ATTENTION ENGINEERING

SOLUTION TO THE ETHICAL ISSUES OF RECOMMENDATION
ALGORITHMS USED BY SOCIAL MEDIA PLATFORMS

FIT1055 IT Professional Practice & Ethics | 28/4/2023
Group Name: Antifragile

Elysia Jewel Lee | 33592861
Kylie Tan Sze Qi |32825544
Lee Zhi Ming | 32377479
Teh Jia Xuan | 32844700
Yap Eng Han | 32140088
Yew Jin Ruo | 33343276
Yoong Qian Xin | 32846517

Background – Problem Statement

In the digital age, people spend a significant amount of time on our phones often without even realizing it. In fact, it's common to see people glued to their screens while walking, waiting in line. While there are some people who are feeling helpless against social media addiction, they claimed that it is irresistible to surf the social media.

In 2017, Harris asserts that there are certain groups of people working behind the scenes at various technological companies to influence people's thoughts. It is commonly known that social media enterprises benefit from user engagement on their platforms. To maximize their profit gained from the extracted data, attention engineers were employed to analyze and determine how to collectively capture people's attention (Dega, 2018). They utilize a technique called attention engineering, which refers to the application of behavioral psychology principles to create products with maximal addictive potential, using cognitive biases to exploit the vulnerabilities of the human mind (Mom, 2022). However, this might be detrimental to the social media users. As the user-generated media shared online rises dramatically, it is challenging to determine which user-generated media can potentially violate users' privacy (Smith et al., 2012).

The actions of these companies towards individuals' personal information have the potential to harm not only the individual, but also others. Ergo, it is crucial that these social media companies are paying close attention to privacy issues, as any leak of personal information could have a significant impact on people's lives. In addition, manipulating people's mindsets and steering their thoughts influencing their thoughts could have a negative impact on our democracy and our ability to engage in the conversations and relationships we want to.

This report is written with the purpose of providing and proposing an ethical solution in the context of social, privacy, and legal aspects arising from the ethical issues surrounding the technology used in attention engineering, focusing more on recommendation algorithms implemented in social media platforms. This report will start with a background to the topic of attention engineering, followed by the methodology, which involves the Ethical Reasoning Framework (ERF), as well as the techniques and methods used for teamwork management. This report will end with the proposed solution for the problem statement and a conclusion. A list of references is also included to aid future research.

Methodology

2.1 ERF

An ethical reasoning framework (ERF) is a collection of principles that you can use for guidance in the development of an ethical algorithm or ethical artificial intelligence (AI) application (Nailah, 2022). In this research, ERF was used to help us structure the concepts and methods covered in this course to find a solution. Below is an ERF figure to help visualize the process.
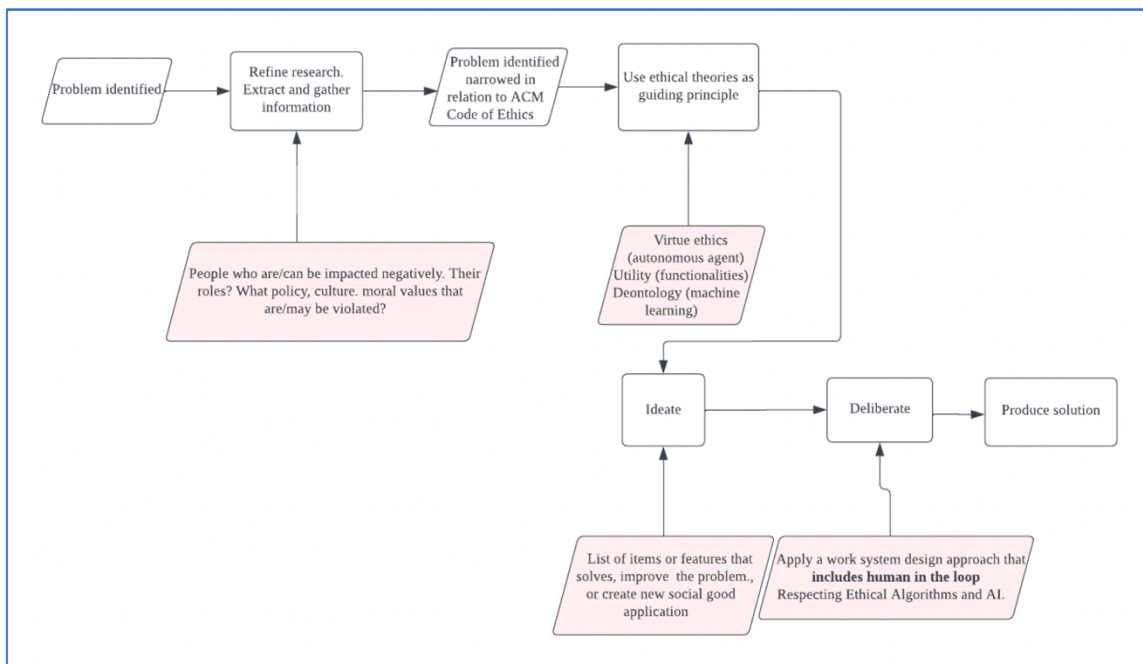

Figure 1: ERF Process Visualization

2.1.1 Results of Research

Following the recognition of the issues, team members were given the duty of conducting further study on which stakeholders were negatively influenced by the existence of attention engineering. The study primarily centered on the manner and reasons in which these stakeholders were affected. During the research, some solutions were identified and documented for the use of subsequent stages in ERF.

The stakeholders that were affected include users of the Internet, and people who own smartphones that range from different age groups.

According to Hari (2022), his son dropped out of school at the age of 15. He appeared to be buzzing at the rate of Snapchat, with his thoughts moving quickly and restlessly. No contemplative or weighty ideas seemed to be able to capture his attention or stick in his mind. His ability to pay attention was cracking and breaking. Users are unaware of the extent of our dependence on technology and concerned that it may be diminishing our ability to concentrate.

The idea of technological companies adding transparency in algorithms built was developed throughout the research. To conduct the research effectively, the strategy of collaboration was employed to divide the group into pairs. Tasks were distributed to gather information related to the stakeholders involved in the matter to eventually lead to the conclusion.

2.1.2 Problem Identification

Research on the identification of the problem was done through the use of the brainstorming technique where all group members gave their thoughts on which of and how the ACM Code of Ethics was breached. Then, the brainwriting technique was used to list out all ideas produced before discussing and finalizing the breached ACM Code of Ethics.

