Information Privacy in the World of Social Media

Submitted by

Michael Thomas Young

Management Information Systems

To

The Honors College

Oakland University

In partial fulfillment of the

requirement to graduate from

The Honors College

Mentor: Xiaodong Deng and Thomas Lauer, Professors of Management Information Systems

Department of Decision and Information Sciences

Oakland University

March 5th, 2023

Abstract

This paper will explore the process in which information is collected, processed, and disseminated by social media platforms. The main concern is how this information can become more secure and private for the user experience. Secondary research will be conducted to assess these concerns brought up by social media users and analyze privacy policies that user agree to when creating an account with a social media platform. Secondary research will be provided by peer-reviewed scholarly articles from the Oakland University library database. This knowledge will help users better understand how their information's integrity is being violated by these companies. The impact of this research will help reduce information privacy scandals and provide a better safeguard for user information, not only in the area of social media, but in other areas of the internet as well. The benefits of this will include enhanced individual understanding, increased privacy, and improved decision making. These benefits will affect every individual who utilizes these social media sites and potentially even affect those in other areas of the internet.

Introduction

Young adults and teenagers commonly use the internet and social media technologies to communicate with each other and share information online. As the usage of social media continues to increase, there are a variety of security concerns related to user information and privacy. Information privacy refers to how user information such as email addresses, locations, and phone numbers, are collected, disseminated, and stored by companies. Information privacy pertains to four main types of harmful activities summed up in the book *Understanding Privacy* by Daniel Solove. These harmful activities include information collection, information processing, information dissemination, and invasion (Solove, 2008). Social media applications such as Facebook have been under much scrutiny after continuous privacy scandals. This is not limited to just Facebook, as other applications such as Path have collected personal user information without consent. One article brings up an important point, which states "such privacy-related incidents that mobile users may experience are one of a few reasons for them to abandon the use of mobile apps" (Degirmenci, 2020). The way in which information is collected and processed has led to many concerns from social media users wondering if their information is in safe hands. Furthermore, users are worried about how their information is disseminated to other parent companies such as Google. This growing concern with no immediate solution in sight has brought a number of people to worry about the privacy of their information.

This project aims to fill the gap of information privacy concerns by providing in-depth research of information privacy in social media and how this information is collected and processed for further use. Current research lacks the immediate implementation of a solution for these privacy concerns. The impact of this research will help reduce information privacy

scandals and provide a better safeguard for user information. The benefits of this will include enhanced individual understanding, increased privacy, and improved decision making.

Methodology

In order to investigate how information is collected by social media companies, a literature search was conducted using the Oakland University database. In addition to this, articles written about wrongful information dissemination from various social media platforms were included to highlight real world scenarios of wrongful dissemination occurring. Lastly, research conducted by Daniel Solove (2008) was included to bridge the connection between privacy and social media.

A literature search was conducted using Oakland University's computer science and engineering database by using search terms such as "information privacy"; "privacy policies"; information privacy in social media"; "social media privacy"; and "information dissemination". The search terms were used to generate results based off of an articles title and abstract search. These key terms narrowed down the list of peer-reviewed scholarly articles to a culmination of roughly 80 results. Once the results were narrowed down, each article was examined to determine if the topics matched the search criteria and if they were useful for the analysis of information privacy in social media. Around 8 of these articles were selected for further review and analysis, based on their relation to the search criteria and the topic of information privacy in social media. Each article offered a unique perspective on information privacy, not just in social media, but across various forms of media.

Furthermore, articles were examined online about recent social media company lawsuits stemming from dissemination of user information. Particular keywords were used to generate these results by using Google and search terms such as "information privacy breach";

"information breach"; and "companies selling user data". These articles will be used to connect real life instances of dissemination to the topic of information privacy and information dissemination. All articles used will include examples of social media companies within the last 10 years. Literary analysis will be required to determine whether or not the instances of dissemination are legal or illegal based on the company's privacy policy. Facebook and Instagram's privacy policies will be included during the analysis of their circumstances of information dissemination.

