

CHAPTER I

INTRODUCTION

1.1 Background

Humanity has recorded its feats since the cavemen era thousands of years ago, and this can be seen as the first form of data. Data has transformed through the ages, from cave markings to hieroglyphs to radio waves, morse code and now the digital era. Humanity now stores much more mundane information in digital forms through our actions, for example, every transaction on a blockchain leaves a digital footprint that will forever be stored in the digital space we call the internet. Technology has developed so far that humanity is able to project any information, be it verbal, oral, in the form of picture or video almost instantly to most parts of the world, this has undoubtedly boosted humanity into a new era, due to the ease of access of information and communication, allowing for global cooperation and collaboration at a much more casual level, greatly enhancing progress in all fields such as academic, corporate, manufacturing and much more.

As the world progressed into the 21st century, the value of data has been exponentially increased as technology has developed, this a result of the convenience brought by digital data that has allowed corporations to grow to a multinational level significantly quicker than before and this has led to the monetary valuation of data increasing exponentially, even surpassing the oil industry which

has held the top spot since humanity found a way to process crude oil into energy.¹ However, such development was not without its faults. The rising power of the internet in the late 1990s, brought about a craze, an almost fanatic belief of the prospective success of companies formed based on the internet that it created a bubble, now more infamously known as the dotcom bubble. Investors and venture capitalists were so afraid of missing out on a big payday, that risks were thrown into the caution, obviously a recipe for failure. This caused dotcom companies to throw more funding towards marketing to establish a brand rather than focusing on a functioning business and inevitably the bubble burst in a few years in late 2001 where investors faced steeped losses. However, this bubble did produce a giant that stands as one of the tallest today, Amazon.²

The top five highest valued companies in 2020 was Apple, Google, Microsoft, Amazon and Facebook, respectively.³ With Apple alone amassing US\$51.41 billion in 2020 alone, this goes to show the sheer size revenue that is generated by these companies, so the question is, how does data contribute to that revenue and to answer that question, it is imperative to learn how data is collected and used in the operations of these corporations and to assess if such utilisation is harmful to the average consumer.

¹ The Economist, economist.com, “The world’s most valuable resource is no longer oil, but data” <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>. Accessed 15/4/22.

² Adam Hayes, Patrice Williams, Investopedia.com., “Dotcom Bubble”, <https://www.investopedia.com/terms/d/dotcom-bubble.asp>. Accessed 28/8/22.

³ Marty Swant, Forbes.com., “The world’s most valuable brands” <https://www.forbes.com/sites/martyswant/2020/07/27/apple-microsoft-and-other-tech-giants-top-forbes-2020-most-valuable-brands-list/?sh=4d2770d13ada>. Accessed 28/8/22.

With the arise of new technology such as smartphones, smart TVs, smart watches, etc. All these gadgets collect data based on how its users act, for example, measuring heart rate on a run, determining the best route for travelling, finding the best deal for a product and many more.⁴ Such actions are all recorded and leave a digital trace and collected by corporations. However, such raw data is not particularly beneficial to a company, it needs to be processed and analysed before it can be utilised and, in most cases, such data serves as the basis for advertising, which contributes to a large proportion of how these companies generate their monumental revenue numbers. Using Google as an example of how data is used to garner profits, it was reported in 2020 that approximately 80% of Google’s revenue which amounts to about US\$144 billion, once again showing the scale of the profits that data brings.⁵ So, how does Google use data to get such great numbers with their ad revenue? First, Google collects data from its users through its applications and services, namely, Google drive, YouTube, Google maps, Gmail, etc. Predominantly, these services are accessed through electronic devices such as the computer or smartphones, as a result, Google is able to collect extensive amounts of data such as but not limited to, location information, contact information (e-mails, physical addresses, names, etc), search histories, browsing histories, User Identifiers (User ID and device ID) and usage data (data from using said services

⁴ Matt Burgess, wired.co.uk., “All the data Google’s apps collect about you and how to stop it” <https://www.wired.co.uk/article/google-app-gmail-chrome-data>. Accessed 4/9/22.

