**Assignment: Investigating and Analysing Ethical Risks of a Digital Engagement application**

**Section 1: Description of application**

TikTok is a social media platform with over 4.5 billion downloads that allows users to create, share and view a variety of short videos, ranging from entertainment and sports news to advice (D’Souza, 2023). The application began as a service where users would share videos of themselves dancing and lip-syncing to music, but it has since branched out to where users can find anything from information on current affairs, individuals advertising their brick-and-mortar businesses and even politicians outlining their policies and stances on contentious topics. It is known for its addictive quality and high engagement and utilises an Ai recommendation algorithm to achieve these qualities. TikTok uses a self-training Ai engine to recommend videos to users to maximise engagement and these recommendations only become more accurate over time. This Ai engine collects information such as the user’s name, age, location, activity on other websites accessed through TikTok and cookies that collect, measure, and analyse which web pages users view most often and how they interact with them (La Jeunesse, 2023). The recommendation system then selects videos from a collection of eligible content and ranks them on how likely the user is to be interested in them (La Jeunesse, 2023). These ranked videos then appear on the user’s unique “For you page”. According to TikTok themselves, the videos that appear on a user’s “For you page” are influenced by three factors: user interactions, video information and user information. User interaction covers topics such as videos the user likes, shares, or comments on, videos the user watches in full or skips and accounts the user follows or is followed by. Video information deals with sounds, hashtags, video views and country of origin. User information considers device and account settings, language preference, user’s country of origin and time zone and device type. A large weight is placed on the user interaction characteristic of watch time. The Ai recommendation algorithm uses initial demographic information to find videos that may potentially interest the user and over time refines the topics down, through self-learning, depending on what types of videos the user watches in full or skips (support.tiktok.com, n.d.)

TikTok’s Ai recommendation algorithm offers benefits to a multitude of parties. Firstly, users benefit as they receive a service that is more likely to keep them engaged. Due to the self-training engine the longer users spend on the service the more engaging the videos are to them. This provides an entertainment factor that users benefit from. Secondly, the corporation itself benefits from the recommendation algorithm keeping users engaged. This results in users spending large quantities of time on the application and available to be subject to advertisements. This is especially important when you consider TikTok’s advertising revenue was $13.2Bn in 2023 alone (Mosby, 2024). The corporation receives large benefits from its Ai recommendation algorithm as it allows them to maintain this ad revenue by keeping users engaged and even allows for the increase in revenue over time due to an increase in user engagement as the recommendations become more accurate, meaning users may spend longer on the application. Thirdly, the information gathered by TikTok’s Ai recommendation algorithm greatly benefits advertisers. The Ai model collects information on user demographics and goes into fine details as mentioned previously, e.g., user’s device and location. This allows advertisers to target their ads to a specific demographic, making their ads more effective while also preventing waste due to the advertiser being able to target solely their target demographic. Finally, the creators themselves benefit from TikTok’s Ai recommendation algorithm. The Ai system helps deliver the videos that creators have shared to the users that are most likely to be engaged with them. This acts as an efficient system for creators to be matched with their target market, leading to views on videos and followers for the creator. Creators benefit from this due to the Creator Fund operated by TikTok, acting as a compensation method for creator’s work. The Creator Fund pays out values ranging from $0.02 - $0.04 per 1000 views for smaller creators and $0.50 - $1.00 for larger creators (McLellan, 2023). As the recommendation algorithm, pushes videos out to the users most likely to be engaged with them, this allows creators to maximise viewership which in turn maximises revenue.

