ESSAY COVER SHEET 2018

NIR605: Critical Data Studies

Full Name: **Paula McMahon**

Student Number: **17185602**

Date Submitted: 6th May 2018

Course Title & Lecturer: NIR605, Critical Data Studies, Prof. Rob Kitchin

Essay/Project title: **The ethics of big data and the data brokerage industry – a critical examination**

## Introduction

Personal data has become a valuable commodity that is bought and sold like any other in the marketplace. Where does this information come from? How is it being collected, and who is collecting it? Who are the organisations that are trading it? And to whom is it being provided? These are questions posed by, among others, the Canadian Internet Policy and Public Interest Clinic (CIPPIC). This essay provides a critical examination of the ethics of big data and the data brokerage industry. In the discussion that follows, I will firstly outline data brokers as they currently exist and who the main players are. Secondly, I will discuss how they collect their data and why they operate as they do. Thirdly, I will talk about how these types of operations affect us in terms of data privacy and data security. Fourthly, I will ask what are the legal implications of data brokers trading our data and lastly, what solutions exist in relation to how we manage and control our privacy as we navigate our way through the big data revolution?

## Discussion

There is no authoritative definition of "data broker" on either side of the Atlantic. Neither the United States nor European policy provides clear guidance. A data broker is a company or business unit, that earns its primary revenue, by supplying data or inferences about people, gathered mainly from sources other than the data subjects themselves (Rieke et al, 2016). Acxiom, DataLogix (now acquired by Oracle), ChoicePoint, Experian, Intelius, eBureau, Equifax: while not widely known to most people, these companies are thriving while hiding in plain sight. Data brokering is not a new phenomenon. Credit bureaus began in the U.S. in the 1950s as small, local companies that helped to provide lenders with better information about prospective borrowers. What has changed in recent years though is the tremendous increase in the volume and quality of digitally recorded data—and the technological advances that have facilitated access to storage, analysis, and sharing of information (U.S. Senate, 2013). Roderick, 2014, informs us that the Acxiom database holds information on consumers' age, race, sex, weight, height, marital status, education level, politics, buying habits, household health concerns, etc., averaging around 1500 data points per person. Kitchin, 2014a, explains to us that vast quantities of data and derived information are being rented, bought and sold daily across a variety of markets – retail, financial, health, tourism, logistics, business intelligence, real estate, private security, political polling. These data concern all facets of everyday life including public administration, communications, consumption of goods and media, travel, leisure, crime, social media interactions, and so on. Selling data to brokers has become a significant stream of revenue for many companies. Kitchin expands further, informing us that retailers often sell on data concerning transactions such as credit card details, customer purchases and store loyalty programmes, customer relationship management, and subscription information. How pervasive then, is the market in Europe? Rieke et al, 2016, tell us that although there are numerous data brokers active in Europe, the European data broker landscape is not comparable to the US market in terms of market size. The European revenues of large data brokers, such as Acxiom, LexisNexis, amount only to a fraction of their overall revenues. The European landscape, is highly fragmented across national European markets, complicating the measurement of revenues. This fragmentation is the result of different national legal regimes and varying availability of data. Each data broker

tends to specialise in different types of data and data products and services. For example, Christl, 2016, tells us that Experian provides data on finances, purchases, mortgages, property and more. Both Visa and MasterCard are listed as data providers, too. Visa, which provides “audience data” based on 16 billion credit-card payments in the US, combines its transactional data with demographic, purchase and other data from Oracle. The company emphasizes that it “aggregates and deidentifies all transactional data output”. Singer, 2012a informs us that eBureau evaluates potential clients on behalf of credit card companies, lenders, insurers and educational institutions, and that Intelius provides people-search services and background checks. Rieke et al, 2016, note that Google or Facebook are not data brokers but do sell their data to brokers. The Federal Trade Commission, (FTC, 2014), broadly welcome the consumer benefits afforded by the work that data brokers do. They say that data broker products help to prevent fraud, improve product offerings, and deliver tailored ads to consumers. Risk mitigation products help to prevent fraudsters from impersonating unsuspecting consumers. Also, marketing products benefit consumers by allowing them to more easily find and enjoy the goods and services they need and prefer.

