The Dark Side of Social Media Data Collection and Retention

Janek Ratnatunga

Abstract

Everyone leaves a data trail behind on the internet. Every time someone creates a new social media account, they provide personal information that can include their name, birthdate, geographic location, and personal interests. Customers understand that their personal information is used to create relevant experiences for them. This is allowing marketers to create and deliver personalised shopping experiences that are exceeding expectations and leveraging customer loyalty. Thus, accessing and mining consumer data has become big business. Sometimes, companies share users' data with third-party entities, often without users' knowledge or consent. This laissez-faire attitude toward data access and mining has encouraged a new class of robber barons to arise. Data brokers, for instance, aim to provide their services from the shadows, while amassing billions and trillions of data points about people worldwide. They justify this by stating that whilst data is everywhere, and generated every second of the day, they are converting it to an asset – by turning it into something of value. This paper argues that by framing data appropriation as a theft of an asset, individuals can collectively lay the groundwork for policies to make it a legal and ethical issue.

Introduction

On September 21, 2022, Australia's second largest Telecommunication company *Optus* announced possibly the largest data breach in Australian history – compromising the personal details and even identity documents such as Medicare cards, passports and driving licences of up to 9.8 million Australians.

Australia's *Cybersecurity Minister Clare O'Neil* delivered a withering assessment of Optus' conduct, with comments that have flipped the narrative from one of a sophisticated raid, as Optus insists, to a "basic" hack; and put the job of chief executive, *Ms*. *Kelly Bayer Rosmarin*, in jeopardy. Questions are also being asked regarding conduct of the company's Singaporean parent, *Singtel*, and its 12-member board.[1]

However, the deeper questions that have gone largely unanswered are:

(1) Why did Optus collect such detailed personal details on their customers?
(2) Why was the data stored up to 6 years after the customer was identified?
(3) Why was the data not destroyed after the individual was no longer a customer of the company?
(4) What did Optus do with their customers' personal details beyond just 'storing' it?

Let us answer these questions in turn.

Legality Issues

[1] Lucas Baird and Mark Di Stefano (2022), "Optus' 'opaque' Singapore owner faces scrutiny over hacking attack", *Australian Financial Review*, Sep 30. https://www.afr.com/companies/telecommunications/optus-opaque-singapore-owner-faces-scrutiny-over-hacking-attack-20220930-p5bm5u

The *first question* is the easiest to answer. Optus had a legitimate need to collect that data – to verify customers were real people and potentially to recover any debts later. This is known as a "know your customer" (or "KYC") requirement.

With regards to the *second question* as to why the data was kept for 6 years, Optus has said it is *legally required* to do so.

This claim is questionable.

It is correct that there is an industry best practice 'code' called the *Telecommunications Consumer Protections Code*, that is overseen by the *Australian Communications and Media Authority*, which requires telecommunication companies to provide customers (or former customers) billing information for "up to six years prior to the date the information is requested".[2]
However, on closer inspection, only a customer's name, address and account reference number should be retained by Optus to meet this requirement - not to keep a customer's passport, driver's licence, or Medicare details for a lengthy period of time. If Optus needed to confirm a customer's identity again, it could simply ask for the documents again.[3]

The only clear 'legal' requirement for Optus to keep "information for identification purposes" comes from the *Telecommunications (Interception and Access) Act 1979*, which requires that identification information and metadata be kept for two years (to assist law enforcement and intelligence agencies).[4]

The *third question* as to why the data was not destroyed opens a *'can of worms'*.
The big problem with Australia's data retention laws is that there's really no limit on how long a company can keep personal data. Australia's Federal *Privacy Act* says only that information must be destroyed *"where the entity no longer needs the information for any purpose for which the information may be used or disclosed by the entity".*[5]

With such a loose requirement, a company could argue it "needs" to keep customer information for anything – such as defending against a civil claim in court, as part of its corporate records, or for *marketing.* This is especially the case when customers have consented to those uses when they sign up for the services, another practice the *Privacy Act* allows.[6] This is a serious weakness with Australia's privacy laws.

