THE UNIVERSITY OF MELBOURNE

State of the Art in Data Tracking Technology

Report written by Dana McKay, with contributions by Yung Ju Chua, Shanton Chang, Suelette Dreyfus, Monica Whitty, Jeannie Marie Paterson, Pan Zhan, Garreth Hanley and Andrew Clausen

November 2019

Funded by

Consumer Policy Research Centre

Acknowledgements

We gratefully acknowledge the support of Consumer Policy Research Centre (CPRC) in Melbourne, Victoria for their support in producing this report. We would particularly like to thank Brigid Richmond and Lauren Solomon at CPRC. Thank you also to Ed Santow, Lauren Perry and Zoe Paleologos at the Australian Human Rights Commission and Amy Grierson at the Office of the Australian Information Commissioner for providing deep background interviews for this report, and to DQUBE Solutions, E. Johnson and L. Ling.

Contents

Introduction

Trading in goods and services has always relied on data; whether it be what path to take to trade with the neighbouring village or which store has the best price on batteries.

While a true free market requires information symmetry [1], we operate in a market where traders hold increasingly more information than customers [2-4]. Traders have long tried to use information about their customers, whether it be innocuous student discounts, or catastrophic red-lining in mortgage markets [3, 5]. Similarly, traders have long studied consumer psychology to manipulate their customers into spending more [6]. What's new is that traders can now target both discrimination and psychological manipulation in a personalised way [7].

This new possibility is the result of traders being able to collect information from browsing history, phones (including location), Google searches, and smart camera technology, among others. Advances in machine learning mean that this data can be exploited on an industrial scale. Like other fields, the rise of big data, sensors, and artificial intelligence has changed the very nature of the relationship between purveyors of goods and services and the consumers who buy them predominantly to the disadvantage of the individual.

This report will outline the technologies used to track, monitor and understand consumers; identify the choices consumers have (or don't have) in the deployment of these technologies and describe how these technologies can harm or disadvantage consumers.

1.1 Our Approach to this Report

This report is based on a literature scan of the academic literature, industry and NGO reports, and news media looking for information. Search strings used include 'technology privacy breach', 'consumer price discrimination technology', 'consumer data price discrimination', 'consumer data artificial intelligence' and 'consumer data sharing'. Once specific examples of data use, harm or technology were identified, we looked for further examples using the terms identified in the original work. We further included materials encountered in the process of our regular news reading, and material provided to us by our social and academic networks. This scan is broad, rather than deep, and non-exhaustive, though we believe it represents the broadest coverage of data technology specifically focusing on consumer harm to date.

In preparing this report, the authors read or scanned more than 40 academic publications, and well over 200 news articles. The examples and papers included in this report are the tip of the iceberg; we have excluded for example a rich seam of academic literature on re-identification, deeply technical academic material on cookies and beacons, and nearly all of the academic underpinnings of artificial intelligence (AI) technologies and the ethics thereof. When choosing specific news publications to include, we looked for the quality of the source and the depth of coverage. Some non-traditional sources, such as Wired and The Verge are provide particularly extensive and thoughtful work in this space. Because the technology and its applications are moving so quickly, we have included such materials in our references.

Further background for this report was sought in conversations with the Australian Human Rights Commission and the Office of the Australian Information Commissioner. An exploratory interview of approximately thirty minutes each was conducted with high-level staff from each of these organisations. These interviews were used as background, and any ideas taken from them have been corroborated and extended with examples and ideas from the literature.

The technology and what it is publicly known to be capable of, has moved on even in the months this report has been in preparation. Example technologies have been chosen because their behaviours are possible (or likely to be possible) in an Australian marketplace.

As a counterexample, we have deliberately excluded the 'Absher' Saudi government app used to control

women [8]. Beyond this, we have looked for particularly novel, well-described, or invasive technologies to illustrate the overall themes of this report: the collection of consumer data and how it may be used to disadvantage consumers.

One example of technology being used for surveillance of consumers in Australia has proven to be a particularly good case study: that of smart advertising boards with embedded cameras such as those commonplace in shopping malls, train stations and movie theatres [9]. We call these smart kiosks, as they often have cameras in them taking images of consumers, with or without their knowledge. The socio-technico-legal ways in which these operate are included as a case study that forms an appendix to this report. This case study is partially based on early empirical work by YJ Chua and Alan Zhan, who interacted extensively with these boards attempting to determine their algorithmic underpinnings.

1.2 The Big Picture: How is Consumer Data Being Used in an E-commerce Setting?

Data on consumers is more readily available than at any time in history. It is being gathered, analysed and used to steer consumer choices at an unprecedented scale. Cookies, in-home audio recordings, in-home cameras on televisions and other Internet of Things (IoT) devices, captured images of consumers' faces merged with identifying data from their phones, consumer location merged with purchase history, and the intimate details of their lives shared on social media are all accessible to and potentially used by marketers and sellers of all kinds. Consumers are aware of and consent to some of these interactions. Trading one's location for recommendations for local restaurants, rather than those in New York City, is a trade-off many would make. Much of what goes on, though, is invisible to consumers. Data collected in public places, or with click-through terms and conditions, is often gathered without consumers being truly aware of it and without any meaningful way to opt out [9].

After the data is gathered, perhaps with terms and conditions which state 'we may share your data with our partners', what happens is even less visible. Data can be gathered, combined and recombined. AI can be applied to the data to learn about individual consumers, and consumers as a class. One example of this is the Hello Barbie, a talking doll for children. By 'conversing' with the child owner, the doll records information about the child, like their favourite food. Conversations with the doll are analysed online by ToyTalk computer servers, which then 'intelligently' provide 'artificial' responses for the doll to communicate with the child.

The ToyTalk doll is an example of essential modern AI: data is gathered and parsed through computers which have been programmed to assess the data and provide responses or suggestions. These learnings can be used in turn to make decisions about and manipulate or persuade consumers. They can select which products and services are promoted to an individual consumer, or determine how these products and services are priced [10, 11]. These uses of data can be explicit (simply not showing some consumers some options) or they can be less obvious. One such practice is 'nudging' [12], which is encouraging consumers toward the 'correct' choice by the subtle alteration of defaults. Another is 'dark patterns', steering consumers to buy or share more than they meant to through the interaction design of online services [13].

There is good evidence that the majority of consumers don't know that this is happening [9, 14, 15]. Even if they did understand it, being able to explain the algorithms underpinning this learning—and often decision making—is a challenge that the technology sector has yet to address [16-18].

There are some clear examples of data combination and differential pricing. Travel companies have been using data gathered about people shopping for a holiday to charge some consumers more, for example [19, 20]. In this particular example those accessing travel websites using iPhones were assumed to be more affluent due to the iPhone's higher price point. iPhone users were shown (and booked) more expensive hotels than those using Android phones—usually about $20 US more per night. The work in this paper tested other industries as well, including e-commerce department stores and hardware stores. There was some price inconsistency between operating systems and browsers when accessing these sites, but not at the level of statistical significance.

The data that is being used, and the ways in which it is being used today, are the beginning of a new world order. New data, and new ways in which to combine it, represents a future where consumers have little choice and where every part of them, their lives, loved ones and behaviour might be used. There is increasing potential for organisations to 'tailor' or 'personalise' based on what they know about individuals based on the data they have gathered. If the inferences they make from that data are incorrect due to incorrect data or individual variation, though, target consumers may find worse options than 'unpersonalised' options would give them. For all consumers, personalisation means not seeing all of their choices, or all of their choices presented in an objective, unedited way.

1.3 Looming Changes to the Data Landscape

There are two major changes on the data landscape horizon. The first is the type of data that is being gathered. The second is the ways in which that data is being combined and used to affect consumer behaviour. We address each of these in turn.

Consumers are already subject to scrutiny in public places. Security cameras and advertising technology capture their movement and their sentiments, how long they pay attention to advertisements and the route taken through physical space [9, 21]. It was recently discovered that aircraft have cameras built into the passenger seat backs [22]. Data from these cameras could potentially be combined with data from the booking system, which issues a unique booking number that can be tracked all the way through a trip. Fridges in stores are gathering data about the types of people who buy drinks [23]. If there were a way to combine this from the data of consumers' home fridges, which are now internet-enabled IoT tools [24], different prices could be displayed to less- and more loyal consumers. Voice-activated systems such as Siri, Alexa and Google are recording (potentially) every move we make in our homes. It is an open question as to when and how law enforcement can access those recordings. Tech media reports suggest that law enforcement have been accessing Amazon Echo data [25, 26]. They further note that the manufacturers of IoT devices are hiding from answering questions about whether their devices spy on end users, and whether this information can be or is passed onto the police [26]. Interestingly, Interpol and the FBI warned the public about the risks to consumers of IoT devices being compromised [27].Whether law enforcement bodies legally can access this data or not, there is some evidence the CIA has used IoT-enabled devices to access sound inside the homes of private citizens [28]. They can certainly be used by the tech giants who have sold us the devices to market to us more effectively, and potentially impose price differentials.

Data commercialisation is extending to our most unique data: our DNA. Commercially available at-home DNA tests were briefly available 10 years ago. They were ultimately withdrawn from the market because they were not accurate enough to meet consumer standards guidelines [29]. These tests proved problematic for Australian consumers, though: once you had learned you had a risk factor in your DNA, the terms and conditions of nearly any life insurance policy in the country required you to disclose it as 'any information that may be relevant' [30]. These tests have not been on the market for some time, but pre-natal genetic testing is on the rise, and the information available is increasing [31]. The time when our DNA is held by commercial organisations under commercial contracts—with the obfuscating language and get-out clauses seen in other privacy policies [14, 32]—is not far away. Losing control of one's own DNA data before one is born via a decision one's mother made is a very real prospect. Pre-natal DNA sequencing could affect a range of experiences for the resultant child, including purchasing health services, health insurance and life insurance.

Society is not so far from the proposed 'smart home' of the future that tells you what is missing from your refrigerator—or better yet orders it for you—and monitors your bodily waste to keep track of your health [33]. This could be used to benefit the health of individuals; more likely, however, is these technologies and the data they generate being used to benefit large companies. While this seems a long time away, internet-enabled refrigerators, bathroom mirrors and toilets are here now [34], and academics are trialling smart home ideas regularly [35]. The policy and legal arrangements made in the near future need to do more just protect consumers in the data landscape we have today; they must anticipate a data landscape where ever more data is available for commercial behavioural analysis.

Beyond just gathering data that is increasingly specific and sensitive, the ways in which companies can combine data are changing. Early attempts to determine individuals' sexualities from images are ongoing [36].

At this stage the technology is fairly inaccurate but this is likely to change over time [37]. This information could be used in a range of discriminatory ways (legal or otherwise), from advertising to denying housing to differentially pricing health insurance [38, 39]. Once data is aggregated or combined, it is no longer 'personal' under Australian consumer law, meaning that individuals have no right to review or correct it [14]. Thus, errors in these types of classification could be very difficult to track down or correct.