Lastly, research about privacy conducted by Daniel Solove in *Understanding Privacy* (2008) will help bridge the gap in understanding information dissemination and the negative effects that it can have on an individual. It will also discuss government legislation that helps protect an individual from unlawful disclosure of information. Furthermore, it will help connect the concept of information privacy in the real world, to instances of information privacy while on social media.

The secondary research will be provided by scholarly articles from Oakland University's Kresge Library. Secondary research will also be found online in heavily consumed media and in published books. In addition, this research will include published research regarding privacy and will incorporate real life events that have transpired. This secondary research will either affirm the findings of this research, or offer a different perspective from the research conducted. Conclusions will then be drawn between the three sources, and recommendations will be provided on how users could better safeguard their information from these social media applications. It will also offer insight into how these social media applications could change the way they collect information, and the secondary ways they use this information.

Literature Review

Information Dissemination

According to "Social Media and Privacy", the explosion of social and professional networking sites, media-sharing sites, blogging sites, and even forum sites, has created an attractive means for both purposeful and accidental dissemination of user information while online (Clark, 2010). Within a 5-year time gap, social media and app revenues has increased from $69.7 billion in 2015 to $188.9 billion in 2020 (App Annie, 2016). The concept of information dissemination refers to "the reveal of personal data or the threat of spreading information" (Solove, 2008). Dissemination has a plethora of categories that fall under it, such as breach of confidentiality, disclosure, exposure, increased accessibility, blackmail, appropriation, and distortion. The primary form of dissemination from social media companies comes in the form of disclosure. Disclosure can have many negative effects for a user. Unethical disclosure threatens a user's security, first amendment rights, and their emotional and physical wellbeing by revealing true information about an individual without their consent.

Disclosure refers to when certain true information about a person is publicly revealed to others (Solove, 2008). This situation can cause a handful of problems when social media companies begin to disclose user information to advertisers or other companies. Users who wish to remain anonymous are now threatened by these companies. As Solove stated, "people want to protect information that can be used by others to harm them physically, emotionally, financially, or reputationally" (Solove, 2008). When a threat could cause any harmful effects such as the ones previously listed, users are more likely to care about their information and how the company collects, disseminates, and distributes the data.

There has been a handful of instances where social media companies disclose user information to other companies or advertisers in exchange for funding. Contrary to public

opinion, the majority of this information dissemination is in fact legal, due to the user agreeing to the company's privacy policy when they sign up for an account. However, due to these lengthy policies, users often opt to agree to the terms of service without even reading or fully opening them. This is the primary reason why many users are unaware that their information is even being disseminated in the first place. A study conducted by Steinfield (2016) surveyed various users on their likeliness to read the terms and conditions when signing up for a website. Of the 64 participants, 50 of them (79.7%) agreed to the terms and conditions without clicking the link to read the policy (Steinfield, 2016). The overwhelming majority of individuals do not bother clicking on or reading these policies, simply because it isn't required and because of how long it takes to fully understand them.

Privacy Policies

Many social media websites contain a "Terms of Service" agreement in which users must agree upon and submit before officially creating an account with the social media platform. Within these Terms of Service, privacy policies dictate how companies view and individual's data and the rights that they are granted when a user signs up for their platform. Privacy policies benefit the users by protecting the right to privacy and what data they choose to share with the platform. However, there have been a variety of concerns of companies breaking these privacy policies and disseminating user data. Even with these risks, users rarely read these privacy policies when choosing to join the platform nor consider the effects of these policies on their data (Pitkänen and Tuunainen, 2012; Acquisti and Gross, 2006; Talib et al., 2014). One of the primary reasons that these policies are rarely read is because of their length and wording (Furnell and Phippen, 2012). There has been a variety of research conducted on the topic of information privacy, policy privacies, and the processes of how information is disseminated. However,

current research lacks an immediate solution on how users can better protect and understand their privacy through privacy policies.

The study conducted by Steinfield (2016) also had an eye tracking experience which would track the users' eye movement as they navigated through an account creation and viewing of the privacy policy. The researcher split participants into two group. One group would automatically have to read the policy when it pops up, and the other group had the option of clicking the link or not clicking the link. A short survey was given to both groups at the end of the experiment to answer certain questions regarding the privacy policy that they had just agreed to. Figure 1 below shows the various questions they were asked in the survey, alongside the question is the correct answer and the percentage of people that provided the correct answer.

