The Illusion of Informational Autonomy: A critique of New Zealand's regulatory framework for data privacy in an age of big data

Louise Wilsdon

A thesis submitted for the degree of
Doctor of Philosophy
at the University of Otago, Dunedin,
New Zealand.

May 2024

Abstract

Privacy legislation premised on the individual is coming under strain as the ability to exercise control over one's information becomes an increasingly difficult task. This thesis considers whether the Privacy Act 2020 is fit for purpose in an age of big data. It also considers other responses to big data and the technologies it enables, both New Zealand's soft law approach and the regulatory responses of the European Union.

A key focus of this thesis is on the collective, public harms of big data. Uses of big data that exacerbate inequalities in society by discriminating on the grounds of race or gender, or by creating feedback loops that contribute to cycles of poverty, set back the public interest in living in thriving, prosperous societies. Very large technology companies exercise significant control over the content we view online; know a great deal about us, and thus, can influence the decisions we make. Vast amounts of personal information in the hands of a few powerful companies creates power imbalances that are harmful to democracies. The predictive power and reach of big data analytics exacerbates this problem.

This thesis draws on social conceptions of privacy that acknowledge the information externalities of personal information use. Ultimately, it concludes that privacy law's focus on the individual is conceptually flawed, inhibiting its ability to respond effectively to the risks, and benefits, of big data.

Preface

My interest in privacy law in the context of big data was stimulated by research I undertook for Professor Colin Gavaghan on algorithmic decision-making and the purported advantages of keeping a human 'in the loop'. Despite a wealth of scholarship from overseas academics, there is a gap in New Zealand research considering social conceptions of privacy as a means of addressing the challenges of big data. I am very grateful to Colin for giving me that research role and for including me in various opportunities to participate in discussions about the law and emerging technologies, and for being incredibly generous with his time and ideas.

I am indebted to both my supervisors, Colin and Dr Jeanne Snelling, for their encouragement, support, and for taking the time to really engage with my research. I could not have asked for kinder, or more intellectually rigorous, supervisors.

I am also eternally grateful to my mother for instilling a love of reading and learning, and for her unfaltering belief, in me. And thank you to Steve, my very patient husband, undoubtedly the most privacy-informed electrical supervisor in the country!

And Annabel and Connor - you can have your mum back now...

Contents

Introduction

Big data typically refers to very large datasets from a range of sources; its size and volume mean that it is most effectively processed by algorithm or artificial intelligence. The processing of big data reveals patterns, associations, and trends in human behaviour and other activity, and can be used to predict the likelihood of future behaviour and events. The rise of big data has been facilitated by developments in information technology and computing, allowing for the exponential collection and processing of information. There are many beneficial uses of big data that sit outside the scope of this thesis because they do not process *personal* information, for example, big data research on vaccines, new antibiotics, anti-viral and anti-fungal drugs[1]; the use of big data to predict earthquakes[2] or to regulate energy supply.[3] This thesis considers the ubiquitous collection and processing of personal information by corporations and governments, raising questions of fairness and power that go to the heart of the type of society we live in.

The processing of big data is prolific and extends across many sectors of society - while the contexts vary, the rationale remains the same: large amounts of personal data are valuable. To remain competitive, many companies generate ever-increasing amounts of data from which to make increasingly detailed profiles and predictions about us.[4] For the big technology companies, the primary purpose of big data is to place people into groups based on their shared characteristics, preferences, and beliefs. It is the relational aspect of big data that gives it its power and reach. Its value lies in the sheer volume of information about people and how they relate to each other.[5]

[1] Carissa Véliz *Privacy Is Power: Why And How You Should Take Back Control Of Your Data* (Penguin Random House, London, 2020) at 165-167.
[2] Omar M. Saad et al "Earthquake Forecasting using Big Data and Artificial Intelligence: A 30-Week Real-Time Case Study in China" (2023) Bulletin of the Seismological Society of America 113 6.
[3] Techstack "Blog: Implementing Big Data Solutions for Transforming the Renewable Energy Sector: Techstack Case Included" (2023) https://tech-stack.com/blog/implementing-big-data-solutions-for-transforming-the-renewable-energy-sector-techstack-case-included/; José Manuel Blanco "Big Data and Energy: A Combination for Success" (2022) https://www.plainconcepts.com/big-data-energy/
[4] Jathan Sadowski, "How 'Smart Tech' Masks an Emerging Era of Corporate Control" (9 March 2020) OneZero.
[5] Salomé Viljoen "A Relational Theory of Data Governance" (2021) Yale Law Journal 131 2 at 573.

This thesis explores the claim that data privacy law's focus on the individual is conceptually flawed in an age of big data. In Chapter One, I provide an overview of the history of privacy law and the rise of big data. Contemporary data privacy frameworks have their roots in information principles created in response to the risks posed by large government computer databases of the 1970s, before the commercialisation of the Internet.[6] The outdated, and narrow, conception of data privacy as the exercise of individual control over one's information, stems from early data privacy laws. Nonetheless, the history of privacy law shows us that privacy is adaptable in response to new technologies and the challenges they present.

Big data enables technologies that can be used to control, influence, and surveil, but also has the potential for increased efficiencies, more targeted provision of services, and greater precision and accuracy. In Chapter Two, I assess the potential of big data in health; an area that has drawn significant attention and resources. Some big data applications in health show potential and are delivering in certain contexts, particularly medical pattern recognition[7] and risk prediction.[8] However, other big data initiatives in healthcare have overpromised and underdelivered, often at the expense of the data subjects concerned and to the advantage of technology companies.[9] Responsible, and socially beneficial, big data initiatives in healthcare are dependent on the effective regulation of personal information.

---

[6] Elizabeth M. Renieris *Beyond Data: Reclaiming Human Rights at the Dawn of the Metaverse* (The MIT Press, Cambridge, Massachusetts, 2023) at 31-36.

[7] Breast cancer screening algorithms show comparable performance with radiologists in reading mammograms and are particularly effective in detecting abnormal cells (Karin Dembrower, Alessio Cripper, Eugenia Colôn, Martin Eklund, Fredrik Strand "Artificial Intelligence for breast cancer detection in screening mammography in Sweden: a prospective, population-based, paired-reader, non-inferiority study" (2023) Lancet 5; Ioannis Sechopoulos and Ritse M. Mann "Stand-alone artificial intelligence - The future of breast cancer screening?" (2021) Special issue: Artificial Intelligence in Breast Cancer Care Science Direct 56; Mattie Salim, Erilk Wahlin, Karen Dembrower, Edward Azavedo, Theodoros Foukakis, Yue Liu, Kevin Smith, Martin Eklund and Fredrik Strand "External Evalution of 3 Commercial Artificial Intelligence Algorithms for Independent Assessment of Screening Mammograms" (2020) JAMA Oncology 6(10)).

[8] The use of big data to screen for non-communicable diseases and chronic conditions is promising. Artificial intelligence has exceeded the performance of experts in referral recommendations for a range of retinal diseases (Jeffrey De Fauw et al. “Clinically applicable deep learning for diagnosis and referral in retinal disease” (2018) 24 Nature Medicine 1342-1350; Nima John Ghadiri "AI in Opthalmology" in Michael F. Byrne (ed.) *AI in Clinical Medicine: A Practical Guide for Healthcare Professionals* (2023, John Wiley & Sons Ltd., Hoboken, USA) at 260).