This new data landscape has the potential to impose significant burdens on consumers, through differential pricing, identity management and lost opportunities to access products and services [11, 18].

2. Historical Data Gathering

Data gathering about consumers pre-dates the consumer internet. Texas International Airlines launched the predecessor to the modern frequent flyer program in 1979. This program necessitated gathering the number of miles flown by each flyer [40]. Modern frequent flyer programs are enormously more complex [41], gathering the number of miles flown, travel patterns, and financial spend in an effort to market directly to consumers. As technology has advanced, the amount and type of data collected has been increasing rapidly.

Early store loyalty programs or 'card linked offers' gathered merely how much consumers spent at various stores. Modern equivalents of such programs gather not just how much is spent, but where it is spent and on what, along with personal data about customers such as their names, emails, and addresses. While gathering this data allows stores to provide 'more useful' offers to customers, it also represents a large privacy risk, particularly if accidentally leaked or sold to other data gatherers and combined with other data [42, 43] .

History has shown us that when significant data is gathered about consumers, they can be disadvantaged. The possible disadvantages include offering no or less service to certain consumer groups, offering services at a higher price to some groups, showing different options to some groups, and consumers knowing less about what is available than purveyors.

An early example of such a disadvantage is 'red lining' in housing lending. This practice used demographic data—such as ethnic background—about homeowners to determine whether a neighbourhood was a good lending investment; and thus effectively lock African Americans and other minority groups out of housing credit [44]. Ironically, red lining as a practice was detected and stopped by the use of data, specifically loans data that was made public through government reporting requirements, in an attempt to balance the rights of consumers of home loans with the rights of lenders [5]. This demonstrates that if policy balances the rights and needs of consumers with those using their data, the symmetry of power in the relationship can be re-asserted.

Other industries have also historically gathered data on consumers and used it in discriminatory and harmful ways. This might be as simple as gathering demographic data based on appearance—for example gender—and charging more for a product such as cars [45]. More recently and less visibly to end users, companies like Google have used voluntarily-provided demographic data in discriminatory ways; showing women fewer highly paid job advertisements than men. This discrimination arises from AI algorithms that have learned that women have lower paid jobs in general, and therefore showing them advertisements for jobs like the ones they have [46]. This demonstrates one of the problems with unquestioningly trusting computers and AI: they replicate human biases.

These historical examples show that even relatively simple data can be misused to harm consumers, by charging them more for products, by refusing to offer products or services, or by offering those products and services in a discriminatory way. Big data, pervasive data collection, data recombination and AI have the propensity to magnify these harms [47]—each piece of data a company has expands their ability to behave in ways that are discriminatory, even inadvertently. These new technologies also make it more difficult for consumers to understand how they are potentially harmed: AI is generally not easily explainable, even by experts, for example [16].

3. Data Collection Today

New technology affords advertisers and purveyors of data more chances to collect data than ever before, and in more locations. Consumers are frequently unaware of this tracking, and even where they are aware they may have little choice but to allow it or forgo participating in public life [9, 14, 32]. Data may be freely disclosed by consumers, it may be collected passively without their full disclosure or knowledge, or it may be inferred from other data—for example the inference that iPhone users are affluent, as discussed earlier. In this section we will address the types of data that are being gathered, and where that gathering is taking place; describe some of the technologies used to do the gathering, and where the potential for error is.

3.1 What (Personally-Identifiable) Data is Being Gathered

Australian privacy law defines personal information as being information from which an individual is reasonably identifiable [48]. Of course, identifiability falls on a spectrum. One's face is always an identifiable data point, and unique identifiers such as license plates on cars or identity document numbers can allow a person to be specifically identified. One's name may be identifiable, depending on how unique it is, but in combination with, say, a birthdate or a home address, it becomes more identifiable. At the other end of the spectrum individual locations or information about individual purchases is unlikely to be personally identifiable, however as little as four locations can be used to identify an individual if they have time data attached [49].

There are many explicit means by which these data are gathered—when we fill out a form to sign up for a loyalty program we understand that we are providing our names and addresses, and consent to providing it. When we stand in front of a camera checking passports at an airport, we understand that our faces are being captured and used to verify our identity. In isolation and where data is clearly being collected, Australian consumers both understand and are (mostly) comfortable with it [14, 15]. Where data is surreptitiously gathered, though, or where it is combined with other data for commercial purposes in unexpected ways, Australians are much more concerned [15].

Face Data

Face data is one of the most highly personally identifiable forms of data there is, and because it often includes biometric data, is classified as 'sensitive' under Australian privacy law. There are many places where it is obvious to consumers that facial data is being used: Smartgates at the airport would be one example of this, though they may still be objectionable to consumers. They also have ethical challenges: facial recognition technology disadvantages women, children or non-white consumers [50], so when—as in some places in the US—it is mandatory [51] these consumers receive a reduced level of service.

Another example of consumer-known use of face data is mobile phone facial recognition and face tracking. The iPhone X's face unlock is one early example of this [52], however many other companies have followed suit. For this unlock to work phones need to be looking for faces continuously [53]. What consumers may not know, or fully understand, though, is that the potential exists for this data to be passed to apps. While this is possible, Apple and Google (Android) have explicitly focused on keeping consumers' facial and other biometric data secure, storing face-map and fingerprint data within a secure enclave on the user's own device [54]. Arguably, though, it is the role of legislation, not individual tech companies to ensure that biometric data is handled ethically [54]. Legislation addressing exactly this issue has been proposed and has bipartisan support in the US, but has not yet been passed [55].

Some ways in which face data are surprising to consumers, though: many consumers would be surprised to know, for example, that their face is being tracked—if not recognised—through shopping centres and public spaces by advertising providers such as oOH! [56] And Quividi [57]. An article in The New York Times shows how easy it is to build a facial recognition engine cheaply and legally [58].

In a relatively extreme example of consumers being surprised by the use of facial data, travellers in China have been disturbed by a 'Creepy Kiosk' that uses facial recognition to provide flight information such as departure gate [50].

Even less likely is any consumer awareness of—for example—fridges in convenience stores which scan customer faces and combine the images with behavioural data to sell more products. This technology is already being implemented in the Walgreens fridge, which collects consumer sentiment from their faces and has the capability to alter what it shows consumers. It is considered so intrusive that it is banned in some US states, including Illinois [23].

Demographic Data

Demographic data, particularly in a high enough concentration (for example name plus birthdate, or name plus address) makes identification of individuals relatively straightforward. There are instances where we offer demographic data knowingly—for example to loyalty card programs, or when we buy something online. What many consumers don't realise, though, is that this data 'anonymised or aggregated' may be passed to 'our partners' for research [59]. Where individuals are especially security conscious, for example domestic violence victims, they may not choose to share an address in exchange for the benefits of a loyalty program (having their address shared has very real psychological and financial costs [60]). Arguably this creates disadvantage for such groups by not being able to access loyalty-targeted prices.

Health Data

There has been considerable public and media attention to the health data being collected by the Government as part of My Health Record. Attached to one's personal details, of course, My Health Record is both sensitive and personal under Australian privacy law. Many consumers are concerned about this—particularly about disclosing sensitive parts of their health information to individual providers [61].

The My Health Record Act also includes the provision to do research on 'de-identified' [62] health data—data from which demographic information has been removed. De-identification offers little protection for many patients; given only a very little health data about a known individual (for example the birthdates of their children, if they are female), it is fairly easy to identify them within a larger health data set [63]. Overseas health data examples have shown consumers are right to be cautious—Singapore has had two breaches in the last year [64].

What may be less obvious to consumers, though, is the that health data they voluntarily share with companies is being provided to other parties. Using a Fitbit (or other movement tracker), a health tracking app (such as a pregnancy app) or doing a commercially available DNA test could result in consumer data being shared with unexpected third parties, for example health insurers or employers [65, 66]. Again, this data may well be de-identified, but re-identification of a known individual is (demonstrably) not difficult [63].

Location Data

Many users share location data with apps for their own personal benefit. Google collects location data to offer personalised search services and its maps platform [67]; the Weather Channel uses location data to provide the weather through their app [68]. Users explicitly consent to and broadly understand these uses of their data, though they may not be aware of how few regular location points are needed to re-identify them [49].

What is not clear to many users, though, is that terms and conditions are used to provide that 'anonymous' location data to third parties; for example the Weather Channel app provides location data to a hedge fund company for monetisation [69]. Given that those providing their location data may well have invested in some of the companies affected by the hedge trading, this means that their data may have been used against them.

Similarly, in 2018 researchers discovered that Google, despite making changes to user data control, was continuing to allow movement and location tracking via its services. This affected people using Android powered mobile phones and Google Maps, even after customers had turned off 'location history'. Vague information and settings deep inside the Google data control interface led to many consumers and technology professionals thinking they had turned off location tracking, when in fact they had not [70]. Google tells customers when they disable location history 'some location data may continue to be saved in other settings, like Web & App Activity, as part of your use of other services, like Search and Maps, even after you turn off Location History' [71].

Often, consumers are unable to reasonably read, comprehend or manage the terms across all the services they use, leaving them vulnerable to entering into unfair or misunderstood contractual agreements which can include location tracking [72]. In 2008, it was estimated that the average internet user in the US would need to set aside 244 hours every year to read the terms and conditions for the websites they visited. In response, many companies have since vowed to make their services easier to use by putting users in control of their data and simplifying usage terms and conditions [73].

The availability of location information is only going to become more prevalent. The introduction of 5G mobile technology will increase these data transfer capabilities even further and enable precision location tracking. Malcolm Johnson, Deputy Secretary General of the International Telecommunication Union, emphasised the expected impacts of 5G mobile and location data during his recent presentation at the 2019 Geospatial World Forum. He said '5G, will act as the connective tissue of tomorrow's digital economy, linking everything from smartphones to wireless sensors to industrial robots and self-driving cars' [74]. This development will lead to data gathering and location data being embedded into the lives of every consumer via their mobile phones, computers, cars, fitness trackers and many other IoT devices. As network speeds and bandwidth capabilities increase, the tendency has always been to collect and transmit more data and it is foreseeable this trend will continue into the future.

Behavioural Data

Users provide behavioural data to a range of social media platforms when they communicate with or broadcast their activities to their 'friends'. This is explicit, though many users were angry when the Cambridge Analytica scandal [75] (or the earlier Facebook Mood research program [76]) revealed that these data were being used for research and commercial or political purposes. This practice is ongoing [77], and hugely beneficial to those who sell products and services; they can increase their profits significantly by using behavioural data to implement price discrimination [78], or even just to target marketing [43]. This targeted marketing must be done with care, though—where Target is selling pregnant women baby products, for example, it is possible to reveal a pregnancy a woman wants to keep hidden, or to make a woman feel 'spied on' and concerned [43].

