

# Analysis on the Control of Data Collection, Privacy, and Usage

Yongzhao Ye

University of Washington Tacoma, Computer Science and Systems, Tacoma, WA, 98404, United States.  
Corresponding author. Email: gaoming@cas-harbour.org

## ABSTRACT

Nowadays, social networking has become the main way of socializing for people. Through the Internet, people can shop, chat, watch videos, check emails, and so on. Most people may not have noticed that every time people use a network device, their browsing records and personal information are transmitted in some ways. This paper is written by reading some literature and related cases on information technology. The purpose of this paper is to discuss the phenomenon of social mass network data being collected today, some of which illustrate how information is collected and how collectors use it. After studying several documents, some methods to prevent information data from being collected are summarized. The paper also reminds data collectors of the potential risks and impact of their actions on the public. The conclusion is drawn as follows: There is a need to strengthen the supervision and punishment for breach of online privacy and security. There is also a need to empower the weak so that they can protect their privacy. Through this paper, the author hopes that the public can have a clear understanding of the social phenomenon of personal information being collected and used.

**Keywords:** Data Collection, Privacy, Policy, Online Information

## 1. INTRODUCTION

Technology refers to a large body of information and technologies that make it easier to use resources, thereby increasing the productivity and creativity of life operations. Technological progress is essential for economic growth and development, and the more advanced the technology, the faster local and global economies can improve. In terms of time savings, such as in the manufacturing of goods or the delivery of services, technology plays a significant role in economic growth. It also improves specialization and labor division while contributing to the production rate efficiently. Technology has demonstrated its impact on business and government activities, and through internet connectivity, information technology has emerged as a critical component in the growth of international trade and information sharing. Through this paper, the author hopes that the public can have a clear understanding of the social phenomenon of personal information being collected and used. The network can bring convenience to people, but also provide "convenience" to technical personnel so that creating a harmonious and fair

network environment is the common goal of our society today.

In the social, economic, and political spheres, technology plays a crucial role. Technology has introduced a lot of functional tools that can help increase work and life quality, and this is mostly because of efficiency. Cars, for example, reduce travel time, televisions provide real-time national news, and workshop technology increases the quality of products produced. Second, it functions as a networking medium, allowing society to engage in meaningful and cross-cultural communication while also boosting economic growth. Third, it allows humanity's future growth to expand. People's lives are becoming more comfortable as technology improves, and there is more research and learning material available at the same time. People will explore the unknown future using current technologies.

## 2. THE DATA THAT TURNED THE WORLD UPSIDE DOWN

Although the value of data is mostly measured in terms of productivity, big data has transformed the political landscape. Big data appeared to have been used in candidate campaigns for the US Presidential

Elections. The article explains how this data use became possible and how insights into voter profiles became important in campaigns that swayed political alignment in favour of specific candidates. Big Data highlights Kosinski's method of profiling individuals, which he developed based on social media strategies used in the campaigns of several candidates in the 2016 United States Presidential Election. The main technique emphasized in the article is 'microtargeting,' which is the use of insights from voter profiles to individualize campaign strategies and thus sway political alignment in a particular candidate's favor.

Personal information was first obtained from a variety of sources, including land registries, vehicle data, shopping data, bonus cards, club memberships, magazines read by the public, and even churches. Personal data is for sale, according to globally involved data brokers such as Acxiom and Experian. If anyone wants to know where Jewish women live, for example, they can simply buy this information, which includes phone numbers. This data is combined with the Republican Party's electoral rolls and online data by a firm called Cambridge Analytica, which produces a Big Five personality profile. Digital fingerprints are transformed into actual people with fears, desires, passions, and addresses [1].

Microtargeting was mentioned as one of the techniques used in generating big data from online sources. This is a marketing strategy that uses consumer data and demographics to identify the interests of individuals or small groups with common characteristics and influence their actions and interests. The goal of microtargeting is to launch an initiative to reach out to a specific audience in order to deliver a message via a preferred communication channel.

During the 2012 US Presidential campaigns, this technique was also used to appeal to and interact with voters. To improve personalization on a large scale with detailed information, campaign managers were able to use predictive analytics to create a model that represents people's sentiments. With this information, campaign leaders were able to go beyond the standard political party-oriented speeches and instead communicate personally with voters on specific topics that had a greater influence on their electoral decision [5].