Figure 1.

---

1. Does the privacy policy allow the researchers to use the study data for research purposes? (Yes – 70%)
2. Does the privacy policy allow the researchers to sell data outside the university for commercial use? (No – 53%)
3. Does the privacy policy allow the researchers to share the study data with others in the university for research purposes? (Yes – 43%)
4. Does the privacy policy allow the researchers to share the study data with others in the university for purposes other than research? (No – 41%)
5. Does the privacy policy allow the researchers to contact you after the study, using the information you provided? (No – 27%)
6. Does the privacy policy allow the researchers to change the privacy policy in the future? (No – 54%)

(Figure 1) – Privacy Policy Survey and Answers

As can be seen above, out of all of the participants that willingly viewed the privacy policy, not even half of them at most times could provide the correct answer for simple privacy questions. This experiment shows how confusing and wordy most privacy policies are when users view them. If users are not aware of what they should even be looking for, they will not want to read over 100 pages of content to find an answer to their question. The experiment from Steinfield (2016) concluded that most users generally forgo reading the privacy document, but when they click on it, they devote much less time than they should to find the answers to their questions.

Cases of Dissemination

In 2018, Facebook became under scrutiny as a result of an information leak scandal occurring from Cambridge-Analytica. Facebook was responsible for allowing the harvesting of user information from a third-party application. This application improperly gained access to over 87 million user accounts information, which included email addresses, names, phone numbers, and more. Cambridge-Analytica then purchased this data in order to help skew election results. At the time, users were largely unaware that their data was being misused because of this dissemination. This singular event was a revolutionary spark in the world of privacy, in which many users became vastly more aware of just how important their information was. Just recently in 2022, Meta, the parent company of Facebook, paid a settlement of over $725 million dollars for their role in the scandal case (BBC, 2022). As these incidents become more common, users are left wondering if their information is truly in the right hands and if it is protected.

Cases of information dissemination are not limited to just Facebook. TikTok collects a user's IP address, a unique identifier, and what a viewer is clicking or searching during their viewing session. According to Melanie Bosselait, a spokesperson from TikTok, data from individuals is not specifically grouped but it is used in reports sent to advertisers (Malwarebytes,

2020). Legally, TikTok has the right to use this data for targeted ads and ad effectiveness when users agree to their privacy policy during their account creation.

Types of Information

Through the agreement of privacy policies, social media companies have the right to collection certain amounts of user information outlined in their policy. This information can be broken down in 5 main subcategories. Of this data, the most valuable to the normal user stems from personal data, which includes age, birthday, first and last name, email address, location, and phone number. Similar to what Solove (2008) mentions, people want to protect information that can be used by others to harm them physically, emotionally, financially, or reputationally. Personal identifiable information is seen as the most valuable type of information because of the number of threats and harmful activities that can result from this information being disseminated. Full names, emails, and phone numbers can be used to mass create bot accounts, take down platforms, skew voting records, and much more. Impersonation and identify theft have steadily increased, with over 5.7 million reports of identity theft and fraud in 2021 (IdentifyTheft, 2022). According to the Cambridge-Analytica scandal that was brought to light in 2018, the majority of the information that was sold from Facebook ended up being personal identifiable information. Thus, this is more than likely the most important information that users end up wanting to protect. Of the remaining categories that fall under user information, these categories include behavioral data, engagement data, personal data, attitudinal data, and preference data (Juicer, 2021).

However, even the other various types of information still pose a threat if misused or placed in the wrong hands. Engagement data and behavioral data can both be used to track how a user navigates a platform, what they engage with, what they like, and what they save. When you

combine all of that information, it becomes very appealing for advertisers. For instance, advertisers pay top dollar to Meta to target users with select advertisements based on their engagement data. Facebook remains the most targeted social media platform for advertisers because there are over 2.9 billion accounts created, which is nearly 1/3rd of the global population. A large number of users find this as a privacy concern because Facebook is always tracking their activity. The real issue they find, is when they search for something on one platform, and receive advertisements for the exact same item on another platform. As more and more users sign up for social media accounts, information privacy continues to be a top concern amongst users.