According to the result of research, attention engineering algorithms are known to have breached several Code of Ethics such as ACM Code of Ethics 1.2, 1.6, 1.7, and 2.9. ACM Code of Ethics 1.2 states that computing professionals should make sure that systems they designed do not bring any form of negative consequences, such as unjustified physical or mental injury and unjustified destruction or disclosure of information (ACM Code of

Ethics and Professional Conduct, 2018). To illustrate further, attention engineering refers to the leveraging of other people's attention as a strategy to foster addictions, where these addictions bring companies profit (Dega, 2018). When users increasingly become more obsessed with the "perfect" lives portrayed by others on social media platforms, this leads to unfavorable social comparisons between users and social media influencers (Braghieri, Levy & Makarin, 2022). These unfavorable comparisons on social media platforms take a toll on users' mental health, violating the fact that users' mental health should be preserved.

Whereas ACM Code of Ethics 1.6 states that computing professionals are responsible of respecting individuals' privacy and collect minimal amounts of private information to use in systems, and ACM Code of Ethics 1.7 states that confidential information, such as trade secrets, client data and research data, should not be disclosed, except in cases where it is evidence of law violation (ACM Code of Ethics and Professional Conduct, 2018). Attention engineering requires data collection in these social media platform interactions, such as browsing history and purchasing habits to implement recommendation algorithms, which would breach ACM Code of Ethics 1.6, if too much personal information is collected for filtering algorithms, and ACM Code of Ethics 1.7, if personal data of users are shared with third-party advertisers by these social media companies. As summarized in a journal report by Tari et al. (2015), where the authors stated that data confidentiality and user authentication are compromised in cloud computing, a method used for analyzing and collecting user data for algorithm implementations. Protecting all private data of a social media account from misuse is an important part of the larger problem, which is controlling access to cloud-based resources (Tari et al., 2015).

ACM Code of Ethics 2.9 states that computing professionals should design and implement systems that are robust and usably secure (ACM Code of Ethics and Professional Conduct, 2018). For example, social media company attention engineers should acknowledge the fact that humans tend to be lazy and skip reading privacy agreements (Meier et al., 2020). Therefore, security measures implemented should be user-friendly to encourage the users to take time and read data privacy statements of social media platforms.

2.1.3 Guiding Principles

After organizing all outcomes of our research, the group decided to use three fundamental types of ethical theories embedded as guiding principles to produce ideas for possible

solutions. These ethical theories include virtue ethics, utilitarianism and deontology (Nailah, 2022). Each theory acts as a guideline for designing ethical attention engineering systems and interfaces.

Virtue ethics emphasizes the importance of developing decisions based on good character traits and virtues for ethical behavior, including basic virtues such as justice, honesty and responsibility (Nailah, 2022). Therefore, ethical decisions should be made based on what a virtuous person would do in a given situation. To raise awareness of the problem, attention engineering algorithms should be designed in mind of honesty, responsibility and accountability, as social media platforms are entrusted with databases full of users' private data. Social media companies should show that they are trustworthy in storage of the data, be responsible for maintaining security of their databases and publicly state their accountability when data leaks happen. These algorithms should also be designed to promote good character traits and virtues that develop a positive impact on society, such as building a respectful community and acknowledging the diversity of cultures, and discourage negative behaviors, such as distraction and manipulation.

Utilitarianism emphasizes the moral obligation to maximize overall utility or well-being for the greatest number of people, which is a common theory chosen in AI because AI is entrusted with operations on large amounts of users' data (Nailah, 2022). This implies that attention engineering algorithms should be implemented in a way to keep users satisfied. Utilitarianism can be applied in recommendation algorithms by using monitoring systems and interfaces that improve users' mental health state. To illustrate further, implementing monitoring systems in attention engineering could be used to filter through a specific user's browsing history and alert them of any excessive amount of a certain topic they are consuming or if they encounter dangerous content on social media, including promotional content of self-harm, racism and segregation. This improves the mental health state of users, keeping them emotionally stable for the sake of their well-being, and also keeps their close friends and family happy as well, proving that it benefits every individual in society.

Deontological theories stress the idea of defining actions as good or bad based on a clear set of rules, as well as identifying the duties of each individual and making decisions accordingly (The Ethics Centre, 2016). A key point identified was to define a clear set of rules in the implementation of attention engineering algorithms to differentiate between good and bad content. However, the limitations of deontology include sensitive content such as posts related to LGBTQ communities, gender inequality, sexuality etc., which spark

up debates on social media platforms often, would be easily classified as bad content as the majority of society frowns upon these topics and communities. Both the users and social media companies hold responsibility in making ethical decisions, as the opinion of the majority affects how the algorithm performs toward these sensitive topics and the social media companies should take accountability for maintaining an ethical online environment for each and every user.

By applying these three ethical theories to attention engineering, these algorithms can be designed to not only be effective and efficient in performing their original tasks but also be ethical and respectful of users' rights and well-being. Therefore, it is important for attention engineering professionals to carefully consider these basic principles in the design and implementation of attention monitoring systems and interfaces.

2.1.4 Ideation

Aiming to obtain maximal efficiency in generating ideas for the solution, the design thinking process was used as a general guideline for the team to understand users of social media, challenge assumptions, redefine ethical problems generated from attention engineering algorithms and create innovative solutions to prototype and perform tests (The Interaction Design Foundation, 2022). Each of the team members who attended the online meeting via Zoom were asked to brainstorm as many ideas as possible to address any of the problems raised upon the ethical concerns of algorithms used in attention engineering. The solution and the proposed prototype will be illustrated further in Section 3 of the report, Proposed Solution.

Summarizing all ideas presented in the meeting held, a few main solutions were generated by compiling the ideas brainstormed together to form a clearer picture. The first solution the group thought of is implementing transparency in attention engineering algorithms. Companies should be more transparent about the algorithms and data they use to make decisions that impact user attention and behavior, making it open-source. An example of companies starting to implement open-source algorithms would be Twitter, which states the core models and features that extract user data from posts and engagements on their platform (Twitter, 2023). The blog post by Twitter (2023) describes the process on how their recommendation algorithms work and what functions are implemented in these recommendation systems that serve to bring relevant content to the preferences of the

user. Making the algorithms used in recommendation systems provide the users insight on how their personal information is being used, and raises awareness on how much of their private data is being operated on by the algorithms used, to prevent them from accidental data leaks.