**Figure 1** Groundgame, an App for election canvassing that integrates voter data with “geospatial visualization technology,” was used by campaigners for Trump and Brexit.[1]

### **3. HOW THE BIG DATA WAS USED FOR THE WRONG PURPOSE**

Big data technologies have been shown to be effective in the business world. Big data analytics improves transaction and financial record efficiency, and the valuable data generated is used in environments that process, manage, and analyze various types of data. Customers' databases, emails, internet records, log files, images, social media posts, and even medical records can be made available to organizations. While businesses use this to increase sales, other organizations also use big data, and this is not limited to businesses; big data is now used for political purposes as well [6].

Following Trump's election as President of the United States, Michael Kosinski, the leading expert behind Psychometrics who has lectured on the dangers of big data as well as the digital revolution, felt that something was wrong with the results and that there was something wrong with the way the data was used in this case as well as the Brexit issue. To demonstrate the scope of big data, consider that anything a person has done online or offline in the last five years has left digital traces. All Google searches, whether conducted on a home computer or a mobile device, are saved in a form that users are unaware of. People's data was collected using psychometrics, a tool that focuses on assessing personality traits based on the Big Five, which is also known by the abbreviation OCEAN. Due to the difficulty of the questionnaire, this strategy is an accurate approach to data collecting, but it has been made easier to access through the internet and Facebook. Openness, Conscientiousness, Extroversion,

Agreeableness, and Neuroticism (OCEAN) are acronyms for the coverage of this technique.

The brains behind this personality assessment technique are Kosinski and his fellow graduate student David Stillwell. It works like a typical Facebook app, where users can try to have their personality assessed by answering questions and allowing Facebook to share access to their profile with other researchers, with the option to share the results of the personality test online. This behaviour on the personality test app was utilized to collect as much data about social media users as possible, particularly on Facebook, which contributed to the Big Data [1].

In response to Kosinski's contemplation on how Big Data was utilized to rig the presidential election results, resulting in Trump's triumph. In her post for The Washington Post, Cohen expressed a similar attitude about social media regulation. People have long been aware of conspiracy theories and misinformation, as well as how their personal information is misused. If consumers pay attention to how they are being targeted by adverts on Facebook today, they will notice that this is because their profile or previous searches led companies to them and recognized them as potential customers. It is not just about the issues they might like to read about or about fake news [2]. Because of the rise of fake news, social media companies such as Facebook, Twitter, and YouTube have been purposefully magnifying their material to keep users interested and putting in stuff that triggers people's basic impulses, such as fear or outrage. People have lost their sense of democracy and equality of rights as they are swayed to express their ideas and rants on certain matters online through their social media accounts.

Based on this review of the use of big data. It draws attention to the fact that this is a normal occurrence in today's culture, even though it is extremely unjust to the majority of people. This, however, is one of the disadvantages of rapid technological advancement. Technology has advanced beyond most people's wildest dreams in the last 20 years, making several previously unthinkable items possible. Many people are struggling to keep up with the technological age's rapid speed. While some people are "using" technology, the majority of people are only doing so in a basic way. Although technology progresses, it is still necessary to focus on technology education for the general public. On the Internet, people can quickly locate any type of information.

People must first understand the designer/user dynamic as contingent—that is, we, the users, must see ourselves as potential designers. Then consider this: if we want an internet that collects data from all of us to increase the power and wealth of those who create automated and gig economy technology for personal gain [3].

#### **4. DATA PRIVACY AND TRANSPARENCY**

Collection of personal data should be committed to the ideals of accountability, valid intent, and proportionality, according to the Data Privacy Act of 2012's Implementing Rules and Regulations. The subject's openness refers to the subject's understanding of the existence, intent, and scope of personal data processing, as well as the risks and protections that are involved. Any information and correspondence related to the processing of personal data should be straightforward. The identity of the data controller, as well as his or her rights as a data subject and how they can be exercised. Any information and correspondence related to the processing of personal data should be simple to find and understand and should be written in plain English. As a result, information processing must be consistent with a stated and defined intent that is not in violation of law, morality, or public policy [2].

A large number of people use social media platforms; they log in to their accounts daily to search for updates, upload and post photos. However, these users are unaware that each time they log into their Facebook, Twitter, YouTube, and other social media accounts, the algorithms of these platforms are intensified, and as a result, impulses, anxiety, and anger, as well as other emotions, are activated as a result of the news shared on these pages.

Many people are unaware that their data is collected and used in their everyday online activities; raising awareness about this information will remind everyone about how their data is collected and used. This paper will inform technical workers at technology companies of the various consequences of their decisions, making them aware of the gravity of the situation.