**Section 2: Identification of shareholder roles involved in the application and its governance**

TikTok as an application has many stakeholders that are involved in its development, use and governance. TikTok, itself, is the Ai developer, an Ai user, and an oversight authority. TikTok’s Ai recommendation algorithm was developed by its engineering team in partnership with parent company Byte Dance. TikTok benefits from developing its own Ai system as there is no recurring licencing fee for usage and it is tailored specifically to the application. Furthermore, TikTok is an Ai user as it utilises its Ai recommendation algorithm as the base for the successful functioning of the application. Without its Ai system, TikTok would fail to maintain the same level of user engagement which is so crucial for its advertising revenue. In addition to these roles, TikTok must operate as an oversight authority through self-regulation of some aspects. In the EU alone, TikTok must self-regulate to ensure that they are operating in line with EU laws such as GDPR and the upcoming AI act. Failing to self-regulate would lead to penalties for TikTok and the potential inability to access the EU market.

The government, both national and EU, act as oversight authorities for applications like TikTok. These authorities enforce data protection laws and Ai regulations to safeguard their constituents. Branches of government concerned with regulation of applications like TikTok include the European Commission, European Data Protection Board, National supervisory authorities e.g., Irish Data Protection Authority, Courts, and Data Protection Officers. Examples of regulation can be seen through a European Commission code of conduct that requires major tech platforms to report misinformation on their networks and their methods of combatting it (O’Leary, 2023.), GDPR which concerns the processing of personal data and where violators can be fined €20 million or 4% of worldwide turnover and the proposed AI act which operates as a risk-based system prohibiting some forms of AI, making other forms permissible once in compliance with strict regulations ,e.g., notifying individuals they are interacting with AI or labelling deepfakes, and offers large penalties of €30 million or 6% of global turnover.

Users of TikTok operate as data providers as by using the application they agree to allow TikTok’s AI recommendation algorithm collect the previously mentioned information on them. This benefits the user by increasing the likelihood that they will view videos that they find engaging based on their physical attributes or interests. Furthermore, this benefits TikTok as they are provided with the necessary data that is required for the effective functioning of their AI system.

Advertisers are stakeholders of TikTok that act as AI info users. The information that TikTok’s AI recommendation system gathers separates individuals into demographics based on physical attributes and interests. Advertisers can then make use of this information to identify who they want to target based on their physical attributes e.g., location, or based on their interests. This allows advertisers to effectively run advertising campaigns and prevent waste as they only advertise specifically to their target market.

An often-overlooked stakeholder comes in the form of family, friends, and colleagues of users. These parties are stakeholders as they act as “affected stakeholders” even if they themselves are not users of TikTok. This is due to the sharing ability on TikTok that allows users to send videos to non-users on platforms such as Snapchat or WhatsApp. This allows non-users to benefit from viewing engaging videos they may not have been able to see before but also opens them up to potential ethical risks that they were separated from previously.

**Section 3: Identification of ethical risks**

There are a variety of ethical risks arising from TikTok as an application that are exacerbated by the platform’s AI recommendation algorithm. To effectively identify these ethical risks, it is necessary to examine them under the social responsibility categories of human rights, labour practices, the environment, fair operating procedures, consumer issues and community involvement and development.

Human Rights

Several ethical risks fall under the category of human rights. These risks pertain to civil and political rights and physical risks. The primary stakeholders faced with these risks are users and family, friends, and colleagues of users. Firstly, civil and political rights are subject to ethical risks due to TikTok and its AI recommendation algorithm. TikTok’s recommendation algorithm analyses videos to understand what users would be most engaged with the content but fails to analyse whether the videos are factual or not. This creates the potential for manipulation and distortion of reality to take place as users are placed in an echo chamber full of misinformation. Furthermore, this affects user’s family, friends and colleagues that receive these videos from the user and fall victim to manipulation and distortion of reality. The probability of this risk is high, and the severity of this risk is also high. The high probability can be seen through examples of fake political ads appearing often on TikTok as the AI recommendation system pushes these videos to those who will be engaged by the topic without considering whether the information is factual. For example, a watchdog in the US submitted fake political ads to TikTok and 90% of these ads made it to the general public (Ortutay, 2022). The severity of this risk is high due to the potential harms resulting from misinformation and distortion of reality. The outcomes range from fake smear campaigns that cause a political figures support to fall to more serious risks, such as riots as seen in the January 6th attacks on the US capitol building due to misinformation spread about the validity of the US presidential election. Thus, the severity is high due to the potential for electoral interference and even violence.

