

Section 4. Sociology

<https://doi.org/10.29013/YSJ-22-1.2-30-45>

Jiaqi Wu,
Shanghai Pinghe School, China

STUDY OF THE ETHICAL ISSUES IN TIKTOK

Abstract. In recent years, TikTok, a globally popular video-sharing smartphone application (app), has gained a reputation for capturing the zeitgeist of culture and redefining the way people use social media. This research paper aims at figuring out the ethical problem in Tiktok. After briefly describing the meaning of algorithms in life and some potential problems of using algorithms, this paper focuses on the social software TikTok, a short video application that has become very popular and controversial in recent years. This research has also used the TikTok dataset and modeling to support the spread of inappropriate text and videos. Through the analysis of the TikTok dataset and some visual data processing methods, how the recommendation system works is better displayed in this paper. By analyzing the data, the paper also makes it clear why the recommendation system brings the threat of spreading inappropriate information, both in text and videos. Additionally, the research outlines to a certain degree some challenges when solving the algorithm bias in TikTok through some examples. Furthermore, this paper analyzes the measures taken by the government and TikTok to face these problems. It emphasized that moral problems with Douyin are inevitable, but in addition to social platforms and the government, this research has concluded that there are better ways for TikTok to alleviate the problem to some extent. This paper recommends further research into more social media apps besides TikTok and further analysis of the potential problems that appeared in these apps or some common problems both share.

Keywords: Tiktok; Algorithm; Recommendation system.

1. Introduction

1.1 Algorithm

On the fast track of development, our society is moving forward by a leap. Technology does more than facilitating communication and ending the days when people used to meet for a chat. Artificial intelligence also equips people with the ability to replace a large proportion of complex and mundane work. Artificial intelligence or algorithm is a set of instructions that take an input, A, and provide an output, B, that changes the data in some processes [1]. It has widespread

use on the daily basis and is still becoming more and more ubiquitous without awareness. In a golden age of data, for a myriad of jobs that people did on their own before, colossal datasets and algorithms can almost completely replace human labor. These sets of instructions for solving problems or completing tasks play an indispensable role in our everyday life: from the Internet to medical care, from education to transportations, algorithms are helping us in far more fields than we can imagine, such as calculating the most efficient routes of delivery, predicting the most profitable

marketing strategy, increase observability of specific behaviors and making a sensible public health policy.

Here is the way how the algorithm can be applied in the dataset. Sifting through masses of data, algorithms provide computers with a successive and precise guide to do specific tasks. Utilization of this can be followed back to one of the famous persons in computational hypothesis. In 1952, Alan Turing, a mathematician, published a set of equations that tried to explain the patterns people see in nature, such as the dappled stripes of zebra, the whorled leaves on a plant stem, and the complex tucking and folding that turns a ball of cells into an organism [2]. Nowadays, simply put, algorithms are also about a set of equations to solve specific problems. The equations work via input and output by applying each step of the algorithm to the data to generate the result we needed which is the output. In recent years, increasing amounts of available data have enabled AI algorithms to be utilized in many fields.

In our daily life, file compression is a process that algorithm runs on data in which the size of a file, as well as the storage, is reduced by re-encoding the data. After downloading a.zip file, operating systems can extract the contents of this.zip file like a normal folder. Rather than extract everything manually just like a decade ago, the operating systems work in the background and dive into the.zip files in a few seconds. This is because compression algorithms can optimize specifically for different files and make them into a usable form. Compression algorithms are a kind of operating system that turns the input,.zip files, into the output, usable files. Long series of bytes that are repeatedly used is found by the algorithms, and it replaces them with a single byte or as short series of bytes as possible that rarely occurs in the file [3]. Therefore, this reduces the amount of memory required to store files [4]. Just like compression algorithms, many other kinds of algorithms are playing important roles on the daily basis and providing people with a richer trove of valuable information. These algorithms are far more complicated than the equation via input and output, but the essential

behind algorithms is to simulate the data processing and decision-making processes of the human brain.

1.2 Concerns of algorithms

These complex algorithms are used in many other aspects that we do not aware when using them, but is affecting us every day, such as business and marketing. “We’re at the beginning of a golden age of AI. Recent advancements have already led to inventions that previously lived in the realm of science fiction – and we’ve only scratched the surface of what’s possible.” Jeff Bezos, Amazon CEO said. Nowadays, almost all the businesses are utilizing and taking advantages of AI. The improvements in algorithms make it possible to automate decision-making based on colossal database. For example, price discrimination algorithm is a practice used by businesses to charge a different price for the same product to different consumers [5]. However, increasing capabilities of price discrimination algorithms is debatable. For businesses, they implement price discrimination using algorithm by profiling consumers using cookies which is commonly used in ecommerce applications [6]. This information can be used for either consumer segmentation with different pricing displayed or personalization, such as product/service recommendation, targeted advertisement, targeted coupons with a flat price, and rank in algorithmic search which influences click-through rates [7]. These can all be approaches to implement price discriminate using algorithms. In this aspect, the outcome of algorithm is negative for consumers. For people who are charged with a higher price, it is somehow inequitable. Additionally, their personal information is gleaned by the business, which will cause some network security problems.

Concerns about algorithm is also present in other fields. With big data, a series of algorithms have significantly improved their usability on Internet and social media. Issues about using algorithm on the Internet and social media bring AI against human ethics. The questions being raised about algorithms are not about algorithms per se, but about how algorithm is used and how data is used [8].