Social networking sites make it simpler and more open for everyone to exchange knowledge, whether true or false, and they hasten the spread of toxic thinking. It is also a good way to spread conspiracies, as well as a way for foreign intelligence to get involved in domestic politics and spread advocacies [4]. Although restrictions have been placed on social media to rein in the free speech tag attached to these sites, drawing the lines which seem simple, filtering social media from the lies and conspiracies that have already invaded the system is a costly endeavour. It will take time to restore security to the unauthorized use of social media users' personal information.

Users' perceptions of data protection and technological advancement have shifted as a result of immunity. Although some have banned political advertisements and others have promised to make reforms, there will still be political advertisements for sale, and no one can stop it. As well as the purchase of users' data, which allows corporations and government officials to exploit data to their benefit.

## 5. DIGITAL FOOTPRINT

One of the reasons that big data can be collected, accessed, and even sold to interested organizations is because of the digital footprint. Each time a person goes online, he contributes to the growing portrait. The portrait in question was originally created to assist businesses in reaching out to their target market/consumers. This is why they use the background information and collaborate with advertisers to track the movements across multiple websites. As a result, whatever a person does online, he unconsciously leaves a trail. It is critical that everyone understands that they leave a trail behind them, that it is impossible to have zero digital footprints, and that the footprints may affect the user in the future. If one wants to effectively manage his digital footprint, he must first learn to limit or manage his digital identity. Every social media post, whether it is a photo or a post, leaves a trail, as does every email. Simple activities such as browsing different websites and reading articles all leave a trace online. People may not be concerned about their digital footprint, but if they learn how it affects them, they may reconsider how they manage their personal information online. It is not a question of whether the information shared is intentional or not, but when this information is gathered by advertisers and companies when people shop online, it contributes to the person's digital footprint [7].

## 6. CONCLUSION

This paper draws the following conclusions: There is a need to strengthen the supervision and punishment for breach of online privacy and security. There is also a need to empower the weak so that they can protect their privacy. Tomorrow's innovations have yet to be developed, and today's can be improved: the road is not "locked in." We need to empower the weak, such as stakeholders. The ethic of active collaboration can counteract negative consequences, such as climate change, that arise when a relentless drive for efficiency and profit pushes other values, such as clean air and water, to the back burner. It cannot expect policymakers and politicians to come up with ways to ensure that online platforms are open, accountable, and provide value to all of us. Too many leaders all over the world are obligated to submit to their richest backers and supporters.

As the number of individuals using social media grows, it is only right to tell them about how their data is utilized and which aspects of their profile are being accessed and gathered. They have a right to know when they sign up for accounts on these social networking platforms. Various online tests or activities that need a user to share access to his profile before receiving a response, whether for fun or serious purposes, should

educate the user about how his data or profile will be used and why it is necessary to share it. Social media platforms should not be let to run on their own; it is past time for the government to extend its Data Privacy Act and rigorously implement the regulations on social media applications to regulate what is disclosed about users' data and how far it can be used. People must also be educated to govern their behaviour so that they do not readily provide access to their data without first understanding how it will be utilized.

Through this paper, the author wants more people to pay attention to the social phenomenon of public information being collected and used, which has gradually penetrated into all parts of our lives. Most people leak their personal information without knowing it, so the author hopes that people can jointly resist this social phenomenon and add strength to a harmonious and fair network environment in the future. If the ideals of fairness, justice, and diversity are sincerely held by those doing the designing, technologies can be built with an eye on their short- and long-term impacts.

## REFERENCES

- [1] H. Grassegger and M. Krogerus. The Data That Turned the World Upside Down. *Vice* 8 January 2017.
- [2] S.B. Cohen. The 'Silicon Six' spread propaganda. It's time to regulate social media sites. *The Washington Post* 25 November 2019.
- [3] R. Srinivasan. *Beyond the Valley*. New York: MIT Press, 2019 285-315.
- [4] R. Srinivasan. Opinion: Automation is likely to eliminate nearly half our jobs in the next 25 years. Here's what to do. *Los Angeles Times*. 2019.
- [5] Tucci, Linda. "Microtargeting." *TechTarget* February 2013. <https://searchcio.techtarget.com/definition/microtargeting>.
- [6] Stedman, Craig. "The ultimate guide to big data for businesses." *TechTarget* 21 May 2021. <https://searchdatamanagement.techtarget.com/The-ultimate-guide-to-big-data-for-businesses>.
- [7] Eriksen, Kristina. "Your Digital Footprint: What Is It and How Can You Manage It?" *Rasmussen University* 16 May 2018. <https://www.rasmussen.edu/student-experience/college-life/what-is-digital-footprint/>.