While the data we provide to social media is in some ways explicit, most consumers do not realise that they provide behavioural data by moving through the world. Logging in to 'free' wireless networks may require users to log in to Facebook or Google allowing a third-party data dealer access to behavioural data [9]. The path of individuals through shopping malls or public spaces can be tracked by recognition cameras [79], and the movement of cars is tracked using parking tools that scan license plates [80]. This can be combined with other behavioural data, such as our shopping habits, which are monitored using loyalty programs—and used to determine who will buy more, and then to market to them [43, 81, 82].

Our movement over the web is increasingly tracked by cookies, and our reaction to certain types of emails is tracked by tagged pixels. Companies such as GlassBox allow online retailers to track mouse movements and behaviour within a site [19]; this in turn can be used to generate differential pricing for more 'desperate' consumers [83]. Similarly keystroke analysis can be used to detect engagement and emotional state; the potential to use keystrokes for sales purposes is readily apparent [84, 85], and in academic settings it is being used to create behavioural-biometric profiles to detect cheating [86].

Voice Technologies

Voice data is now being used as a form of identification by the Australian Government, so represents a new class of personally identifiable data [87]. A recent court summons of a user's Alexa data, in the hope of securing audio that had been inadvertently captured from the background of a possible crime scene, shines a light on how voice-based control might lead to significant privacy challenges [25]. Recently, it has been revealed that some of the voice recordings collected by digital assistants are also being reviewed by employees and contractors of the technology companies which provide them [88]. The companies have responded by saying they de-identify the recordings, and consumers are also able to adjust settings to stop any recordings being reviewed by humans [89]. The companies say this data is used for service improvement, but the privacy implications of collecting and retaining voice data have not been deeply engaged by tech giants [25].

3.2 Data Gathering Technologies

There are a range of technologies being deployed to gather data that didn't exist 20, 10, or even 5 years ago. These include mobile-phone based technologies, voice-based technologies, web-based technologies, and public scanning technologies. Some of these technologies have been mentioned in the previous section, but they will be summarised here.

Mobile Phone-Based Technologies

There are two major mobile technologies that are used for data gathering: apps and bluetooth beacons. The data collected by apps is only limited by the imagination of developers. Bluetooth beacons, without app support, can collect detailed movement data. Adding an application that interfaces between the beacon and a smart phone greatly increases the data it is possible to collect; subsequently using a pre-existing social media login to link that app to a person's social media account increases the data available to a developer exponentially [90].

Bluetooth beacons can track users to a relatively fine grain of detail through indoor- or outdoor space. This can be used to detect shopping patterns or other behaviours. Sometimes, as with the Auckland Airport study [91] in advance of the Rugby World Cup in 2011, this is used to benefit consumers, but again, this consumer tracking and data collection is being used passively and without consent from users.

Apps, dependent on the platform, may gather anything about a phone's user including their photos, their web browsing and communication history, and their location. Apps are required to disclose what they are gathering, but by using click-wrapped terms and conditions users cannot—and nearly always do not—give meaningful consent [14, 15]. In 2016 the Norwegian Consumer Agency, Forbruker Rådet, assessed the terms and conditions reading requirement for an average mobile phone user in Norway to be 250,000 words [72].

Many apps available to Australian consumers are developed and managed by companies overseas. A popular Chinese messaging app, WeChat, provides two versions of its user terms, one in English and another in Chinese. WeChat's parent company TenCent also has a separate website that is exclusively in Mandarin that highlights the full capabilities for WeChat itself. This website does not have an English version and non-Mandarin speakers would be completely unaware of the existence of such a website.

TenCent's Mandarin-language website promotes the advertising capabilities of the app and explains how the app collects the following data to 'learn' about its users; geo-location to determine resident or travelling status, gender, age, marriage status and education level. The app also builds an 'interest and behaviour' profile based on internet usage; this includes internet browsing, searching and reading of news articles and online shopping and 'related behaviours'. It uses activity frequency, type and time of interaction, and how long the behaviour has been occurring to generate user profiles. WeChat links this data to create a profile which also includes the device a person uses, the cost of the device, the network they use to access the internet, and device operating system. The TenCent website offers an example (in Mandarin) to illustrate how this data is being used to affect prices shown to the app's users; a translation of this is available in Appendix 1.

Voice-Based Technologies

Connected voice-based technologies such as Alexa, Google home and even toys such as talking Barbie are almost certainly constantly 'listening' for their activation commands (such as 'Alexa' or 'Hey Google'). Technologically, it has already been shown that when they operate in error, they record everything that happens in a home; early trial versions of Google Home did this. This data is gathered by the companies that own the devices, and used for service improvements. However, this data might also be used by law enforcement: this is still a legal question [25].

Web Based Technologies

There are a range of technologies that companies use to track behaviour on the internet. These include cookies, behaviour tracking and tracking pixels.

Cookies are not a new technology, but they are pervasive. When they were first used, they were specifically to assist users in retaining context and entering less data manually. However, they are much more commonly used now to track users and observe their behaviour. The problematic cookies from a privacy standpoint are 'third party cookies' which follow a user from one website to another [92].

Attempts were made to limit their use with do not track rules, however these rules are not enforceable, and as such many companies (including, for example, LinkedIn [93]) do not observe them. Tracking pixels are similar to cookies, but may be included in email as well as on web pages; they are predominantly invisible to the users of the content that contains them. They may also provide more information than cookies do to those gathering the data, including screen resolution.

Physical Presence Data

Many companies are now commercially tracking shoppers in public and semi-public places. Closed circuit television (CCTV) systems, Wi-Fi base stations and cameras in advertising screens are being used ever more frequently by shopping centre operators and marketers to track and analyse consumer biometrics and behaviour [9]. In 2018 one Australian shopping centre operator identified 12.3 million unique devices in their centres. The same operator logged 2.2 million devices connecting through their Wi-Fi network; the remaining 11.1 million devices being scanned for their unique identifiers without consumers connecting to Wi-Fi [94].

Physical presence tracking is used to segment and target audiences, determine the impact of advertising based on segment (for example age or gender), and watch the flow of human movement through shopping centre spaces. This segmentation is expected to drive differential advertising [94]. Some of these companies explicitly claim to be compliant with Europe's new privacy laws because they do facial tracking rather than recognition [95], but others make no such statement. This physical presence data should also be considered as a type of online data for the purposes of consumer protection. While it requires the consumer to be physically present, the unseen data gathering may be automatically sent offsite for data matching, merging and analysis.

Our team has conducted preliminary investigations with some of these technologies by standing near smart advertising boards, noting the ads served, and watching the advertisements they serve to others. Our conclusions are that these technology-enhanced billboards can operate in 'smart' mode, where the advertisements they serve are based on the audience segment engaging with them, linger time, and interest or 'dumb' mode where they merely serve up advertisements without regard for who is looking.

The suspicion is that the boards operate in dumb mode when the location is either very busy, or very quiet; in these circumstances tracking individuals is difficult. When there are a reasonable number of consumers, but not too many, smart mode may be active and serving tailored advertising. To further investigate these hypotheses, actors of a range of ages, genders and ethnicities would need to interact with the boards over a period of at least a day. Some of these boards also have Near Field Communication (NFC) devices; our team did not investigate what the outcome of using these devices might be.

3.3 Potential for Error

Any of these communication technologies are subject to the usual forms of error; data entry mistakes, technical errors due to mechanical or electronic failure. Consumers themselves also deliberately introduce data errors when they use false information with online services for privacy reasons [96]. The new data environment, though, gives rise to two novel types of error: combination errors and AI mistakes.

Combination errors occur when data matching with a range of sources takes place, and data from separate individuals is inadvertently ascribed to a single individual. Examples of this include some of the 'Robo-debt' mistakes; where individuals were identified as having debts to the government that did not exist by data matching [97], and mistaken traffic fines on the basis of incorrect information about who owns which license plate [98].

Far more difficult to combat or identify are errors introduced by AI, where attributes are ascribed to individuals that simply don't apply. AI is based on human assumptions, and reflects human prejudices [17, 99]. This means AI is more likely to discriminate against already disadvantaged groups. One example of this is the poor performance of facial recognition with women and ethnic minorities [51, 100], another is the poor performance of voice recognition with women's voices [101]. Similarly, the bodyscanners used in many airports are more likely to flag overweight passengers, or non-white passengers as requiring additional screening [102]. Because AI classifications come from a computer we are much more likely to assess and trust them as 'objective' [17]. AI is increasingly being used to target consumers and set prices; errors in these assumptions are difficult to explain to consumers [16], much less correct.

3.4 Data Merging

Gathering each of these individual types of data has the potential to be harmful to consumers but merging them has the potential to reveal more about consumers than many of them realise, and this is one area where major future risks lie. Companies like oOh!media Insights are already merging data to provide 'better insights' [103]. In a highly competitive and possibly skewed marketplace, though, 'better insights' often means 'less privacy', implemented in a way that may well 'spook' consumers [43]. It is also the merging of these datasets—which, where de-identified, is not protected by Australian privacy law—that has the potential to engender the greatest disadvantage for consumers. A low credit score (financial data), buying habits of cheap, unhealthy foods (purchasing data), and a low step count (health data) may well indicate poverty, for example; poverty is a known risk factor for a range of illnesses and short life span [104]. Given that re-identification is not difficult given enough information [49, 63], it is theoretically possible to determine much more about a consumer than it is legal to ask in an insurance application or job interview. Consumers showing the trends outlined above could well be further disadvantaged by more expensive life insurance by not being offered higher paying jobs. This might seem fantastical, but is in fact a rapidly approaching reality [105]. While more expensive health insurance is one problem, this technology is being used to identify Muslim cab drivers in the US, and gay men in Uganda—both of which represent a threat to those individuals' safety [47].

This data sharing is already happening commercially. Facebook has recently been in court over its practices with respect to advertisers: certain advertisers, primarily those who spent a lot of money with Facebook, were given much greater access to consumer data than Facebook users were led to believe [106]. This is an example of data sharing, but data amalgamation can happen through data sharing—when a single entity collects and matches data from a range of sources. For a good technical description of how consumer data is merged, see [107].

4 Consumer Self-Help

Australia has a comprehensive statutory consumer protection regime and privacy laws [14, 108, 109]. Nonetheless, the options for consumers who try to protect themselves from the collection, use and sharing of their data are limited. This is because the current mechanisms used by companies for obtaining consumer consent do not seek genuine, informed consent but are instead premised on a 'click-to accept' approach and reliant on technical or distracting terminology [32, 72]. The alternative is opting out of many aspects of modern life that we consider normal [83, 110]. None of these strategies allow full control, however, either alone or in combination.

