content moderation actors.
Tiktok as a popular platform: Indonesia ranks second in the world for active users

As societies around the world are increasingly digitized, political life and cultural activities are increasingly being reconfigured into a global ecosystem with a global platform (Keskin, 2018). As a social media platform, TikTok's existence in the global market has experienced several events, as required (Zeng & Kaye, 2022, Ridley, 2021), banned (Ellis-Petersen, 2020) and forced to sell (Allyn, 2020). On TikTok, geopolitical factors played an important role in causing several controversies for this platform from China, especially TikTok's failure to regulate the content shared on its platform, which was also a crucial factor (Gray, 2021).

Despite the controversy, the existence of TikTok is still considered and has even skyrocketed. At the global level, as reported by We Are Social in January 2023, this short video application has 1.05 billion users worldwide, placing TikTok as the social media application with the sixth most users. In the same year, TikTok users worldwide increased by 18.8% compared to the previous year. The active users of this application come from the United States. Then followed by Indonesia in second place with the number of TikTok users reaching 109.9 million. Followed by Brazil and Mexico, which each reached 82.21 million users and 57.51 million users.

Judging from the Play Store application distribution platform in June 2023, the number of TikTok application downloads exceeded 500 million. A number of feedbacks were provided by TikTok users by providing comments and ratings based on their experiences when using TikTok. Data from the Play Store shows user reviews on mobile media of 16.7 million, 476 thousand tablets, 1.04 thousand Chromebooks, and 882 TV reviews. The existence of reviews on various types of communication tools shows the popularity and wide adoption of TikTok. At the same time, TikTok on the App Store occupied the No. 1 entertainment app with 14.9 million reviews. This position shows the high popularity of TikTok among iOS users.

As one of the popular social media platforms, user-generated content or UGC has a significant impact on the TikTok platform, where TikTok is heavily dependent on UGC. User-generated content is one of the main attractions of this platform (Zeng & Kaye, 2022), wherein TikTok enables users to create and share creative and entertaining short videos, which other users on the platform can then view and participate in. UGC on TikTok includes various types of content, such as dance, lip-syncing, comedy and challenges, all created by users. TikTok also has features that allow users to interact with content and other users, such as liking, commenting and sharing content (Zeng & Kaye, 2022). This allows TikTok to grow as a vibrant and dynamic community, which allows users to actively participate in expressing themselves, sharing their creativity, and engaging with others. However, along with its positive aspects, user-generated content also brings certain challenges and implications, especially in implementing TikTok's moderation policy.

Content Moderation Part of Platform Governance

Like other social media platforms, TikTok also has a source of regulation in the form of a governance platform. This refers to the way social media platforms
regulate and manage UGC, which includes aspects of community policy, algorithms, including content moderation processes (Zeng & Kaye, 2022). Moderation is carried out by deciding and filtering content that is deemed in accordance with applicable policies, legal requirements, and cultural norms. The content moderation process involves cooperation between human and non-human actors, as well as negotiations between platform epistemic authorities and user agencies to carry out resistance. Content moderation also involves competition between corporate interests and regulatory obligations. Regulatory obligations refer to the legal responsibilities that social media platforms, including TikTok, must comply with. This regulatory obligation can come from various parties, such as the government, regulatory agencies, or non-governmental organizations. These parties may set certain requirements and standards that must be complied with by social media platforms in carrying out content moderation, such as removing content that violates the law or harms others. Therefore, social media platforms must pay attention to this regulatory obligation in carrying out content moderation so that they do not violate the law and can maintain the trust of users and the public.

In practice, content moderation on social media, including on TikTok, can vary from country to country. This is caused by differences in language, culture, and laws that apply in each country (Zeng & Kaye, 2022). For example, content that is considered to violate policy in one country may not be considered to be in violation in another. Therefore, TikTok has a content moderation team in each country to ensure that content uploaded by users complies with community policies and applicable laws in that country (Zeng & Kaye, 2022). Content moderation teams in each country can also monitor and handle content that violates community policies and applicable laws in that country, including on Tiktok Indonesia. Therefore, they have content moderation in each country to ensure that user-uploaded content complies with community policies and applicable laws in that country. The content moderation team in each country is responsible for monitoring uploaded content, handling reports of violations, and taking appropriate action if the content violates community policies and applicable laws in that country.