This problem is “the social dilemma”, meaning social media applications or websites manipulate people by using algorithms. Nowadays, technology companies offer a lot of free products, Facebook, WeChat, Tiktok, Twitter, for which people don’t pay to use. This is attributed to that the advertisers are people who pay. The advertisers are the customers of the social applications and websites, and the users and their personal information are the goods to be sold. For example, Google’s system scans the emails being sent and received by personal account and gain the content of these emails. By using information gleaned from a user’s email combined with data from their Google profile, Google display what it considers are relevant advertisements [9]. Because users will only see the advertisements they are interested in, they are likely to click on it and it is harder to determine whether the messages presented by advertising are truthful or not. More often than not, by exaggerating some attributes of the product and downplaying the product’s drawbacks to people who are interested in it. This also gives companies an opportunity to encourage endless purchase requests.

Surveillance capitalists sell certainty to companies that want to know with certainty what people do and what people will do. In 1986, 1% of the world’s information was digitised. In 2013, it was 98% [10] and this is where the certainty comes from. By using digitized information, algorithm determine the things you receive in those websites and applications. Engineers at companies like Facebook and Google spend huge chunk of time to create new algorithms and tweaking old ones [11]. It is a matter of fact that today’s internet is ruled by algorithms to some extent. Social media is no longer a bicycle waiting to be ridden, or a car to be driven. It is self-learning and self-analyzing. To some extents, artificial intelligence has gradually evolved into a tool for soliciting, luring, and even manipulating and profiting from human beings.

1.3 Background information of TikTok

Artificial intelligence is also commonly used in other social media websites and applications, includ-

ing the popular video-sharing application–TikTok. TikTok is a globally popular smartphone application of sharing and creating short videos. It is owned by Byte Dance Ltd., a privately held company headquartered in Beijing, China and owned many other popular applications in China [12]. In recent years, it has gained a reputation as capturing the zeitgeist of culture and redefining the way people use social media. Initially launch in China by Byte Dance, the app become well-known around the world after renaming to TikTok. In November 2017, the previous rival of TikTok, Musical.ly, was acquired by the company for \$1 billion, and the user accounts of both are consolidated [13]. Now, it has 689 million users internationally every month [14] and has downloads estimated at between 850 million and 987 million, excluding Chinese third-party downloads, in 2020 [15]. Allowing users to create, watch, and share short entertaining videos set to various music and sound effects, TikTok is notable for its high level of engagement. The personalized feeds of videos and the addictive quality further drives TikTok’s popularity. By collaborating on content with other users, having online live shows, and easily adding effects like filters, stickers, and lip sync sound effects, TikTok attracts mounting number of creators and becomes the second most downloaded iPhone app in 2020 [16].

Since the app is free, it seems mysterious how the application makes money. The truth is that the main value of TikTok is the users or the gargantuan user base. With more than 2 billion users, the advertisements run by TikTok are the most obvious way it makes money. Different types of advertisements, including In-Feed ads, Brand Takeovers, and Branded Hashtag Challenges, allow TikTok to earn a large chunk of money from people or businesses who want to display their advertisements or new hashtags. When we are scrolling through the For You page, receiving an advertising videos, and seeing the newly shown up hashtags, branded custom stickers, or augmented-reality filters and lenses, TikTok is earning money and selling the products–the users. Additionally, the in-app purchase is another income

for TikTok. Using money to purchase virtual coins, people over eighteen years old are allowed to use this as in-app currency and send virtual gifts to others.

2. Ethical problem in TikTok

2.1 Newly Appeared Ethical Problem

TikTok's appeal relies on addictive quality and high levels of engagement. After opening this application, the For You page is what the users see first. Equipped with a powerful recommendation system, TikTok can automatically provide videos curated to people's interests. In the For You page, the short-form video uniquely tailored for every user is played one-by-one instantaneously as the application is being opened.

In fact, the For You page relies on algorithms to provide a myriad of videos which is different for every one of the 689 million monthly active users worldwide [17], and the success of TikTok also has a close relationship with this powerful intelligent algorithm behind it. People were speculating about this complex and mysterious For You page and how recommendations are delivered to their feed, attributed to that TikTok had been secretive about the platform's algorithms in the past. TikTok finally revealed how the For You Page recommendations are generated several years ago. However, at the same time when people know more about the algorithms. There are newly appeared ethic problems in TikTok. The data from the users is the main basis of this application, but the way of using data and feature videos makes people concern. As how videos get featured is almost only based on the For You algorithm, but lack manual review, the videos that is ethically wrong or the videos with misleading information will easily get spread. Under this circumstance, TikTok becomes a great place for malicious people to hide nude images and objectionable contents. Its lax security and control have allowed it to become a place filled with pedophiles, profanity, crime, violence and extremism and will lead to many fatal consequences on not only TikTok users, but also all people online [18].

This paper focuses on TikTok's video recommendation system and the ethical problem arising from it.

By using the TikTok video dataset, the paper further analyzes why the ethical problem is caused by the recommendation algorithm. The TikTok dataset is processed by using Apache Spark to formulate a series of queries, and some analytics are extracted based on the data contained within this dataset. Furthermore, this paper analyzes the solutions to those ethical problems in terms of the algorithm bias. This research can be useful for people to have a deeper understanding of the TikTok's video recommendation system and figure out the way of improving this technology.

2.2 Technical development of video recommendation system

As mentioned before, not like other application, the addictive quality led by algorithm in recommendation system is the kernel of TikTok's success. Bytedance is focusing on the artificial intelligence in various products. An AI Lab is established in 2016, focusing on the development of innovative technologies that serve the purposes of ByteDance's content platforms [19]. In TikTok, it is also this powerful algorithm technology that leads to its popularity. TikTok is an algorithm-driven, content-oriented product. Besides the basic, industrialized recommendation engine need a robust backend and architecture design for integration and huge database to support the algorithm [20].