4.1 Consumer Awareness and Consent

For consumers to control data collection two preconditions need to be met. First, they must be able to understand what is being gathered and why. Second, the facility to choose to opt out of elements of the product or service that require data collection must be available.

The first point—being able to understand what is being collected and why—is a major impediment to consumer control currently. In an attempt to show how difficult it is to begin to understand the terms and conditions imposed by digital retailers, an artist read Amazon's terms and conditions aloud. This took 9 hours [111]. While consumers are concerned about online privacy, fewer than half read privacy policies and terms and conditions all or most of the time; claiming they are too difficult, too lengthy, too difficult to understand or too hard to read as provided [14, 15, 32]. This means that consumers rarely understand what data is being collected and why: most consumers would not be comfortable with data being used for research purposes, for example, and all of the major online platforms use data this way [14].

Even where consumers do read information about privacy consents, they often cannot choose which parts of an agreement to consent to—consents are all-or-nothing 'click-wrapped' agreements [14]. If they want to use a service, they have to agree to the privacy policy and the terms and conditions. Most Australian consumers would prefer more granular control than this [15], but it is simply not offered by many online services.

Early work in this area has advanced a typology of consent for data gathering [107]:

- Forced consent: A clickwrap ‘take it or leave it’ approach
- Unforced consent: Fully informed, and where consumers have some control over their own data
- No consent: No information or opportunity to opt out is provided to users of a space or service.

To have any control over their data, consumers must be able to give unforced or genuine informed consent.

4.2 Consumer Control

There are some forms of data consumers can control—they are not (yet) required to use biometric markers such as their face or fingerprints to unlock their phones for example; nor do they have to have Wi-Fi or location switched on when they use their phones. In these situations, they may make lower-risk choices, but these choices are likely to come at a cost to convenience. Equally some services will allow users to access them without storing data—checkout as guest is the most common of these.

Where data is collected in a public or semi-public space such as a shopping mall, consumers have limited opportunity for control. If data is being collected by advertising boards or overhead cameras, consumers may not be aware of it, and there are certainly no readily accessible terms and conditions or opportunities for opting out. If consumers wish to enter the space, their consent is inferred from the very act of entry. In fully public spaces such as the street there currently there is little effort by companies to flag the use of the technology or provide options for consent. Canada, however, has strict consumer protection laws which has led to Google’s sister company Sidewalk Labs designing signage to show what surveillance technology is in use. This does not allow users of the public space to opt out, but does at least tell them what is going on [21].

Asking consumers to control their data by opting out of mobile phone service, of public spaces and of the internet entirely is simply not reasonable, though some commentators think this is the only way forward for privacy [110]. Mobile phones are one of the most affordable ways to access the internet [112], and some countries have declared internet access a human right [113].

Consumers can nonetheless take active steps to protect anonymity; the use of incognito browsing, VPNs or anonymous search services such as Tor is driven by a desire for privacy. The data on how many Australians use such services and why is limited. Use of VPNs rose substantially amid privacy concerns in 2015, to around 16% [114]; though most Australians who reported using a VPN last year were using it to access entertainment, rather than preserve their privacy [115]. The Australian Government has, in response to this, considered making the use of VPN technology illegal [116]. This recommendation was the opposite of what the Productivity Commission had suggested—enshrining the right of Australian consumers to access media at a reasonable price in law [117].

4.3 Anonymous Data

There are three ways to be anonymous with big data gatherers: when data is designed to be anonymous, when users can opt explicitly to be anonymous, or when users act pseudonymously outside of the terms and conditions of the service.

Designed anonymity

There are two types of anonymous or pseudonymous data: Born anonymous and de-identified.

It is possible to engage with some services in a born-anonymous manner. Electronic health records in Australia will permit anonymity for some healthcare users, though the rules for this are not made public [62]. Other services, such as physical presence services may collect identifiable information, but they claim it is never accessed in identifiable form; only as aggregate data that provides information (such as fixation on ads) needed by their clients [95].

Other data—such as the data Facebook provides to advertisers [118], or the information provided to researchers by the electronic health record system—is de-identified, or stripped of all information that could in theory lead to the identification of a single individual.

As we saw earlier though, the protection afforded by de-identification is limited, and its failings are discovered by identifying individuals [63]. Once an individual is identified it is too late—their privacy has already been lost.

Opting for Anonymity

Some services, such as the electronic health record scheme in Australia will allow users to opt for anonymity under certain circumstances [62]. This is a requirement of Australian privacy law [108], though what circumstances allow for it with health records are not clear. Equally, many e-commerce sites allow consumers to check out as guests. This has the benefit of improving conversion rates, largely on the back of privacy concerns, but does not entirely eliminate the collection of identifiable data [119].

Data collection in public space or semi-public spaces such as shopping malls does not facilitate user choice for anonymity in any way; this choice is made for consumers.

Using a Fake Name

Many consumers elect to use fake names on social media or with other online services [120]. This approach is prohibited under the terms and conditions of most online service providers (see for example Facebook's terms and conditions [121]). Nonetheless it is may be considered a social good to benefit particularly women and other disadvantaged groups [122]. It is a choice that many consumers may be making, although there are no reliable statistics on how many. The right to use a fake name has been protected by a German court; this could be a ground-breaking change in the right to anonymity [123].

4.4 Passive Data Gathering

There are many data gathering practices that consumers are unaware of or completely unable to control. These practices are those where data generated by consumers' everyday activities—such as surfing the web—are gathered without their consent. Internet service providers, mobile phone providers and websites gather a large amount of this data; and consumers have little or no ability to opt out. Cookies are not obliged to respect do not track [92], and ISPs are often legally obliged (including in Australia) to collect significant usage data [124]. While data gatherers may use this data to their advantage, service users have a low awareness of this type of data gathering. Even where they are aware, many users do not use technologies such as clearing cookies regularly or incognito browsing [15].

5. Where is the Harm?

It would be easy to say that there is little risk of harm to consumers in all this data gathering, and that it is reasonable for users to have data gathered about them. Such a cavalier approach, though, would ignore the values and preferences of consumers [14, 15, 125] and the real economic harms that may arise from this kind of data gathering. One major potential for harm is the same as where anything of value is gathered in a single location: theft, in this case identity theft. Crime against consumers is outside the remit of this report, though, and is not considered further here. There are three other kinds of potential harm that arise from the use of human data explained in this report: economic harm, denial of service or opportunity, and the undermining of fundamental rights.

5.1 Economic Harm

There are two kinds of direct economic harm that can arise from data gathering and ownership. The first, identified by the ACCC is lack of competition [14]. A competitive economic environment is perceived to be a de facto good for consumers, and having a lot of data about consumers gives major players a huge advantage over new entrants to the market [126].

The second kind of economic harm is differential pricing. This has been a long-standing issue in areas of geography, for example the high prices for poor quality fresh foods in 'food deserts' that are more likely to be occupied by low-SES and ethnic minority people [127]. With big data it has further been demonstrated in travel sites and hotel sites; iPhone users have been demonstrably charged more, for example [20].

Differential pricing is not always harmful [11]: some companies choose to charge more for those on higher incomes to make (for example) news accessible to all [128]. Equally the entire insurance industry is based on differential pricing according to risk, though recently we have decided that some kinds of discrimination (such as charging male drivers more) are not acceptable in some societies. Arguably some of these decisions entrench, rather than address inequality, but they demonstrate a legislative response to the issues [129]. What is certain is that for differential pricing to be ethical consumers should be aware of it, and the reason for the price differential should be clear to them [11].

5.2 Denial of Service or Opportunity

As in the red-lining scandal or the job advertising scandal outlined in Sections 2 and 3, sometimes big data will be used in a way that could result in some users being denied services or opportunities. This risk is particularly insidious when it is the result of derived or combined data. Recently scientists have claimed to be able to detect (for example) sexuality from face data [36]. While this has since been shown to be inaccurate [37], such data—if it were available—would be incorrect in some cases. It could nonetheless be used to discriminate in (for example) health insurance or other health services [130]. There would be little opportunity for consumers to even understand where this came from, let alone correct it.

More prosaically, the mere existence of big data allows companies to discriminate against those who do not have the finances or inclination to purchase a fitness tracker, or who live in exercise-unfriendly circumstances [65]. This is an early example of a practice that—without care—could become wider-spread and more commonplace. Is it within society's expectations of what is acceptable, for example, for insurance companies to offer a discount to those who were willing to supply the results of genetic testing? In Australia consumers are obliged to tell life insurers about genetic data if they know it about themselves. If foetal DNA sequencing becomes widespread and commercial [31], insurers may be able to access this information even about people who don't know it about themselves, possibly resulting in some individuals being uninsurable [30, 131].

A further and more concerning example of harm is the practices of Facebook that are the subject of a lawsuit brought by the US Department of Housing. Facebook's housing advertising is alleged to be inherently discriminatory in a form of modern-day red-lining: advertisements for higher quality housing are not shown to people of all incomes or ethnic backgrounds, even where advertisers specifically request that they are [132]. While we have no evidence that this sort of harm is occurring in Australia, we do have a rental market where renters feel discouraged from exercising their rights [133, 134]; in this climate rental discrimination would be easy to implement in Australia.

Finally, one of the major influences that this type of targeted advertising can have is on election outcomes. A significant part of the Cambridge Analytica scandal was the role it may have had in affecting the outcome of the US election. There is evidence that similar tactics are now being used here in Australia [135]. This case study illustrates the power of such collected information to manipulate end-users without their knowledge.

5.3 Challenges to Fundamental Human Rights

There are three kinds of challenging human rights harms that can come from the use of big data: discrimination, undermining of political and intellectual freedoms, and privacy breach.

Discrimination disadvantages the already-disadvantaged, and when technology companies are staffed predominantly by young white males, there is considerable discrimination inadvertently built into for example face recognition or voice recognition algorithms [99]. This can result in a differential level of service for those with accents or non-white, non-male faces [99, 100, 136-138]. This is particularly pernicious in a culture of early-and-often release of software, where testing may have been incomplete or lightweight, with a view that any problems would be identified and fixed post-release [17].

Once software becomes more accurate, it can then be used to further entrench discriminatory practices: once Microsoft improved its facial recognition software, social justice commentators pointed out it could be used unethically by law enforcement to target non-white Americans [139]. Other examples include the discrimination in which job advertisements get shown to women or ethnic minorities (see an example from Facebook here [140]); or the fact that African American defendants get harsher sentences than white defendants when sentenced using artificial intelligence [141]. This is clearly discriminatory and affects lives significantly beyond the scope of just a job advertisement or a sentence: women are paid less than men, and harsher sentences carry harsher social penalties in addition to the judicial ones.