In the context of TikTok Indonesia, the content moderation team in Indonesia will monitor and handle content that violates community policies and applicable laws in Indonesia. They will follow the guidelines and content moderation guidelines set by TikTok and consider cultural and legal aspects that apply in Indonesia in determining the necessary moderation actions. This aims to maintain the integrity of the platform, protect users from prohibited content, and ensure a safe and positive experience for TikTok users in Indonesia.

**RESEARCH METHOD**

This study uses a qualitative research methodology with a case study approach. Primary data was collected through in-depth interviews with a member of the Moderator Team who works for company "X" which is a Commercial Content Moderation (CCM), namely as a third party appointed by Tiktok Indonesia to moderate content. The identity of the informant remains confidential in accordance with the code of ethics of the content moderator profession and to
maintain the confidentiality and security of the informant. Interviews were conducted to gain a direct understanding of the content moderation process and the challenges they face based on the informants' subjective views and experiences.

Apart from primary data, this study will also use secondary data to support the analysis. Secondary data will be obtained from various sources, including data on the trend of the Tiktok platform, documents on the official TikTok website, news and related literature. The use of secondary data as an additional source can also provide more complete and in-depth information about the context of content moderation in Indonesia. Moreover, the research aims to explore the complexities and unique aspects of Indonesian Tiktok content moderation and its implications for the development of UGC.

RESULT AND DISCUSSION

The existence of TikTok in Indonesia is experiencing its own dynamics, which are marked by warnings and temporary blocking by the Ministry of Communication and Information of the Republic of Indonesia (Kominfo RI) for content that is considered to violate applicable rules and norms. Several issues that have colored the dynamics of Tiktok Indonesia in the last six years include women's activities which are considered vulgar, content that is not child-friendly (Daon001(a), 2018; Daonn001(d), 2019), and content that exploits senior citizens in the form of activities "begging online" (Clinton, 2023). In the Indonesian context, the Ministry of Communication and Informatics plays a role in overseeing and regulating content circulating on digital platforms such as TikTok, so that its use remains in accordance with Law No. 11 of 2018 concerning Information and Electronic Transactions. In accordance with these provisions, there are twelve groups of content that are categorized as negative content, including pornography, gambling, extortion, fraud, violence, defamation, SARA, terrorism/electronic radicalism which violate other laws. Apart from the Government, regulatory obligations can also come from supervisory bodies, or non-governmental organizations (Zeng & Kaye, 2022). These parties may set certain requirements and standards that must be complied with by social media platforms in carrying out content moderation.

First related to content containing women's activities. In 2018 Kominfo temporarily blocked the Tiktok platform in response to the many negative reports that came in, especially from the Ministry of Women's Empowerment and Child Protection, the Child Protection Commission, as well as reports from society (Daon001(a), 2018). Reporting from its official website, Kominfo RI announced on Tuesday 3 July 2018 that the Government officially temporarily blocked Tik Tok (Daon001(a), 2018), where the Director General of Informatics Applications for Kominfo, Semuel Abrijani stated that the blocking was carried out in relation to violations found on the platform, among others, pornographic content, immoral content, religious harassment content obtained from data collection from the AIS Team or Kominfo internet scavenging machines (Daon001(a), 2018).

From the perspective of informants as content moderators, this is also a challenge in itself where AI technology is still difficult to distinguish whether certain content contains exposure to the body, certain activities that contain
pornography, immorality or not. Tiktok enables users to create and share creative and entertaining short videos, which can then be viewed and followed by other users on the platform. UGC on TikTok includes various types of content, such as dance, lip-sync, comedy, and challenges, all of which are user-generated (Zeng & Kaye, 2022). It is often found that the content of a woman doing a dance where IA on Tiktok detects that the video contains vulgar or pornographic content. This might happen when IA translates a dance move performed by a person, perhaps referring to a body exposure that is considered private.

In the context of Indonesia which is multi-ethnic, religious and racial, content moderators need to have the wisdom to address this. For example, videos of regional dances wearing certain traditional attire should not violate certain norms because the definition is regional art content. Different cultures and societies have different views on this body exposure. In several cultures in Indonesia, there are social norms and policies that regulate body exposure and nudity in order to maintain decency, privacy, and protect existing cultural values. Body exposure limits are also implemented in some legal contexts to protect individuals from harm or exploitation.