2.2.1 How TikTok recommends videos

According to TikTok: "The system recommends content by ranking videos based on a combination of factors – starting from interests you express as a new user and adjusting for things you indicate you're not interested in, too" [21]. From hashtags used to user interactions and your settings– they can all influence the algorithm and the videos recommended to you. With this recommendation system, the application has the capability to eliminate almost all the initiative searching of its users. It enables users to easily get the videos they like and enables videos to pass to only people who will like them. TikTok revolutionize the way for people to search. It greatly alleviates the effort people pay to obtain useful and interested

information from massive amounts of other information. With thoroughly understanding of the recommendation algorithm principles, we can better figure out the ethical problems in TikTok and the solutions. The technical development of video recommendation system in TikTok is explained by how specific users and videos are matched to each other.

In the blog posted by TikTok on June 18, 2020, "How TikTok Recommends Videos #ForYou." According to this blog, a number of factors are used in the recommendation system, some are not just considered equally important but individually weighted by TikTok's For You algorithm. The specific factors, according to TikTok, are listed below [21]:

1. User interactions such as the videos you like or share, accounts you follow, comments you post, and content you create.

2. Video information, which might include details like captions, sounds, and hashtags.

3. Device and account settings like your language preference, country setting, and device type. These factors are included to make sure the system is optimized for performance, but they receive lower weight in the recommendation system relative to other data points we measure since users don't actively express these as preferences.

According to TikTok,

All these factors are processed by our recommendation system and weighted based on their value to a user. A strong indicator of interest, such as whether a user finish watching a longer video from beginning to end, would receive greater weight than a weak indicator, such as whether the video's viewer and creator are both in the same country. Videos are then ranked to determine the likelihood of a user's interest in a piece of content and delivered to each unique For You feed [21].

This blog, to some extents, provides transparency to the public and let people know more about the recommender system. Using the factors listed above, TikTok recommend the content people love. If one clicks on a singing video, TikTok will customize according to this video and the mechanism will analyze

further to trace users' future behaviors and offer precise recommendations. Without any initiative selection or search, the user passively accepts the system's personalized recommendation content [19].

2.2.2 How videos get featured

From the aspect of how videos get featured, the platform's algorithm can be explained in another way. First, the algorithm needs draw user's persona and understands the users well, even better than themselves do. By sorting the user's information and classifying the user according to their interest characteristics, the algorithm can deliver the videos accurately according to users' persona and the classification, which is User Clustering formed by the algorithm. Second, once a video is uploaded, TikTok will automatically extract features of videos in the context of recommender systems. It first exploits critical engineered features that would be derived and feed-in into the recommendation system [20]. After this, by showing it to a small number of users who are classified to be more likely to engage with it. This process of choosing a small subset of users is based on the users' persona and the likelihood of a user's interest in this piece of content. If these users watch the video in full, share it, give a like or follow the creator, which means they respond favorably, the video will be shown to more people that is also considered to have the similar interests. That same process then repeats itself in just a few seconds, and if this positive feedback loop happens continuously, the video can go viral in TikTok [22]. This process is done almost completely by the system's algorithm.

3. Analysis of the ethical problem using TikTok video dataset

3.1 Ethical problems caused by powerful algorithm in TikTok

"A Global Interest Discovery Recommendation Method and Device" is a patent ByteDance applied for how the company systematically recommends contents to users who are interested in them. As mentioned before, the powerful algorithm adopted by the recommendation system help to build a hi-

erarchical interest label set according to the topic of the content, and calculate the correlation degree of each interest label in the set in order to customize videos to the users [19]. Therefore, with this powerful algorithm, TikTok successfully attract many users and creators.

Mr. Han (translated by Erik Butler) wrote in his book: “Under the de-medialization trend in the digital media era, users are no longer just passive receivers and consumers of information, but also active senders and producers” [23]. He also said that the shitstorm represents an authentic phenomenon of digital communication [23].

People now finds TikTok not only a place to receive information, but also a platform for them to enjoy the right of expression. Even ordinary people can speak out, show their life, use this free, easy-to-use application to record everything. Therefore, with more senders and producer, TikTok is now covering all area of interests with its miscellaneous content materials. When people want to see current news, they open TikTok rather than newspaper; when they want to learn special skills, such as cooking, they open TikTok rather than books; when they want to have rest, they open TikTok, too, to watch mini self-made sitcoms, rather than open the television.

With such a huge content material on TikTok platform accurately match the corresponding audience, interesting things can quickly reach to the audience and can be spread in just a few minutes. Different from the connection system in real life between circles acquaintances, friends, and families, information in TikTok is spread without this limitation, without the narrow circle in people’s real life. People have the opportunities to get access to any information they want, and, for the same, information can be spread to anyone who may be interested in it. However, TikTok not only creates the platform that dilutes the significance of social connection, but also makes it a way for malicious people to take advantage. Although there is the sophisticated and logical algorithm, which is the foundation of recommenda-

tion system, and the model used to check the content of videos, there must be unethical things being spread. TikTok embrace all the interests of users and video creators, so people are likely to be interested in and misled by fake information, objectionable content, and inappropriate material.