Gathering big data can undermine political and intellectual freedom, particularly for more data-aware citizens. A data-aware citizen will be well aware, particularly given recent actions against (for example) the ABC that whistle-blower protections can be undermined by large scale data gathering [142]. Equally data-aware citizens are self-censoring online [143], and data-aware librarians are practicing censorship in selecting collection materials [144]. This means potentially not reading materials such as ‘how to get into university’ in communities where university attendance is not the norm. Each of these is a limitation to personal freedoms, both in expression and in information access, that would not happen were this data not being collected.

It can be argued that privacy is a fundamental right which is compromised by unauthorised data collection practices [125]. Moreover, the Office of the Australian Information Commissioner has identified distress as a clear side effect of data breaches, alongside financial loss and harm [145], as has early legal work in this space [125]. Gathering data, recombining it and making it available are all risk factors for private data being made public, or making individuals feel spied on and distressed [43, 125]. Legal regimes are only just beginning to come to terms with this kind of harm and how redress may be provided.

6. Conclusions

The new data landscape presents a range of new potential risks to consumers and challenges to regulators. Consumer data is gathered without consent or (often) even awareness, and without any process for anonymisation or deletion. Once gathered, this data can be used—and sometimes is used—to benefit commercial entities, potentially in discriminatory ways that are sometimes at direct odds with the needs and interests of consumers. This can happen even with the commercial entities being deliberately discriminatory; AI algorithms are making decisions that may be directly at odds with the values of some companies. This report has provided a summary of some of these types of data gathering, and the harms it can cause.

Any new regulatory response developed to protect consumers from discrimination and harm should take in not just the current data landscape, but reasonably expected future changes and challenges. This means it will have to address AI and the use of de-identified data, pricing discrimination and discriminatory algorithms. To enable sufficient regulatory agility, legislators should aim to future-proof legislation by describing principles, rather than technologies and rights and responsibilities rather than specific practices.

References

1. E Roy Weintraub, 'Neoclassical Economics', in David R Henderson (ed), *The Concise Encyclopaedia of Economics* (Liberty Fund, 2010).
2. Oren Bar-Gill, 'Seduction by Plastic' [2003] (4) *Northwestern University Law Review* 1373.
3. Oren Bar-Gill, *Seduction by Contract: Law, Economics, and Psychology in Consumer Markets* (2012, Oxford University Press).
4. Jon D Hanson and Douglas A Kysar, 'Taking Behavioralism Seriously: The Problem of Market Manipulation' (1999) 74 *New York University Law Review* 630.
5. Kirk Hallahan, 'The Mortgage Redlining Crisis, 1972-75' (Conference Paper, Proceedings of the Annual Meeting of the Association for Education in Journalism and Mass Communication, 5 August 1992).
6. Paco Underhill, *Why we Buy: The Science of Shopping* (2008, Simon & Schuster).
7. Eliza Mik, 'The Erosion of Autonomy in Online Consumer Transactions' (2016) 8(1) *Law, Innovation and Technology* 1, 1—38.
8. Bill Bostock, 'Saudi Arabia Runs a Huge, Sinister Online Database of Women That Men Use to Track Them and Stop Them From Running Away', *Insider* (online, 2 August 2019) <https://www.insider.com/absher-saudi-website-men-control-women-stop-escape-2019-1>.
9. Garreth Hanley, 'If You Go Down to the Mall Today, You're Watched by a Thousand Eyes', *Sydney Morning Herald* (online, 15 December 2017) <https://www.smh.com.au/technology/if-you-go-down-to-the-mall-today-youre-watched-by-a-thousand-eyes-20171211-h02h9q.html>.
10. Bernard Marr, '27 Incredible Examples Of AI And Machine Learning In Practice', *Forbes* (online, 30April 2018) <https://www.forbes.com/sites/bernardmarr/2018/04/30/27-incredible-examples-of-ai-and-machine-learning-in-practice/#28c457057502>.
11. Neil Levy, 'Online Sales and Difference Pricing', *Practical Ethics: Ethics in the News* (Blog Post, 29 October 2018) <http://blog.practicalethics.ox.ac.uk/2018/10/online-sales-and-differential-pricing/>.
12. Richard H Thaler and Cass R Sunstein, *Nudge: Improving Decisions about Health, Wealth and Happiness* (Yale University Press, 2008).
13. Thomas Germain, 'How to Spot Manipulative 'Dark Patterns' Online', Consumer Reports (online,30 January 2019) <https://www.consumerreports.org/privacy/how-to-spot-manipulative-dark-patterns-online/>.
14. Australian Competition and Consumer Commission, *Digital Platforms Inquiry: Preliminary Report* (December 2018).
15. Jayne van Souwe et al, *Australian Community Attitudes to Privacy Survey* (Report, May 2017)<https://www.oaic.gov.au/assets/engage-with-us/research/acaps-2017/acaps-2017-report.pdf>.
16. Patrick Ferris, 'An Introduction to Explainable AI and why we Need it' *freeCodeCamp* (Blog Post,28 August 2018) <https://medium.freecodecamp.org/an-introduction-to-explainable-ai-and-why-we-need-it-a326417dd000>
17. Claire Cain Miller 'When Algorithms Discriminate', *New York Times* (online, 9 July 2015)<https://www.nytimes.com/2015/07/10/upshot/when-algorithms-discriminate.html>.
18. Pekka Ala-Pietilä et al, *Draft Ethics Guidelines for Trustworthy AI* (Working Document, EuropeanCommission, 18 December 2018) <https://ec.europa.eu/digital-single-market/en/news/draft-ethics-guidelines-trustworthy-ai>.
19. Kathleen F, Major Companies Like Singapore Airlines are Recording Every tap and Swipe you Make on their iPhone Apps' *The Online Citizen* (online, 7 February 2019)<https://www.theonlinecitizen.com/2019/02/07/major-companies-like-singapore-airlines-are-recording-every-tap-and-swipe-you-make-on-their-iphone-apps/>.
20. Aniko Hannak et al, 'Measuring Price Discrimination and Steering on E-Commerce Web Sites' in *Proceedings of the 2014 Conference on Internet Measurement Conference - IMC '14* (ACM Press,2014) 305 <https://mislove.org/publications/Ecommerce-IMC.pdf>.
21. Katharine Schwab, 'These Street Signs Explain How You're Being Watched by Smart City Tech' *FastCompany* (online, 22 April 2019) <https://www.fastcompany.com/90337467/welcome-to-the-surveillance-city-these-signs-explain-how-youre-being-watched>.
22. Cydney Henderson, 'Some Airlines Have Seat-Back Cameras: Here's What you Need to Know' *USA Today* (online, 1 March 2019)<https://www.usatoday.com/story/travel/news/2019/03/01/some-airlines-have-cameras-installed-back-passengers-seats/3022742002/>.
23. Bruce Brown, 'Walgreens' Smart Fridges Scan Your Face and Remember Your Behavior', *Digital Trends* (online, 31 January 2019) <https://www.digitaltrends.com/home/walgreens-smart-fridge-spys-on-you/>.
24. Swapnil Bhartiya, 'Your Smart Fridge May Kill You: The Dark Side of IoT', *InfoWorld* (online, 3March 2017) <https://www.infoworld.com/article/3176673/your-smart-fridge-may-kill-you-the-dark-side-of-iot.html>.
25. James Vlahos, 'Smart Talking: Are Our Devices Threatening Our Privacy?', *The Guardian* (online,26 March 2019) <https://www.theguardian.com/technology/2019/mar/26/smart-talking-are-our-devices-threatening-our-privacy>.
26. Zack Whittaker, 'Echo Is Listening, but Amazon's Not Talking', *ZDNet* (online, 16 January 2018)<https://www.zdnet.com/article/amazon-the-least-transparent-tech-company/>.
27. Danny Palmer, 'IoT Security Warning: Your Hacked Devices Are Being Used for Cybercrime Says FBI', *ZDNet* (online, 3 August 2018) <https://www.zdnet.com/article/iot-security-warning-your-hacked-devices-are-being-used-for-cyber-crime-says-fbi/>.

28. Craig Timberg, Ellen Nakashima and Elizabeth Dwoskin, ‘Wikileaks Opens 'Vault 7', Claims CIA is Using Your TVs, Smartphones and Cars for Spying’ *Sydney Morning Herald* (online, 8 March2017) <https://www.smh.com.au/technology/wikileaks-opens-vault-7-claims-cia-is-using-your-tvs-smartphones-and-cars-for-spying-20170308-gut0gp.html>.

29. Amanda Holpuch, ‘FDA Orders Genetics Company 23andMe to Cease Marketing of Screening Service’ *The Guardian* (online, 26 November 2013) <https://www.theguardian.com/science/2013/nov/25/genetics-23andme-fda-marketing-pgs-screening>.

30. Margaret Otlowski et al, ‘Genetic Testing and Insurance in Australia’ (2019) 48(3) *Australian Journal of General Practice* 96, 96—99.

31. Megan Molteni, ‘How Much Prenatal Genetic Information Do You Actually Want’, *Wired* (online, 27March 2019) <https://www.wired.com/story/how-we-reproduce-testing/>.

32. Phuong Nguyen and Lauren Solomon, Consumer Policy Research Centre, *Consumer Data and the Digital Economy: Emerging Issues in Data Collection, Use and Sharing* (Report, 17 July 2018)<https://cprc.org.au/wp-content/uploads/Full_Data_Report_A4_FIN.pdf>.

33. NanoDTC Cambridge, ‘Intelligent Toilet?’ (YouTube, 6 June 2018) <https://www.youtube.com/watch?v=OLqOXN0OKv4>.

34. Andrew London, ‘The Smart Home Tech of CES 2018: From Heated Toilets to Connected Fridges’, *TechRadar* (online, 11 January 2018) <https://www.techradar.com/au/news/the-smart-home-tech-of-ces-2018-from-heated-toilets-to-connected-fridges>.

35. Kirsten Gram-Hanssen and Sarah J Darby, ‘“Home Is Where the Smart Is”? Evaluating Smart Home Research and Approaches against the Concept of Home’ (2018) 37 *Energy Research & Social Science* 94, 94—101.

36. Paul Lewis, ‘“I was Shocked it was so Easy”: Meet the Professor who Says Facial Recognition Can Tell if You're Gay’, *The Guardian* (online, 7 July 2018)<https://www.theguardian.com/technology/2018/jul/07/artificial-intelligence-can-tell-your-sexuality-politics-surveillance-paul-lewis>.

37. Alan Burdick, ‘The A.I. “Gaydar” Study and the Real Dangers of Big Data’, *The New Yorker* (online,15 September 2017) <https://www.newyorker.com/news/daily-comment/the-ai-gaydar-study-and-the-real-dangers-of-big-data>.

38. András Tilcsik, ‘Pride and Prejudice: Employment Discrimination against Openly Gay Men in the United States’ (2011) 117(2) *American Journal of Sociology* 586, 586—626.