Furthermore, physical risks arise as a result of TikTok and its AI recommendation algorithm. In a similar manner to how TikTok’s AI recommendation algorithm analyses political videos, it fails to identify whether videos relating to physical and mental health are promoting factual information or not. This affects users and their family, friends, and colleagues due to the ability to share videos. The algorithm will simply push these videos to those who it believes will be engaged by the video and fails to consider whether the advice in the videos they are recommending is harmful or not. This creates the potential for individuals, especially at-risk individuals, to be fed misinformation about their physical and mental health and to be pushed towards not seeking medical advice or even self-harm. The probability of this risk is high, and the severity is equally high. There are countless examples of TikTok promoting harmful content related to physical and mental health. For example, technical research carried out by Algorithmic Transparency Institute and AI Forensics uncovered that between 3 and 20 minutes into their manual research that more than half the videos in their “For you page” were related to mental health struggles, with multiple videos romanticising and even encouraging suicide (Amnesty International, 2023). These harrowing statistics highlight the high probability of this risk. Furthermore, the probability is increased as the videos are not only limited to users but also their family, friends, and colleagues, who may not even have the application, due to the ability to share videos. Needless to say, the severity is high on this risk. TikTok’s AI recommendation algorithm is recommending content that promotes suicide and self-harm, potentially leading to deaths. The severity only increases when you consider the fact that majority of users on TikTok are teenagers, an age group that are quite impressionable and could be influenced to take harmful action due to a constant stream of these types of videos.

Labour Practices

There exists a potential ethical risk for the corporation in social responsibility area of labour practices. This risk relates to worker dissatisfaction and worker mental health and arises due to the effects of TikTok’s AI recommendation algorithm. There exists the potential that employees will become dissatisfied with their work and may even suffer with their mental health declining because of the harms, to civil and political rights and to physical health of user, being perpetuated by their employer through its AI recommendation algorithm. Employees want to work in a corporation that does not cause harms to the public and by TikTok failing to satisfy this, the organisation will suffer from dissatisfaction and declining mental health. The severity of this risk is medium; however, the probability is quite high. The high probability arises due to 76% of employees wanting to work in companies that are socially responsible (Cox, 2023). With over ¾ of employees concerned with working in a socially responsible business, the risk of dissatisfaction and mental health decline are quite high. The severity of this risk is medium, in comparison to the high risk under the human rights category, due to the potential for political interference and harm to or death of individuals. This severity is highlighted by a study stating that employees who report their employer being socially responsible, on average have better mental health (Allen, 2021). TikTok, due to acting socially irresponsibly, may cause its employees to become dissatisfied with their work and have deteriorating mental health, impacting the functioning of the corporation and quality of life of the employees. Thus, the severity is placed at a medium level as while it does cause great harm it does not lead to loss of life as the risk under human rights does.

The Environment

TikTok’s AI recommendation algorithm leads to ethical risk for a multitude of stakeholders under the social responsibility category of the environment. Firstly, there is an environmental risk due to the operation and monitoring of this AI system. As it is a self-training AI system, it is constantly training itself and harvesting new information and there is increased pollution associated with this. This affects the corporation as it is responsible for all emissions because of its AI system. Furthermore, the government, both national and EU, is affected due to the necessary monitoring of AI systems, like TikTok’s, to ensure that regulations are satisfactory in a field that sees constant development. This constant monitoring also leads to increased emissions for the government. The probability of this risk is high; however, severity is quite low but has potential to increase. Probability is high for both stakeholders due to the necessity of the operation of the AI system for the corporation and the monitoring for the government. It is inevitable that carbo emissions will result from these activities. The inevitable emissions of TikTok’s complex AI system are seen as TikTok emits 3.3 times more carbon emissions per minute of use than comparable websites like Facebook (L, 2023). The severity of the environmental risk of AI operation, development and monitoring is relatively low since AI currently only makes up 1% of global carbon emissions (Xing and Monck, 2023). However, there is potential for this severity to increase as AI models become more prevalent and TikTok eventually must replace or update its AI system due to increased competition, further leading to more emissions in the monitoring on behalf of the government. The potential increase in severity can be seen as the training of a single AI model can emit an estimated 284 tonnes of Co2 (Xing and Monck, 2023). It is clear to see how the severity of TikTok’s environmental impact can increase in future upon considering the emissions for a single AI model and how complex TikTok’s AI recommendation algorithm is required to be.