[2] Communications Alliance Ltd (2022), *Industry Code C628:201 9 Telecommunications Consumer Protections Code Incorporating Variation No.1/2022 C628:2019 June 20,*
https://www.commsalliance.com.au/__data/assets/pdf_file/0011/64784/TCP-C628_2019-incorporating-variation-no.1-2022.pdf

[3] Brendan Walker-Munro (2022), "Optus says it needed to keep identity data for six years. But did it really?", The Conversation, September 30, https://theconversation.com/optus-says-it-needed-to-keep-identity-data-for-six-years-but-did-it-really-191498

[4] Office of the Australian Information Commissioner (2015), *Telecommunications service providers' obligations arising under the Privacy Act 1988 as a result of Part 5-1A of the Telecommunications (Interception and Access) Act 1979,* 29 July. https://www.oaic.gov.au/privacy/guidance-and-advice/telecommunications-service-providers-obligations-arising-under-the-privacy-act-1988-as-a-result-of-part-5-1a-of-the-telecommunications-interception-and-access-act-1979

[5] Office of the Australian Information Commissioner (2018), *Guide to securing personal information*, 5 June https://www.oaic.gov.au/privacy/guidance-and-advice/guide-to-securing-personal-information

[6] Op Cit., Munro (2022).

This brings us to the *fourth question*, and the focus of this article. Did Optus use our personal data for social media and targeted marketing purposes, either directly or indirectly?

Social Media Data

Today, customers understand that their personal information is used to create *relevant experiences* for them. This is allowing marketers to create and deliver personalised shopping experiences that are exceeding expectations and leveraging customer loyalty. Thus, accessing and mining consumer data has become big business. Data-driven analytic platforms are the main sources of collecting information on us: credit card companies, retailers, social media likes, shares and actions, web browsing history, emails, *Google Analytics*, and Ads. These are a few but all of some of the more effective ways that can, as per Forbes magazine, collect up to 1,500 pieces of data on each one of us.[7]

It comes as no surprise that most of us are not even aware of the amount of data being collected on us. It is important to note that consumers are now demanding more control over how their data is collected and used.[8]

Almost all internet users have noticed how some of the ads on the sites they visit seem to be a perfect match to their current interests. This is obviously not a coincidence. Advertisers would do just about anything the online environment allows them to do – even if it means breaking online privacy laws – to develop new ways to promote products. And the easiest way for them to find out an individual's likes and habits is keeping a close eye on their social media behaviour. As such, companies are collecting – and keeping – much more personal information than they need without a truly legitimate commercial or legal purpose.

Since the arrival of social networking sites in the early 2000s, online social networking platforms have expanded exponentially, with some the biggest names in social media in 2022 being *Facebook, YouTube, WhatsApp, Instagram, WeChat, TikTok, Sina Weibo, Twitter, Snapchat, and LinkedIn.* The massive influx of personal information that has become available online and stored in the cloud has put user privacy at the forefront of discussion regarding the database's ability to safely store such personal information. The extent to which users and social media platform administrators can access user profiles has become a new topic of ethical consideration, and the legality, awareness, and boundaries of subsequent privacy violations are critical concerns in advance of the technological age

Data Access Methods

Everyone leaves a data trail behind on the internet. Every time someone creates a new social media account, they provide personal information that can include their name, birthdate, geographic location, and personal interests. In addition, companies collect data on user behaviours: when, where, and how users interact with their platform. All of this data is stored and leveraged by companies to better target advertising to their users. Sometimes, companies share users' data with

[7] Kalev Leetaru (2018), "The data brokers so powerful even facebook bought their data - but they got me wildly wrong", *Forbes Magazine*, April 5, https://www.forbes.com/sites/kalevleetaru/2018/04/05/the-data-brokers-so-powerful-even-facebook-bought-their-data-but-they-got-me-wildly-wrong/?sh=65d82ed43107

[8] Kudzanai Makanza (2021), "How are social media platforms collecting and using your data?" Yellow Door Collective, November 10. https://yellowdoorcollective.com/blog/how-are-social-media-platforms-collecting-and-using-your-data/

third-party entities, often without users' knowledge or consent. Often companies collect so much data in so many ways that the companies themselves lose track of all the ways they are tracking us.[9]

There are several ways for third parties to access user information without their knowledge. Social media has opened up an entirely new realm for researchers (and hackers) to get information from normal posts and messages.

Take *Facebook* for example. Facebook's privacy policy says they can provide "*any of the non-personally identifiable attributes we have collected to advertisers*." However, applications on Facebook itself violate this policy in a number of ways by providing personally identifiable attributes to advertisers. For example, if a user clicked a specific ad in a page, Facebook will send the user address of this page to advertisers, which will directly lead to a profile page.

Facebook has also been scrutinized for the collection of users' data by *Cambridge Analytica*, which initially was collecting data from Facebook users after they agreed to take a psychology questionnaire. However, not only could Cambridge Analytica access the data of the person who took the survey, they could also access all of the data of that person's Facebook friends. This data was then used to hopefully sway people's' beliefs in hopes that they would vote for a certain politician in 2020.