39. Nathanael Lauster and Adam Easterbrook, ‘No Room for New Families? A Field Experiment Measuring Rental Discrimination against Same-Sex Couples and Single Parents’ (2011) 58(3) *Social Problems* 389.

40. Dawna Rhoades, *Evolution of International Aviation: Phoenix Rising* (Aldershot, 2nd ed, 2008).

41. Leslie Josephs, ‘Frequent-flyer Programs are too Complicated to Understand, the US Governmenthas Concluded’, *Quartz* (online, 22 June 2016) <https://qz.com/712665/frequent-flyer-miles-programs-are-too-complicated-to-explain-to-consumers-us-government-finds/>.

42. Daniel Graham, ‘Loyalty Programs and Privacy Law: What are Businesses Allowed to do with yourPersonal Information’ *Choice* (Blog Post, 15 January 2016)<https://www.choice.com.au/shopping/consumer-rights-and-advice/your-rights/articles/loyalty-programs-data-collection-privacy-law>.

43. Charles Duhigg, ‘How Companies Learn Your Secrets’, *The New York Times Magazine* (online, 16February 2012) <https://www.nytimes.com/2012/02/19/magazine/shopping-habits.html>.

44. Alexis Madrigal, ‘The Racist Housing Policy That Made Your Neighborhood’, *The Atlantic* (online,22 May 2014) <https://www.theatlantic.com/business/archive/2014/05/the-racist-housing-policy-that-made-your-neighborhood/371439/>.

45. David Harless and George Hoffer, ‘Do Women Pay More for New Vehicles? Evidence from Transaction Price Data’, (2002) 92(1) *American Economic Review* 270.

46. Amit Datta, Michael Tschantz and Anupam Datta, ‘Automated Experiments on Ad Privacy Settings’ [2015] (1) *Proceedings on Privacy Enhancing Technologies* 92.

47. Nathaniel Raymond, ‘Safeguards for Human Studies Can’t Cope with Big Data’ [2019] 568 (7753) *Nature* 277.

48. Privacy Fact Sheet 17: Australian Privacy Principles. [cited Available from:https://www.oaic.gov.au/individuals/privacy-fact-sheets/general/privacy-fact-sheet-17-australian-privacy-principles.

49. Yves-Alexandre de Montjoye et al, ‘Unique in the Crowd: The Privacy Bounds of Human Mobility’(2013) 3(1) *Scientific Reports* 1.

50. Fatema Patrawala, Is China’s Facial Recognition Powered Airport Kiosks an Attempt to InvadePrivacy Via an Easy Flight Experience?, *Packt* (online, 26 March 2019)<https://hub.packtpub.com/chinas-facial-recognition-powered-airport-kiosks-an-attempt-to-invade-privacy/>.

51. Harrison Rudolph, ‘DHS Is Starting to Scan Americans Faces Before They Get on InternationalFlights’, Slate (online, 21 June 2017) <https://slate.com/technology/2017/06/dhss-biometric-exit-program-is-starting-to-scan-americans-faces-before-they-get-on-international-flights.html>.

52. Arielle Pardes, ‘Facial Recognition Tech is Ready for Its Post-Phone Future’, *Wired* (online, 10September 2018) <https://www.wired.com/story/future-of-facial-recognition-technology/>.

53. Jack Morse, ‘Why the iPhone X's Facial Recognition Could be a Privacy Disaster’, *Mashable* (online, 28 August 2017) <https://mashable.com/2017/08/28/trouble-facial-recognition-technology-smartphones/>.

54. Kalev Leetaru, ‘Could Apple Ban Unethical Facial Recognition and Become The Patron Saint Of Privacy?’, *Forbes* (online, 2 February 2019)<https://www.forbes.com/sites/kalevleetaru/2019/02/02/could-apple-ban-unethical-facial-recognition-and-become-the-patron-saint-of-privacy/#5f1274ea30d0>.

55. Mariella Moon, ‘Face Recognition Privacy Act aims to Protect Your Identifying Info’, *Engadget* (online, 15 March 2019) <https://www.engadget.com/2019/03/15/face-recognition-privacy-act/>.

56. oOH! Media. [cited 4 April 2019]. Available from: https://www.oohmedia.com.au/.

57. Quividi. [cited 5 April 2019]. Available from: https://quividi.com/.

58. Sahil Chinoy, ‘We Built an ‘Unbelievable’ (but Legal) Facial Recognition Machine’ *New York Times* (online, 16 April 2019) <https://www.nytimes.com/interactive/2019/04/16/opinion/facial-recognition-new-york-city.html>.

59. Woolworths Rewards Privacy Policy. [cited 7 March 2019]. Available from:https://www.woolworthsrewards.com.au/privacy.html.

60. Ben Smee, ‘Queensland Police “Breached Privacy” of Domestic Violence Victim by Leaking Her Details’, *The Guardian* (online, 27 March 2019) <https://www.theguardian.com/australia-news/2019/mar/27/queensland-police-breached-privacy-of-domestic-violence-victim-by-leaking-her-details>.

61. Sophie Scott, Ariel Bogle and Laura Gartry, ‘My Health Record Deadline Looms, with Privacy Experts and Government at Odds’, *ABC News* (online, 30 January 2019)<https://www.abc.net.au/news/2019-01-30/my-health-record-deadline-looms-jan-31/10759956>.

62. My Health Record Privacy Policy. [cited 10 April 2019]. Available from:https://www.myhealthrecord.gov.au/about/privacy-policy.

63. Vanessa Teague, Chris Culnane and Ben Rubinstein, ‘The Simple Process of Re-Identifying Patients in Public Health Records’ *Pursuit* (online, 18 December 2017) <https://pursuit.unimelb.edu.au/articles/the-simple-process-of-re-identifying-patients-in-public-health-records>.

64. Eileen Yu, ‘Singapore Suffers “Most Serious” Data Breach, Affecting 1.5M Healthcare Patients Including Prime Minister’, *ZDNet* (online, 20 July 2018) <https://www.zdnet.com/article/singapore-suffers-most-serious-data-breach-affecting-1-5m-healthcare-patients-including-pri>.

65. Suzanne Barlyn, ‘Strap on the Fitbit: John Hancock to sell only Interactive Life Insurance’, *Reuters* (online, 20 September 2018) <https://www.reuters.com/article/us-manulife-financi-john-hancock-lifeins/strap-on-the-fitbit-john-hancock-to-sell-only-interactive-life-insurance-idUSKCN1LZ1WL>.

66. Drew Harwell, ‘Is Your Pregnancy App Sharing Your Intimate Data with Your Boss?, *Washington Post* (online, 10 April 2019) <https://www.washingtonpost.com/technology/2019/04/10/tracking-your-pregnancy-an-app-may-be-more-public-than-you-think/>.

67. How Google Uses Location Data. [cited 28 June 2019]. Available from:https://policies.google.com/technologies/location-data.

68. Weather Channel Terms and Conditions. [cited 28 June 2019]. Available from:https://weather.com/legal.

69. Makena Kelly, ‘The Weather Channel app Unlawfully Obtained User Location Data’ *The Verge* (online, 4 January 2019) <https://www.theverge.com/2019/1/4/18168373/los-angeles-weather-channel-app-user-location-data>.

70. Emily Dreyfuss, ‘Google Tracks You Even If Location History’s Off: Here's How to Stop It’, *Wired* (online, 13 August 2018) <https://www.wired.com/story/google-location-tracking-turn-off/>.

71. Manage Your Location History. [cited 25 June 2019]. Available from: https://support.google.com/accounts/answer/3118687?hl=en.

72. 250,000 Words of App Terms and Conditions’, *Forbrukeårrdet* (Web Page, 24 May 2016).https://www.forbrukerradet.no/side/250000-words-of-app-terms-and-conditions/.

73. Florian Schaub, ‘Nobody Reads Privacy Policies – Here’s How to Fix That’, *The Conversation* (online, 10 October 2017) <http://theconversation.com/nobody-reads-privacy-policies-heres-how-to-fix-that-81932>.

74. Anusaya Datta, ‘How are 5G and Geospatial Powering Future Cities?’, *Geospatial World* (online, 2April 2019) <https://www.geospatialworld.net/blogs/how-is-5g-and-geospatial-powering-future-cities/>.

75. The Guardian, *The Cambridge Analytica Files* (Web Page)<https://www.theguardian.com/news/series/cambridge-analytica-files>.

76. Robinson Meyer, ‘Everything We Know About Facebook's Secret Mood Manipulation Experiment’, *The Atlantic* (online, 28 June 2014) <https://www.theatlantic.com/technology/archive/2014/06/everything-we-know-about-facebooks-secret-mood-manipulation-experiment/373648/>.

77. Sue Halpern, ‘Why the UK Condemned Facebook for Fuelling Fake News’, *The New Yorker* (online, 22 February 2019) <https://www.newyorker.com/tech/annals-of-technology/why-the-uk-condemned-facebook-for-fuelling-fake-news>.

78. Anna Bernasek and DT Mongan, ‘Big Data is Coming for Your Purchase History - To Charge You More Money’, *The Guardian (*online, 29 May 2015) <https://www.theguardian.com/commentisfree/2015/may/29/big-data-purchase-history-charge-you-more-money>.

79. Quividi, Westfield USA (Web Page) <https://quividi.com/solution-spotlight/westfield-usa/>.

80. Allie Coyne, ‘Westfield Ditches SMS Feature Over Privacy Issues’, *ITNews* (online, 3 February2016) <https://www.itnews.com.au/news/westfield-ditches-sms-feature-over-privacy-issues-413991>.

81. Quantium, ‘Q Shopper’ (Web Page) <https://www.quantium.com/q-shopper/>.

82. Bernard Wilson, ‘You Don’t Need a Loyalty Program’, *Quantium Insights* (Blog Post) <https://www.quantium.com/insights/you-dont-need-a-loyalty-program/>.

83. Shoshana Zuboff, *The Age of Surveillance Capitalism: Fight for a Human Future at the New Frontier of Power* (Profile Books, 2019).

84. Martin Pielot, 'When Attention Is Not Scarce - Detecting Boredom from Mobile Phone Usage' in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and UbiquitousComputing - UbiComp '15* (ACM Press, 2015).

85. Mihai Dascalu, 'Predicting Student Performance and Differences in Learning Styles Based on Textual Complexity Indices Applied on Blog and Microblog Posts: A Preliminary Study' in Katrien Verbert, Mike Sharples and Tomaz Klobucar (eds), *11th European Conference on Technology Enhanced Learning (EC-TEL 2016): Adaptive and Adaptable Learning* (Springer, 2016) 184.

86. Hare, J., 'Start-Up Cadmus Set To Wage War On Cheats', *NewsLimited* (2016).

87. Australian Tax Office Online Services Voice Authentication. [cited 18 April 2019]. Available from:https://www.ato.gov.au/General/Online-services/Voice-authentication/.