Therefore, not only efficient extraction of the information but also a high-quality review system is a necessary requirement. However, the review of manual editors cannot be sufficient for the rapidly changing content and a myriad of new information. With insufficient manual review, the videos that is ethically wrong or the videos with misleading information will easily get spread, even in a few seconds. It is difficult for TikTok to maintain a healthy community and eco-system with simply today’s algorithm. User-generated images and videos can harm the platform and offense the users.

For example, a disturbing video of suicide was reportedly live-streamed on TikTok. The video first showed up on other social media platforms and is re-uploaded and went viral in TikTok quickly. It was not long before people starting to post the graphic video of a man live-streaming his suicide hidden in innocent contents to shock other users. Malicious people are now hiding the video in other videos to make it look like just a normal video or anything else [24]. They share it in comments, hide it in other contents, and edit it into the innocent dog and cat videos. Since this suicide video is causing negative repercussions on millions of users, the application is flooded with warnings about this viral trend, with many people warning others to exit any video pop up showing a man with long hair and a beard. Some also suggested to stay off TikTok to avoid seeing these disturbing and dangerous contents.

Theo Bertram, the European director of public policy of TikTok responded to this and said: “Through our investigations, we learned that groups operating on the dark web made plans to raid social media platforms including TikTok, in order to spread the video across the internet. What we saw

was a group of users who were repeatedly attempting to upload the video to our platform” [25].

Just like this suicide video, many other videos with inappropriate content have spread on TikTok once. Although some are removed quickly, there are many cases that the content causes a negative impact and even breaks the law. This “spreadability” is both the key draw and an ethical problem of TikTok. By facilitating the share of a new video based on the users who previously watched it and respond favorably, it is impossible for platforms to moderate the video as soon as it is posted and stop the spread in time. Users will continue to post disturbing and inappropriate content, including inappropriate language (discriminatory, racist, insults, swear words...), objectionable content, and explicit or suggestive content. There are many other cases that this kind of inappropriate material is hidden in normal contents and becomes viral in TikTok.

3.2 TikTok data analysis

3.2.1 TikTok video dataset

Analysis of the TikTok video dataset can help to get a deeper understanding of the recommendation system and extract useful and unknown patterns, relationships, and information in this recommendation algorithm. The dataset [26] used in this paper includes 165111 comments in TikTok collected from 2020.04 to 2020.10. The information included in the dataset is all the 165111 comment ID, comment text, video ID,

create time, like count, status, author unique ID, author nickname, whether the author is private, author language, author signature, author custom verification, author UID, author second UID, author avatar thumb, author region, author ins id, author youtube channel title, author youtube channel id, and author Twitter id. This TikTok dataset is worldwide data, but most in the US. The aim of using the data analysis is to show the basic rule in recommendation system. Therefore, it can support that hiding the inappropriate material into popular hashtags and terms can make it spread quickly.

3.2.2 Methodology

The data processing method is the use of Apache Spark in Python. Apache Spark is an open-source, distributed processing system that utilizes in-memory caching, and optimized query execution for fast analytic queries against data of any size [27]. It provides development APIs Python [27]. Using Apache Spark and Python, an open-source, distributed processing system used for big data workloads, some analytics about Tik Tok videos and comments are extracted by formulating a series of queries based on the data contained within the dataset.

3.2.3 Data processing

In the dataset, the text of comment and the like counts are collected and the top 5 most adopted terms by each author are depicted using the TikTok dataset and are shown in table 1.

Table 1. – Top 5 most adopted terms by each author (only showing top 10 rows)

author.unique_id	TOP5WORDS
the_stomping_lan_chop	['Hahahahahahaha!!!!']
jxst_vxbing	['😂😂😂']
mahakhamidd	['lying', 'tho', 'not', 'he']
noneyabusness1994	['he', 'a', 'Tesla', 'Can', 'get']
taylorjwoodward	['but', 'where', 'his', 'is', 'Tesla']
thomasmalcolm04	['@willyg04']
jillpickle31	['a', 'could', 'How', 'keep', 'you']
salma.lazraq	['not', 'him', 'getting', 'or', 'a']
someonekami17	['😂😂😂']
zacharywiseman21	['have', 'ate!!', 'David', 'for', 'a']

people sleepy. According to the US Food and Drug Administration, the health problem caused by too much medication includes serious heart problems, seizures, coma or even death [31].

B. Inappropriate video content

In addition to the text appear in the TikTok video and comment, the video content will also contain inappropriate things, which are even more difficult to be detected. For example, there was another trend in TikTok called the Silhouette challenge. In this trend, many female users will pose sexy in the video, and some will dance to the music rhythm. The mysterious shadow that seduces and cleverly shows off the body lines is shown with the backlit red light [31]. In some videos, underage users come across swearing, scantily clad adults, and suggestive dancing in TikTok. Many people even take advantage of this trend to edit the clip to hunt for nude images from TikTok's underage users.

Since these videos seem to have no association with nude trading, TikTok has failed to moderate them. For instance, 'Tradefortrade', a hashtag once spread in TikTok, connects to a profile with a male stroking themselves suggestively [32]. A comment on that post is, in fact, associated with trading nude images. One user's video asked for nude photos. Finally, the accounts are found to be connected to a community with lots of accounts centred around for nudity [32].

Additionally, there are some popular hashtags linked to dangerous behaviors like self-harm, suicide, and jumping off a building. Even choosing to filter out inappropriate content, one option in TikTok's Digital Wellbeing feature, cannot eliminate all the dangerous things. The filters can never catch everything since the hashtags change frequently with new creative spelling and hidden words or images can easily make the inappropriate content bypass filters. Additionally, hashtags can be edited to target a particular community, attributed to that the hashtag can help the videos send to people who have shown interest in it. When people swipe and watch the videos on the For You Page, the hashtag with specific

words or content gradually reaches the particular community.