Fair Operating Procedure

TikTok’s AI recommendation algorithm causes potential risks under the social responsibility category of fair operating procedure. This is due to the ability to game the algorithm to undermine the political process. This risk affects users and their family, friends and colleagues, as affected stakeholders. It is possible for influence operations to occur in order to undermine the political process. Groups can excessively lobby, push out misinformation or even manipulative information to cause social divide and sway votes. Furthermore, deepfakes can be used to deceive individuals into thinking a political figure has made promises that in reality they have not. These influence operations are possible due to TikTok’s AI system recommending videos to users based on perceived engagement without considering the validity of these videos. This affects users as they are recommended these manipulative videos on their “For you page” but also affects family, friends, and colleagues of users as they may share the videos to them believing them to be accurate. The probability of this risk is high, and the severity is equally high. The high probability of influence operations to undermine the fair political process can be seen with numerous examples in US politics but also on a local level, where TikTok recently dismantled a covert influence operation targeting the Irish public with divisive content intended to intensify social conflict. This influence operation was only dismantled after amassing 94,743 followers and 2,440,995 views (O’Leary, 2023). The high severity relates to the potential for these influence operations to go undiscovered and to seriously undermine the fair political process by manipulating and providing misinformation to users and their peers. This severity is added to by the lack of fact-checking in TikTok’s AI system and how quickly it recommends these videos to millions of users, allowing influence operations to flood TikTok with videos that are spread rapidly before any validity can be established. This may result in extremist candidates being elected or others losing support due to false claims.

Consumer Issues

TikTok’s AI recommendation algorithm has created ethical risks under the social responsibility heading of consumer issues. These risks affect users primarily. The risks arise due to the ability to create videos on TikTok that promote a product or service but without any checks as to the validity of these claims. It is possible to create a video advertising a product full of fake claims and to include an external link to a website to buy this fake product. This is due to TikTok promoting videos to users the algorithm feels will be engaged by the video and users may feel safe to purchase the product as TikTok has recommended it to them by placing it on their “For you page”. The probability of this risk is at a medium level and the severity is also at a medium level. The probability being a medium level is due to the main users of TikTok being children and teenagers who may not be able to identify whether an ad is legitimate or not, however it is expected that a reasonable adult would be able to distinguish this. Furthermore, countless examples of fake or misleading ads can be found on TikTok, with one report highlighting numerous fake mobile apps, diet pills and other fake products and services appearing on the app (Wong, 2020). The medium level severity arises because users are being scammed out of their money by purchasing products that are not as advertised but there is no threat of death or immense financial harm. The products are often lower cost in comparison to if a fake car or house was being sold. However, there remains a health risk as products are unlikely to have undergone safety tests and some cosmetic items may contain toxic ingredients leading to a medium severity rating of the risk.

Community involvement and development

TikTok’s AI recommendation algorithm has created ethical risks for users and their family, friends and colleagues under the social responsibility heading of community involvement and development. This risk is due to the ability of the recommendation algorithm to spread misinformation and hateful views designed to tear communities apart. The algorithm pushes videos that it feels will garner engagement and fails to account for whether the information in a video is accurate or designed to spread hate. This poses a risk to community involvement and development as TikTok can perpetuate harmful views against community members causing a divide in the community. The probability of this risk is high, and the severity is equally high. The high probability can be seen in an ISD (Institute for Strategic Dialogue) report on anti-immigrant videos on TikTok highlighted how followers are not necessary for high view counts as of 20 accounts in the study, followers ranged from 15 to 10,000 but views ranged from 4,000 to 1,200,000 (ISD, 2023). Furthermore, the severity is high due to the possibility of divide within communities but also of violence due to hateful views and misinformation. An example being seen in a video of a migrant entering a hospital, with overlaid text accusing him of committing sexual crimes against children, parents, and staff, which led to the innocent man being identified and assaulted on the streets of Dublin (ISD, 2023).