[9] In Chapter Three I examine the collaboration between DeepMind and the Royal Free London NHS Foundation Trust to develop a smartphone app to detect acute kidney injury (AKI). Five years' worth of patient admission, discharge and transfer data was transferred from the Royal Free to DeepMind, including health data unrelated to AKI and its causes. Ultimately the app was decommissioned. (Julia Powles and Hal Hodson “Google DeepMind and healthcare in an age of algorithms” (2017) 7 Health Technology 351-357; Natasha Lomas "Google completes controversial takeover of DeepMind Health"

Consequently, one of the questions this thesis addresses is whether New Zealand's regulatory framework for data privacy law is properly equipped to address the harms, and harness the benefits, of big data. There is nothing inherently wrong with big data nor is there anything necessarily problematic about big data analytics: the predictive models, algorithms, and artificial intelligence that process big data. References to big data in this thesis include big data analytics and the technologies enabled by big data. Individual data subjects can benefit from big data as well as be disadvantaged by its use; however, the focus of this thesis is on collective harms to society overall. In Chapter Three, I categorise the social privacy harms of big data into two broad categories. First, the public harm of increasing inequality caused by the cumulative impacts of unjustified discrimination on the grounds of race, gender, and socioeconomic status. I consider the use of big data in significant, human decision-making contexts and its impact on the indigent and marginalised. I examine how big data can exacerbate existing inequities in societies and entrench them.

The second category of social privacy harms includes the harm to democracy facilitated by vast amounts of personal information held by a handful of powerful corporations. Very large online platforms and very large online search engines exercise significant control over the information we view online. They also know a great deal about us; information which is used to sell targeted advertising spaces to companies, lobbying or interest groups, individuals, and political parties that seek to influence the decisions we make. Too much personal information in the hands of a few powerful corporations creates power asymmetries that are harmful to democracies.[10] The issue of power being concentrated in the hands of a few is not a problem created by big data; however, it is exacerbated by its predictive power and reach.

(20 September 2019) Techcrunch https://techcrunch.com/2019/09/19/google-completes-controversial-takeover-of-deepmind-health/).

[10] Shoshana Zuboff *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (Profile Books Limited, London, 2019); Safiya Umoja Noble "Algorithms of Oppression: How Search Engines Reinforce Racism" (New York University Press, New York, 2018); Véliz, above n 1; Cathy O'Neill *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (Penguin Random House, United Kingdom, 2016).

Harms resulting from big data are not always easily identifiable or attributable to a particular action, or even to a specific individual or agency. Social privacy harms can be cumulative, such as the effects of unfair algorithms deployed at scale to allocate scarce social resources or determine eligibility for entitlements. Big data harms can evade remedy through a regulatory framework that relies predominantly on the individual to identify the discrete harm caused to them by a particular privacy interference. In Chapter Four, I outline various conceptions of privacy and make the argument that privacy is primarily a social concept. Our expectations of privacy are based partly on the law but also stem from the social norms that govern our relationships and interactions with other people and agencies. This thesis questions whether the collective harms of big data can be adequately addressed by an individualistic, complaints-based privacy framework.

The Privacy Act 2020 embodies an individual rights-based approach to data privacy, which is typical of data privacy legislation. In Chapter Five, I assess the effectiveness of the Privacy Act in addressing big data. I compare it to the European Union's General Data Protection Regulation,[11] which goes further than New Zealand's Privacy Act in enhancing the rights of individual data subjects. However, both instruments, premised on the individual, are an incomplete response to the harms of big data, particularly social privacy harms.

In Chapter Six, I consider some of the uses of big data in New Zealand, in particular, the increasing use of surveillance technologies by law enforcement and the private sector.[12] Big data enables increasingly invasive technologies, for example facial recognition, and live tracking of vehicles through automatic number plate recognition supported by networks of CCTV cameras.[13] New Zealand's responses to the potential harms and benefits of big data beyond the Privacy Act embody a soft law approach. I outline the

[11] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1.
[12] Foodstuffs is currently undergoing a trial of facial recognition technology in response a rise in retail crime (Foodstuffs "Foodstuffs North Island begins trialling facial recognition in select stores as part of its commitment to keep teams and customers safe by keeping previous offenders out" (8 February 2024). https://www.foodstuffs.co.nz/news-room/2024/Foodstuffs-North-Island-begins-trialling-facial-recognition-in-select-stores).
[13] Police use of automatic number plate recognition (ANPR) through ANPR platforms has come under legal challenge (Phil Pennington "Legal challenges to Police use of automated number plate recognition cameras" (26 October 2023) RNZ https://www.rnz.co.nz/news/national/501012/legal-challenges-to-police-use-of-automated-number-plate-recognition-cameras).

increasing uptake of surveillance technologies and the use of predictive algorithms in the public service in the absence of effective regulation to mitigate the risks of their use.

This thesis questions whether New Zealand's data privacy framework is fit for purpose in an age of big data. It also considers whether privacy law is the right response and if other approaches and regulatory measures should also be considered. Critically, this thesis addresses the question of whether the narrowly-defined conception of data privacy as an individual right of control over one's information, is inhibiting privacy law from providing an effective response to the social privacy harms of big data.

Chapter One - The Development of Privacy Law and the Rise of Big Data

Introduction

The background to New Zealand's regulatory framework for data privacy is multifaceted and complex, influenced by international developments in privacy law and the development of new computing and information technology, as well as New Zealand's unique political and cultural history.

Societies have relied on the collection and use of personal information to manage relationships and generate revenue for thousands of years (the Ancient Egyptians used bone tokens to record owners of inventory for tax collection from as early as 3000 BC[1]). Today governments not only routinely collect and process personal information to tax populations, but also to inform policy, provide social services and enforce laws. In commerce, personal information is used to keep track of customers' orders, payments, and activity online; to ascertain our preferences and predict our behaviour. Many facets of our lives are influenced by the processing of personal information - from the jobs we see advertised online, to the posts Facebook prioritises in our newsfeed, and the prices we pay for goods and services.

Developments in computing technology and the capacity for the storage of unprecedented amounts of information have facilitated the exponential collection, generation and processing of personal information, giving rise to an age of big data. Corporations collect massive amounts of personal information because they can - the storage of big data is cheap, and its processing profitable.

While the collection and use of personal information are necessary for cohesive and organised societies, equally important are appropriate limits on the collection and use of personal information for free and democratic societies. The history of privacy law helps

[1] Reckon "Ancient Egyptians - the original bookkeepers" (2018) https://www.reckon.com/reckon-blog/ancient-egyptians-the-original-bookkeepers1/

to explain contemporary data privacy regimes' shortcomings in addressing the social privacy harms of big data - the harms that ripple beyond the individual data subject, impacting the type of society we live in.

The social privacy harms of big data include the effects of unjustified discrimination on the grounds of race, gender, or socio-economic status. The harm of increasing inequality as a result of the cumulative impacts of unfair or unethical information practices sets back the public interest in living in healthy, happy, and prosperous communities. Additionally, too much personal information in the hands of a few powerful corporations creates power asymmetries that are harmful to democracies.

Thus, the term 'big data' in this thesis encompasses not only the vast volumes of personal information held by the big tech companies and governments but also a varied range of technologies that collect or generate information about us and use it in ways that impact our lives. A wide range of technologies sit under the umbrella of 'big data' – from predictive algorithms used in significant decision-making contexts, to surveillance technologies powered by artificial intelligence that collect and generate personal information. The scope of the technologies considered in this thesis is necessarily broad - they are all connected by their potential to contribute to social privacy harms. Because big data can convey social and relational meaning, it is capable of benefitting and harming others beyond the data subject to whom the information relates. These information externalities are not accounted for under an individualistic complaints-based privacy framework.

Remedies for data privacy breaches in New Zealand, as in other countries, are premised on individual rights. Contemporary data privacy frameworks have their roots in data privacy principles of the late 1970s to 1980s, created in response to the risks to individual citizens posed by large government computer databases. However, the social privacy harms of big data are predominantly caused by sorting or classifying people into groups, and the inferences or assumptions that are enabled by those classifications. The relational aspect of big data is inadequately addressed by contemporary data privacy frameworks, including New Zealand's. This is in large part due to the focus on the individual's right of control over their information, which was the rationale underlying early data privacy regimes and still persists today.