The Advent of Data Brokers

Any discussion of how companies secretly acquire our personal private data without our informed consent must start with the *data brokers*. This shadowy world buys and sells our most intimate private information every day and we have no right to demand to know what the companies hold on us. The sheer size of the data broker industry means it would be nearly impossible even to ask each of the myriad companies individually if they have data on us. A tiny company we have never even heard of on the other side of the world may have based its entire business on selling our data every day that we never gave them permission to have, let alone sell. Making matters worse, an overwhelming majority of the data that several of the largest brokers hold on individuals is most often wrong – a dangerous situation as companies increasingly go beyond mere ad targeting towards basing real decisions on it.[10]

Common ways that data brokers facilitate assess individual personalised data by the invasion of their privacy are:

Data Scraping: This involves tracking people's activities online and harvesting personal data and conversations from social media, job websites and online forums. Usually, research companies are the harvesters, and sell the compiled data via data brokers to other companies that in turn, use these details to design targeted ad campaigns for their products and services.

Online social tracking: We all use the "Like", "Tweet", "+1", and other buttons to share content with our friends. But these social widgets are also valuable tracking tools for social media websites. They work with cookies – small files stored on a computer that enable tracking the user across different sites – that social websites place in browsers when you create an account or log in, and together they allow the social websites to recognize you on any site that uses these widgets. Thus, your

[9] Kalev Leetaru (2018), "Social media companies collect so much data even they can't remember all the ways they surveil us" *Forbes Magazine*, October 25. https://www.forbes.com/sites/kalevleetaru/2018/10/25/social-media-companies-collect-so-much-data-even-they-cant-remember-all-the-ways-they-surveil-us/

[10] Op. Cit., Leetaru (April 2018)

interests and online shopping behaviour can be easily tracked, and your internet privacy rudely invaded.

API Protocols: *Application Programming Interface (API)* is a set of routines, protocols, and tools for building software applications that allow software to "speak with other software'. The *Optus* breach is supposedly due to lax API protocols that enabled the hackers to access personal customer information. Thus, companies and individual researchers can collect and provide information that is not publicly accessible via API routines and protocols.

API data is supposedly anonymous when personal information was collected in mass, with no ability to match this information with specific people. However, the scandal between *Facebook* and the political consulting firm, *Cambridge Analytica* which was able to exploit a loophole to gather information on not only people who used the app but all their friends — without them knowing. All of this does not indicate that Optus *knowingly* gave their customers' private information in a usable form to data brokers and other third parties. But the very fact that they held onto customer information for 6-years, (when they were only legally required to do so for two-years); and in level of detail not required by their own Industry code, requires Optus to answer questions as to what their motives were for doing so.

The fact is that Optus was hacked, and therefore let us examine what the consequences that such a breach could entail.

Invading Social Media Privacy

With any service that puts a premium on personal information, there will be risks that individual data will be exposed whether by accident or through security loopholes. Once private data is obtained — via hacking or sale — there are several ways advertisers can invade an individual's social media privacy, take advantage of their data and make them a target for their ads.

Here are 8 reasons why social media is bad for an individual's personal data.

Closed groups and discrimination: There are specific social groups on Facebook — some of them are based on medical issues and addictions — where users share their experiences and issues freely in these groups, believing that a "closed group" affords them some anonymity. Unfortunately, a recent investigation found that membership lists of these groups can easily be found by researchers.[11] Personal information on these lists could, for example, be passed onto insurance companies and employers who could use it against individuals to either drop them from insurance coverage or fire them from their current position.[12]

Fake profiles and impersonation: Online criminals target social platforms because individual accounts are rife with personal information that they can use for a variety of purposes The information gathered can be used against you via blackmail or to impersonate you for reasons such as to obtain financial advantage.[13]

[11] Susan Morrow (2018), "5 social media site privacy issues you should worry about', *Infosec Institute*, January 30. https://resources.infosecinstitute.com/topic/5-social-media-site-privacy-issues-worry/#gref

[12] Roy Maurer (2018), "Screening candidates' social media may lead to TMI, discrimination claims", SHRM, April 23, https://www.shrm.org/resourcesandtools/hr-topics/talent-acquisition/pages/screening-social-media-discrimination-claims.aspx

[13] Norton (2022), "Top five social media privacy concerns', *Norton Reputation Defender*, January 12. https://www.reputationdefender.com/blog/privacy/top-five-social-media-privacy-concerns

Spam, viruses, and malware: Social media is a better, faster way to spread malicious content like scams and malware - more so than the run-of-the-mill spam emails one sees in one's inbox asking to help out a Nigerian prince. This is because, when someone gets a link from their friend or a social media contact, they're much more likely to click the link as they have no reason not to trust it.