Another solution is to adopt ethical advertising practices that prioritize user privacy and consent, such as providing clear opt-in options for targeted advertising and limiting the use of invasive tracking technologies. Providing users with greater control over their online experiences, such as the ability to customize their news feeds, be in control of their own personal information and opt-out of targeted advertising, can help promote user autonomy and privacy. To further address ethical concerns surrounding attention engineering algorithms, companies can take additional measures, such as limiting the amount of data that applications can access to protect user privacy. This can be achieved by implementing stricter data access controls, such as requiring user consent every time before allowing apps to collect certain types of data. For instance, Apple obtains user consent before allowing apps to collect personal data. The official developer webpage of the company provides detailed insight on describing how these applications use the data obtained from users, types of tracking included in the privacy agreement that users are consenting to, and an addition of answers to frequently asked questions on issues related to user data privacy (Apple Inc., n.d.).

A different approach used as a solution is to perform optimizations on the interface design of social media platforms, which increases the usability of the social media application and creates a secure virtual environment for users to share information on. In particular, making consent buttons for user data collection more visible, such as placing it at the top of the page, instead of sitting at the bottom of the preferences settings where users tend to miss out on, to prevent accidental data leaks. Moreover, companies should provide a shorter, concise version of privacy agreements alongside the regular-lengthed ones, and allow users to choose between which one they prefer to read and include a disclaimer informing users that the shorter version may be less detailed. A short and precise context of privacy regulations is able to resolve problems caused by human factors, where users tend to skip reading lengthy privacy agreements due to laziness (The Paypers, 2020).

The final solution proposed is to add features in the recommendation systems which provide limitations to the amount of time spent on social media platforms and enlighten the users on their problem of having an unhealthy amount of social media interactions, especially towards children and adolescents. Limiting the time spent on social media

applications for children by using parental controls, allow parents to monitor and restrict their children's use of digital devices and prevent excessive screen time. Another strategy is to encourage people to take regular breaks from their digital devices to promote healthy digital habits. This can be achieved by implementing known features such as YouTube's "Remind me to take a break" feature and the ability to set your own sleeping time and get reminders to log off the platform implemented by TikTok (Wong, 2022), where these mechanisms improve the digital wellbeing of users.

2.1.5 Deliberation

Using the ideas our group produced in the Ideation step of ERF, each member of our group who attended the meeting session held via Zoom came up with scenarios of the use of social media applications in real-world settings considering that advertisement mechanisms and recommendation algorithms are already implemented in most of the popularized social media platforms, such as YouTube, Facebook, Instagram and many more (Bustamante, 2022).

Scenario	Description	Related Ethical Problems	Solutions
Teenager who has intentions on buying a new phone when sensing that their current phone would fail to work in the near future	- Teenagers intends to buy a new phone and searches online for information, such as prices, models etc. - Recommendation systems of other applications, such as Facebook gain this information and uses this information to amend their recommendations to the teenager - Every time the teenager logs onto Facebook, the recommendation system shows	- Influences the teenager to buy a new phone, as there is an ongoing promotion - Creates a sense of urgency or dissatisfaction to the teenager, making the teenager think that they should buy it now to get the discounted price - Pushing users to buy more expensive products than the products they intended to buy, e.g. User wanted to buy a mid-priced phone, but the algorithm keeps recommending expensive and branded phone products which are on sale	- Companies should be transparent on their algorithms, e.g. how recommendation outcomes are generated - user can evaluate the fairness of the outcome themselves

	advertisements of the iPhone 14 Pro Max, which is an expensive product and relevant ongoing promotions, e.g. 5.5 Shopee Super Brand Day Sales	- A form of manipulation as it leads to the teenager buying something that they do not currently need - Develops a 'Fear of Missing Out' (FOMO) on the teenager to hold on to trends - Promotes materialism, which is harmful to society, secluding the people who cannot afford high-end products	
Individuals who suffer from mental health issues	- The promotion of perfectly curated lives on social media - Algorithms recommend posts based on what the user likes to scroll - Promoting unrealistic trends online that influence the mental state of teenagers	- Social media anxiety disorder, 'Compare and Despair' - As algorithms are not built to be considering a user's health issues, the teenager with mental health issues might face a deterioration in their mental health from constant exposure to social media, e.g. overattentive about engagement on their own posts - Degrading confidence of teenagers, as they deem their own lives not perfect	- Algorithms used to build recommendation systems that impact user behavior should be guided by a health and well-being mechanism - Be alert of how much content of these "perfect" lives is being searched up by the user and notify close family members if there is a strange or dangerous trend in the search history of the user - Regular tests for user's health and avoid recommending specific topics - Privacy may be an issue in data collection - Companies should be transparent about the data the algorithm collects

			and uses to make decisions - The user must be aware of the data they are exposing to social media companies
An elderly person who does not keep up with recent news and has less knowledge on certain topics	- Social media platform news pages use interesting titles or graphics to attract users to click on links - However, news content does not relate to the title, or may be fake news - News may be biased content	- Misinformation, misleading certain groups of members in society who have limited access to certain topics - Manipulates the public mindset - Harmful towards those who do not have the ability to differentiate fake news from legitimate news, e.g. children - May spread false allegations on certain minorities	- Social media companies should implement authentication mechanisms to check if fake news is promoted on their platforms - Use a flag to let users know about news pages which are self-promoted and may not be trustworthy
A political party which wants to gather votes during voting season	- During election season, political parties buy advertisements and pay for promotion by influencers on social media, e.g. TikTok and Facebook - Their motive is to influence and target first-time voters, who are teenagers ranging from 18 to 21 years old, to vote for the specific political party	- Brainwashes the public to their propaganda, violating the rights of an individual to make choices independently - When the political party wins the election, it becomes an example of attention engineering misused by the government for its own political agenda	Companies should implement transparency in their recommendation algorithms to let users know which posts are advertised by political parties for their own gain, and let users make their own decisions on whether to believe the posts or not.
A university student who is a Snapchat user and busy with assignments and exams	- Snapchat uses the concept of "streaks" and notifications to keep the users engaged - When users actively communicate every day with a particular friend	- This would make users of the platform who are busy with real-life matters, e.g. homework, projects, work etc., feel anxious that they might lose these "streaks"	- Transparency in the algorithms should be implemented in these applications to avoid manipulation

	on the platform, a "streak" indicating the number of days they interacted non-stop would appear beside their usernames	that they took time to maintain - These "streaks" become a form of emotional stress as they manipulate the user into neglecting more important matters, e.g. research has shown that users of younger ages give their Snapchat accounts to at least 4 friends when they go on holiday to maintain these "streaks' - Psychologically manipulate users to make them engage with their platform at least once a day - users gain the fear of missing out psychological behavior (FOMO)	- Snapchat should issue a statement to educate their younger users that "streaks" are not as important as

The brainstorming technique was used to come up with different scenarios where the implementation of attention engineering would cause ethical issues and how mechanisms and implementation of the solutions thought up during the Ideation process could help resolve these ethical problems. Brainstorming for the deliberation part was conducted online via Zoom, with each participant illustrating their own scenarios and ideas to everyone else. After much discussion, it was finalized that the solution of implementing transparency in attention engineering algorithms was the most effective choice for our solution, as it resolves most problems in scenarios stated in the table above.