Historical factors such as 9/11 (the terrorist attacks that took United States' intelligence agencies by surprise) and the desire to increase efficiencies and cut expenditures, have resulted in increased surveillance by governments.[2] Accompanying this loss of privacy is a resigned acceptance by populations whose citizens disclose their personal information in exchange for 'free' online services. The notion that privacy is an outdated value and that 'if you have nothing to hide then you have nothing to fear' are fallacies promulgated by big tech and governments. A core claim made in this thesis is that we should all be concerned about the ubiquitous collection and processing of personal information, and the largely unregulated industry of online service providers.

The social privacy harms of big data were not contemplated when contemporary data privacy regimes were in their embryonic stages. The discourse of individual control over one's data, exercised through consent and the rights of access to, and correction of, information about oneself, has its roots in early data protection laws,[3] drafted before the commercialisation of the Internet and the development of big data analytics.[4] I argue that the discourse of individual control is conceptually flawed in an age of big data; it is promoted by technology companies *because* it is ineffective at addressing information asymmetries underlying the power of big tech.[5]

While the outlook for privacy may appear bleak, increasing surveillance and rampant data collection are not inevitable. Privacy law has proven adaptable in responding to challenges that new technologies present. In this Chapter, I describe how remedies in tort were developed in response to intrusions by the media into private life that were facilitated by the invention of mechanical shutter photography from the late nineteenth century. Similarly, in the period after the Second World War, human rights treaties were

[2] Carissa Véliz *Privacy Is Power: Why And How You Should Take Back Control Of Your Data* (Penguin Random House, London, 2020) at 36-37.
[3] Elizabeth M. Renieris *Beyond Data: Reclaiming Human Rights at the Dawn of the Metaverse* (The MIT Press, Cambridge, Massachusetts, 2023) at 31.
[4] At 31.
[5] Ari Ezra Waldman argues that the discourse of privacy as control works in the interests of big tech because it is the discourse of self-governance, which is "a sham" because it allows tech companies to continue to extract ever-increasing amounts of personal information but profess that they care about privacy (Ari Ezra Waldman *Industry Unbound: the inside story of privacy, data, and corporate power* (Cambridge University Press, Cambridge, 2021) at 52).

adopted in response to the atrocities committed by the Nazi regime; I outline the development of privacy as a civil and political right.

It is important to acknowledge that the impetus for data privacy regulation has been driven not just by a desire to protect human rights and civil liberties, but also to facilitate the exchange of personal information. Regulatory frameworks that do not address the information asymmetries and power imbalances between big tech and individual data subjects, promote the interests of a few powerful individuals, governments, and corporations. These stakeholders have a vested interest in maintaining the status quo.

This chapter provides an overview of some of the more significant developments in privacy law and sets the international scene in which New Zealand's Privacy Act was enacted in 1993, before its repeal and replacement in 2020. The inherent tension of contemporary data privacy frameworks in attempting to balance the free flow of personal information with the protection of individual privacy, is coming under increasing strain in New Zealand and around the world. The transatlantic divergence in attitudes to privacy and, accordingly, the divergence in legal protections ascribed to it, has implications for how New Zealand chooses to respond to the increasing commodification of personal information.

The birth of privacy law

In their groundbreaking article, “The Right to Privacy”, published in the Harvard Law Review in 1890, Samuel D. Warren and Louis D. Brandeis responded to the emergence of a ‘gutter press’ and the publishing of photos and salacious details of high society parties. Warren and Brandeis illustrate how the common law grows to meet the demands of society, and they apply the right “to be let alone” to the protection of private life from the tabloid media.[6] Warren's and Brandeis’ argument is that “the press is overstepping in every direction the obvious bounds of propriety and of decency".[7] “Gossip has become a trade”[8] and is an assault on morality and responsible for a lowering of social standards.

[6] Judge Cooley refers to the right “to be let alone” in Cooley on Torts (2nd ed.) at 29 in Samuel D. Warren and Louis D. Brandeis “The Right To Privacy” (1890) Harvard Law Review IV 5 at 195.
[7] Samuel D. Warren and Louis D. Brandeis “The Right To Privacy” (1890) Harvard Law Review IV 5 at 196.
[8] At 196.

Warren and Brandeis set out their argument that the right to privacy is “the right to one’s personality”.[9] The existing law establishes a principle which can be invoked to protect the privacy of an individual from invasion by the press:[10]

> The principle which protects personal writings and all other personal productions, not against theft and physical appropriation, but against publication in any form, is in reality not the principle of private property, but that of an inviolate personality.

With the development of cameras with mechanical shutter mechanisms, photos can be taken quickly and surreptitiously. Previously, someone’s picture could only be taken if they sat for it, hence the laws of trust and contract provided adequate protection for a "prudent man".[11] Warren and Brandeis explain that with the development of technology that allows photos to be taken without the subject’s knowledge, the laws of contract and trust are inadequate, and the law of tort must be relied on.[12] Warren and Brandeis conclude their argument with a direction to the courts to respond to privacy incursions enabled by this new technology:[13]

> The common law has always recognized a man’s house as his castle, impregnable, often, even to its own officers engaged in the execution of its commands. Shall the courts thus close the front entrance to constituted authority, and open wide the back door to idle or prurient curiosity?

The threat of the exposure of one's private life posed by the first paparazzi in the late nineteenth century was limited to the wealthy and connected. The ubiquitous collection and processing of personal information today has a much greater reach, impacting the nature of the society we live in. Today the need for privacy law to adapt to address the challenges of emergent technologies is much more pressing than it was at the time of its birth in 1890.

---

[9] At 207.
[10] At 205.
[11] At 211.
[12] At 211.
[13] At 200.

International Privacy Law

A comprehensive assessment of New Zealand's regulatory framework for data privacy cannot be undertaken without consideration of the international context in which New Zealand operates. International codes and agreements play an important role in influencing national regulatory regimes. In his seminal book, *Data Privacy Law: An International Perspective,*[14] international data privacy law expert, Lee Bygrave, notes that the number of international agreements on privacy has increased, and their provisions have become more detailed. He states that "the overall result of this growth in regulatory density is a decreasing capacity of states to adopt data privacy regimes as they alone see fit".[15]

The creation of New Zealand's data privacy framework was influenced significantly by international developments in privacy law. Today's 'Internet of Things' - the proliferation of smart devices and appliances all connected to the Internet, along with the corresponding ease with which personal data is transferred across jurisdictional boundaries, means that international data privacy rules and norms will only grow in influence.

Universal Declaration of Human Rights 1948

Privacy as a human right is an increasingly important aspect of privacy law. After World War II, governments around the world attempted to foster peace and prevent conflict. The United Nations was established in 1945 and the Universal Declaration of Human Rights was drafted three years later, setting out the rights and freedoms of all people, regardless of race, religion, or gender. The earliest declaration of the right to privacy is set out in article 12 of the Universal Declaration of Human Rights 1948:

> No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks.

[14] Lee Bygrave *Data Privacy Law: An International Perspective* (Oxford University Press, Oxford, 2014).
[15] At xxvi.

Misuse of personal information enabled the Nazi regime to execute its multitude of atrocities with greater efficiency and lethality. The 1933 census was an important mechanism for the German government to identify Jews, Gypsies, and other ethnic minorities it considered 'undesirable'. One of the questions in the census was: "Is either of your grandparents a Jew?" This information was recorded through the census system and identity cards were issued with a "J" if Jewish. In his explosive book, *IBM and the Holocaust,*[16] Edwin Black meticulously outlines IBM's relationship with the Nazi regime, and how the machine tabulation of the 1933 census data facilitated the systematic identification, tracking, and extermination of Jews. IBM's punch card system was a new technology in the early 1930s, pre-dating the computer.[17] Nazi concentration camps had their own Hollerith-Abteilung (Hollerith Departments) to keep track of inmates using IBM's punch card machines.[18] Black states that "mankind barely noticed when the concept of *massively organized information* quietly emerged to become a means of social control, a weapon of war, and a roadmap for group destruction".[19]

Black's words are a chilling caution for today's digital world where personal information is collected and processed at an exponential rate. The atrocities of the Nazi regime were accelerated by the processing of census data through IBM's punch card system, highlighting the risks inherent in large datasets of personal information.