⁵ Statista.com., “Distribution of Google segment revenues from 2017 to 2021” <https://www.statista.com/statistics/1093781/distribution-of-googles-revenues-by-segment/>. Accessed 29/8/22.

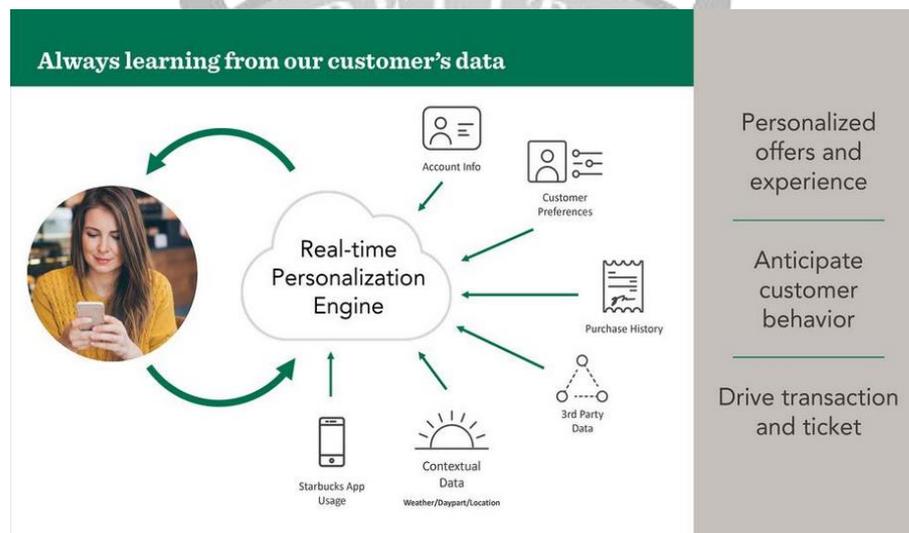
as well as advertising data).⁶ Google processes such data and distributes data into different categories, certain data is distributed to third parties such as usage data or location information while the rest are usually kept internally. As mentioned, Google uses such data to personalise and target advertising towards its consumers and this function is through algorithms and machine learning, where Google learns about the wants of consumers through their actions, for example, a person looking for a new running route would signal to Google that running shoes would be a good product to be shown to the consumer and as such personalises advertisements to cater to the consumers' needs. On a side note, it must also be mentioned that the data is also used to ensure that Google's services are functioning as intended through crash diagnostics which allows Google to fix any unintended problems that might have arisen.

One cannot deny the benefits data has brought to the world, it has definitely greatly benefitted the economy, because companies are able to better cater to consumers' needs while the consumers use the services provided by the company through several avenues such as machine learning. Furthermore, the collection of data would allow for companies to collect customer feedback on a much larger scale that would provide much more justifiable information than conducting research with focus groups instead. A great example of data collection can be seen in the 2016 Starbucks program, where Starbucks started using AI to send personalised promotions to its customers via E-mail, the customised promotions come from

⁶ Matt Burgess, *Op.cit.*

collecting data from their loyalty card and app programs to collect and analyse customer data, including where and when purchases are made.⁷ Below we can see how Starbucks illustrates the process to which they push out promotions to their consumers.

Annex 1.1 2016 Starbucks Advertising Program⁸



With this, it can be seen that Starbucks collects data in many ways, using a combination of first- and second-party data to enhance their marketing. As it stands, if each company is collecting data in this way and then going onto buying and selling such data to each other, it would bring about a collective increase in resource efficiency in the entire economy, bringing benefit on a much larger scale. In this regard, the benefit of trading of such data can be seen but everyone must remain

⁷ Terence Mills, forbes.com., “Five Benefits of Big Data Analytics And How Companies Can Get Started” <https://www.forbes.com/sites/forbestechcouncil/2019/11/06/five-benefits-of-big-data-analytics-and-how-companies-can-get-started/?sh=4d63967317e4>. Accessed 19/7/22.