3.2 Challenges when solving algorithm bias

There are reigniting debates about what TikTok need to do when facing the ethical problem. It is critically important for TikTok to suppress content that involves violence, scamming, pornography, flatulence, and weighing in facts, high-quality content like news [20]. To reach this goal, TikTok is trying to create a border control frame needs to be defined beyond quantifiable model objectives (Content Audit system) [20]. However, challenges exist when solving algorithm bias. Rather than the quantifiable objectives in the recommendation system, such as click-through rate, reading time, likes, comments, and reposts, there are many things cannot be easily measured. This huge chunk of factors that cannot be easily evaluated or checked by the system cause many challenges when solving algorithm bias. This is due to that these intangible objectives, such as the content without hashtags or clear signals, cannot be identified by quantifiable indicators.

3.2.1 Content moderation

As mentioned before, one of the main challenges is the insufficient of content moderation, both text moderation and video moderation. Content moderation is the most used method to remove inappropriate content in recommendation system. However, content moderation is hard and cannot be perfect. TikTok must put a lot of work into making sure the algorithms catch any objectionable content. Content moderation includes pre-moderation and post-moderation. This means the content can be vetted before being posted or can be analyzed shortly after being posted.

Content moderation should not harm user interactions. When choosing to pre-review or post-review, TikTok needs to consider that having user videos and comments appear quickly is important to ensure a lively exchange of ideas and opinions, which is the social media platform's ultimate goal. If the likelihood of receiving problematic content is low, post-moderation may suffice for most of the

content [33]. Nevertheless, in TikTok, the likelihood of receiving problematic content is high.

Without completely pre-moderation, there are chances for malicious people to pose the worst content. Even with adequate post-moderation, either by using machine learning systems, or reviewing by human moderators, the worst content will get chances to be spread in ten seconds, ten minutes, or ten hours. Even within a short period, the power of spreadability cannot be alleviated. Cutting the video, editing it within seemingly harmless material, putting filters on it, or distorting the audio are all ways to spread the content [34]. Millions of variations for each unwanted word, including letter replacements, character insertions, phonetic variations, leetspeak, obfuscation, hyphenation, special characters, embeddings, etc, are hard to be eliminated. The possibility may be small, but once an inappropriate video gets spread, the impact is big. Machine learning and algorithm is now advancing but it can never be perfect, even in the future [35].

Therefore, post-moderation alone will not be timely and effective. The hybrid model is adopted to solve the problem to some extent. Ideally, this model includes software that blocks videos or comments from the moment they first appear online until reviewers evaluate them and make a final decision on whether to remove them. Yet even if the software is used to highlight questionable posts, TikTok should aim to read all videos and comments. There are not completely fixed and uniform standards to measure, and even personal feelings will be involved sometimes. Many people have said that the same video is deleted at first, but when posting it again, the video successfully pass verification.

3.2.2 No requirement for the video to go viral

As mentioned in the previous part, the recommendation system features the videos based on the users' interests. This means that a mass number of followers is not the requirement for the video to go viral. Although TikTok will ban the users who have spread inappropriate content in the past. There are ways that

new users also spread these things. As TikTok has moved from local to global, more users are present, a larger possibility of having malicious people is present, and a bigger influence the content of videos will have.

The auto-play recommendation system in the For You Page making it easy to let thousands of users watch the video in just a few minutes. Without the limitation of based on who you follow, how many people you are following, TikTok finds it far more difficult to solve the ethical problem than other news feeds and social networking. For the users, after watching a video contain inappropriate material or disturbing content, the For You Page will start to recommend more to them again and again, even when they do not leave a single comment, interact with others, or follow the creator.

Furthermore, live streaming, a way to simultaneously record and broadcast in real-time to all the viewers around the world, is even harder to moderate. With an internet enabled device, such as a smartphone or tablet, and the broadcast platform, TikTok, almost any users can share anything in live streaming, which means the viewers will be exposed to any content the live streamer shows. The content in live streaming is even more difficult to moderate than videos and comments.

3.2.3 Migrate across to other platforms

As mentioned before, the suicide video once went viral in TikTok. The representative of TikTok once responded to the ethical problem:

"Our systems have been automatically detecting and flagging these clips for violating our policies against content that displays, praises, glorifies, or promotes suicide. We are banning accounts that repeatedly try to upload clips, and we appreciate our community members who've reported content and warned others against watching, engaging, or sharing such videos on any platform out of respect for the person and their family" [12].

With just a few taps users and reviews can remove the video showing disturbing content. Videos can be re-uploaded easily and migrated across other platforms quickly. To minimize the chance of detection

by the content moderation machine learning system, once the video is uploaded, bad actors are prone to download it, modulate it, and share it across multiple platforms like Reddit, Instagram, Facebook, and more [18]. It can very quickly become a cross-platform problem affecting many social media sites and thus millions of people around the world [35]. When the inappropriate videos are out in the wild and shown in other platforms, it will far more difficult for TikTok to remove them completely. Some people even move onto another app after making contact in TikTok and create bigger community to do illegal things. In this case, it will be hard to not only avoid the footage on For You Page when people scrolling through it, but also limit the spread of videos in only one platform.

4. Data governance and management solution

4.1 Solutions and effect

TikTok has announced a set of Community Guidelines to strengthen its existing policies focusing on the community's well-being and relieve social pressure. In the Community Guidelines, TikTok states how it processes inappropriate content and some other reactions to the ethical problem.