Although the Universal Declaration is not legally binding, it has significant global influence and has been incorporated into many other international treaties. The fundamental freedoms of the Universal Declaration of Human Rights are reaffirmed in the European Convention on Human Rights 1950.

---

[16] Edwin Black *IBM and the Holocaust: The Strategic Alliance Between Nazi Germany and America's Most Powerful Corporation* (Three Rivers Press, New York, 2001).
[17] A punch card is a thin, stiff piece of cardboard with columns and rows of tiny, punched holes that correspond to data points. In the late 1880s, Herman Hollerith invented a tabulating machine that could read and summarise the data stored on punch cards. IBM's punch card system was the first automated information storage device. (https://www.ibm.com/history/punched-card)
[18] IBM concentration camp codes show a number assigned to various prisoner types - '3' for homosexual, '8' for Jew, '12' for Gypsy. There were also numbers assigned to inmate death - '3' for natural causes, '4' for execution, '5' for suicide, '6' for "special treatment" in gas chambers (Edwin Black "IBM's Role in the Holocaust - What the New Documents Reveal" (27.2.12, updated 17.3.15) Huffington Post https://www.huffpost.com/entry/ibm-holocaust_b_1301691).
[19] Black, above n 16, at 7.

The European Convention on Human Rights 1950

The European Convention on Human Rights is an international agreement between the 47 states of the Council of Europe. Individual complaints about alleged breaches of human rights, including the right to privacy, are heard by the European Court of Human Rights (ECHR). The European Convention on Human Rights has been influential in shaping case law in the EU, as the member states are all signatories of the European Convention on Human Rights.[20] Article 8 of the European Convention on Human Rights states that:

> 1. Everyone has the right to respect for his private and family life, his home and his correspondence.
> 2. There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others.

The right to "respect for his private and family life, his home and his correspondence" extends to relationships with colleagues and people in the community. It includes personal development, how one chooses to live (for example, the right of Roma to live nomadically), the right to be free of environmental pollution and excessive noise, the right to keep the name you were born with.[21] Arguably, the right to respect for private and family life in the European Convention on Human Rights is broader than the right to privacy enshrined in the International Covenant on Civil and Political Rights 1966, to which New Zealand is a party.

[20] Aysem Dike Vanberg "Informational Privacy Post GDPR – end of the road or the start of a long journey?" (2020) International Journal of Human Rights DOI:10.1080/13642987.2020.1789109 at 2.
[21] European Court of Human Rights "Guide on Article 8 of the European Convention on Human Rights - Right to respect for private and family life, home and correspondence" (Council of Europe, 2020).

The International Covenant on Civil and Political Rights 1966

The International Covenant on Civil and Political Rights[22] (ICCPR) was adopted and opened for signature, accession, and ratification by the United Nations General Assembly in 1966. It came into force in 1976. The ICCPR is a treaty and therefore binding on all the 172 countries that have ratified it, including New Zealand, which ratified it in 1978.[23]

Article 17 of the ICCPR states that:

> 1. No one shall be subjected to arbitrary or unlawful interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation.
> 2. Everyone has the right to the protection of the law against such interference or attacks.

One of the purposes of New Zealand's Privacy Act 2020 is to give effect to New Zealand's obligations under the ICCPR.[24] The long title to the New Zealand Bill of Rights Act 1990 also affirms New Zealand's commitment to the ICCPR. On 26 May 1989, New Zealand ratified the first optional protocol to the ICCPR which provides for a complaints process.[25] If domestic remedies have first been exhausted, people in New Zealand can lodge communications with the UN Human Rights Committee, which has the power to form views on alleged violations of ICCPR rights. The expectation is that a state will comply with these views unless it has a legitimate reason not to. To date there have been no article 17 complaints against New Zealand, possibly because there are remedies available in New Zealand law for breaches of article 17 of the ICCPR; the most obvious being the right of an individual to make a complaint to the Privacy Commissioner if they believe an action of an agency has interfered with their privacy.[26]

---

[22] International Covenant on Civil and Political Rights, *adopted* 16 Dec. 1966, G.A. Res. 2200 (XXI), U.N. GAOR, 21st Sess., Supp. No. 16, U.N. Doc. A/6316 (1966), 999 U.N.T.S. 171 *(entered into force* 23 March 1976).
[23] Ministry of Justice "Constitutional issues and human rights: International Covenant on Civil and Political Rights" https://www.justice.govt.nz/justice-sector-policy/constitutional-issues-and-human-rights/human-rights/international-human-rights/international-covenant-on-civil-and-political-rights/
[24] Privacy Act 2020, s 3.
[25] Optional Protocol to the International Covenant on Civil and Political Rights https://treaties.un.org/doc/Treaties/1976/03/19760323%2007-37%20AM/Ch_IV_5p.pdf
[26] Privacy Act 2020, Part 5, Subpart 1.

From the foregoing discussion, it is apparent that from the 1940s to the late 1960s, the importance of privacy as a human right was being discussed and affirmed on the world stage, setting the scene for the developments in computing technology and privacy law that followed later in the twentieth century.

Developments in computing prompt national data privacy laws from the 1970s

From the early 1970s, governments around the world utilised advances in computing to generate digital databases of personal information. The Council of Europe, the Organisation for Economic Co-operation and Development (OECD) and the European Union (EU) drafted the bulk of the international instruments that have been the most influential in shaping national data privacy frameworks.[27] To date, 137 countries have enacted privacy legislation.[28]

From the 1970s onwards there was a proliferation of national data privacy laws, partly in response to the threats population-wide databases posed. Elizabeth Renieris argues that early data protection laws, conceived of before advances in household and commercial computing and the commercialisation of the Internet, have proven "woefully inadequate".[29] Renieris states that:[30]

> Before the rise of big data and sophisticated analytics, the database was a mere digital representation of its analog predecessor and presented largely the same risks, apart from additional concerns about data security, storage and protection from hacking.

Contemporary data privacy regimes have their roots in these early data protection laws. Renieris states that "unfortunately, this decades-old logic, encoded in our early approaches to data protection, still persists in large measure today, even as it is no longer fit for purpose".[31] This *decades-old logic* includes the concept of privacy as individual control over one's data. In the late 1970s and 1980s, basic information rights, including

[27] Bygrave, above n 14, at 31.
[28] United Nations Conference on Trade and Development https://unctad.org/page/data-protection-and-privacy-legislation-worldwide
[29] Renieris, above n 3, at 31.
[30] At 36.
[31] At 31.

rights of access to, and correction of, personal information; the principles of purpose specification, collection minimisation, data quality, and security safeguards, addressed many of the concerns about digitalised government databases. Renieris argues that the emergence of the World Wide Web in 1989, "built on top of the internet's networking infrastructure, would transition us from the database age into the information age without the laws to match".[32] In Chapter Three, I describe the social privacy harms of big data that highlight the need for new laws, and in Chapter Four, I make the case for adopting a social conception of privacy as the first step in addressing social privacy harms.