⁸ Coral Garnick, bizjournals.com., “With artificial intelligence, Starbucks hopes to upgrade your next mobile coffee order” <https://www.bizjournals.com/seattle/news/2016/12/22/artificial-intelligence-starbucks-starbucks.html>. Accessed 19/7/22.

vigilant as such unrestricted access has brought large incidents as mentioned previously.

Clearly, data is viewed both lightly and severely by different parties, lightly by the consumers and heavily by corporations, such a stark perception is an ominous sign for the safety for consumers. It is clear that any rational corporation would push boundaries to further their bottom line and seeing that consumers either do not truly understand the severity of giving away such data or they understand the severity yet feel that the convenience brought by such services outweigh the danger of third parties exploiting such information, it is much too easy for corporations to collect such information. There are many arguments for such a lack of apathy towards data security, the inability to understand the technicality of certain functions such as cookies, or perhaps just ignorance or laziness to try to understand and acceptance that the data collection is inevitable and that it is hopeless to prevent such collection. According to the 2021 Deloitte Digital Consumer Trends Survey, which polled around 33,150 respondents in 20 countries about their online activities and their views on data privacy, shows that many surveyed consumers are well aware of the risks—but are continuing to expand their online activity nonetheless.⁹ This fault does not only lie on the consumer, as corporations disguise intention to exploit consumer data through vague description in their terms and conditions as well as little to almost no mention of how data is

⁹ Paul Lee, Joanna Conway, Roxanna Corduneanu, Suhas Raviprakash, deloitte.com., “Are data privacy concerns driving consumer behaviour? Not yet.” <https://www2.deloitte.com/xe/en/insights/industry/technology/protecting-consumer-data.html>. Accessed 29/8/22.

used after collection. Such lack of care has definitely impacted the world greatly, with many attributing huge scandals as a result of a lack of protection for consumer data such as the Brexit vote as well as the United States General Election. At the source of such a scandal was Cambridge Analytica, and Facebook, now known as Meta. Cambridge Analytica was officially known as a consulting company that specialized in using data science methodologies to support political campaigns, through the use of social media platforms such as Facebook to run personalised advertisements based on certain data parameters and was most infamously known for their assistance with political campaigning.¹⁰ Simplifying the issue, a testimony from an employee of Cambridge Analytica admitted that the company was hired to “persuade” several members of the British community through online political messages to vote for the Brexit. Following such a huge success, the company was hired to assist in Donald Trump’s campaign in the 2016 General Elections.¹¹ Riding the success of the Brexit campaign, the company applied similar methods and bought user data in the form of personal psychological profiles of approximately 250,000 Facebook users from a British academic. However, what was different this time was due to Meta’s privacy policy at the time, Cambridge Analytica was able to access the data of the friends of the original 250,000 users, which allowed them to collect and analyse the data of nearly 90 million users, however, these users did

¹⁰ Jason Fernando, Somer Anderson and Katrina Munichiello, Investopedia.com., “Cambridge Analytica” <https://www.investopedia.com/terms/c/cambridge-analytica.asp#citation-11>. Accessed 30/8/22.

¹¹ Mark Scott, politico.eu., “Cambridge Analytica did work for Brexit groups, says ex-staffer” <https://www.politico.eu/article/cambridge-analytica-leave-eu-ukip-brexit-facebook/>. Accessed 30/8/22.

not have knowledge of Cambridge Analytica taking such data and certainly did not give their consent for the company to be able to target advertisements to “sway” voters to further their purpose to make Trump the President.¹²