Users' Information Privacy Concerns

According to Degirmenci (2020), users' privacy concerns are directly correlated with three main attributes of users which include prior privacy experience, computer anxiety, and perceived control. Degermenci conducted research on how these attributes can directly influence how users view their privacy in mobile applications. Within this survey, 775 participants were chosen to give their thoughts on mobile application privacy. This was done by having the participants fill out a 7-point rating scale and a 7-point Likert scale to answer the survey questions. From these results, Degirmenci conducted partial least square structural equation modeling to analyze the collected data (Degirmenci, 2020). This allowed them to create a latent variable correlation matrix, in which they could analyze which variables directly correlated with users' privacy concerns. Figure 2 highlights the results of this survey when shrunken down into a structural model for viewing. A negative or positive path coefficient value above .05 will note that the path is significant, thus the variables are correlated. The lower the path coefficient, the more correlated they are. For instance, a path coefficient of .185 will be more statistically significant than a path coefficient of .283. A negative path coefficient indicates an indirect

correlation, which means as one value increases, the other decreases accordingly. When the path value is positive, it means that as one value increases, the other increases accordingly.

Figure 2.

Effect	Model 1 (without APC)			Model 2 (with APC)		
	R^2	*Path coefficient*	*Effect size (f^2)*	R^2	*Path coefficient*	*Effect size (f^2)*
Mobile users' information privacy concerns (MUIPC)	**0.185**			**0.617**		
Prior privacy experience (PPE)		0.185***	0.039		0.137***	0.044
Computer anxiety (CA)		0.283***	0.092		0.157***	0.057
Perceived control (PC)		−0.164***	0.032		−0.051*	0.005
App permission concerns (APC)					0.686***	1.128
Intention to accept app permissions (INT)	**0.270**			**0.270**		
MUIPC		−0.520***			−0.520***	

Notes: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; Cohen's f^2-statistics = [R^2incl. – R^2excl.] / [1- R^2incl.] with $f^2 \geq 0.02$ = small effect size, $f^2 \geq 0.15$ = medium effect size, and $f^2 \geq 0.35$ = large effect size (Cohen, 1988).

(Figure 2) – Structural viewing model for path coefficients and effect sizes

From further analysis, it can be concluded that these attributes all directly correlate with mobile user information privacy concerns (MUIPC). It was discovered through analysis, that there was a direct positive correlation between computer anxiety and privacy concern. Users who portrayed traits of higher computer anxiety, also portrayed higher traits of privacy concern. Another study conducted by Smith et al. (1996), concluded that those who have been the victim

of previous personal information misuses have much higher anxiety and concerns regarding information privacy. In addition, users who have had prior privacy experience had a direct positive effect on information privacy concerns. This means that the more experience and knowledgeable a user is with information privacy, the more likely they were to be concerned about it. On the other hand, there was a direct negative correlation between perceived control and information privacy concern. The more that the user felt they were in control of their information, the less they were concerned for the information privacy and vice-versa.

These findings imply that app permission and policy concerns have an important effect on mobile users' concerns for information privacy (Degirmenci, 2020). Furthermore, Degirmenci (2020) backs up the findings from Solove (2008) in regards to how users' want to protect and understand how their information is being used. This is in hopes that users can protect themselves against unwanted damage that can harm them physically, emotionally, or financially. With these findings, researchers have been conducting studies on how information privacy can be improved through various techniques of interpreting privacy policies.