It is important to note that a country's overall data privacy regime consists not just of its legislative enactments. Also significant are the information systems, industry codes and standards,[33] and cultural or social practices and expectations. Bygrave believes that customs, attitudes, corporate and administrative culture are all factors that can be misinterpreted or forgotten in the assessment of data privacy regimes.[34]

OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data 1980

The Council of Europe collaborated closely with the OECD on its work on data privacy.[35] The Guidelines on the Protection of Privacy and Transborder Flows of Personal Data 1980 (OECD Guidelines) are not binding but are supposed to be taken into consideration by Member countries when drafting domestic legislation on data privacy. They provide Member countries with a common framework from which to create national data privacy laws. The OECD Guidelines have influenced the development of data privacy law in New Zealand as well as other countries outside of Europe, including Japan, Australia, Canada and Hong Kong.[36] Commercial interests were the primary motivator behind the OECD Guidelines. The assumption underlining the Guidelines is that transborder flows of personal data are of fundamental importance and value.[37] The creation of the OECD Guidelines was motivated by commercial concerns; they were designed to prevent data

[32] At 36.
[33] Bygrave, above n 14, at xxvi.
[34] At xxvi.
[35] At 43.
[36] At 50.
[37] "Member countries have a common interest in protecting privacy and individual liberties, and in reconciling fundamental but competing values such as privacy and the free flow of information". OECD Council Recommendation, OECD Guidelines 1980, at 1.

protection legislation from being used as a trade barrier. It was motivated, at least in part, by a concern that member countries would enact their own privacy legislation as a form of economic protectionism.[38] Bygrave refers to an OECD Council Recommendation accompanying the OECD Guidelines that states that: "member countries should 'endeavour to remove or avoid creating, in the name of privacy protection, unjustified obstacles to transborder data flows of personal data'".[39]

Privacy Principles in the OECD Guidelines

There are eight data privacy principles in Part Two of the OECD Guidelines which apply to manual and automated processing of personal data in both the public and private sectors. They are:[40]

> **7. Collection Limitation Principle**
> There should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject.
>
> **8. Data Quality Principle**
> Personal data should be relevant to the purposes for which they are to be used, and, to the extent necessary for those purposes, should be accurate, complete and kept up to date.
>
> **9. Purpose Specification Principle**
> The purposes for which personal data are collected should be specified not later than at the time of data collection and the subsequent use limited to the fulfilment of those purposes or such others as not incompatible with those purposes and as are specified on each occasion of change of purpose.
>
> **10. Use Limitation Principle**

[38] Bygrave, above n 14, at 44.
[39] Recommendation of the Council concerning Guidelines governing the Protection of Privacy and Transborder Flows of Personal Data (adopted 23 September 1980; (C(80)58/FINAL) in Bygrave, above n 9, at 43.
[40] OECD Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data 1980, paragraphs 7-14.

Personal data should not be disclosed, made available or otherwise used for purposes other than those specified in accordance with paragraph 9 except:

a) with consent of the data subject; or

b) by the authority of law

11. Security Safeguards Principle

Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorised access, destruction, use, modification or disclosure of data.

12. Openness Principle

There should be a general policy of openness about developments, practices and policies with respect to personal data. Means should be readily available of establishing the existence and nature of personal data, and the main purpose of their use, as well as the identity and usual residence of the data controller.

13. Individual Participation Principle

An individual should have the right:

a) to obtain from a data controller, or otherwise, confirmation of whether or not the data controller has information relating to them;

b) to have communicated to him, data relating to them

i. within a reasonable time;

ii. at a charge, if any, that is not excessive;

iii. in a reasonable manner; and

iv. in a form that is readily intelligible to them;

c) to be given reasons if a request made under subparagraphs (a) and (b) is denied, and to be able to challenge such a denial; and

d) to challenge data relating to them and, if the challenge is successful to have the data erased, rectified, completed or amended.

14. Accountability Principle

A data controller should be accountable for complying with measures which give effect to the principles stated above.

The privacy principles of the OECD Guidelines reflect basic information rights recognised worldwide. In the Explanatory Memorandum to the 1980 Guidelines the

‘Individual Participation Principle’ - the right of individuals to access and challenge personal data - is described by the Expert Group as “perhaps the most important privacy protection safeguard”.[41]

However, with the uses of data becoming more obscure and complex, a core claim of this thesis is that the significance of individual control becomes less compelling and less effective as a check on the processing of personal data. If individuals are unaware that an agency holds information about them then this mechanism is an ineffective safeguard. This is acknowledged in the *OECD Privacy Framework* document titled "The Evolving Privacy Landscape: 30 Years After the OECD Privacy Guidelines" which influenced many of the subsequent revisions to the 1980 Guidelines.[42]

Asia-Pacific Economic Cooperation

The OECD Guidelines were used as a model for the Information Privacy Principles of the 2005 Asia-Pacific Economic Cooperation (APEC) Privacy Framework, of which New Zealand is a member. APEC initiatives in respect of data privacy have been designed to promote confidence in business and encourage Asian countries to adopt privacy legislation. Described as "OECD lite",[43] the predominant motivation behind the 2005 Privacy Framework was to facilitate commercial interests rather than to protect fundamental human rights.[44] The APEC Privacy Framework was updated in 2015, drawing on concepts introduced in the updated OECD Guidelines. Its focus continues to be "effective privacy protections that avoid barriers to information flows and ensure continued trade and economic growth in the APEC region".[45]

Amendments to the OECD Guidelines

Amendments made to the OECD Guidelines in 2013 introduced new concepts such as National Privacy Strategies to be coordinated at the highest level of government, Privacy

---

[41] Original Explanatory Memorandum to the OECD Guidelines, paragraph 58.
[42] OECD “The Evolving Privacy Landscape: 30 Years After the OECD Privacy Guidelines” in OECD *The OECD Privacy Framework* (2013) at 67.
[43] Graham Greenleaf “Australia’s APEC privacy initiative: the pros and cons of ‘OECD lite’” [2003] Privacy Law and Policy Reporter 1.
[44] Bygrave, above n 14, at 80.
[45] Asia-Pacific Economic Cooperation *APEC Privacy Framework* (2015) at 2.

Management Programs as the central mechanism through which organisations implement privacy protection, and Data Security Breach Notification, which was adopted in amendments to New Zealand's Privacy Act in 2020.[46] The amended Guidelines also make it more difficult for member countries to restrict transborder flows of personal information between OECD member states.[47] Despite the momentous developments in information technology since 1980, the principles of the OECD Guidelines were not revised or added to in the 2013 amendments. Bygrave states that:[48]

> The decision to leave the core principles untouched is remarkable given that the principles are over 30 years old, were formulated in an environment that was markedly different from today's, and could arguably be fine-tuned or otherwise improved in light of such change.

There was a fear on the part of some member countries that privacy rules could be exploited as a means of economic protectionism. Bygrave suggests that the reluctance to revise the principles on the part of the Expert Group responsible for drafting the 2013 revisions may have been due to a concern that reopening the debate could have resulted in a reduction in protection offered by the existing principles.[49] Bygrave notes that the OECD Guidelines stipulate the need for member countries to adopt privacy laws with the Guidelines to "be regarded as minimum standards which can be supplemented by additional measures for the protection of privacy and individual liberties".[50] The OECD has adopted other instruments that relate to data privacy since 1980, including guidelines dealing with information security, cryptography policy, and consumer protection in electronic commerce. They are all consistent with, and uphold, the principles of the 1980 Guidelines.[51]

[46] Privacy Act 2020, Part 6, Subpart 1 - Notifiable privacy breaches.
[47] Bygrave, above n 14, at 48.
[48] At 44.
[49] At 44.
[50] OECD Guidelines, paragraph 6 in Bygrave, above n 14, at 45.
[51] Bygrave, above n 14, at 51.

Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data (Convention 108), 2018

The Council of Europe is responsible for drafting a number of data privacy instruments. The 1981 Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data (Convention 108) was the first legally binding international instrument governing automated and manual processing of personal data in the public and private sectors. The objectives of Convention 108 are to provide greater harmonisation between Member States by strengthening data privacy and also to ensure the free flow of personal data across borders. It stipulates fair and lawful collection of personal information and for information to not be used inconsistently with the purpose for which it was collected.[52] It is concerned with the quality of data, its accuracy and proportionality,[53] as well as ensuring rights of access and correction.[54]

Convention 108 was modernised in 2018. The modernised Convention 108 maintains its 'technologically neutral', principles-based approach while aiming to address challenges to the protection of personal data posed by new technologies and practices. It also aims to ensure consistency and compatibility with other data protection frameworks, particularly the EU's, and is designed to provide a potentially universal standard for data protection.