Another example of consumer data being exploited is the popular controversy regarding the sale of consumer data by Twitter. On two separate instances in 2013 and 2019, the Federal Trade Commission received complaints that Twitter sold their users’ data and after investigation, it was a result of new policy in 2013 that required users to include their email addresses and phone numbers onto the platform. The purpose on the surface was to enhance security and implement a two-factor authentication. However, it was discovered that Twitter was selling such information to third parties to aid advertisers to target their preferred audience, allowing advertisers to target specific groups of Twitter users, by matching the telephone numbers and email addresses to the advertisers’ own lists of telephone numbers and email addresses. As a result, the Federal Trade Commission fined Twitter an amount of US\$150 million for misleading their consumers for obtaining data from users on the pretext of harnessing for security purposes while using the data to target users with advertisement, boosting Twitter’s primary source of revenue.¹³

Following the events mentioned previously, it was evidently very clear to the world that the data industry needed to be regulated and supervised carefully due

¹² Jason Fernando, Somer Anderson and Katrina Munichiello, *Op.cit.*

¹³ Dan Milmo, theguardian.com., “Twitter fined \$150m for handing users’ contact details to advertisers” <https://www.theguardian.com/technology/2022/may/25/twitter-user-data-advertising-settlement>. Accessed 30/8/22.

to the significant influence data could bring at multiple fields, hence, the governments of the world adapted new laws to existing agencies or even created new agencies to better protect their citizens. Some examples of new regulations include the California Consumer Privacy Act from the State of California from the United States and the General Data Protection Regulation from the European Union. Such regulations will be explored in further detail later but a general summary is that it provided a much-needed overhaul for consumer protection, mandating corporations to be more transparent in how data is collected and giving freedom back to the consumers to be able choose which type of data is shared.

In the context of Indonesia, there have been much less evidence of companies exploiting consumer data due to several reasons, mainly that the tech industry in Indonesia is still in its infancy and therefore there is less opportunity to exploit and due to its relatively new state, Indonesia's tech sector is also definitely underregulated. However, due to Indonesia's huge population, Indonesia has a massive potential for tech companies. Based on a survey conducted by the Indonesian Internet Providers Association (APJII), it was revealed that internet users in Indonesia have increased from 175 million to 220 million users over the period from 2020 to 2022.¹⁴ This naturally conveys that a large majority of the Indonesian population actively utilises the internet and a large consumer market means that business can have access to larger economies of scale and as such Indonesia is a great prospect for the tech industry. As such, the purpose of this paper

¹⁴ Mahinda Arkyasa and Dewi Elvia Muthiariny, tempo.co., "Indonesian Internet Users Jumped by 45 million due to Pandemic: Survey" <https://en.tempo.co/read/1600176/indonesian-internet-users-jumped-by-45-million-due-to-pandemic-survey>. Accessed 8/9/22.

is to compare the current status of legal regulation concerning data protection in the United States with Indonesia's regulations in attempt to come up with suggestions to improve the structure and framework of such laws. This paper will focus more on the legislative and executive measures, exploring how the laws are written, and how the law is applied and executed by the respective Authorities or government agencies. The reason that the Californian law was chosen for this paper, was primarily because the area is generally considered to be the global centre for innovation and development in the tech industry and the CCPA was also one of the earliest data protection laws that set the standard of Consumer Protection with regards to data privacy issues.¹⁵

The Author does not deny the utility and benefit that comes from companies utilising consumer data to further cater to consumer needs and that governments are making the effort to protect its citizens from exploitation by corporations but through this paper, the Author will explore further how governments regulate consumer protection with respects to consumer data, exploring the terms and conditions of corporations and breaking down the obstacles that impede the protection of consumer data. The Author wishes to make a stand with respect to first party data trading, in which the Author believes that first party data collection cannot and should not be prevented but governments should step in to restrict the ability for first party data collectors to sell such data to other parties and only allow them to only be processed internally.

¹⁵ Troy Segal, Thomas Beck and Amanda Jackson, Investopedia.com., "Silicon Valley: Definition, Where Is It and What It's Famous For" <https://www.investopedia.com/terms/s/siliconvalley.asp>. Accessed 22/12/22.