According to TikTok

We will remove any content – including video, audio, livestream, images, comments, and text – that violates our Community Guidelines. Individuals are notified of our decisions and can appeal if they believe no violation has occurred. We will suspend or ban accounts and/or devices that are involved in severe or repeated violations; we will consider information available on other platforms and offline in these decisions. When warranted, we will report the accounts to relevant legal authorities [36].

This Community Guidelines made it clearer about what kinds of content is banned in TikTok, but many users and reports were still criticizing TikTok, on the ground that this do not completely solve the problem. On February 24, 2021, TikTok published its latest transparency report. The report covers from the beginning of July to the end of December. According to this report, 89132938 videos were removed globally for violating the Community

Guidelines. Moreover, TikTok states that it identified and removed 92.4% before a user reported them, 83.3% before they received any views, and 93.5% within 24 hours of being posted [37].

But removing videos like this will never completely ensure the health of the community. As stated in the report, there is no way to delete videos every single time. We can't deny TikTok's efforts and changes, but we can't completely solve the problems and challenges mentioned above. What TikTok needs is to make sure that they can't send the video out and spread it before it has an impact.

Another way to address the problem is to protect vulnerable groups, especially the underage users. TikTok has long had a series of protection and prohibition measures for underage users and children under the age of 13. For TikTok, it is also a way to deal with social criticism and skepticism. Children and adolescents are very vulnerable to the inappropriate content in TikTok. According to TikTok, registration for users who are below the age of 13 is not allowed.

In its Community Guidelines, TikTok claims "users must meet the minimum age requirements to use TikTok, as stipulated in our Terms of Service. When underage account holders are identified, TikTok will remove those accounts" [36]. It's true, however, that TikTok's primary audience is children (including those under 13) and teenagers. The fact is that it is easy for children and teenagers to register since the age verification process is so relaxed that only self-declaration is required [38]. Of the viral video-sharing app's 49 million daily users in the United States in July, about 18 million were 14 or younger, according to data obtained by The New York Times [39].

4.2 Grow in strict supervision

Under regulatory scrutiny, TikTok is facing threats of being banned in the US and many other countries as for data privacy and immoral content. TikTok was temporarily banned in Indonesia in 2018 and in Pakistan in 2020 [40]. In recent years, some countries or regions, especially the European Union, have also stepped up supervision and regulation of social me-

dia platforms in many aspects such as data protection and cyber hazards. Other countries have also imposed heavy fines and even prosecution on social media executives to effectively remove harmful content [41]. For example, in 2019, Sharing of Abhorrent Violent Material Act is passed in Australia. Technology executives can be sentenced to up to three years in prison and fined up to 10 percent of a company's global turnover [41]. The development of these legislation has put pressure on platforms to become more self-policing in terms of content censorship. Rather than ban TikTok, the regulatory plan is more efficient in that it offers a long-term, rules-based approach to addressing concerns about this social media platform.

5. Conclusion

5.1 TikTok data ethical dilemmas are inevitable

The problem of "immoral" content raises a kind of difficulty for the freedom of speech and is a dilemma that can hardly be solved [42]. The TikTok data ethical dilemma is that the Internet makes it easy for bad actors, ranging from trolls to spammers to malicious hackers, to deter or frustrate speech within online channels [43]. Nowadays, there are many problems and challenges in solving ethical problems. Even though both Tiktok and the government have taken steps to ensure the safety of users and the well-being of the community, there are still some potential problems. Many measures are being taken to mitigate the problem as much as possible, but they are fundamentally unavoidable. As long as there are people on Tiktok who want to take advantage of bad information or profit from it, there will be more ways for them to circumvent censorship and laws, and there will always be a certain chance that the recommendation system will spread the content, because enough hiding can make it past the screening and surveillance system. In general, the existence of these moral problems is undeniable and inevitable. The TikTok data ethical dilemma now becomes a permanent social problem sewn into the logic of the Internet [42]. Minimizing harm and potential problems is the best that Tiktok, the government, and users can do.

5.2 Three ways can help

Although the TikTok ethical dilemma is inevitable, there are three ways to help.

Firstly, for the users, protecting themselves is of great importance since it is the most direct and efficient way to maximize their security and privacy. Ensuring the TikTok is private or not allowing other users to find you and interact with you essentially means that others cannot follow, like, comment on, duet with you, find your account via a search engine, access to your videos and likes without approval. This can be done easily by turning the option on in the Safety section of the Privacy page, attributed to that TikTok share users' contents by featuring them on the For You Page by default. By blocking interactions and preventing others from viewing your content, the possibility that you receive inappropriate content or harassment under your videos and messages will be lower. Additionally, using restricted mode for children's accounts is another effective way. Restricted mode which stops inappropriate content cannot eliminate all the contents that are not suitable for underage users, such as false or misleading information, violence, sexually explicit material, hateful or offensive material, but it can help to some extent to alleviate the problem of seeing those dangerous, misleading content.

Secondly, government plays an indispensable role in managing social media. As mentioned before, banning social media is not the primary solution. For instance, since 2020, the UK government has outlined new powers for the media regulator Ofcom to regulate social media platforms [44]. Rather than rely largely on the self-governance of those platforms and let them make their own rules about removing inappropriate content, the government now forces them to ensure the harmful and objectionable content is removed in the first place. This means companies are responsible for this as well rather than only the person who posted it is at risk of prosecution [44].