The principles in Convention 108 are broadly similar to those in the OECD Guidelines, despite the different objectives of each - the latter being more concerned with facilitating commerce and the former primarily with human rights protection. Bygrave attributes the similarities to the extent of cooperation between the two bodies responsible for their drafting.[55]

[52] Convention 108, art 5.
[53] Article 5.
[54] Article 8.
[55] Bygrave, above n 14, at 45.

EU General Data Protection Regulation (GDPR) 2018

In contrast to the pragmatic approach of the OECD, the EU is responsible for creating the most extensive and comprehensive data protection instruments albeit, motivated in part by economic considerations. Renieris states that:[56]

> By the mid-1990s, it was clear that the commercialisation of the internet and web had ushered in a new era of cross-border digital technologies and data flows. And as the European Union grew, it also quickly became apparent that the fragmented approach to privacy and data protection across various member states ... was creating barriers to free trade and the cross-border flow of data, threatening the completion of the internal market.

Consequently, in 1995, two years after New Zealand passed its first Privacy Act, the EU adopted Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, OJ 1995 L 281/31 (Data Protection Directive) with the objectives of harmonising Member States' data privacy laws and protecting fundamental human rights. EU directives, however, are not directly applicable in Member States and need to be transposed into national law. Ultimately, the Data Protection Directive failed in its objective to align the Member States' data protection frameworks and hence Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1 (GDPR) was adopted to replace it. The GDPR took four years to draft and involved tough negotiations and numerous amendments to the original text.[57]

Nevertheless, the same objectives and principles of the Data Protection Directive are carried through to the GDPR; the protection of personal data, considered in relation to its function in society, and balanced against other fundamental rights including freedom of

[56] Renieris, above n 3, at 39.
[57] European Data Protection Supervisor "The history of the General Data Protection Regulation" https://edps.europa.eu/data-protection/data-protection/legislation/history-general-data-protection-regulation_en

thought, conscience and religion, freedom of expression and information, and freedom to conduct a business.[58] The GDPR came into force in 2018. It sets the world standard in data protection legislation; however, it is no panacea to the threats of emergent digital technologies and pervasive surveillance. The GDPR, like New Zealand's Privacy Act 2020, is based on an individual consent model of data privacy protection. Its focus is on strengthening individual privacy rights and enabling the individual to better manage their own data. Data privacy models premised on the individual's right to control their data, are coming under increasing strain with the exponential rise in the amount of personal information generated and processed, and the number of agencies that hold personal information. The strengths of the GDPR, and the weaknesses inherent in individualistic models of data privacy, will be discussed in more detail in Chapter Five.

The GDPR has significant global influence because it prohibits (with some qualifications) the transfer of personal data to countries outside of the EU, unless they provide adequate levels of data privacy.[59] New Zealand is one of those countries outside of the EU considered to have GDPR adequacy.[60] However, the basis for this finding, and the degree of reassurance we can draw from it in regards to the adequacy in real terms of New Zealand's data privacy regime, will be considered later in this chapter.

Charter of Fundamental Rights of the European Union 2000

The Charter of Fundamental Rights of the European Union (2000/C 364/01) (the Charter) enshrines the full range of civil, economic, social, and political rights of EU citizens and residents.[61] It was introduced to clarify and bring consistency to the rights established at different times and in different ways in individual EU Member States. The Charter incorporates the rights in the European Convention on Human Rights; the constitutional traditions of the EU Member States that exist in common law and constitutional law; the Council of Europe's Social Charter; the Community Charter of Fundamental Social

[58] General Data Protection Regulation [GDPR], Recital 4.
[59] GDPR, Chapter V.
[60] European Commission - Press release "Commission finds that EU personal data flows can continue with 11 countries and territories" (15 January 2024) https://ec.europa.eu/commission/presscorner/detail/en/ip_24_161
[61] Preamble to the Charter of Fundamental Rights of the European Union.

Rights of Workers; and other conventions to which the EU or its Member States are parties.[62]

The Charter was drafted by the EU and proclaimed by the European Parliament, Council of Ministers, and European Commission in 2000. It obtained full legal effect when the Treaty of Lisbon came into force in 2009[63] and was amended in 2012.

The Equality and Human Rights Commission states that: [64]

> The Charter can be seen as the overarching framework for human rights in the EU, of which the European Convention on Human Rights forms only one part, albeit an important one.

The European Convention on Human Rights, while containing overlapping provisions with the Charter, operates within a separate legal framework. The European Convention on Human Rights was drafted by the Council of Europe and is interpreted by the European Court of Human Rights. The Charter is interpreted by the Court of Justice of the European Union.

Article 7 of the Charter states that: "Everyone has the right to respect for his or her private and family life, home and communications". Article 8 states:

1. Everyone has the right to protection of personal data concerning him or her.

2. Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified.

---

[62] Equality and Human Rights Commission "What is the Charter of Fundamental Rights of the European Union?" https://www.equalityhumanrights.com/en/what-are-human-rights/how-are-your-rights-protected/what-charter-fundamental-rights-european-union

[63] Prior to this its legal status was unclear.

[64] Equality and Human Rights Commission, above n 62.

3. Compliance with these rules shall be subject to control by an independent authority.

The right to protection of personal data, a fundamental value of the EU enshrined in the Charter, has been affirmed in the *Schrems* cases. Nonetheless, the following analysis highlights the difficulties faced in holding very large online platforms to account for privacy-invasive practices, irrespective of legal judgments confirming their illegality.

Schrems I - *Maximillian Schrems v Data Protection Commissioner*

Maximillian Schrems is an Austrian national, residing in Austria, and a Facebook user. Like other Facebook users in the EU, Schrems' personal information is transferred by Facebook Ireland to servers belonging to Facebook Inc. based in the United States, where it is processed. Facebook, like other U.S. companies, is subject to National Security Agency PRISM surveillance programs in the United States. In 2011, Schrems lodged a complaint with the Irish Data Protection Commissioner, essentially to stop Facebook's transfers of personal information to the United States.

The Safe Harbor Agreement of 2000 facilitated the legal transfer of personal data by U.S. companies from EU member states to the United States. It provided a mechanism by which U.S. companies could meet an "adequate level of protection" as required under the Data Protection Directive. A company could self-certify that it was compliant with the principles and requirements of the Safe Harbor Agreement. Participation was available to companies subject to regulation by the Federal Trade Commission. Persistent failure to comply with the principles and requirements would result in withdrawal of Safe Harbor status. There were approximately 4,500 companies on the Safe Harbor list, including Facebook Ireland.[65]

Schrems challenged the adequacy of protection against access by government authorities to personal data transferred to the United States by Facebook Ireland. Schrems' complaint was dismissed by the Irish supervisory authority on the ground that it had no basis to evaluate the complaint because Facebook Ireland had adhered to the Safe Harbor

[65] Martin A. Weiss and Kristin Archick *U.S.-EU data Privacy: From Safe Harbor to Privacy Shield* (Congressional Research Service, United States Congress, 2016) at 5-6.

Agreement. The Irish supervisory authority maintained that it was bound by 2000/520/EC: Commission Decision of 26 July 2000 pursuant to Directive 95/46/EC of the European Parliament and of the Council on the adequacy of the protection provided by the safe harbour privacy principles and related frequently asked questions issued by the US Department of Commerce (notified under document number C(2000) 2441) (Decision 2000/520), in which the Commission had ruled that the Safe Harbour Agreement provided "an adequate level of protection" as required by the Data Protection Directive.[66] Schrems appealed the decision in the Irish High Court, which then requested the Court of Justice of the European Union (CJEU) to consider whether the Irish supervisory authority could investigate Facebook's information practices or whether it must defer to the Commission's earlier decision.[67]

In Case C-362/14 *Maximillian Schrems v Data Protection Commissioner* [2015] ECR 398 (*Schrems I)*, the CJEU made several findings. It ruled that Decision 2000/520 was invalid. It held that third-party countries can only receive personal information from the EU if those countries guarantee adequate data protection safeguards. Adequacy includes the fundamental values on which the EU is founded, including Articles 7 and 8 of the Charter. PRISM[68] trawls the personal information of all Facebook users: mass surveillance violates Article 8 of the Charter by compromising the fundamental right to respect for private life. The CJEU stated that:[69]

> ... legislation permitting the public authorities to have access on a generalised basis to the content of electronic communications must be regarded as compromising the essence of the fundamental right to respect for private life, as guaranteed by Article 7 of the Charter.