**Section 4: Discussion of Mitigation Measures for Risk**

The most poignant risk arising from TikTok’s AI recommendation algorithm is responsible for many ethical risks under the social responsibility headings and is that of the recommendation algorithms inability to identify misinformation and the spread of hateful views within videos. This is a severe risk that affects a variety of areas from political rights, individual’s mental and physical health, and harmony within communities. Therefore, this is the risk that I believe is the most important to address.

My proposed mitigation measure involves collaboration from the government, as a regulatory body, the corporation itself and users. Firstly, the government must outline high-risk categories that require additional screening before being able to be shared with the public. A move which would reduce the effectiveness of influence operations by preventing the rapid sharing of misinformation and harmful views. These categories would cover the areas that have the potential to have the largest negative impact. Potential categories to be outlined could be mental and physical health, news stories, politics and deepfakes. Additionally, the government would need to establish an effective regulatory system to encourage compliance. Primarily, this involves government departments tasked with testing the levels of misinformation or harmful views or advice on each social media platform or the ability to spread misinformation on the site, work already being done by watchdog NGOs. The government could establish potential fines greater than those currently in place to act as a greater deterrence against poor self-regulation from TikTok. Furthermore, a rewarding system may encourage greater and stricter compliance. Rewards could consist of tax incentives for those social media companies that prevent misinformation and harmful views most effectively.

Secondly, the corporation plays a significant role once the high-risk categories are outlined. The mitigation measure involves the development of a similar AI model to that of the self-training recommendation system that would also scan the content of the videos but instead of using that information to recommend videos to users, it would flag these videos for additional checks to verify information or scan for potential harmful advice or views e.g., romanticising suicide or anti-migrant views. This helps mitigate the risk as all videos on high-risk topics undergo additional screening which prevents the rapid spread of misinformation and harmful views or advice. Furthermore, it is practical for the corporation as not every video would require additional screening but only those that are flagged as being high-risk by the AI model, allowing for the continuation of the recommendation algorithm while removing harmful videos at the source. I understand that the development of the screening AI model would be costly, but the hope is that the prevention of fines and the potential tax incentives outweigh this cost.

Finally, the user is necessary for the effective functioning of the self-training AI detection model. The corporation needs to provide an effective user reporting system where the user can report any videos they feel may have slipped through the cracks. These videos would then undergo additional screening. Further preventing the spread of misinformation and harmful views or advice as any videos that slip by the rigorous pre-upload checks can be reported by users before amassing large view counts. These mitigation measures from the government, corporation and users would help prevent the risks that are currently seen under many social responsibility headings.

**References**

Section 1

* D’souza, D. (2023). *TikTok: What It Is, How It Works, and Why It’s Popular*. [online] Investopedia. Available at: <https://www.investopedia.com/what-is-tiktok-4588933>. (Accessed: 29/02/2024)
* Mosby, A. (2023). *TikTok Ad Revenue From 2020 To 2025 (Reports & Analysis)*. [online] Yaguara. Available at: <https://www.yaguara.co/tiktok-ad-revenue/>. (Accessed 29/02/2024)
* McLellan, K. (2023). *How Much Does TikTok Pay You? (+Free TikTok Money Calculator)*. [online] The Leap. Available at: <https://www.theleap.co/blog/how-much-does-tiktok-pay/>. (Accessed 29/02/2024)
* ‌La Jeunesse, W. (2023). *Here’s the data that TikTok collects on its users*. [online] FOXBusiness. Available at: <https://www.foxbusiness.com/technology/heres-data-tiktok-collects-its-users>. (Accessed 01/03/2024)
* support.tiktok.com. (n.d.). *How TikTok recommends content | TikTok Help Centre*. [online] Available at: <https://support.tiktok.com/en/using-tiktok/exploring-videos/how-tiktok-recommends-content>. (Accessed 01/03/2024)