Finally, it is a matter of fact that social media companies should be responsible for the inappropriate content and should aim at removing all the contents

that will potentially cause negative repercussions. In sharp contrast, the users, those bad actors in particular, won't get rid of their responsibility. The anonymity and free speech should not be the excuses for users to say whatever they want. Not only do social media companies need to be regulated, but the public should also be limited by laws. Legal restriction on content censorship standards is the prerequisite of eliminating the inappropriate contents but not removing them.

5.3 Further study aspects

The study suggests that future research should focus on how social software and short video apps deal with inappropriate information. Nowadays, many ethical issues are unavoidable, and there are still many potential problems and challenges. How to deal with and deal with these problems from the aspects of the software itself, government regulation, self-protection, and so on is very important. At the same time, it is hoped that there will be more profound research on how to improve the video rec-

ommendation system. Through such research, this software can better bring positive effects to users and eliminate inherent hidden dangers and dangers. At the same time, for a better network environment as well as the protection of teenagers' network use

The study suggests that there should be more research into the moral problems of TikTok in addition to its inappropriate content. It cannot be denied that TikTok has contributed in many ways and brought convenience and entertainment to people. But criticism of TikTok has been widespread across platforms and places. Only by further studying the remaining aspects of the problem, can we finally achieve the optimization of this social software as far as possible.

The study suggests more research on other social software in the future. Other social software also has some ethical issues to some extent. More research can be done on the problems that exist in other aspects as well as TikTok. Only by discovering these problems can we better improve them.

References:

1. Y. What Are Computer Algorithms, and How Do They Work? How-To Geek. (2016. September 28). URL: <https://www.howtogeek.com/howto/44052/htg-explains-what-are-computer-algorithms-and-how-do-they-work>
2. Mosaic A. K. The Powerful Equations That Explain The Patterns We See In Nature. Gizmodo Australia. (2020. August 25). URL: <https://www.gizmodo.com.au/2014/08/the-powerful-equations-that-explain-the-patterns-we-see-in-nature>
3. Sweigart A. How Does Compression Work? – The Invent with Python Blog. The Invent with Python Blog. (2012. August 17). URL: <https://inventwithpython.com/blog/2012/08/17/how-does-compression-work>
4. Note M. Managing Image Collections: A Practical Guide (Chandos Information Professional Series) (1st ed.). Chandos Publishing. (2011).
5. Tragakes E. Economics for the IB Diploma with CD-ROM (2nd ed.). Cambridge Univ. Press. (2011).
6. Cahn A., Alfeld S., Barford P., & Muthukrishnan S. An empirical study of web cookies. In: Proceedings of the 25th international conference on world wide web. 2016. – P. 891–901. International World Wide Web Conferences Steering Committee.
7. Ghose A., & Yang S. An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management Science*, – 55(10). 2009. – P. 1605–1622.
8. DeAngelis S. E. F. S. Artificial Intelligence: How Algorithms Make Systems Smart. (2015, August 7). WIRED. URL: <https://www.wired.com/insights/2014/09/artificial-intelligence-algorithms-2>
9. Gibbs S. Gmail does scan all emails, new Google terms clarify. The Guardian. (2020, July 1). URL: <https://www.theguardian.com/technology/2014/apr/15/gmail-scans-all-emails-new-google-terms-clarify>

10. Kavenna J. Shoshana Zuboff: 'Surveillance capitalism is an assault on human autonomy.' *The Guardian*. (2019, October 29). URL: <https://www.theguardian.com/books/2019/oct/04/shoshana-zuboff-surveillance-capitalism-assault-human-autonomy-digital-privacy>
11. Finley K. Wanna Build Your Own Google? Visit the App Store for Algorithms. *Wired*. (2014 b, August 11). URL: <https://www.wired.com/2014/08/algorithmia>
12. Congressional Research Service. TikTok: Technology Overview and Issues. (2020, January). URL: https://www.everycrsreport.com/files/2020-10-01_R46543_657b0e11b8b8cc8e9dfa307a961025825318883e.pdf
13. D. What is TikTok? Why Is It So Popular? *Brandastic*. (2020, November 27). URL: <https://brandastic.com/blog/what-is-tiktok-and-why-is-it-so-popular>
14. TikTok Revenue and Usage Statistics (2021). (2021, August 4). *Business of Apps*. URL: <https://www.businessofapps.com/data/tik-tok-statistics/>
15. Blacker A. Worldwide & US Download Leaders 2020. *Apptopia*. (2021, January 7). URL: <https://blog.apptopia.com/worldwide-us-download-leaders-2020>
16. Geysler W. TikTok Statistics – Revenue, Users & Engagement Stats (2021). *Influencer Marketing Hub*. (2021, July 8). URL: <https://influencermarketinghub.com/tiktok-stats>
17. Mohsin M. 10 TikTok Statistics You Need to Know in 2021 [March data]. *Oberlo*. (2021, July 1). URL: <https://www.oberlo.com/blog/tiktok-statistics>
18. Weimann G., & Masri N. Research Note: Spreading Hate on TikTok. *Studies in Conflict & Terrorism*. Published. (2020). URL: <https://doi.org/10.1080/1057610X.2020.1780027>
19. Zhao Z. Analysis on the “Douyin (Tiktok) Mania” Phenomenon Based on Recommendation Algorithms. (2020). *E3S Web of Conferences* 235, 03029 (2021). *NETID2020*. Published. URL: <https://doi.org/10.1051/e3sconf/202123503029NETID2020>
20. Wang C. Why TikTok made its user so obsessive? The AI Algorithm that got you hooked. *Medium*. (2020, June 7). URL: <https://towardsdatascience.com/why-tiktok-made-its-user-so-obsessive-the-ai-algorithm-that-got-you-hooked-7895bb1ab423>
21. T. How TikTok recommends videos (2020b, November 5). #ForYou. *Newsroom | TikTok*. <https://newsroom.tiktok.com/en-us/how-tiktok-recommends-videos-for-you>
22. Matsakis L. How TikTok’s “For You” Algorithm Works. *Wired*. (2020, June 18). URL: <https://www.wired.com/story/tiktok-finally-explains-for-you-algorithm-works>
23. Han B., & Butler E. *In the Swarm: Digital Prospects (Untimely Meditations)*. The MIT Press. (2017).
24. Warnock C. TikTok Works to Take Down Disturbing Suicide Video as Users Share Warnings. *Heavy*. *Com*. (2020, September 8). URL: <https://heavy.com/news/2020/09/tiktok-viral-suicide-video>
25. Associated Press. TikTok says coordinated attack behind suicide clip uploads. *Washington Post*. (2020, September 22). URL: https://www.washingtonpost.com/business/technology/tiktok-says-coordinated-attack-behind-suicide-clip-uploads/2020/09/22/5c0b9a98-fcdb-11ea-b0e4-350e4e60cc91_story.html
26. Hosn J. TikTok Video Comments – David Dobrik’s top videos: Scraped video comments on David Dobrik’s top videos on TikTok. *Kaggle*. (2020).
27. Databricks. Apache Spark™ – What is Spark. (2020, April 14). URL: <https://databricks.com/spark/about>
28. Leander K. M. Critical literacy for a posthuman world: When people read, and become, with machines. *British Educational Research Association*. (2020, July 1). URL: <https://bera-journals.onlinelibrary.wiley.com/doi/abs/10.1111/bjet.12924>