Opening credit cards and bank accounts: If a cybercriminal gets enough personal information about an individual, he or she can open a credit card or bank account in that person's name. This is one of the grave dangers of the Optus data breach, and why the company is being pressured to pay for the reissue of all passports and driver's licenses of those affected.

Business fraud: If cybercriminals get a hold of a business account, they can believably trick people into thinking they are legitimate businesses. Then they can funnel money straight into their pockets by persuading unsuspecting parties to provide credit card information for products that they will never receive.[14]

Access tokens and third-party apps: *Cambridge Analytica* took the data of at least 87 million users without their knowledge after harvesting that data from people who partook in a third-party quiz app (referred to earlier in this article). *But 87 million people didn't take the quiz!* The app took advantage of a Facebook loophole that allowed it to get the information from the quiz but also all of their friends' data and information as well.

Location-based apps: Most people's smartphones already automatically track and collect location data. Social media apps are especially interested in their data because it gives insight into their habits and whereabouts which advertisers can use to target ads to them at certain times of the day.[15]

Invasive privacy agreements: Most of us agree to privacy terms and conditions with reading them. Most do not realise that privacy agreements on many social media apps state that the content users upload including pictures, videos, and messages are owned by the platform, even if you decide to delete your account.

Legal Systems and the Modern Surveillance Society

The most troubling element of our modern surveillance society, however, is that under most international jurisdictions individuals have no legal right to know just how much they are being surveilled. Data-driven policing technologies are shielded by trade secrets, which prevent the public from knowing what data the analytics crunch and how it influences police activity.[16]

Companies are free to issue whatever public statements they like or to decline to offer details under the guise that doing so would "help bad actors". This is despite the fact that most western countries believe transparency is so sacrosanct that it outweighs the risks of so-called 'bad actors'.
People often do not even know how their data is taken and used, let alone how to give meaningful consent. When companies do seek consent, it is typically through terms of service agreements – overly long contracts are full of dense legal language that users are expected to "agree" to without

[14] Trulioo (2021), "How to verify legitimate businesses and merchants", *Truiloo*, March 31. https://www.trulioo.com/blog/verify-legitimate-businesses

[15] David Nield (2018), "All the Ways Your Smartphone and Its Apps Can Track You, *Gizmodo*, January 1. https://gizmodo.com/all-the-ways-your-smartphone-and-its-apps-can-track-you-1821213704

[16] Jathan Sadowski (2016), "Companies are making money from our personal data – but at what cost?", *The Guardian*, August 31. https://www.theguardian.com/technology/2016/aug/31/personal-data-corporate-use-google-amazon

understanding. It is a remarkable victory for the data appropriators that acquiescence has become the standard model for obtaining "consent".

Data as an Asset

This laissez-faire attitude toward data access and mining has encouraged a new class of robber barons to arise. Data brokers, for instance, aim to provide their services from the shadows, while amassing billions and trillions of data points about people worldwide. They justify this by stating that whilst data is everywhere, and generated every second of the day, they are converting it to an asset – by turning it into something of value.[17]

Thus, rather than allow them to unscrupulously take, trade and hoard our data, individuals must collectively reclaim their ill-gotten gains, by *changing the narrative.*
Data appropriation must be framed as a form of *exploitation* because companies use data to create value without providing people with comparable compensation. While some might argue that *Google* and *Facebook* pay us for our data with "free" services, this still does not account for the multitude of data appropriators that have no intention to provide some kind of mutual benefit to those whose data they possess.

Thus, by framing data appropriation as a theft of an asset, individuals can collectively lay the groundwork for policies to make it a legal and ethical issue. We need new models of data ownership, protection and compensation that reflect the role information has in society. This is where management accountants can be of most value not only to companies, but also to society.

Summary

Today, the management accounting profession has evolved beyond costing and budgeting. The profession is taking a leadership role in *environmental, societal and governance (ESG)* issues. Through mechanisms such as the 'Strategic Audit', companies already undertake *information security audits* to safeguard its data. Such audits should now be extended to *value* this information and *compensate* those who are providing it. After all, if an artist who has a song on *Spotify* can be compensated every time that song is downloaded, there is no reason that an algorithm cannot be developed to compensate those in society (individually or collectively) for the use of data taken from them by invading their privacy.

Unfortunately, as the Optus' storing of customer data for an unreasonable amount of time without their consent indicates, in today's world privacy is no longer viewed as a right, but instead it is a privilege of luxury no longer accessible to the majority of the world's population.

[17] ibid