Additionally, the lack of legislation in the United States allowing for an individual to pursue legal remedies in relation to access to data about himself, or to request its

[66] Under article 45 of the GDPR, transfers of personal data to a third country may take place where the Commission has decided that that third country offers an adequate level of protection.
[67] Weiss and Archick, above n 65, at 6-7.
[68] PRISM is a code name for the program under which the U.S. National Security Agency collects internet communications from ISPs. (T.C. Sottek and Janus Kopfstein "Everything You Need To Know About PRISM: A Cheat Sheet for the NSA's unprecedented surveillance programs" (17 July 2013) The Verge https://www.theverge.com/2013/7/17/4517480/nsa-spying-prism-surveillance-cheat-sheet).
[69] Case C-362/14 *Maximillian Schrems v Data Protection Commissioner* [2015] ECR 398 at 94.

correction, compromises the right to effective judicial protection.[70] The CJEU also held that the existence of a Commission decision finding that a third-party country guarantees an adequate level of protection does not limit the powers of national DPAs under the Charter and the Data Protection Directive.[71]

Another illustration of the transatlantic divergence in privacy standards is Edward Snowden's whistleblowing, which brought the transnational power of the United States in influencing privacy practices to the forefront of the minds of the international community. It highlighted that not only do the actions of the United States government and big tech companies impact the privacy of U.S. citizens but they also have far-reaching implications for individuals across the world.

Edward Snowden's leak of classified documents

The extent of numerous global surveillance programs, many of them run by the United States' National Security Agency (NSA) and the Five Eyes Intelligence alliance was dramatically revealed by Edward Snowden in June of 2013. Snowden was a former technical assistant for the United States Central Intelligence Agency (CIA) and had been working at the NSA as an employee of various contractors for four years prior to leaking thousands of classified NSA documents to journalists. His whistleblowing prompted international public debate over national security and privacy.[72] Snowden maintains that he acted in the public interest to highlight the erosion of privacy by the Intelligence Services. He argues that the NSA's ubiquitous surveillance is undertaken with no public oversight and poses an "existential threat to democracy".[73] In August 2013, U.S. president Barack Obama created an independent panel to examine the U.S. government's surveillance practices. In December 2013, the panel made a number of recommendations, including that the mass collection of personal information from internet service providers

[70] At 95.
[71] At 99.
[72] The Guardian US was awarded the Pulitzer prize in 2014 for its aggressive reporting on the Snowden leaks, sparking a public debate on the issues of security and privacy and the relationship between the public and the U.S government (https://www.pulitzer.org/winners/guardian-us); the Washington Post was also awarded the Pulitzer prize in 2014 for its revelation of widespread secret surveillance by the National Security Agency and assisting the public in understanding how the Snowden leaks fit into the broader context of national security (https://www.pulitzer.org/winners/washington-post-1)
[73] Glenn Greenwald "Edward Snowden: the whistleblower behind the NSA surveillance revelations" *The Guardian* (online ed, United Kingdom, 11 June 2013) https://www.theguardian.com/world/2013/jun/09/edward-snowden-nsa-whistleblower-surveillance

and phone carriers be suspended and greater oversight taken in relation to other sensitive surveillance programs, many of which were acted on.[74]

UN General Assembly Resolution 68/167. The right to privacy in the digital age

The Right to Privacy in the Digital Age, G.A. Res. 68/167, U.N. Doc. A/RES/68/167 (Dec. 18, 2013) was adopted just six months after Edward Snowden's leak of thousands of classified NSA documents. The UN General Assembly adopted a consensus resolution calling on all countries to adopt measures to address unlawful or arbitrary surveillance, unlawful or arbitrary interception of communications, and unlawful or arbitrary collection of personal data that "violate the rights to privacy and to freedom of expression and may contradict the tenets of a democratic society".[75] It affirmed the right to privacy as set out in article 12 of the UN Declaration of Human Rights and article 17 of the ICCPR in light of technological advances that have enhanced the capacity of corporations, governments and individuals to undertake surveillance, interception and data collection.[76] The UN recognised that new surveillance technologies and data collection techniques posed a significant threat to the human right to privacy. - A threat that remains just as pressing today.

Schrems II - *Data Protection Commissioner v Facebook Ireland and Maximillian Schrems*

Following the ruling in *Schrems I,* Schrems reformulated his complaint to the Irish Data Protection Commissioner in light of the declaration by the CJEU that Decision 2000/520 (the Safe Harbour decision) was invalid. In the meantime, the Commission adopted Commission Implementing Decision (EU) 2016/1250 of 12 July 2016 pursuant to Directive 95/46/EC of the European Parliament and of the Council on the adequacy of the protection provided by the EU-U.S. Privacy Shield (notified under document C (2016) 4176) (Decision 2016/1250), which stated that the protection offered by the EU-US Privacy Shield was adequate.

---

[74] Britannica "Edward Snowden" https://www.britannica.com/biography/Edward-Snowden; National Whistleblower Centre "Edward Snowden" https://www.whistleblowers.org/whistleblowers/edward-snowden/

[75] UN General Assembly Resolution 68/167. The right to privacy in the digital age at 2.

[76] At 1.

Schrems' reformulated complaint claimed that the U.S. does not offer sufficient protection of data transferred to it (basing his allegation on the information that had come to light as a result of Edward Snowden's leaks). Schrems argued that U.S. law requires Facebook to make personal data transferred to it available to the National Security Agency and the Federal Bureau of Investigation, which use it in the context of several surveillance programs. This is incompatible with Articles, 7, 8 and 47 of the Charter.[77]

In Case C-311/18 *Data Protection Commissioner v Facebook Ireland and Maximillian Schrems* [2020] ECR 297 (Schrems II), the CJEU declared Decision 2016/1250 (on the adequacy of protection offered by the EU-US Privacy Shield) invalid. Again, the Court ruled that data subjects whose personal information is transferred to a third-party country must be guaranteed a level of protection essentially equivalent to that guaranteed in the EU by the GDPR, read in light of the Charter.[78]

The CJEU held that Decision 2016/1250 (the Privacy Shield decision), like Decision 2000/520 (the Safe Harbour decision), gave priority to requirements of U.S. national security and law enforcement, condoning interference with the fundamental rights of the EU.[79] Surveillance programs based on those provisions are not limited to what is strictly necessary and are, therefore, disproportionate.[80] The CJEU also held that the provisions failed to provide data subjects with legal rights against U.S. authorities.[81]

Alan Butler, Electronic Information Privacy Centre (EPIC) Interim Executive Director and General Counsel,[82] stated in response to the judgment that: [83]

> This is another landmark ruling for privacy rights by the Court of Justice, and a clear signal that the United States needs to reform its surveillance laws or risk losing its

[77] At 55.
[78] At 104 and 191.
[79] At 185.
[80] At 184.
[81] At 192.
[82] The Electronic Information Privacy Centre (EPIC), acted as amicus curiae in the *Schrems* cases.
[83] EPIC.org Press Release https://epic.org/privacy/intl/dpc-v-facebook/cjeu/RELEASE-EPIC-CJEU-July2020.pdf

position as a global technology leader. Congress should act quickly to bring US law in line with international human rights standards.