Regarding this limitation of body exposure, TikTok has regulated in the Community Guidelines, which was launched from the TikTok page and updated in March 2023 where TikTok claims to be a platform that enables users to discuss and learn about sexuality, sex, or reproductive health. However, TikTok understands that certain content may not be suitable for youth, may offend some people, or has the potential to exploit youth. TikTok expressly does not allow sexual activities or services such as sex, sexual arousal, fetishism or other abnormal sexual behavior, as well as seeking or offering sexual services.

TikTok understands that people's approaches to clothing or body exposure to each other can be different, therefore TikTok needs to understand the cultural norms that exist in the region. Exposure of the body in culturally acceptable contexts such as in athletic wear and swimwear at the beach or festivals.

While the concept of body exposure can be highly subjective, in general, TikTok does not allow content that shows the genitals, buttocks, and nipples and areolas of women and girls, including when wearing sheer, see-through clothing. TikTok also doesn't allow content that displays significant body parts of teenagers. It is excluded in certain contexts such as medical treatment, educational purposes, or as part of local cultural practices.

Apart from the issue of body exposure to women, the Moderator Team is also often faced with content about women related to religious attributes. As a country with the largest Muslim population in the world, content moderators also often find videos of women dancing with Islamic religious attributes, in this case the hijab. IA will automatically detect such content as harassment or blasphemy. As a content moderator on a platform like TikTok, handling videos of women dancing while wearing hijab can be a very ambiguous situation. This is caused by various interpretations and assessments of these actions, especially in the context of religion and culture.

In several cases that were found by informants, videos of women dancing while wearing the hijab were not considered to violate existing policies or
guidelines, especially if the dance was not considered to be harassing or offensive to religion. Here the moderator will consider the context, intent and message conveyed in the video to make a moderation decision. Several factors are considered, such as whether the video still respects religious values, whether the video is truly artistic or promotes Indonesia's cultural diversity or not.

In practice, moderator decisions are often not in accordance with user expectations, especially if the content is considered a violation which results in the removal of content and the suspension of the user's account. Talking about the role of content moderators, researchers are interested in seeing popular hashtag searches related to the keywords "Tiktok and Moderators" which then obtain the top five data as shown in Figure 1.

![Figure 1. Searching for hashtags with the keywords “Tiktok and Moderators”](source: Tiktok, June 2023)

Of the five popular hashtags that contain the keywords "Tiktok and Moderators" are then arranged in a ranking according to the number of views, which can be seen in the following table:

<table>
<thead>
<tr>
<th>Rating</th>
<th>Hashtags</th>
<th>Views</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>#tiktokmoderators</td>
<td>1.2M</td>
</tr>
<tr>
<td>2</td>
<td>#tiktokmoderatorssuck</td>
<td>924.4K</td>
</tr>
<tr>
<td>3</td>
<td>#tiktokmoderator</td>
<td>293.2K</td>
</tr>
<tr>
<td>4</td>
<td>#hittiktokmoderators</td>
<td>603</td>
</tr>
<tr>
<td>5</td>
<td>#aderuptiktok</td>
<td>4</td>
</tr>
</tbody>
</table>

Source: Tiktok, June 2023

From the search results for these keywords in June 2023, data was obtained that #tiktokmoderators received 1.2 million views, followed by #tiktokmoderatorssuck with 924.4 thousand views, and #tiktokmoderator with 293.2 thousand views. Hashtags on social media whether TikTok, Instagram, or Twitter, "views" refers to the number of times content using these hashtags has been displayed or seen by other users. Views hashtags can provide an indication of the extent to which content with these hashtags gets attention and interaction from other users.
users. The higher the number of views on a hashtag, the more people will see or find content related to that hashtag.

In this finding, #tiktokmoderatorssuck got the second highest views after #tiktokmoderator. This shows that the hashtag is quite popular and widely used by TikTok users which can be interpreted as an expression of disappointment or dissatisfaction with the TikTok moderator. In this context, hashtag users are used to express users' opinions or complaints about moderator decisions that are considered unfair.