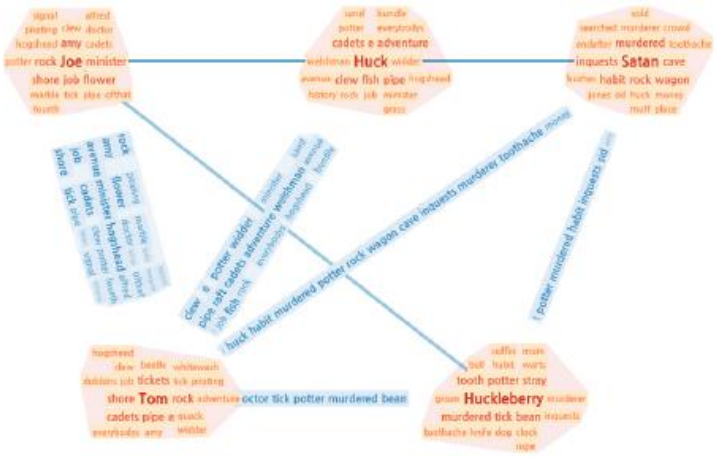
Visualization Techniques

In 2020, research was conducted on how various visualization techniques can help contribute to user privacy awareness rather than using traditional privacy policies. Visualization revolves around taking something that is primarily text-based, and transforming it into a different media for better understanding and analysis. This sometimes involves making cards with images, PowerPoint slides with images and limited text, or creative videos detailing a process or statement. Soumelidou et al. (2020) conducted an experiment to determine which visualization technique was the most effective in creating awareness of possible threats related to the disclosure of personal information on social media platforms. In order to do this, two

visualization techniques were chosen to analyze the privacy policy of a popular social media platform called Instagram. The two techniques were the WordBridge technique and the Document Cards technique. Alongside these techniques, a control group was established to compare the effect of having a visualization technique versus reading a privacy policy normally.

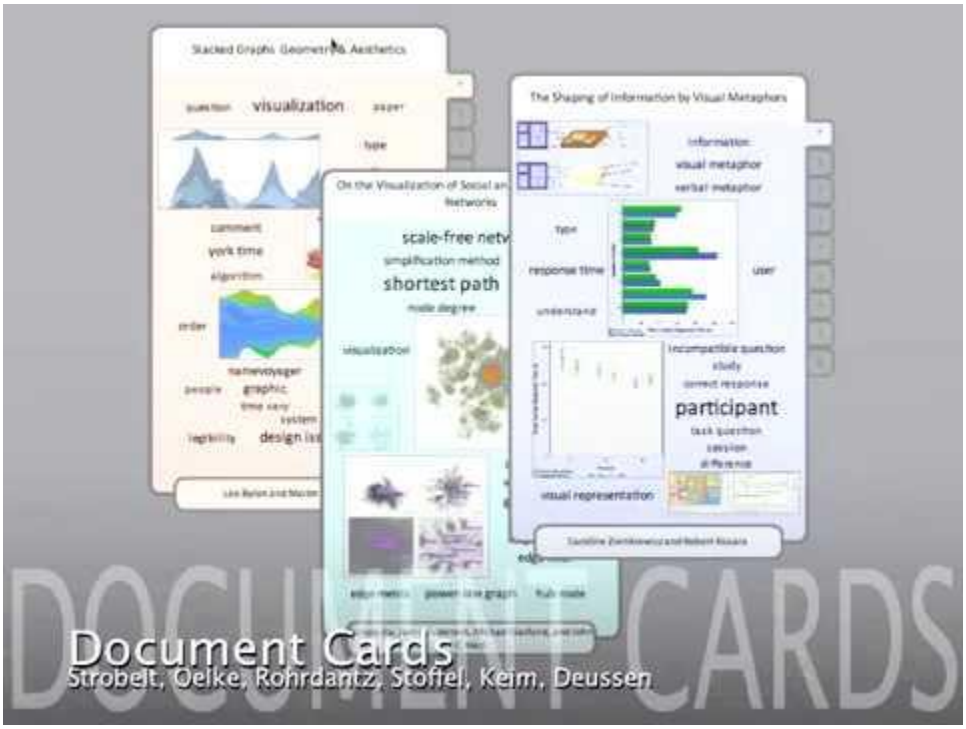
The WordBridge technique involves extracting keywords from the privacy policy which are then weighted and highlighted on a PowerPoint slide to show meaning. Two different types of tags, red and blue, were used to demonstrate this. Red cloud tags represented an entity that a privacy policy statement included, while the blue cloud tags showed the relationship between these entities. For the other technique, the Document Cards technique involves taking privacy policy statements and placing them on cards, each with an image, key term, and representative title. For instance, one card could be used for Instagram to show age limitations. Both techniques offer various ways for users to visualize these policy statements to better understand them. Figures 3 and 4 below outline examples of various WordBridge and Document Card techniques. These figures could be further applied to privacy policies outlined in social media applications.