This solution fulfills the Ethical Algorithms and AI principles of human centered values, privacy protection and security, and reliability and safety. Our system would respect the human rights of an individual to make their own decisions. To illustrate further, companies influence users' choices with recommendations of the same relevant topic repeatedly. The solution of implementing transparency in algorithms is to give the user control over what posts they see on their news feeds. For example, introducing customizable preferences that allow the recommendation algorithm to work on. The system proposed would also ensure that users' data is safely stored in databases and not shared to third parties. As types of

data collected by companies are revealed to the public, people know which part of their data is being used and clear privacy statements are used to minimize legal issues in data breaches. The solution given also enhances users' experience on social media apps, as the alert systems to be implemented in the solution is the key method of maintaining the mental health state of users. Known transparency initiatives, such as Google's Ad Transparency Report, which provides users with information on how advertisements are recommended to them, increases transparency and accountability in attention engineering algorithms, reducing the risk of manipulation or abuse (Edelson et al., 2021).

Hence, the idea of implementing transparency into recommendation algorithms is the best idea our group can pursue that solves the major problems in attention engineering.

2.1.6 Produce Solution

The final solution was made by having all members carefully review and scrutinize the proposed solutions, and then vote for the one that is more effective. Open-source algorithms to implement transparency in attention engineering was voted to be the main solution for the topic. The major problem faced by attention engineering is that it brings negative influence towards people who own a smartphone regardless of their age (Torngren, 2022). For example, without the transparency in their recommendation algorithms, the algorithm keeps promoting unrealistic trends online, thus influencing the mental state of teenagers. This is detrimental to individuals who suffer from mental health issues.

Therefore, having transparency in algorithms would allow more ethical and fair decisions to be made. It can also prevent manipulation of users if the algorithm in the application leads to a form of emotional stress as they manipulate the user into neglecting more important matters.

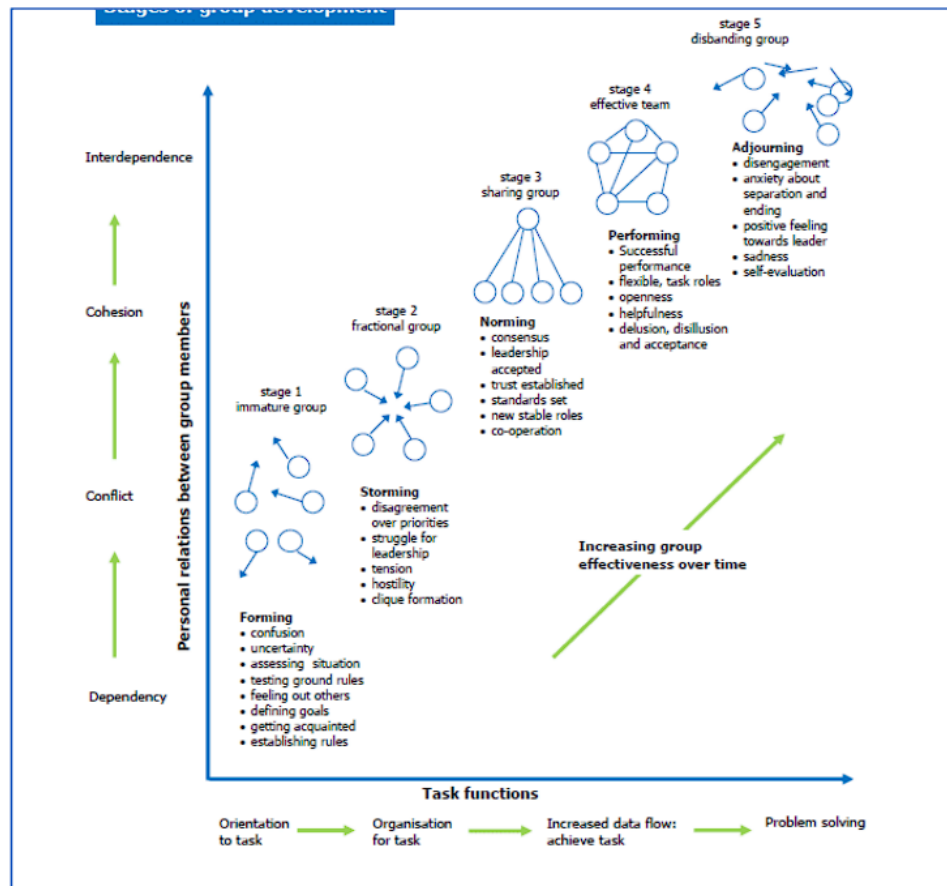## 2.2 Techniques / Methods for Teamwork

Figure 1: Stages of group development (Nailah, 2022)

Tuckman's theory is a model that outlines the sequential stages that a group typically goes through as they are working towards a shared objective (Tuckman, 1965, as cited in Nailah, 2022). Tuckman claims that teams always process through specific stages before becoming goal-oriented teams, commonly known as Forming, Storming, Norming, Performing and Adjourning.

During the forming stage, the members started to familiarize themselves with one another and started to introduce themselves as everyone came from different backgrounds. In order to enhance the chemical bonding of the group, an ice-breaking session was conducted with online games, such as 'draw.io' and 'word bomb'. The session was conducted during the start of the first tutorial session after the group was formed, which led to everyone being comfortable with each other. Furthermore, a personality test was done for a better understanding of the personalities of the team members, and the leader and other roles were assigned by voting based on their personality test. At the end of the

day, the group members exchanged social media accounts and personal phone numbers to provide access to a WhatsApp group for communication, a shared Word file for the group report and a shared Google Drive file for storing the recordings of all our meeting sessions for the group members who were absent. These collaboration tools were utilized for cooperation among team members.