However, Tiktok actually opens a space for dialogue between moderators and users where users can take advantage of the "report" feature to appeal against the moderator team's decision. Reported on the Tiktok website, it is stated that users whose content has been deleted by the Moderator Team can appeal the decision. If the appeal is approved then the content or the user's account will be restored, and the strike will be removed from the user's account. Even TikTok provides a firm note that removing infringing content does not mean removing the warning that TikTok has sent to the user, even TikTok can still issue a warning again if infringing content is still found.

In this case the informant admits that dialogue space exists between the user and the Moderator Team but is not wide open. Because defining safe content will also be very subjective, it can even take up discussion time, thus hindering the efficiency of moderation work in the midst of heavy Tiktok content. The informant described that one Tiktok content moderator must be prepared for the possibility of entering 70 thousand content that must be followed up in one working day, which means that to determine whether one content passes or does not pass, the moderator only has 20 seconds. This was done because of the rapid production of UCG using Tiktok. So that the moderator is required to really work in an effective and efficient time.

Tiktok content moderators must also face the challenge of digital literacy in Indonesian society, which is still in the medium category, as stated in a report released by the Indonesian Ministry of Communication and Informatics in 2023 (Agustini, 2023). Even though at the same time, Indonesia is included in a fairly high internet penetration rate according to the Internet Service Users Association (APJII) through a survey conducted in 2023, where the number of people connected to the internet reaches 215.62 million people out of a total population of 275.77 million people. Indonesian residents.

The informant gave an example that there are still frequent reports from users who do not comply with reporting rules where the reported content does not indicate a violation of regulations or community policies, but because the content concerns female figures who are considered too beautiful. As stated by an informant that:

"I often get reports that sometimes make me think differently, like someone reporting a video just because the video isn't there, or just because the creator of the video is too beautiful so it's reported", Informant, June 2023.

Second, child related content. Tiktok has actually improved for the protection of children, especially after the temporary blocking incident carried out by the Indonesian Ministry of Information and Communication in July 2018 (Daon001(a),
where the Government only opened access to this application made by ByteDance for children over the age of 13 (Rahmi, 2023). In the context of this user age limit, TikTok does have a strict policy.

Reported on the TikTok website, the age of children who can use TikTok is 13 years old (or 14 years for South Korea and Indonesia), in this case TikTok is committed to providing an age-appropriate experience, so it is important for teenagers to provide their actual date of birth.

If it is found that video content only shows children without the presence of adults, this can be considered a violation, namely the use of the platform by minors. TikTok also uses a machine-based moderation system to detect age-restricted content. If the content is identified as content involving unauthorized underage users, then moderation actions such as deleting or updating access can be carried out by TikTok like the experience experienced by the user in Figure 2.

When a user’s account is blocked, the user will receive push notifications and in-app notifications when opening the application again. If an underage user submits an appeal, the user affected by the account blocking can appeal and need to confirm the user’s age and domicile location. Nonetheless the TikTok moderators will decide whether to unban a user's account or not, because TikTok has a legitimate interest in keeping the TikTok community safe and preventing misuse of the service accordingly Privacy Policy those platforms. For example, for users under the age of 16, TikTok does not allow these users to do things such as send or receive direct messages and allow others to download their videos, and make stich or duet videos to their content. Moderators can influence content dissemination and user visibility.

Still in relation to children, Tiktok Indonesia’s content moderators also face challenges regarding the limits of protecting children's privacy that are safe in the context of sharenting. The term sharenting is a combination of the two words share and parenting, referring to the trend in cyberspace that parents do when sharing detailed information about their children on social media (Marasli et al, 2016). In the context of Indonesia’s digital society there is a sharenting study which found that a number of uploads by parents found on Social Networking Sites (SNS) indicate a lack of awareness of parents in maintaining the privacy of their children in their uploads on Instagram (Dwiarsianti, 2022). Parents may intentionally share content of young children who are bathing and are considered not to violate the limits of pornographic content on children, but Tiktok Indonesia is very firm in
handling it. One of the experiences of a user who has received a violation strike for uploading content of a child taking a bath is shown in Figure 3.