Figure 3.



(Figure 2) – Example of a sample WordBridge technique

Figure 4.



(Figure 4) – Example of a sample Document Cards technique

Soumelidou et al. (2020) concluded that a visualized privacy policy was much more effective in contributing to a user's privacy awareness while on Instagram. This research raises concerns on what social media platforms could be doing better when conveying privacy policies. This concept of visualization is just one method that social media platforms could use to better convey their policies. Through these visualization techniques, they can become applicable to various privacy policies found within social media applications such as Instagram or Facebook. Visualization techniques will help increase users' privacy awareness and understanding. This increase in awareness will help impact users' prior privacy experience, computer anxiety, and perceived control as previously mentioned by Degirmenci (2020).

Implications

One of the main objectives of this research was to shed light on how companies disseminate user information through the agreement of their privacy policy when user create an account on their platform. The results of this research show an increasing number of users being concerned about the dissemination of their information on social media platforms. Through the analyzation of various studies conducted by researchers such as Solove (2008), Degirmenci (2020), and Steinfield (2020), it was found that the majority of users are concerned for the privacy of their information. Alongside this, it was also discovered that even with this concern for privacy, many users opt to skip reading the Terms of Service when signing up for a social media platform. With this information available, potential solutions to making privacy policies easier to read can be analyzed. There has been a recent wave of research done in hopes to figure out the best method for users to decipher privacy policies and better understand how social media companies can use their data. One of the most popular ways to currently decipher policies is called the visualization method. The visualization method takes the policy and converts it into easy-to-understand pictures with simple words. More research should be conducted with visualization methods in mind, to determine the most effective method.

From a corporate perspective, social media companies should be made aware of these flaws in their privacy policies. Lengthy documents with confusing jargon are not easy for a normal non-technical person to read and understand. Furthermore, reading the privacy policy during the Terms of Agreement should be required, rather than optional. A few social media websites already have this implemented; however, it should be absolutely mandatory on every single platform. These corporations should consider creating user-friendly approaches for users to more easily understand these policies. User-friendly approaches such as the visualization

method should be looked at further to determine its effectiveness. More research will be needed to be done in order to find alternatives to the visualization method.

Furthermore, future research should expand upon how not just companies can disseminate user data, but how users can disseminate other user's data. There have been numerous circumstances where individuals have used bots to scrape information off of public accounts and sell that information to third party companies. Any social media user is able to follow an account, publicly view their data, and then create another account to impersonate them. With the ever looming threat of identity theft growing each year, this is becoming even more of a problem which can cause negative effects financially, emotionally, and in some cases, physically.

Conclusion

In conclusion, the purpose of this paper is to raise awareness about how social media companies disseminate user information through privacy policies when a user creates an account on their platform. Information dissemination occurs when personal data is revealed or there is a threat of spreading information. Social media companies are able to collect certain types of information from user permission when they agree to the company's privacy policy. This information sometimes includes phone numbers, email addresses, first and last names, locations, and more. For the literature review, articles and research were chosen to connect the bridge between information privacy, information dissemination, privacy policies, and visualization methods. With more users signing up for social media each and every year, information privacy continues to be a top concern amongst users when using these applications. The studies in this paper helped reveal that most users forgo reviewing the privacy policy and, in some cases, when they view the policy, they can not find the correct information that they are looking for.

However, researchers have begun developing various techniques and methods to help individuals better understand privacy policies. A key method being used is called visualization, which helps transform policy jargon into recognizable words and pictures. Various studies have shown that visualization should be incorporated by social media platforms to help users better understand their rights.

This study may be used in the future to help users better understand their right to information privacy and how to better navigate privacy policies. The study may also be used to help expand upon various visualization methods to determine which is most effective for these types of documents. The study also opens up the potential for more research to be done in regards to other types of methods for understanding user information privacy.

Acknowledgments

I want to thank my faculty advisers and professors, Dr. Xiadong Deng and Dr. Thomas Lauer, for their support and guidance throughout this project. I also want to thank my friends, family, and coworkers for continually guiding me through school and motivating me to graduate from the Honors College. Lastly, I want to thank Health Alliance Plan for giving me the tools to succeed and a pathway for beginning my career once I have graduated.