How do data brokers operate and what is their motivation? Kitchin, 2014a, tells us that data brokers can build up extensive datasets that can be bundled and used to produce new derived data which provide more insights than any one source of data. In addition to these privately sourced data, data brokers also gather together public datasets relating to both individuals and aggregates (e.g., groups, places) such as property and census records. Why are they doing all of this? Kitchin further explains that what data brokers and analysis companies desire are a wide variety of data (both small and big), relating to as large a segment of the population as possible. The more data a broker can source and integrate, the more likely their products work optimally and successfully, and they gain competitive advantage over their rivals. By gathering data together and structuring them appropriately they can create derived data, individual and area profiles, and undertake predictive modelling as to what individuals might do under different circumstances and in different places (Singer, 2012a). From Hoofnagle, 2004, we know that in the USA, state security agencies have used private sector data brokers to profile individuals and compile terrorist watch lists. According to Rieke at al, 2016, Police in both the United States and Europe purchase corporate assistance in order to profile residents based on personal data. In the U.S., prospective employers routinely turn to data brokers to purchase criminal history reports regarding job candidates (reports that are notoriously error-prone). Furthermore, Rieke at al tell us that many data brokers thrive by providing data and predictions about consumers that help businesses optimize their commercial offerings. These data brokers often work hand-in-hand with large online advertising platforms, such as Facebook and Google, to help target advertisements. Christl, 2017, outlines for us how one of the major developments in recent years is that companies can now address, identify, and recognize consumers on an individual level across a growing number of disparate situations in their lives. Therefore, they increasingly aggregate data suitable for combining, linking, and cross-referencing profile data from different sources, such as email addresses and phone numbers. Some large data companies such as Acxiom, Experian, and Oracle have introduced their own proprietary identifiers for people, which are used to link their extensive consumer

profile information with data managed by other companies, and then to link it with the advertising data infrastructure around the globe.

Now, let us address the ethics of big data, in terms of privacy, keeping the data brokerage industry firmly in focus. In the "About Us" section on the Acxiom website, it says that it provides the critical services, products and technology that companies need to honour customer preferences and maintain good customer relationships and that Acxiom globally pledge to conduct their business according to certain principles. These are; Notice, Access and Choice; Compliance; Ethical Relationships; Accuracy; Security and Consumer Value. Interpreting this, it would seem that Acxiom are passing on the responsibility of implementing these principals to their marketer customers. In the 1980s, seven global privacy principals known as Fair Information Practice Principles or FIPPs were accepted worldwide. With these principals; Notice, Choice, Consent, Security, Integrity, Access, Accountability - individuals are, in theory, given control over their personal data and provide consent to others with regards to it. However, it could be argued that these FIPPs are redundant in the era of big data. There is certainly evidence to support this argument. Kitchin (2016) gives particular mention to the futility of notice and consent as (i) people do not read privacy policies; (ii) if people read them, they do not understand them; (iii) if people read and understand them, they often lack enough background knowledge to make an informed choice. CIPPIC, 2006, explains to us that individuals give up their personal data, wittingly or unwittingly, in various capacities: as purchasers, subscribers, registrants, members, cardholders, donors, contest entrants, survey respondents, and even mere inquirers. Kitchin, 2014a, informs us that former chairperson of the FTC in the U.S., Edith Ramirez, expressed worry that data brokers practise a form of 'data determinism' in which individuals are not profiled and judged just on the basis of what they have done, but on the prediction of what they might do in the future using algorithms that are far from perfect, which may hold inbuilt biases relating to race, ethnicity, gender and sexuality, and yet are black-boxed, and use data that are often low in quality and thus prone to error. Kitchin also tells us that in late 2012, the FTC subpoenaed nine data brokers to discover more about what data and derived information they generate and collate about people and how the data are employed and sold, as well as issuing a report calling for "privacy by design" i.e. that privacy be the default mode of operation and all data remains private unless the consumer explicitly says otherwise. This philosophy was first mooted by Ann Cavoukian in the 1990's (Cavoukian, 2009). These calls have for the most part been rejected. The Federal Trade Commission, (FTC, 2014), report that data brokers are resistant to sharing details about their data sources, citing confidentiality clauses in their contracts, and concerns about putting themselves at a competitive disadvantage. Christl and Spiekermann, 2016, say that there is strong evidence to suggest that online shops already show differently priced products to different consumers, or even different prices for the same products, based on individual characteristics and past behaviours. Furthermore, they report that in 2010 that large insurer Aviva, together with the consulting firm Deloitte, predicted individual health risks for e.g. diabetes, cancer, high blood pressure, and depression for 60,000 insurance

applicants based on marketing consumer data that they had purchased from data broker Equifax.