The storming stage is where team members begin to argue and disagree with one another (Nailah, 2022). This happened when we were choosing a main topic, as all group members chose different topics in the first assignment. As everyone wants their topic to be on the table, to ensure that the final topic is chosen with fairness, the team leader suggested voting, in which each group member briefly explained the underlying potential of their chosen topic to see which topic sounded the most convincing and interesting to implement a solution on. This allowed for a more thoughtful and well-informed voting process and ensured the quality of the decision made. After a thorough discussion and voting, we decided to go with Attention Engineering as it is the most relevant topic to all group members, who are all active users of social media.

In the norming stage, team members start to actively contribute by offering ideas and suggestions. The members often speak out on their preferences and opinions on the structure of subtasks of the report and how workload distributions should be. It happened after the group decided on the topic and started to allocate tasks to each member of the group. With the guidance of the leader, every task completed during meetings is recorded in detail in the Meeting Minutes file for group members to refer back to and to ensure progress is made at every meeting. Apart from that, all members came to an agreement that in order to ensure the efficiency of the brainstorming process, it would be conducted face-to-face to avoid problems like technical difficulties, shyness to voice out, no responses etc.. During the brainstorming session, each subtask was given a time restriction for group members to write down their opinions and ideas roughly. After that, group members were all given the opportunity to present their ideas and persuade the group members to use these ideas they gave. An interesting discovery was made during norming, where we found out the effectiveness of employing a causal chain approach in conveying ideas.

By using this concept, it enables us to showcase how our idea links to another and how the cumulative effect of these linked elements would help us to solve the problem. In the end, we voted for the best idea and used it as our solution.

In terms of the members' performances, the team members displayed a spectrum of abilities, from commendable to satisfactory. A number of effective features and solutions were suggested, and these concepts were strengthened by the ideas of each team member. Moreover, some team members made extra efforts to gather all the relevant research papers to ensure the quality of the reports. However, some issues arose, such as no responses were given when the group leader asked for preferences, being absent during a meeting without notifying the group, a few members staying silent during meetings etc. Due to the impending assignment and interview tests around the corner, this led to a delay in progress, in which the group decided to organize more meetings and consultations with our lecturer, Ms Arini, to make sure the group is on track.

Finally, the adjourning step is where the group is disbanded after the assignment is complete. However, the group will be working together on the next assignment, so the team was not disbanded.

Proposed Solution

We suggest that addressing the problems associated with attention engineering requires implementing transparency in algorithms, specifically through open-source algorithmic solutions. The root of the issue with attention engineering algorithms lies in creating products that aim to seize users' attention and retain it for as long as possible, which could lead to influencing user behavior and even encouraging addictive usage patterns. By making algorithms open source, we aim to increase transparency and foster ethical behavior in the development and use of such algorithms.

The approach of using open-source algorithms can be effective for a number of reasons. Firstly, it provides transparency by making an application's internal workings visible to all users. This means that the tactics used to hold users' attention would be revealed to the world, reducing the potential for manipulating user behavior. Secondly, open-source algorithms create accountability for firms, as they would be held responsible for any negative impacts resulting from their algorithms. Lastly, open-source algorithms encourage collaboration and innovation, as developers and researchers can work together from all around the world to improve algorithms and ensure they are used in an ethical manner.

An example of such a successful project would be AutoGPT, where developers use GPT-4 for non-malicious and for legal purposes. These gates would be flooded right away if the developers of the application are not scrutinized well enough. This would prioritize user welfare and ethical concerns before user engagement. Open-source algorithms would also be utilized as educational tools, assisting nascent academics and developers in comprehending how specific algorithms operate and inspiring a new generation of moral, user-centered designers and developers.

To successfully implement open-source practices for algorithms, the IT sector needs to offer its support. This may require regulatory requirements or incentives to encourage adoption. Additionally, an auditing and review on methods can help ensure that open-source algorithms are beneficial to society.

However, there might be challenges. One example would be intellectual property. Since business strategies might rely on proprietary algorithms, there would be security risks associated with making such algorithms open source. Another challenge would be the lack

of sufficient expertise among the public to fully understand and scrutinize the code. Striking a balance between these challenges is critical and requires careful consideration. It is crucial to identify if such a solution benefits your use case.

During the deliberation step, two information flow diagrams were created to illustrate the user experience of scrolling through a social media page, with Twitter serving as an example.
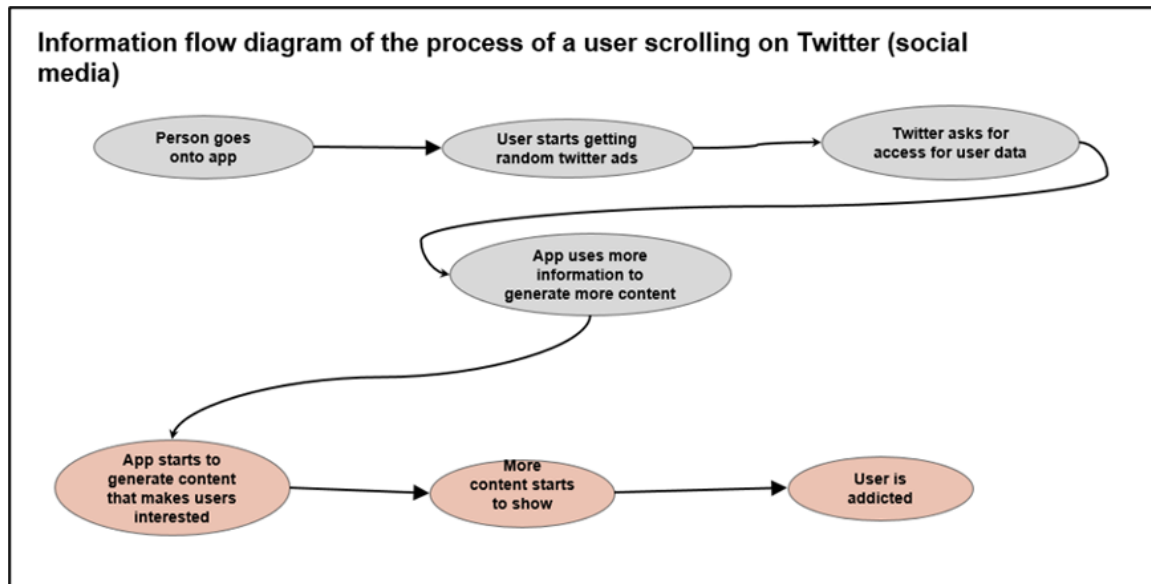

Diagram 3.1: Flow diagram of a user scrolling through social media

Having said that, Diagram 3.1 illustrates the sequence of events when individuals engage in leisure activities on social media without the implementation of transparency in attention engineering algorithms. The diagram portrays how users are guided by recommendation AI towards developing an addiction to social media, which greatly affects public mental health. Another problem that should be prioritized is the manipulation of human thinking, as recommendations repeatedly showing topics that the users consume most does not provide diversity in content, leading to a closed-minded mentality and violates the fact that algorithms should not interfere with humans' decision making.