![Figure 3. One of the user reviews affected by the issue of unfriendly content for children](Source: Tiktok, June 2023)

When it comes to kid-friendly content, social media platforms like TikTok keep updating their community guidelines or usage policies regularly. This continues to be done to adapt its policies to the latest developments. The aim is to address new challenges as they arise, such as privacy, security, harmful content or other issues related to the use of the platform. Related to this, informants have unique experiences:

"For example, the habit of Indonesian people who like to ride their children on a motorbike in front of them. Maybe in the last three years, content like that has been considered endangered for the safety of children. But because it responds to customs in Indonesia, it's common to carry a child in front of you, so content like that is now safe."

To handle content management involving children and adolescents, TikTok has also provided a guardian's guide that can be accessed on the TikTok website to help keep the TikTok community safe, as well as general information about internet safety issues that children and their parents or guardians can learn together.

**Third**, content that exploits the elderly. It's still fresh to remember that in early January 2023 Indonesia was horrified by TikTok's content which exploited the elderly for the purpose of raising funds or "begging online" which was carried out by several users" (Clinton, 2023). This "begging online" activity for informants as content moderators is also a challenge in itself. First, where there are no “clear” violations of community policy found in the activity, moderators may have limited ability to take direct action on the content. Even though ethically, the informant views the content as inhumane and not educational.

Second, in this case the position of the moderator team is the content moderation service provider appointed by TikTok so that in certain cases it cannot be separated from the interests of Tiktok as a corporation. Corporate interests refer to the profits and business interests pursued by social media platforms such as TikTok, where corporate interests can influence how social media platforms carry out content moderation (Zeng & Kaye, 2022). For example, a social media platform may prioritize content deemed popular or controversial in order to increase the number of users and engagement on that platform. This can affect how social media platforms carry out content moderation, such as allowing content that violates policies to remain on the platform to increase the number of users and engagement.
In addition, corporate interests can also affect how social media platforms manage user data and maintain user privacy. Therefore, it is important for social media platforms to pay attention to corporate interests and regulatory obligations in conducting content moderation in order to maintain a balance between business profits and public interests.

On the Government side, the phenomenon of "online ngemis" is also seen in a gray position. In terms of regulation, both ITE Law No. 11/2008 and Government Regulation No. 71 of 2019 concerning the Implementation of Electronic Systems and Transactions do not regulate begging activities as prohibited content such as pornography, terrorism and radicalism. Therefore, Kominfo RI does not use the ITE Law as a basis for handling online begging activities, but the Government through the Ministry of Social Affairs (Kemensos RI) issued Circular Number 2 of 2023 concerning Control of Exploitation and/or Begging Activities that Utilize the Elderly, Children, People with Disabilities, and/or Other Vulnerable Groups. This circular is addressed to local governments as an effort to prohibit the exploitation of the elderly and protect other vulnerable groups from begging practices that take advantage of them. This Circular Letter is issued as an effort to address exploitation and/or begging activities that utilize vulnerable groups such as the elderly, children, persons with disabilities, and other vulnerable groups. In the circular, local governments are instructed to prohibit the practice of exploiting the elderly and protect other vulnerable groups from begging practices that take advantage of them. This Circular is then the basis for banning "ngemis online" activities which are then followed by the deletion of accounts that previously carried out these activities.

By looking at content moderation practices carried out by TikTok Indonesia in the last seven years, one can see the dynamics of the "power relations" that occur between moderators and platform users. TikTok moderators have an important role in overseeing and managing content uploaded by users so that they have the power to determine what is allowed or not allowed on the platform. In this context, there are several aspects related to power relations that need attention.

The first is related to the determination of platform policies. TikTok Moderators are responsible for establishing and enforcing platform policies regarding permitted or prohibited content. They have the authority to remove content that violates policies, impose sanctions, or even block user accounts.

Second, regarding content filtration. TikTok moderators use algorithms and moderation tools to detect and remove content that violates platform policies. In this process, they may influence content dissemination and visibility of certain users based on their policies and judgments.

Third, communication with users. TikTok moderators also interact with users by sending messages or providing explanations regarding content being removed or actions taken against users' accounts. In this interaction, power dynamics occur where the moderator has the authority to explain or defend his actions.

The fourth is related to transparency and accountability. Power relations are also related to transparency and accountability of TikTok moderators to users. Users are interested in understanding the policies and reasons behind moderator actions. Therefore moderators have a responsibility to explain their decisions and maintain a fair relationship with users.
The power relationship between the moderator team and users of the TikTok platform was also found by several users who felt that moderators were using their power unfairly and could be very subjective. Therefore, it is important for moderators to strike a balance between giving users freedom of expression while maintaining the integrity and security of the TikTok platform.