The security of our data is of paramount concern, so, how safe is our private data? Rieke et al, 2016, say that with modern technologies, law enforcement agencies particularly in the U.S. have access to public and private records far beyond the agencies' traditional reach. Agencies share data with each other and have many avenues to gain access to data from the private sector. Agencies in the US can subscribe to commercial data broker products that are purpose-built for law enforcement. Christl, 2016, relates that most marketing data brokers also trade many kinds of sensitive information about consumers, including about their financial situation. Acxiom, for example, provides data and scores about someone's income, net worth, economic stability, socioeconomic status, loans, banking and insurance policies for marketing purposes. It is therefore, pertinent to ask - is our data safe from being breached? The sensational Cambridge Analytica story has been known about for two years but has only recently gained traction in the headlines. Granville, 2018, writing in the New York Times, reports that Cambridge Analytica, a political firm, hired by the Trump campaign, acquired access to private data on 50 million Facebook users. The firm offered tools that could identify the personalities of American voters and influence their behaviour. Cambridge Analytica has been largely funded by Robert Mercer, a wealthy Republican donor, and Steve Bannon, a former adviser to the president. The data included details on users' identities, friend networks and "likes". The idea was to map personality traits based on what people had "liked" on Facebook, and then use that information to target audiences with customised digital ads. Private information was scraped from profiles and those of their friends via a personality app (activity that Facebook permitted at the time). Those that used the app (a tiny percentage of the 50 million) consented to having their data harvested for "academic" use. Facebook's lack of disclosure on the harvesting of data could violate privacy laws in several states and EU countries. Kang, 2018, reports in the New York Times that the social networking giant is also facing an investigation by the FTC, which is looking into whether Facebook violated an agreement with the consumer protection agency. This relates to a settlement reached with Facebook in 2011 after finding that the company had told users that third-party apps on the social media site, i.e. games, would not be allowed to access their data. Nevertheless, the apps, the FTC found, were able to obtain almost all personal information about a user (FTC, 2011).

What are the legalities surrounding our personal data and the way in which it is being traded and utilised? Let's consider the differences in law between Canada versus the U.S. versus Europe? Gill and Law, 1989, discuss how the power held by consumer data brokers cannot exist without the application of political authority, legitimation, and a system of coercion on the part of the state. CIPPIC, 2016, advises that under Canadian law, any commercial organisation collecting, using or disclosing "personal information" must obtain the individual's knowledge and consent to such activities. Data protection in the private sector is governed by the federal Personal Information Protection and Electronic Documents

Act (PIPEDA). These statutes regulate the collection, use, and disclosure of personal information by private sector organisations. Anonymous data does not constitute "personal information" and is therefore not subject to the restrictions in the Act. Put simply, a number of data brokers take the position that PIPEDA does not apply to them because they do not collect, use or disclose "personal information". What of the U.S. laws? Rieke et al, 2016 outlines for us that in the United States, a sector-by-sector approach leaves vast swathes of consumer data largely unregulated. However, uses of data in key areas, such as credit, employment, insurance, and housing are subject to some restrictions. The FTC is the closest thing the U.S. has to a general-purpose data protection authority, but its powers are limited. What about the laws within the European Union? Rieke at al, 2016, tells us, by contrast, that privacy and data protection are treated as fundamental rights and the EU has a broad regulatory framework that attaches data protection safeguards to the processing of personal data by any entity in the private sector, including data brokers. However, those rules are developed and debated in largely abstract terms, and every sector from civil to commercial, struggle to understand and apply the rules in concrete situations. The DPD (Data Protection Directive), a precursor to GDPR (General Data Protection Regulation) harmonised data privacy laws across the EU, but implementations varied considerably. The EU defines and regulates profiling explicitly in its upcoming GDPR laws which could certainly have a significant impact on data brokers. The EU has powerful principles, but uncertain impacts. There is currently no authoritative report about how European data protection law applies to data brokers, and no coordinated enforcement action against data brokers at the EU level. Rieke et al, 2016, predict that GDPR will have a major impact on the EU data protection landscape—including on data brokers. It will further harmonize data protection law across the EU. They comment however, that despite the seemingly significant legal differences between the EU and the U.S., for some data brokers, e.g. in the marketing and credit sectors, operating in the EU is possible and the differences may not ultimately be that material, save for a substantially higher overhead compliance cost on the EU side. Christl, 2016, advises that GDPR might ban or at least slow down the most irresponsible and invasive practices of third-party dataveillance and open up ways to make corporate data practices more transparent and accountable. Furthermore, its impact will most likely reach beyond the European Union. The US legal and regulatory framework, in contrast, has enabled the growth of today's data-driven world without any effective consumer safeguards and there is little in sight that will bring about a fundamental change.

In light of all that has been put forward; the privacy concerns, the security fears, the quagmire that envelopes the legalities - are there any tangible solutions? Kitchin, 2016, in writing about a solution for data privacy with respect to smart cities, says that the approach advocated is multi-pronged and that a suite of solutions is needed. These would be market driven (self-regulation from within the industry), and technical in nature (using privacy enhancement tools). Other approaches should be more governance and management orientated (with respect to delivery and compliance), and some more policy, regulatory and legally focused (FIPPs, privacy-by-design, security-by-design). Together these form a broad assemblage of protections including: accountability and transparency; law, regulation and

independent oversight; privacy by design; market forces; education and awareness; audit and control; data security; and fair information practices.