88. Brigid Richmond and Consumer Policy Research Centre, *A Day in the Life of Data: Removing the Opacity Surrounding the Data Collection, Sharing and use Environment in Australia* (Report, 24 May 2019) 1 <http://cprc.org.au/wp-content/uploads/CPRC-Research-Report_A-Day-in-the-Life-of-Data_final-full-report.pdf>.

89. 'Smart Speaker Recordings Reviewed by Humans' *BBC News* (online, 11 April 2019) <https://www.bbc.com/news/technology-47893082>.

90. H O Maycotte, 'What Data Can a Beacon Actually Collect?' *MultiChannelMerchant* (online, 9 April 2019) <https://multichannelmerchant.com/marketing/data-can-beacon-actually-collect/>.

91. Auckland Airport, 'New Passenger Tracking Technology Trialled' (Blog Post, 22 August 2010) <https://corporate.aucklandairport.co.nz/news/latest-media/2010/new-passenger-tracking-technology-trialled>.

92. Chris Hoofnagle et al, 'Behavioral Advertising: The Offer You Can't Refuse' (2012) 6(2) *Harvard Law and Policy Review* 273.

93. LinkedIn, Cookie Policy. [cited 20 March 2019]. Available from:https://www.linkedin.com/legal/cookie-policy.

94. Vicinity Centres, Creating Beautiful Places (Annual Report, 2018) <https://www.vicinity.com.au/media/763327/vicinity-annual-report-2018.pdf>

95. Quividi Privacy Information. [cited 17 March 2019]. Available from: https://quividi.com/privacy/.

96. BBC Newsbeat, 'This is Why Some People Change their Facebook Names' (online, 17 December 2015) <http://www.bbc.co.uk/newsbeat/article/35112297/this-is-why-some-people-change-their-facebook-names>.

97. Rowan McRae, '"I Have a Duty" – We're Challenging Centrelink's Robo-debt Process', *Victoria Legal Aid* (online, 6 February 2019) <https://www.legalaid.vic.gov.au/about-us/news/i-have-duty-were-challenging-centrelinks-robo-debt-process>.

98. Rebecca Moore, 'Invercargill Man Wrongfully Fined by Auckland Transport', *Southland Times* (online, 24 July 2018) <https://www.stuff.co.nz/southland-times/news/105558044/invercargill-man-wrongfully-fined-by-auckland-transport>.

99. Jeannie Marie Patterson and Yvette Maker, 'Why Does Artificial Intelligence Discriminate?', *Pursuit* (online, 24 October 2018) <https://pursuit.unimelb.edu.au/articles/why-does-artificial-intelligence-discriminate>.

100. James Vincent, Gender and Racial Bias Found in Amazon's Facial Recognition Technology (Again)', *The Verge* (online, 25 January 2019) <https://www.theverge.com/2019/1/25/18197137/amazon-rekognition-facial-recognition-bias-race-gender>.

101. Selena Larson, 'Research Shows Gender Bias in Google's Voice Recognition', *The Daily Dot* (online, 15 July 2016) <https://www.dailydot.com/debug/google-voice-recognition-gender-bias/>.

102. Benjamin Zhang, 'TSA Body Scanners May Be More Likely to Trigger False Alarms if You're Black or Overweight', *Business Insider Australia* (online, 20 April 2019) <https://www.businessinsider.my/tsa-body-scanners-may-be-likely-trigger-false-alarms-black-overweight-2019-4/>.

103. oOh!Media Tools. [cited 26 April 2019]. Available from:https://www.oohmedia.com.au/insights/tools/.

104. M G Marmot et al, 'Health Inequalities Among British Civil Servants: The Whitehall II Study' (1991) 337(8574) *The Lancet* 1387—1393.

105. Sarah Jeong, 'Insurers Want to Know How Many Steps You Took Today', New York Times (online, 10 April 2019) <https://www.nytimes.com/2019/04/10/opinion/insurance-ai.html?curator=TechREDEF&login=email&auth=login-email>.

106. Gabriel Dance, Michael LaForgia and Nicholas Confessore, 'As Facebook Raised a Privacy Wall, It Carved an Opening for Tech Giants', *New York Times* (18 December 2018) <https://www.nytimes.com/2018/12/18/technology/facebook-privacy.html>.

107. Yung Ju Chua, 'Privacy and Discrimination of Consumers in a World of Ubiquitous Computing' (unpublished manuscript on file with author, 2019, University of Melbourne).

108. Australian Privacy Principles. [cited 28 June 2019]. Available from:https://www.oaic.gov.au/privacy-law/privacy-act/australian-privacy-principles.

109. 'Consumers and the ACL', *Australian Consumer Law* (Web Page) <http://consumerlaw.gov.au/consumers-and-the-acl/>.

110. Ross Douthat, 'The Only Answer Is Less Internet', *New York Times* (online, 13 April 2019) <https://www.nytimes.com/2019/04/13/opinion/china-internet-privacy.html>.

111. The Editorial Board, 'How Silicon Valley Puts the "Con" in Consent', *New York Times* (online, 2 February 2019) <https://www.nytimes.com/2019/02/02/opinion/internet-facebook-google-consent.html>.

112. Paul Mozur, 'Inside China's Dystopian Dreams: A.I., Shame and Lots of Cameras', *New York Times* (online, 8 July 2018) <https://www.nytimes.com/2018/07/08/business/china-surveillance-technology.html>.

113. Saeed Ahmed, ‘Fast Internet Access Becomes a Legal Right in Finland’, *DigitalBiz* (online, 15 October 2009) <http://edition.cnn.com/2009/TECH/10/15/finland.internet.rights/index.html>.

114. Claire Reilly, ‘VPN use Skyrockets in Australia Amid Privacy Concerns’, *CNet News* (online, 13 April 2015) <https://www.cnet.com/news/vpn-use-increases-in-australia-amid-data-retention-and-piracy-concerns/>.

115. Olivia Valentine, ‘VPNs are Primarily Used to Access Entertainment’, *GlobalWebIndex* (online, 6 July 2018) <https://blog.globalwebindex.com/chart-of-the-day/vpns-are-primarily-used-to-access-entertainment/>.

116. Catharine Logan, ‘Is It Legal to Use a VPN in Australia’, *LegalVision* (Blog Post, 6 September 2016) <https://legalvision.com.au/is-it-legal-to-use-a-vpn-in-australia/>.

117. Karl Quinn, ‘Australians' Rights to Use VPNs Should be Enshrined in Law: Report’, *Sydney Morning Herald* (online, 4 May 2016) <https://www.smh.com.au/entertainment/tv-and-radio/australians-rights-to-use-vpns-should-be-enshrined-in-law-report-20160504-gom863.html>.

118. About Facebook Ads. [cited 12 June 2019]. Available from:https://www.facebook.com/ads/about/?entry_product=ad_preferences.

119. Enrico Pavan, ‘From Analytics to Conversion Rate Optimisation’ (2018) 4(1) *Applied Marketing Analytics* 63—78.

120. Adrianne Jeffries, ‘Facebook's Fake-Name Fight Grows as Users Skirt the Rules’, *The Verge* (online, 17 September 2012) <https://www.theverge.com/2012/9/17/3322436/facebook-fake-name-pseudonym-middle-name>.

121. Facebook Terms and Conditions. [cited 3 March 2019]. Available from:https://www.facebook.com/terms.php.

122. Danah Boyd, ‘None of This Is Real: Identity and Participation in Friendster’ in Joe Karaganis (ed),*Structures of Participation in Digital Culture* (Social Science Research Council, 2007) 133, 133—157.

123. Jacob Kastrenakes, ‘German Court Says Facebook’s Real Name Policy is Illegal’, *The Verge* (online, 12 February 2018) <https://www.theverge.com/2018/2/12/17005746/facebook-real-name-policy-illegal-german-court-rules>.

124. Alex Chorros, ‘Mandatory Data Retention: What you Need to Know’, *WhistleOut* (online, 13 April2017) <https://www.whistleout.com.au/Broadband/Guides/mandatory-data-retention-what-you-need-to-know>.

125. Daniel J Solove and Danielle Keats Citron, ‘Risk and Anxiety: A Theory of Data Breach Harms’(2018) 96 *Texas Law Review* 737, 737—786.

126. Michael Porter, ‘The Five Competitive Forces That Shape Strategy’ [2008] (January) *Harvard Business Review*, 25—40.

127. Bonnie Ghosh-Dastidar et al, ‘Distance to Store, Food Prices, and Obesity in Urban Food Deserts’ (2014) 47(5) *American Journal of Preventive Medicine* 587, 587—595.

128. Christine Schmidt, ‘”The Widest Shoulders carry the Heaviest Load”: A Danish Socialist OutletCharges Membership Fees Based on Personal Income’, *NiemanLab* (online, 18 March 2019)<https://www.niemanlab.org/2019/03/the-widest-shoulders-carry-the-heaviest-load-a-danish-socialist-outlet-charges-membership-fees-based-on-personal-income/>.

129. Patrick Collinson, ‘How an EU Gender Equality Ruling Widened Inequality’, *The Guardian* (online,14 January 2017) <https://www.theguardian.com/money/blog/2017/jan/14/eu-gender-ruling-car-insurance-inequality-worse>.

130. Shabab Mirza and Caitlin Rooney, ‘Discrimination Prevents LGBTQ People from Accessing Health Care’, *Center for American Progress* (online, 18 January 2018)<https://www.americanprogress.org/issues/lgbt/news/2018/01/18/445130/discrimination-prevents-lgbtq-people-accessing-health-care/>.

131. Sarah West, Meredith Whittaker and Kate Crawford, *Discriminating Systems: Gender, Race and Power in AI* (Report, AI Now Institute, April 2019) <https://ainowinstitute.org/discriminatingsystems.pdf>.

132. Worden, J., K.M. Pennington, and A.R. Weiss, United States Department of Housing and Urban Development vs Facebook Inc Charge of Discrimination. 2018, United States of America Department of Housing and Urban Development Office of Administrative Law Judges: Washington DC.

133. Steven Curry, *The Renter's Journey: A Consumer-Centred Approach to Understanding the Dynamics of Australia’s Private Rental Market* (Report, Consumer Policy Research Centre, 27 February 2019) <https://cprc.org.au/wp-content/uploads/The-Renters-Journey_Full-Report_FINAL_27Feb19.pdf>.

134. Erin Turner, ‘It’s Time to Make Renting a Safe and Fair Experience for Everyone’, *Choice* (online, 25March 2019) <https://www.choice.com.au/money/property/renting/articles/how-to-make-the-private-rental-market-work-better>.