References

Aygin, D., & Gül, A. (2020). Management of Patient Information and Privacy Protection. Is Ahlakı Dergisi, 13(1), 92-99. doi: http://dx.doi.org.huaryu.kl.oakland.edu/10.12711/tjbe.2020.13.1.0144

Clark. (2010). Social Media and Privacy. Air Medical Journal, 29(3), 104–107. https://doi.org/10.1016/j.amj.2010.02.005

Degirmenci, K. (2020). Mobile users' information privacy concerns and the role of app permission requests. International Journal of Information Management, 50, 261–272. https://doi.org/10.1016/j.ijinfomgt.2019.05.010

Gebre-Mariam, M., & Bygstad, B. (2019). Digitalization Mechanisms of Health Management Information Systems in Developing Countries. Information and Organization, 29(1), 1–22. https://doi.org/10.1016/j.infoandorg.2018.12.002

Juicer. (2023, January 13). *Your Social Media Data: What's collected and how is it used? - juicer social*. Social Media Feeds for your website. - Juicer Social. Retrieved March 13, 2023, from https://www.juicer.io/blog/your-social-media-data-what-s-collected-and-how-is-it-used#:~:text=The%20types%20of%20social%20media,Attitudinal%20Data%2C%20and%20Preference%20Data.

Labs, M., & ABOUT THE AUTHOR Malwarebytes Labs. (n.d.). *Tiktok's secret operation tracks you even if you don't use it*. Malwarebytes. Retrieved March 13, 2023, from https://www.malwarebytes.com/blog/news/2022/10/tiktoks-secret-operation-tracks-you-even-if-you-dont-use-

it#:~:text=While%20the%20data%20is%20used,aggregated%20reports%20sent%20to%20advertisers.

Lapowsky, I. (2019, March 17). *How Cambridge Analytica sparked the Great Privacy Awakening*. Wired. Retrieved March 13, 2023, from https://www.wired.com/story/cambridge-analytica-facebook-privacy-awakening/

McCallum, S. (2022, December 23). *Meta settles Cambridge Analytica scandal case for $725M*. BBC News. Retrieved March 13, 2023, from https://www.bbc.com/news/technology-64075067

Mutimukwe, C., Kolkowska, E., & Grönlund, Å. (2020). Information privacy in e-service: Effect of organizational privacy assurances on individual privacy concerns, perceptions, trust and self-disclosure behavior. Government Information Quarterly, 37(1), 101413–. https://doi.org/10.1016/j.giq.2019.101413

Solove, D. J. (2008). *Understanding privacy*. Cambridge, Mass: Harvard University Press.

Steinfeld. (2016). "I agree to the terms and conditions": (How) do users read privacy policies online? An eye-tracking experiment. Computers in Human Behavior, 55, 992–1000. https://doi.org/10.1016/j.chb.2015.09.038

Soumelidou, & Tsohou, A. (2020). Effects of privacy policy visualization on users' information privacy awareness level: The case of Instagram. Information Technology & People (West Linn, Or.), 33(2), 502–534. https://doi.org/10.1108/ITP-08-2017-0241

Vigderman, A. (2023, January 23). *How much would you sell your social media data for?* Security.org. Retrieved March 13, 2023, from https://www.security.org/blog/how-much-would-you-sell-your-social-media-data-for/#:~:text=You%20may%20not%20know%20this,third%20parties%20for%20targeted%20advertisements.

Zhou, T. (2020). The Effect of Information Privacy Concern on Users' Social Shopping Intention. Online Information Review, 44(5), 1119-1133. doi: http://dx.doi.org.huaryu.kl.oakland.edu/10.1108/OIR-09-2019-0298

Zuboff, S., & Schwandt, K. (2019). *The age of surveillance capitalism: the fight for a human future at the new frontier of power*. Profile Books.

Appendix A

Instagram and Facebook Privacy Policies

Instagram privacy policy: https://help.instagram.com/155833707900388

Facebook privacy policy: https://www.facebook.com/privacy/policy/