1.2. Formulation of issue

In regard to the topic of this thesis, the Author will discuss the following formulation of issues:

1. How do the Californian and Indonesian governments regulate the commodification and trading of first party data?
2. How should the Indonesian Government address the dangers of unrestricted access and capability to trade first party data, particularly by Amazon and Tokopedia?

1.3. Purpose of Research

The Author's purpose of drafting this thesis is to answer the formulation of issues stipulated above, namely:

1. To identify the existence of legal protection with regards to consumer data.
2. To raise awareness with regards to the dangers of exploitation of consumer data.
3. To suggest potential methods to mitigate such dangers.

1.4. Benefits of Research

1.4.1. Theoretical Benefit

The theoretical benefit this paper would bring is to attempt to reform the society's perception towards their own personal data and to expose how

companies track people's activities and collect data from it and use said data to generate revenue without consent. It would allow for greater support for anti-trust laws and introducing the topic of data collection in the discussion of consumer protection. Much like how governments have started taxing MNCs, the Author hopes that the government will act in favour of its citizens and prevent private companies from exploiting them.

1.4.2. Practical Benefit

This paper can be used as the basis of consideration for the creation of new laws that would serve to better protect individuals from the manipulation from private corporations. The Author is hopeful that this paper would be beneficial for the general public and political parties to ensure the public's welfare is maintained and furthering of the public's interest.

1.5. Systematics of Writing

This thesis will be split into 5 (five) main chapters, that would allow the readers to fully comprehend the discussion presented within.

CHAPTER I: INTRODUCTION

As the first chapter, it will serve to introduce the concept of consumer data, provide some background on the topic and how it relates to the underlying issue wherein the trading of consumer data conducted by private corporations could potentially lead to the exploitation of consumers. This

chapter also raises the issues to be addressed, the purpose of the paper, the benefits this paper would bring and the framework of research.

CHAPTER II: LITERATURE REVIEW

This chapter will serve to establish the theoretical and legal framework with regards to the issue raised in the paper. Some of the theories include, the policies of tech corporations, the definitions of consumer data, the regulations concerning consumer protection. This chapter will explore some opinions from experts with regards to both the benefits and dangers of data collection. This chapter will also explore the dangers of exploitation of consumer data and assess how this has affected existing concepts of consumer protection.

CHAPTER III: RESEARCH METHODS

This chapter will formulate the methods of legal research used in the analysis of the issues. Such methods would involve the type of research used, a presentation of data collected, the process in which the research was collected and the course of action taken to deeply analyse the issues raised in the thesis. To elucidate, the processes include reviewing privacy policies of specific corporations in the United States and Indonesia, reviewing the respective regulations as well as observing the actions of the government agencies from the two aforementioned countries.

CHAPTER IV: RESEARCH RESULTS AND ANALYSIS

This chapter will be used by the Author to attempt to resolve the issues raised earlier. The primary method would be to use the information gathered

from primary and secondary sources, extrapolate and deduce conclusions from the data presented in the previous chapter. From these deductions, the Author hopes to bring across the dangers that consumer data can bring in the hands of private corporations and discuss whether existing regulations have provided enough protection. It will also introduce the idea of government intervention, firstly in the form of supervision of corporations' abilities to utilise consumer data and secondly, exploring the notion of regulating the trading of consumer data. This chapter will also include the cross-examination between Indonesia's and the United States' regulations to compare the regulatory functions and approach that the respective judicial systems took to protect consumer data. With the focus on the legislative and executive aspect, this chapter will highlight observations regarding the efficiency of investigations, the severity of punishments and other key aspects.

CHAPTER V: CLOSING

This is the final chapter that will conclude the findings of the previous chapters. This chapter will summarise the conclusions drawn from the previous analysis and give suggestions on the approach would serve to improve the corporate climate of Indonesia. It will also attempt to give practical recommendations to enhance the legal certainty and protection of consumer data as is the right of all Indonesian citizens.