The second diagram illustrates the method of implementing transparency in social media recommendation systems using various mechanisms listed in the diagram below.
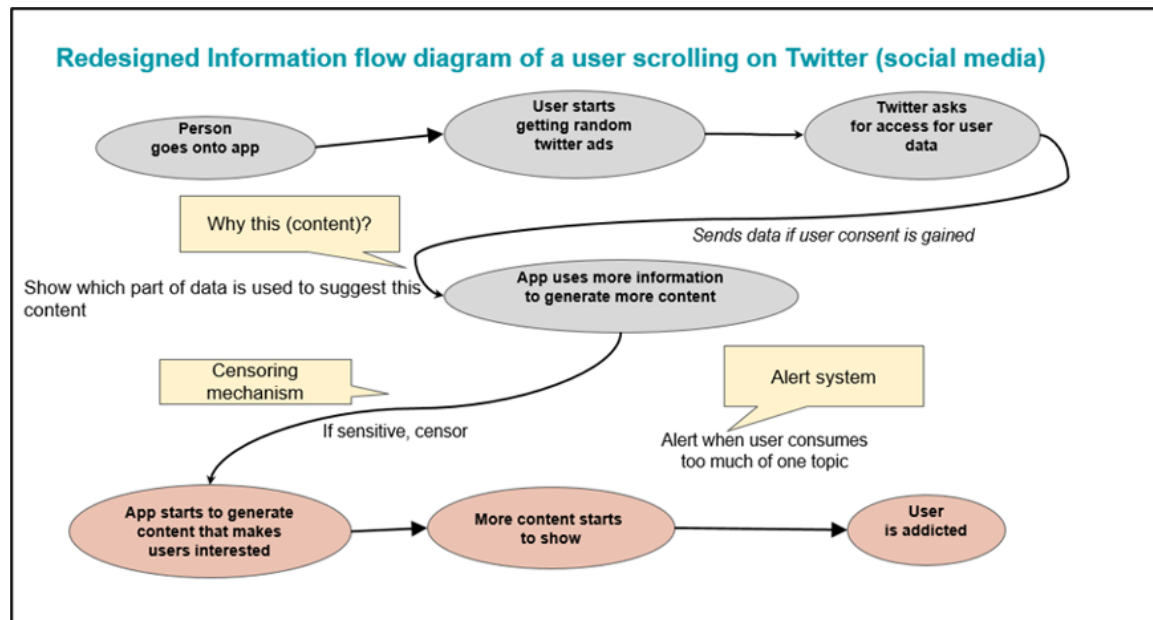
Diagram 3.2: Flow diagram after the solution involving transparency mechanisms is implemented

According to Diagram 3.2, the group has come up with the addition of three different mechanisms to implement in the process of the user scrolling through their social media page. Ideally, after going on the app, the user would start to see advertised content on their news feed pages. To provide personally-tailored content recommendations to the user, the app's side would then ask for access to user data and will start processing user data if consent is granted by the user. The first mechanism used is a content transparency feature to show users which part of their data is used to suggest that piece of content.

Following that, the group proposed a solution to address sensitive content by implementing a censoring mechanism. The censoring mechanism is used to detect sensitive content and take appropriate action before it is presented to the user. As the generative recommendation algorithm operates in a transparent manner, the users can enjoy the content recommended by the AI algorithm without being exposed to explicit material.

The last part of the solution is implementing an alert system that is integrated into the content generation algorithm. When specific content is shown to the user, the system would track the topics consumed by the user. Users will be notified if they have been consuming too much content on a certain topic, which can help prevent negative consequences, such as social media addiction and use of social media to research on dangerous actions, including self-harm. By incorporating these features, user experience can be enhanced and a healthy relationship with social media is promoted to users,

eliminating the ethical problems of privacy issues, personal data violations and the implications on society well-being.

Open sourcing algorithms could increase transparency and accountability of social media platforms, which could lead to better regulation and policies to address issues such as addiction. It could also enable users to have more control over their social media experience by allowing them to customize their own algorithms or choose to use alternative algorithms designed to promote healthier content consumption habits.

However, there are still limitations to the solution as overly detailed transparency reports used in algorithms of major social media platforms may increase the risks of malicious cybersecurity attacks on the platform's databases as too much information is revealed to the public, which contradicts with the ethical problem that user data should be secure. Technical challenges and the high use of computing resources in implementing this system as it requires robust Natural Language Processing (NLP) techniques to identify topics that need to be warned to users, and high-level security measures to protect the large amounts of data used by these mechanisms.

Below, we break down the pros and cons of our decided solution into a table.

Solution	Features	Pros	Cons
Transparency in attention engineering algorithms	Implementing censorship mechanisms in attention engineering algorithms of social media platforms	Raise awareness among the public about the presence of attention engineering techniques when censoring labels occur or pop out	Raises concerns about freedom of speech and human rights (public-deemed sensitive content is not shown)
		Users can make choices on the content shown in their social media feed from the warnings given from the censorship feature of the platform	User resistance: users may ignore the warnings sent to them about attention engineering especially if it interferes with their usage of the app and their online experience
	Implementing an alert system which acts like a time management tool	Promote healthier lifestyles by reducing user time spent on social media platforms	Does not effectively solve the issue of attention engineering but mitigates the effects of it on users

	on social media platform (time limit or screen time feature)		
		The alert system can also notify users if they have been reading excessively on a certain topic (enhanced user protection to avoid manipulation)	Technical challenges and resources in implementing this system as it requires robust algorithms to identify topics that needs to be warned to users
		Alerts can also cause behavioral changes on users (reconsider their engagement with false news or manipulative contents)	Algorithms used may contain confidential information so the requirements for a transparent algorithm must also consider user privacy
	Transparency reports which will show why the app shows or recommends these contents to each individual users	Gives users the chance to understand why certain content is recommended to them and empowers them to decide which topic they want to engage in and what to avoid.	Detailed transparency of such algorithms used in large social media platforms may increase higher risk to malicious attacks and activities happening.
		Helps in identifying biases or fake news within algorithms and contents	
		Make sure platforms practice ethical attention engineering practices to live up to public expectations	

Although the solution proposed has certain flaws, it is important to acknowledge that the solution and its features align with the ACM code of ethics. Transparency regarding the algorithms in attention engineering is crucial as it provides insight into how our data is used to generate content to individual users. This is essential as it obeys ACM 1.6 which emphasizes "Respect Privacy", by understanding how our private data is utilized and being informed about what context is generated from specific data is our privacy rights (Arik et

al., 2015). This enables us to know whether our privacy rights are being preserved and this allows trust to be built between users and the platform. Hence, when we have a clear understanding of which specific personal data are being used to generate content, we can be more confident that our privacy is being respected.