Section 2

* O’Leary, N. “TikTok Discovers “Covert Influence Operation” Targeting Ireland.” *The Irish Times*, 28 Sept. 2024, [www.irishtimes.com/technology/2023/09/28/tiktok-discovers-covert-influence-operation-targeting-ireland/](http://www.irishtimes.com/technology/2023/09/28/tiktok-discovers-covert-influence-operation-targeting-ireland/). (Accessed 06/03/2024)
* Lewis, D. (2024). ‘Ethical Digital Engagement: Lecture 2’ [PowerPoint Presentation] *TEU00062 :Engaging in The Digital World: Today and Tomorrow* (Accessed 06/03/2024)

Section 3

* Ortutay, B. (2022). *Report: TikTok bad at culling US election misinformation ads*. [online] AP NEWS. Available at: <https://apnews.com/article/technology-business-government-and-politics-misinformation-a18a35a2a2542eadfc3893983a818511>. (Accessed 11/03/2024)
* Amnesty International (2023). *Global: TikTok’s ‘For You’ feed risks pushing children and young people towards harmful mental health content*. [online] Amnesty International. Available at: <https://www.amnesty.org/en/latest/news/2023/11/tiktok-risks-pushing-children-towards-harmful-content/>. (Accessed 11/03/2024)
* Cox, J. (2023). *Despite Economic Headwinds, Employees Still Want To Work For Companies That Care*. [online] Forbes. Available at: <https://www.forbes.com/sites/josiecox/2023/02/15/despite-economic-headwinds-employees-still-want-to-work-for-companies-that-care/?sh=3db79e466d20>. (Accessed 11/03/2024)
* Allen, P. (2021). *Does working for a sustainable company make you happier?* [online] Corporate Knights. Available at: <https://www.corporateknights.com/health-and-lifestyle/does-working-for-a-sustainable-company-make-you-happier/>. (Accessed 11/03/2024)
* L, J. (2023). *TikTok Dances toward Net Zero Emissions by 2030*. [online] Carbon Credits. Available at: <https://carboncredits.com/tiktok-dances-toward-net-zero-emissions-by-2030/>. (Accessed 12/03/2024)
* Xing, E. and Monck, A. (2023). *AI emissions are fueling a new doomerism. This time it’s climate change*. [online] Fortune. Available at: <https://fortune.com/2023/12/12/ai-emissions-doomerism-climate-change-environment-tech/>. (Accessed 12/03/2024)
* O’Leary, N. “TikTok Discovers “Covert Influence Operation” Targeting Ireland.” *The Irish Times*, 28 Sept. 2024, [www.irishtimes.com/technology/2023/09/28/tiktok-discovers-covert-influence-operation-targeting-ireland/](http://www.irishtimes.com/technology/2023/09/28/tiktok-discovers-covert-influence-operation-targeting-ireland/). (Accessed 14/03/2024)
* Wong, Q. (2020). *TikTok ads have pushed scams about apps, diet pills, other products, report says*. [online] CNET. Available at: <https://www.cnet.com/news/privacy/tiktok-ads-used-to-push-scams-about-apps-diet-pills-and-other-products-report-says/>. (Accessed 14/03/2024)

‌

* ISD. (2023). *TikTok videos of migrants being attacked or harassed trend in Ireland*. [online] Available at: <https://www.isdglobal.org/isd-in-the-news/tiktok-videos-of-migrants-being-attacked-or-harassed-trend-in-ireland/> (Accessed 18/03/2024)

‌

‌

‌