29. Statista. TikTok MAU user ratio in the U.S. 2019, by age group & gender. (2021, June 7). URL: <https://www.statista.com/statistics/1095196/tiktok-us-age-gender-reach/#statisticContainer>
30. Benadryl. Drugs.Com. (2020, December 8). URL: <https://www.drugs.com/benadryl.html>
31. Nguyen B. Don't let your kids try these 9 dangerous TikTok trends. CyberPurify. (2021, August 19). URL: <https://cyberpurify.com/knowledge/9-dangerous-tiktok-trends>
32. Cox J. TikTok, the App Super Popular With Kids, Has a Nudes Problem. Vice. (2018, December 6). URL: <https://www.vice.com/en/article/j5zbxm/tiktok-the-app-super-popular-with-kids-has-a-nudes-problem>
33. Pre- and post-moderation. (2019, June 18). Ontheline. URL: <https://newsrooms-ontheline.ipi.media/measures/pre-and-post-moderation/>
34. Productions B., Productions B., Productions B., Productions B., Productions B., Productions B., & Productions B. Blog: ANALYSIS – TikTok suicide video: it's time platforms collaborated to limit disturbing content – AdNews. (2020, September 9). Best Soundcloud Rappers 2019. URL: <https://btoktiktok.com/2020/09/09/blog-analysis-tiktok-suicide-video-its-time-platforms-collaborated-to-limit-disturbing-content-adnews>
35. ABC News. TikTok suicide video: It's time social media platforms collaborated to limit disturbing content. (2020, September 9). URL: <https://www.abc.net.au/news/2020-09-09/why-so-hard-tiktok-remove-disturbing-content-suicide-video/12643832>
36. TikTok Community Guidelines. (n.d.). TikTok. Retrieved August 21, 2021. From URL: <https://www.tiktok.com/community-guidelines?lang=en>
37. A. (2021). TikTok Transparency Report 2020 H1. TikTok. <https://www.tiktok.com/safety/resources/transparency-report-2020-1?lang=en>
38. BEUC The European Consumer Organisation. (2021). TikTok without filters. BEUC The European Consumer Organisation. Published. URL: https://www.beuc.eu/publications/beuc-x-2021-012_tiktok_without_filters.pdf
39. Zhong R., & Frenkel S. A Third of TikTok's U.S. Users May Be 14 or Under, Raising Safety Questions. The New York Times. (2020, September 18). URL: <https://www.nytimes.com/2020/08/14/technology/tiktok-underage-users-ftc.html>
40. Wang D. J. From Banning to Regulating TikTok: Addressing concerns of national security, privacy, and online harms. The Foundation for Law Justice and Society. (2021). Published. URL: <https://www.fljs.org/sites/default/files/migrated/publications/From%20Banning%20to%20Regulating%20TikTok.pdf>
41. Criminal Code Amendment (Sharing of Abhorrent Violent Material) Act 2019. (2019). Federal Register of Legislation. URL: <https://www.legislation.gov.au/Details/C2019A00038>; URL: <https://www.kaggle.com/jhosn13/tiktok-video-comments-david-dobriks-top-videos>
42. Langvardt K. Regulating Online Content Moderation. The Georgetown Law Journal. (2018). Published. URL: <https://www.law.georgetown.edu/georgetown-law-journal/wp-content/uploads/sites/26/2018/07/Regulating-Online-Content-Moderation.pdf>
43. Lidsky L. B. Public Forum 2.0. UF Law Scholarship Repository. (2011). Published. URL: <http://scholarship.law.ufl.edu/facultypub/155>
44. Team B. R. C. Social media: How do other governments regulate it? BBC News. (2020, February 12). URL: <https://www.bbc.com/news/technology-47135058>