Moreover, showing which personal data is used to generate content plays a vital role in preventing breaches of ACM Code 1.6, Respect Privacy. When users are being informed which specific data points are used to generate the content. Users can exercise more control over their personal information and decide what information they are willing to disclose (Ouyang, 2019). Since they are aware of the exact data points that go into the creation of content. By providing the information, users are able to identify any potential misuse or unauthorized access to their data. Users are more equipped to hold platform providers responsible for their data practices thanks to this awareness, and they may also take the necessary precautions to protect their privacy.

Furthermore, transparency in algorithms is closely related to the principle of honor confidentiality outlined in ACM 1.7. Transparency in algorithms in terms of confidentiality requires being open and honest about how sensitive information is handled (Ouyang, 2019). It means clearly describing to users how the confidentiality of their information is maintained in their algorithms and under what conditions it may be disclosed. Tech companies should be transparent regarding any circumstances where confidentiality may be compromised, such as those involved in law, organizational policies etc. (Association for Computing Machinery, 2018). These terms and the proper authorities to whom the information may be given should be made clear to users. Hence, confidentiality can be preserved by tech companies if they make the algorithm transparent.

Countless people have argued over the effect of excessive electronic gadget use on children's health since the technology boom of the 90s (Clements, 1998) and as we discussed in prior meetings, we think that the source of the problem stems from our inability to keep our hands off social media platforms. While the platforms' founding objectives were not to bring harm to humans, they have been getting closer and closer to breaching the ACM Code 1.2 over the years. This is because the ACM Code of Ethics states that computing professionals are obliged to mitigate harms—unintended or otherwise—brought on by the systems they built, and "harm" here includes adverse effects on our health as well (Association for Computing Machinery, 2018). Hence, one possible way for social media companies to reduce said effects is by mandating the implementation of an alert system that reminds users to take breaks on a regular basis.

For teenagers, a healthy screen time could help to reduce anxiety from the "FOMO" phenomena and their desire to gain attention of the public and obtain validation from the members of society (Yang, 2023). By staying away from social media, they start focusing on improving their real-life situation instead of concerning themselves with getting as many likes as possible on social media. For adults, limiting social media use could result in better sleep quality and positive emotions throughout the day (Graham et al., 2020). Furthermore, spending less time on the internet means that social media companies gain less chances of harvesting personal data, so it is a small step to avoid breach of ACM Code of Ethics 1.6 from implementing human control at the users' side.

Regardless, prompting too frequently for users to log off social media may build user resistance, whereby users become increasingly likely to ignore the given warnings. For example, pop-up advertisements on websites irritate users most of the time. It could also be emotionally stressful to the point where it could potentially lead to users boycotting the platform or switching to other platforms which do not implement such alert systems. Therefore, more creative ways of implementing alert systems should be used to increase the efficiency of the platform to avoid user resistance. This will be demonstrated in greater detail in our prototype later on.

While it may be an impossible task to develop a fully secure and privacy-preserving system due to inherent design conflicts between the two (Zhang et al., 2010), the proposed solution aims to achieve both to the best of its abilities and limitations. This is in line with ACM Code of Ethics 2.9 that requires all systems to be robustly and usably secure. An ideal security system should be intuitive, easy-to-use, and not confusing (Association for Computing Machinery, 2018).

It is important for social media companies to be transparent and specific about the data they collect and how it will be used. By providing clear and detailed information to users, companies can empower users to make informed decisions about their data, implementing human control over their own personal information. Thus, creating a securely usable system.

Presenting users with illustrations or visual representations of how their data is being shared can also be a helpful approach. Visual aids can make it easier for users to understand the extent of data sharing and the potential implications. This kind of transparency can foster a sense of trust and accountability between social media companies and their users. Furthermore, it is crucial to encourage users to develop a broader understanding of the implications of sharing their data. Social media companies

can play a role in educating users about privacy, data security, and the potential risks associated with sharing personal information. By promoting digital literacy and providing resources for users to make sound decisions, companies can empower users to take control of their own data and privacy.

Overall, the solution proposed can help create a more secure and informed environment for social media users. It is important for companies to prioritize user privacy and work towards building a safe and secure platform for users to share information and engage with others on.

Conclusion

In this report, we proposed a solution to the problems caused by attention engineering, which is implementing transparency mechanisms in recommendation algorithms, specifically through open-source algorithmic solutions.

By making algorithms open-source, we can increase transparency, discourage unethical behavior, provide consumers with greater control over their information, and reduce the tendency for these applications to manipulate user behavior. After much deliberation, the final features of the solution were borned.

The solution proposed adheres to the ACM Code of Ethics, which is further amplified throughout the text. In particular – ACM code 1.2 (avoid harm), 2.9 (designing robust and usable secure systems), 1.6 (respecting privacy), and 1.7 (honoring confidentiality).

Furthermore, our approach offers several further strategies for reducing social media addiction such as user education, behavioral interventions, technological advancements, and implementing support systems. However, companies must exercise caution to avoid any potential backlash or boycotts. Also, the implementation of such open-source algorithms should require the backing of the IT industry, as well as certain legal mandates or incentives.

In conclusion, open-source algorithms would be a great solution to ethical issues caused by technology used in attention engineering, as it adheres to ethical standards, respects user privacy, and averts any potential drawbacks despite the difficulties that come with it.

References

Apple Inc. (n.d.). *User Privacy and Data Use - App Store - Apple Developer*. Apple
    Developer. https://developer.apple.com/app-store/user-privacy-and-data-use/

Association for Computing Machinery. (2018, June 22). *ACM Code of Ethics and
    Professional Conduct*. Association for Computing Machinery.
    https://www.acm.org/code-of-ethics

Braghieri, L., Levy, R., & Makarin, A. (2022, November 11). *Social Media and Mental Health.*
    The American Economic Review, *112*(11), 3660–3693.
    https://doi.org/10.1257/aer.20211218.