## Conclusions

To speak about the impact of data brokers in the context of global business would produce a very evenly balanced debate with plenty of positives. To skew that context by focusing on ethics, privacy and security and that debate, in my opinion, becomes undoubtedly negative. The biggest of the players, Acxiom, say they will assist their customers in adhering to privacy principles without seeming to adhere to them themselves. CIPPIC, 2006, put it succinctly when they say that people have ever fewer options to resist the power of the data ecosystem; opting out of pervasive tracking and profiling has essentially become synonymous with opting out of much of modern life. Christl, 2006, concludes that preventing the dominant data platforms and conglomerates from abusing the unprecedented data power that they have consolidated based on extensive behavioural information on billions of people presents a major challenge. kitchin, 2016, concludes that what cannot be allowed to happen is a continuation of the ad hoc and arbitrary approach taken to date and the harms regarding data privacy, data protection and data security to multiply unchecked. How then, can we trust our private data to remain so? The scepticism of the data brokerage industry is unanimous from the various writers I have referenced in this essay and they are all heavily immersed in documenting and researching big data, ethics and privacy laws. The EU is admirably leading the way in bringing forth a sea-change in our data laws. One must hope that the imminent arrival of GDPR, the adoption of the multi-pronged measures discussed in the previous section, heavy handed penalties meted out to those who flout the laws and a mindset change by consumers to stand up to social media giants, may, hopefully, effect a seismic shift globally. The data brokerage industry though, is a wild animal that won't like being tamed.

## References

Acxiom (2018): Privacy Principles: https://www.acxiom.com/about-us/privacy/privacy-principles

Cavoukian, A., (2009): Privacy by Design: The 7 Foundational Principles. [Accessed 30th April, 2018] Available from the World Wide Web: https://www.ipc.on.ca/wp-content/uploads/Resources/7foundationalprinciples.pdf

Christl, W. (2017): Corporate Surveillance in everyday life: How Companies Collect, Combine, Analyze, Trade, and Use Personal Data on Billions.

Christl, W, Spiekermann, S. (2016): Networks of Control. A Report on Corporate Surveillance, Digital Tracking, Big Data & Privacy. Facultas, 2016. Available at: http://crackedlabs.org/en/networksofcontrol

CIPPIC (2006). On the Data Trail: How detailed information about you gets into the hands of organizations with whom you have no relationship. A Report on the Canadian Data Brokerage Industry.

FTC, Federal Trade Commission (2014): Data Brokers: A Call for Transparency and Accountability, https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf

FTC, Federal Trade Commission (2011): Facebook Settles FTC Charges That It Deceived Consumers By Failing To Keep Privacy Promises. [Accessed 4th May 2018] Available from the World Wide Web: https://www.ftc.gov/news-events/press-releases/2011/11/facebook-settles-ftc-charges-it-deceived-consumers-failing-keep

Gill S and Law D (1989) Global hegemony and the structural power of capital. International Studies Quarterly 36(December): 475–499.

Granville, K. (2018). Facebook and Cambridge Analytica: What You Need to Know as Fallout Widens. [Accessed 30th April 2018] Available from World Wide Web: https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html

Hoofnagle, C. (2004). Big Brother's Little Helpers: How Choicepoint and other commercial data brokers collect and package your data for law enforcement. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=582302

Kang, C. (2018). Facebook faces growing pressure over data and privacy inquiries. [Accessed 1st May 2018] Available from the World Wide Web: https://www.nytimes.com/2018/03/20/business/ftc-facebook-privacy-investigation.html

Kitchin, R. (2014a). The data revolution: Big data, open data, data infrastructures and their consequences. London: Sage.

Kitchin, R. (2016) Getting smarter about smart cities: Improving data privacy and data security. Data Protection Unit, Department of the Taoiseach, Dublin, Ireland.

Singer, N. (2012a). Mapping, and Sharing, the Consumer Genome. New York Times, 17th June 2016. [Accessed 3rd May, 2018] Available from the World Wide Web: https://www.nytimes.com/2012/06/17/technology/acxiom-the-quiet-giant-of-consumer-database-marketing.html

Rieke, A., Yu, H., Robinson, D., von Hoboken, J. (2016). Data Brokers in an Open Society. London: Open society foundation, London.

Roderick, L., (2014). Critical Sociology. Discipline and Power in the Digital Age. London: Sage.

U.S. Senate (2013). A Review of the Data Broker Industry: Collection, Use, and Sale of Consumer Data for Marketing Purposes. Committee on Commerce, Science, and Transportation.