135. Michael Workman, Ariel Bogle and Elise Worthington, ‘The app that helped Donald Trump win is targeting Australian voters in 2019 election’, *ABC News* (online, 25 April 2019)<https://www.abc.net.au/news/2019-04-25/the-trump-style-gameification-of-political-campaigning-comes-to-/11043072>.

136. Buolamwini, J. and T. Gebru. Gender shades: Intersectional accuracy disparities in commercialgender classification. in *Conference on Fairness, Accountability and Transparency*. 2018.

137. Johnson, ‘In the World of Voice-Recognition, Not All Accents are Equal’, *The Economist* (online, 15 February 2018) <https://www.economist.com/books-and-arts/2018/02/15/in-the-world-of-voice-recognition-not-all-accents-are-equal>.

138. Steve Lohr, ‘Facial Recognition is Accurate, if You're a White Guy’, *New York Times* (online, 9 February 2018) <https://www.nytimes.com/2018/02/09/technology/facial-recognition-race-artificial-intelligence.html>.

139. Drew Harwell, ‘Facial Recognition Technology is Finally More Accurate in Identifying People of Color’, *Washington Post* (online, 28 June 2018) <https://www.washingtonpost.com/technology/2018/06/28/facial-recognition-technology-is-finally-more-accurate-identifying-people-color-could-that-be-used-against-immigrants/?noredirect=on>.

140. Karen Hao, ‘Facebook’s Ad-serving Algorithm Discriminates by Gender and Race’, *MIT Technology Review* (online, 5 April 2019) <https://www.technologyreview.com/s/613274/facebook-algorithm-discriminates-ai-bias/>.

141. Matthew Desmond, Evicted: Poverty and Profit in the American City (2016, Broadway Books).

142. Craig McMurtrie, ‘Why the ABC is Going to Court Over Police Raids’, *ABC News* (online, 27 June2019) <https://www.abc.net.au/news/about/backstory/news-talk/2019-06-26/why-the-abc-is-going-to-court-over-police-raids/11249092>.

143. Alex Hern, ‘Data Collection Leads to Discrimination and Self-Censorship, MPs told’, *The Guardian* (online, 20 June 2019) <https://www.theguardian.com/technology/2019/jun/19/data-collection-leads-to-discrimination-and-self-censorship-mps-told>.

144. Andrea Jamison, ‘Librarians Beware: Self-Censorship’, *Intellectual Freedom Blog* (Blog Post, 8May 2018) <https://www.oif.ala.org/oif/?p=13550>.

145. OAIC Data Breach Guidance. [cited 22 April 2019]. Available from:https://www.oaic.gov.au/individuals/data-breach-guidance.

146. Yung Ju Chua, 'Privacy and Discrimination of Consumers in a World of Ubiquitous Computing' (unpublished manuscript on file with author, 2019, University of Melbourne)

147. Luke Anscombe, ‘Westfield is using facial detection software to watch how you shop’, *News.com.au* (online, 19 October, 2017). <https://www.news.com.au/finance/business/retail/westfield-isusing-facial-detection-software-to-watch-how-you-shop/newsstory/7d0653eb21fe1b07be51d508bfe46262>

148. Chris Culnane, Ben Rubinstein, & Vanessa Teague. (2017). Health Data in an Open World. *arXiv eprints* (December, 2017) <https://ui.adsabs.harvard.edu/\#abs/2017arXiv171205627C>

Appendix 1

Translation of WeChat's targeted marketing flow chart

The following contents can be found at https://ad.weixin.qq.com/guide/152 in Mandarin. It should be noted that Tencent (parent company of WeChat) owns the qq.com domain, and 'weixin' means 'WeChat' in Mandarin.

1. Targeting Ability

See section 3 privacy policy for more details about what information WeChat collects.

The most important point about these abilities is that a lot of "facts" about you are not provided by yourself, but "learned" by WeChat.

1.1 Region

In a nutshell, the ability to target users based on their geo-Location shows that WeChat is constantly tracking users' location.

1.1.1 Resident (domestic)

The data comes from the information (location) of the WeChat users in the past one month. It supports different level of accuracies such as provinces, cities, districts and business districts.

1.1.2 LBS (domestic)

More accurate location-based service: e.g. ads can be seen by users within 500-5000 meters.

1.1.3 Travelling Users (domestic)

1.1.4 Travelling Users (overseas)

1.1.5 Overseas Resident

Targeting at the users living overseas.

1.2 Population Characteristics

1.2.1 Gender

1.2.2 Age group

1.2.3 Marriage

The state of marriage is not something the users mark as part of their profile. The identification comes from analysis on the basis of social media posts.

1.2.4 Education

1.3 Interest & Behaviour

WeChat tracks users' behaviour and analyses what they might be interested in.

Behaviour analysis:

1) Behaviour

- Downloading app
- Searching for or reading news articles
- Online shopping related behaviours

2) Time

How long the behaviour has been.

3) Strength

In terms of frequency, time and type of interaction. Take the car industry as an instance, you can push ads to users that searched for cars for the past 30 days more than once.

Particularly in this category, WeChat gives an example which shows the fact that they are showing different prices to different kinds of users. For example, if you happen to be a WeChat user that is recently helping your family members or friends deciding which car to purchase, they tend to show you a lower price.

1.4 Mobile Device

1.4.1 Operating System

IOS or Android

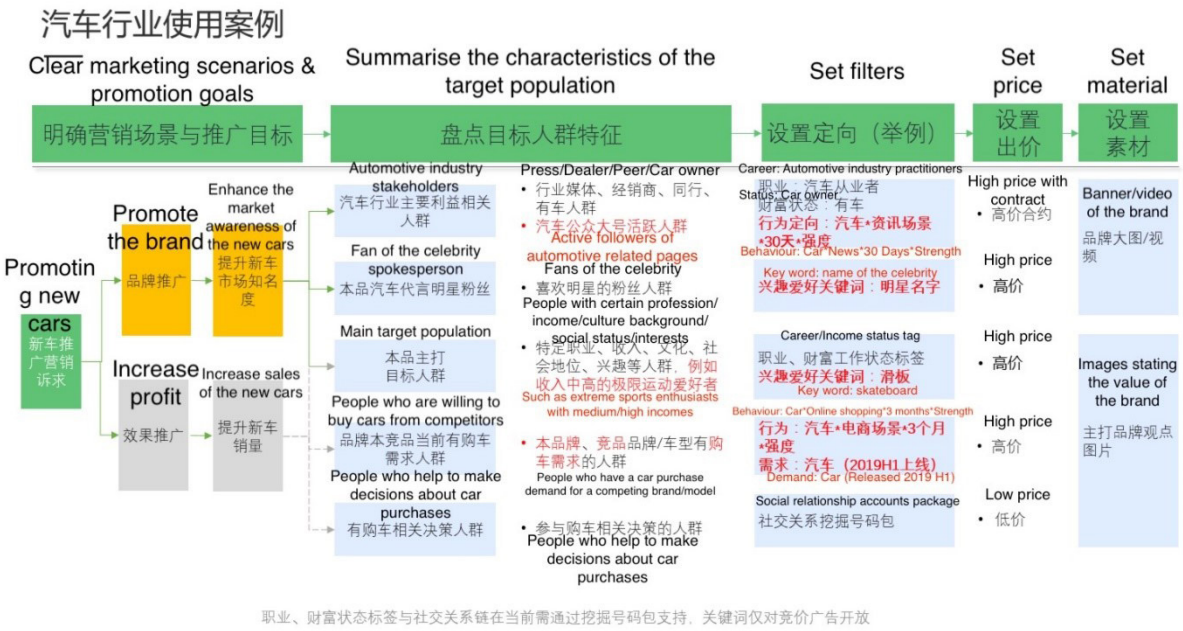
1.4.2 Network

Users' current network environment, i.e. Wi-Fi/4G/3G

1.4.3 Price of the Device

1.4.4 Brand/Model of the Device

1.4.5 Carrier



Appendix 2

Case Study – Shopping centre and marketing facial recognition tools

One area of concern, due to legislative gaps, in the consumer data tracking landscape is the increasing prevalence of facial recognition, mood analysis, and movement tracking of consumers in private-public spaces.

Petrol stations and shopping centers are installing digital signage systems which monitor people's attention and their emotional responses to advertisements while recording demographic information via facial recognition technology. Mobile-phone location tracking is increasingly being employed to monitor consumer and employee movements within shopping centers. This technology is being implemented with many people entering into agreements to be tracked, physically or biometrically, through either 'forced consent', or with 'no consent' [146].

This new data collection and consumer profiling technology aims to transmute the well-developed online advertising and profiling model into a system with enhanced capabilities enabled by data gathered in the real-world. It is possible to merge website visits and social media likes to cameras with emotional response tracking and movement information from mobile phones. These new sensor technologies are increasingly being used to monitor and collect data about people in the real world, under privacy and terms of service agreements which include clauses that signal an intent to link of the real-world information with data collected online. Media reports confirm the intention of the technology being used to collect data with an intention of sharing consumer data between advertisers [147].

Privacy policies provide the terms which consumers agree to when visiting properties owned by two of Australia's largest shopping centre operators. Consumers agree to their movement data, images, demographic and other information being collected for a variety of purposes including; provision of service, marketing and security. The policies are broad, non-specific and provide no option for opting out. Consumers can request information being held about them under the requirements of the Australian Privacy Principles, however this information is limited to data which is currently defined as 'Personally Identifiable information'.

Consumers are informed that data-sharing of information collected may occur between the centre operators and their partners. This type of data sharing creates real possibilities for individualised consumer profiling. When matched with publicly available data and any privately held data, like mobile phone IMEI numbers (held by partner companies who operate in shopping centres) and transaction details (held by any rewards or credit provider), emotional responses to specific advertisements can provide detailed insight into consumers.

This type of information is a marketer's philosophers stone. Knowing what people are interested in, when they are interested, where they are located, how they feel, and their travel paths is of immense commercial value, as evidenced by the proliferation and commercial success of online targeted advertising. The growth trend in IOT devices, capable of many similar functions, will see an expansion in the same concerns as are raised here.

The only solace provided to consumers about much of the non-personal data being collected on their emotions, movement, behaviors and biometrics is an assurance the data is de-identified. However, re-identification of de-identified or anonymized' data is a real concern. It has been shown that data can be re-identified.

Culnane, Rubinstein, & Teague re-identified sensitive medical data from de-identified publicly available [148]. The ability to identify individuals increases with rich data-sets which include many points of longitudinal information that can be matched.

Currently, consumers have little control over how their behaviors, emotions, and shopping habits are being monitored. As these technologies are deployed eventually becoming ubiquitous, entering a shopping centre, filling-up at a petrol station and other routine activities will expose consumers to scanning and data gathering. Understanding the extent of these technologies, how they could influence people, and regulatory responses to protect consumers, is of vital importance.

THE UNIVERSITY OF MELBOURNE