Clements, D. (1998, February). *Young Children and Technology*. Forum on Early Childhood
    Science, Mathematics, and Technology Education.
    https://files.eric.ed.gov/fulltext/ED416991.pdf

Dega, M. (2018, August 08). *Attention engineering and the addictive power of social media.*
    LinkedIn.
    https://www.linkedin.com/pulse/attention-engineering-addictive-power-social-me
    dia-mohan-/

Edelson, L., Chuang, J., Fowler, E. F., Franz, M., & Ridout, T. N. (2021, August 03). *Universal
    Digital Ad Transparency*. Social Science Research Network.
    https://doi.org/10.2139/ssrn.3898214.

 Friedman, Arik, Knijnenburg, Bart P., Vanhecke, Kris, Martens, Luc, & Berkovsky, Shlomo.
    (2015). Privacy Aspects of Recommender Systems. In Recommender Systems
    Handbook (pp. 649–688). *Springer US*.
    https://doi.org/10.1007/978-1-4899-7637-6_19

Graham, S., Mason, A., Riordan, B., Winter, T., & Scarf, D. (2021, June 14). Taking a Break
    from Social Media Improves Wellbeing Through Sleep Quality. *Cyberpsychology,
    Behavior, and Social Networking*, *24*(6). https://doi.org/10.1089/cyber.2020.0217

Hari, J. (2022, January 02). *Your attention didn't collapse. It was stolen*. The Guardian.
    https://www.theguardian.com/science/2022/jan/02/attention-span-focus-screens-a
    pps-smartphones-social-media

Meier, Yannic, Schaewel, Johanna, & Kraemer, Nicole. (2020, June 23). *The Shorter the Better? Effects of Privacy Policy Length on Online Privacy Decision-Making.* Media and Communication (Lisboa), *8*(2), 291-301. https://doi.org/10.17645/mac.v8i2.2846

Mom, T. (2022, October 20). *Become Your Own Attention Engineer: How to Hack Bad Digital Habits with Behavior Design*. TechDetox Box. https://www.techdetoxbox.com/digital-wellbeing/become-your-own-attention-eng ineer-how-to-hack-bad-digital-habits-with-behavior-design/#:~:text=Attention%20 engineers%20behind%20social%20media,products%20for%20maximum%20addict ive%20potential

Nailah. (2022). *FIT1055 IT PROFESSIONAL PRACTICE & ETHICS*. Moodle Monash. https://d3cgwrxphzofqu.cloudfront.net/c4/5d/c45d0ebf4f1411d12390ae8c670f830e12 65f93e?response-content-disposition=inline%3Bfilename%3D%22FIT1055%20IT%2 0professional%20practice%20and%20ethics%20textbook%20v2.pdf%22&response-content-type=application%2Fpdf&Expires=1683991800&Signature=TRHk-Nrnpsn8g 7Vy4zZp5~RBbk9RZQ6jDDl~gmb6qTOEBQQnaR~on2L~dQ4d6nHtKelv4haL9ed11 tVdS9IlgGD-nR1hRGbCoIyEt2-qq-pP22b9A0STxg60eAKhi0EeWHNxkE9HqTvfWp EIeT7BviWpHEM2hkPeLRWjjdmx8fUFr1m20R2~v7bDRJlxkEBnGZoU-pHz2lmwjTf SBfin-HqzUrtxC3GX93QK~v50vMNWLzmSdl6P1kxEe44Dxbf~HXCV8trV9KM5VZf ihI~pH5Th-0YlQ96zAcbIkP-S9713udfg-pG3VhoRYlrUtZynWgir88PbHYn2b24Me7 MKPw__&Key-Pair-Id=APKAJRIEZFHR4FGFTJHA

Ouyang, W. (2020, January 09). *Research on the Role of Algorithm Transparency in Algorithm Accountability.* Atlantis Press. http://doi.org/10.2991/assehr.k.191221.055

Smith, M., Szongott, C., Henne, B., & von Voigt, G. (2012, June 18). Big data privacy issues in public social media. *2012 6th IEEE International Conference on Digital Ecosystems and Technologies (DEST)*, 1–6. https://ieeexplore.ieee.org/document/6227909

Tari, Zahir, Xun Yi, Premarathne, Uthpala S., Bertok, Peter, & Khalil, Ibrahim. (2015, June 02). *Security and Privacy in Cloud Computing: Vision, Trends, and Challenges.* IEEE Cloud Computing, *2*(2), 30-38. https://ieeexplore.ieee.org/document/7116473

The Ethics Centre. (2016, February 18). *Ethics Explainer: Deontology.* THE ETHICS CENTRE. https://ethics.org.au/ethics-explainer-deontology/#:~:text=Deontology%20is%20a n%20ethical%20theory,don't%20aren't.

The Interaction Design Foundation (2022). *What is Design Thinking?* The Interaction Design Foundation.

https://www.interaction-design.org/literature/topics/design-thinking#:~:text=Desi gn%20thinking%20is%20a%20non,are%20 ill%2d Defined%20or%20 unknown.

The Paypers. (2020, November 25). *Many People Accept Privacy Policies Without Reading Them, Study Reveals.* The Paypers. https://thepaypers.com/payments-general/many-people-accept-privacy-policies-wi thout-reading-them-study-reveals--1245901.

Torngren, M. (2022, October 18). *Attention Engineers, Dopamine and Social Media – Here's how they control your life.* LifeBlonder. https://lifebonder.com/blog/2022/10/18/attention-engineers-dopamine-and-social- media-heres-how-they-control-your-life/

Twitter. (2023, March 31). *Twitter's Recommendation Algorithm.* https://blog.twitter.com/engineering/en_us/topics/open-source/2023/twitter-reco mmendation-algorithm

Wong, Q. (2022, June 9). *TikTok Wants to Remind You to Take Breaks From the App.* CNET. https://www.cnet.com/news/social-media/tiktok-wants-to-remind-you-to-take-bre aks-from-the-app/

Yang, Y. (2023, January 12). *Reasons for Teenagers' Habitual Use of Social Media: A Case Study of TikTok.* SHS Web of Conferences, *155*, 02006. https://doi.org/10.1051/shsconf/202315502006

Zhang, C., Sun, J., Zhu, X., & Fang, Y. (2010, July 19). *Privacy and security for online social networks: challenges and opportunities.* IEEE Network, *24*(4), 13–18. https://doi.org/10.1109/mnet.2010.5510913