CONCLUSION

As a digital platform, TikTok uses algorithms and moderation tools to detect and remove content that violates platform policies. However, in practice there is some content that requires more contextual interpretation, especially for vulnerable groups such as women, children, and the elderly. In these situations, the moderator team has an important role to play in verifying content and taking appropriate moderation decisions. Although the standard of moderation ultimately becomes subjective and debatable because it involves the human senses of the Moderator Team members.

However, the involvement of content moderators is very important because they are able to translate ambiguous situations. For example, what often happens to women's groups is whether content is considered as an expression of cultural art or exposure of the body that ends up being pornographic. TikTok moderation is also faced with digital literacy challenges when dealing with groups of children. The limits of protecting the safety and privacy of children on TikTok are very subjective for parents, but it is the duty of the moderator to decide whether to unblock the account of an underage user or not. Or when child-related content is deemed to endanger the child's privacy. TikTok has a legitimate interest in keeping the TikTok community safe and preventing misuse of the service accordingly.

Finally, the elderly group is vulnerable to exploitation. At first, the moderator team was in a position that could not decide when the content did not clearly violate the platform's community policies, even though ethically and humanely the content was inappropriate. So that in a certain position, the practice of content moderation by humans will eventually be faced with the interests of platform owners, namely when TikTok as a corporation also has certain strategies to increase user traffic by utilizing content that is considered "sellable". Although in the end the role of community groups and the government can also push for TikTok's decision to stop the content.

It is therefore important to find a balance between giving users freedom of expression and maintaining the integrity and security of the TikTok platform. By highlighting the challenges faced by TikTok Indonesia, this research offers recommendations for improving content moderation practices on the platform, especially for the protection of vulnerable groups such as children, women, and the elderly.

First, develop a policy of moderation that is clear and continuously updated following the developments and needs of the community. In the phenomenon of "begging online" experienced by vulnerable groups of the elderly for example. In this situation, neither the Government nor TikTok initially had devices that could block this content. However, at the insistence of the public through the Institute,
they urged that there be a follow-up to the incident, and ended with the deletion of the content and user accounts for "begging online".

Second, the need for transparency. Submission of an outline of the content moderation policy to users is necessary. Searches with the hashtag #tiktokmoderatorssuck show that the public feels the need to convey things related to more balanced Tiktok content moderation. TikTok should actively engage users in the content moderation process by providing easily accessible reporting mechanisms and providing users with open feedback about actions taken on infringing content. By communicating it will certainly help users understand the limits of expression and reduce violating content.

Third, collaboration with external parties related to vulnerable groups. Here, TikTok can work with external parties, such as child and women's protection agencies, human rights organizations, and related community groups, to get diverse input and views on content moderation practices. This collaboration can help improve TikTok content moderation policies and practices that are proportional to vulnerable groups such as women, children and the elderly. This research also found that the moderator team also often faced the problem of mastering digital literacy in Indonesian society which was still in the moderate category. So it is important for the TikTok platform to synergize with related parties to jointly contribute to increasing the digital literacy of the Indonesian people.

Finally, this research is still limited to describing content moderation practices carried out by humans based on the subjective experience of informants. Although informants provide valuable insights, that single perspective cannot represent all content moderation practices on TikTok Indonesia. Future research on content moderation on TikTok Indonesia can broaden its scope by delving deeper into the geopolitical factors that influence content moderation policies. Factors such as government regulations, social norms, and political pressures can be important considerations in determining content moderation policies. In the Indonesian context, the existence of unique geopolitical aspects can influence how TikTok Indonesia moderates its content. Research can also consider a qualitative approach involving interviews with various stakeholders, such as content moderators, TikTok Indonesia representatives, government, and community groups. By involving perspectives from various parties, research can provide a more complete insight into content moderation practices and the geopolitical factors that play a role in the Indonesian context.

REFERENCES


Kumar, A., Bezawada, R., Rishika, R., Janakiraman, R., & Kannan, P. K. (2016). From social to sale: The effects of firm-generated content in social media on