Decision of the Data Protection Commission in the matter of Meta Platforms Ireland Limited

The Decision of the Data Protection Commission made pursuant to Section 111 of the Data Protection Act 2018 and Articles 60 and 65 of the General Data Protection Regulation in the matter of Meta Platforms Ireland Limited (previously known as Facebook Ireland Limited) (Data Protection Commission, Ireland, 12 May 2023) concludes the inquiry of the Irish Data Protection Commission into Meta Platforms Ireland Limited and the basis on which it transfers personal data from the EU to the U.S. After the judgment in *Schrems II,* the Data Protection Commissioner decided to commence an "own volition" inquiry under the Data Protection Act 2018. The inquiry began in August 2020 but was stayed by Order of the High Court of Ireland, pending the outcome of a series of legal proceedings, until 20 May 2021. On the conclusion of its investigation, the Data Protection Commission issued a draft decision on 6 July 2022, in which it found that the data transfers in question were being carried out in breach of Article 46(1) GDPR and that consequently, the data transfers should be suspended.[84]

Under Article 60 of the GDPR, which mandates the sharing of draft decisions with other supervisory authorities in the EU, also known as concerned supervisory authorities (CSAs), the Data Protection Commission shared its draft decision with all of the other supervisory authorities, as CSAs. The CSAs all agreed with the Data Protection Commissioner's decision on Meta Ireland's non-compliance with the GDPR and the proposal to make an order to suspend the data transfers.[85] However, four of the 47 CSAs believed that Meta Ireland should be subject to an administrative fine, and two of the CSAs believed that Meta Ireland should be ordered to address the personal data that it held as a consequence of the unlawful transfers.[86] The DPC disagreed, maintaining that

[84] Data Protection Commission "Data Protection Commission announces conclusion of inquiry into Meta Ireland" (22 May 2023) https://www.dataprotection.ie/en/news-media/press-releases/Data-Protection-Commission-announces-conclusion-of-inquiry-into-Meta-Ireland
[85] Data Protection Commission, above n 84.
[86] Data Protection Commission, above n 84.

exercising these powers would go beyond what is "appropriate, proportionate and necessary" to address the breach of Article 46(1) GDPR. A consensus could not be reached, therefore the DPC referred the objections to the European Data Protection Board (EDPB) in accordance with the Article 65 dispute resolution mechanism.[87] The DPC issued its final decision on the basis of the EDPB's decision adopted on 13 April 2023. On 12 May 2023, the DPC issued the following corrective powers in its decision:[88]

> (1) Under Article 58(2)(j) GDPR, to require Meta Ireland to suspend the data transfers in accordance with a timeline set out in a suspension order;
>
> (2) Under Article 58(2)(d) GDPR, to cease the processing and storage of the personal data transferred in violation of the GDPR within 6 months of being notified of this decision;
>
> (3) Under Article 58(2)(i) GDPR, an administrative fine of 1.2 billion Euro.

This is the biggest ever fine issued for breaching the GDPR. The ECJ gave Meta a grace period of five months to implement the suspension of data transfers and six months to stop the unlawful processing. On 22 May 2023, Meta described its frustration with the DPC decision, its intention to appeal the ruling and stated it felt Facebook had been targeted unfairly.[89] It reassured customers that there would not be an immediate effect on its services and that policy makers in the EU and the U.S. were working towards a new EU-U.S. Data Privacy Framework.[90]

[87] Data Protection Commission, above n 84.
[88] Decision of the Data Protection Commission made pursuant to Section 111 of the Data Protection Act 2018 and Articles 60 and 65 of the General Data Protection Regulation in the matter of Meta Platforms Ireland Limited (previously known as Facebook Ireland Limited) (Data Protection Commission, Ireland, 12 May 2023) at 10.3.
[89] Meta "Our response to the decision on Facebook's EU-US data transfers" (22.5.23) https://about.fb.com/news/2023/05/our-response-to-the-decision-on-facebooks-eu-us-data-transfers/.
[90] Meta, above n 89.

EU-U.S. Data Privacy Framework

On 10 July 2023, the Commission adopted an adequacy decision on the EU-U.S. Data Privacy Framework (the Framework).[91] This new adequacy decision allows companies participating in the Framework to transfer personal data from the EU to the U.S. without having to implement additional data protection safeguards. The Commission states that:[92]

> The EU-U.S. Data Privacy Framework introduces new binding safeguards to address all the concerns raised by the European Court of Justice, including limiting access to EU data by US intelligence services to what is necessary and proportionate, and establishing a Data Protection Review Court (DPRC) to which EU individuals will have access.

Commission President, Ursula von der Leyen, stated that "the US has implemented unprecedented commitments to establish the new framework".[93] The Framework will be subject to periodic reviews by the Commission, representatives of European data protection authorities and competent U.S. authorities.[94] It remains to be seen if the European data protection authorities and U.S. authorities agree on the parameters of "necessary and proportionate" access to EU data by U.S. intelligence services. The first review is due by July 2024.

Irrespective of the ability of the EU-U.S. Data Privacy Framework to withstand a review by the CJEU, the years of litigation raise concerns. The Irish Data Protection Commissioner, the independent authority responsible for upholding the fundamental right of individuals to have their data protected, exhibited a marked reluctance to hold Meta to account. Meta displayed a blatant disregard for the CJEU; ignoring decisions that were unfavourable to it. The rulings in both *Schrems I* and *Schrems II* declared the transfers of personal information from the EU to the U.S. illegal, but Facebook continued its data

[91] Article 45(3) of the GDPR grants the Commission the power, after assessing the adequacy of the level of protection, to decide that a third country offers an adequate level of protection.
[92] European Commission press release "European Commission adopts new adequacy decision for safe and trusted EU-US data Flows" (Brussels, 10 July 2023) https://ec.europa.eu/commission/presscorner/detail/en/ip_23_3721
[93] European Commission, above n 92.
[94] Article 45(3) of the GDPR states that an implementing act shall provide a mechanism for periodic review.

transfers. It raises the question of whether Meta is too big and too influential to be effectively held to account for violations of the law.

The rise of Big Data

Over the last two decades, developments in information technology and computing have facilitated the exponential collection and processing of personal information by algorithm and artificial intelligence, giving rise to an age of big data. Significant concerns generated by the ubiquitous collection and processing of personal information include questions of accuracy of the information collected, its re-use and transfer across jurisdictional boundaries, and the fairness and justice implications of processing information by solely or partly automated means. Corporations' and governments' use of big data can reveal detailed pictures of the preferences and behaviours of individuals and groups, but they often lack the same level of transparency when it comes to disclosing their own information practices. Some of the benefits, and the harms, of the processing of big data by artificial intelligence will be discussed in more detail in Chapters Two and Three, respectively.

Background to today's contemporary predictive models

In *A Citizen's Guide to Artificial Intelligence,* John Zerilli et al emphasise that the modelling techniques that are utilised by big tech, academia, and governments today are extensions of predictive models that have been in use since the start of computing, and in some instances, well before that.[95] Zerilli et al state that:[96]

> Mathematical methods have been in use for centuries to guess an unknown variable by consulting a database of known facts. The first serious applications were in the insurance industry. The Lloyd's register, which developed in 1688 and assessed the likely risk of shipping ventures, is a well-known early example.
>
> ...

[95] John Zerilli, John Danaher, James Maclaurin, Colin Gavaghan, Alistair Knott, Joy Liddicoat, Merel Noorman *A Citizen's Guide to Artificial Intelligence* (The MIT Press, Cambridge, Massachussetts, 2021) at 16.
[96] At 4-5.

> To begin with, developing predictive models involved calculations done by hand, using databases stored in written ledgers. Computers can help in two ways: they facilitate the storage of large amounts of data, and they can perform calculations automatically. Predictive models are now routinely implemented as computer programs that consult databases held in computer memory.

The first computers were used only by governments and large corporations because they were prohibitively expensive: [97]

> Both companies and governments started to develop computer-based predictive models almost as soon as computers were invented. For instance, the US government used computers to predict missile trajectories in the 1940s, to predict weather in the 1950s, and to predict suitability of military personnel for missions in the 1960s. In industry, the FICO corporation in the United States, which specializes in predicting credit risk, produced its first computerized model of risk scores in 1958.

Today the predictive models used by the big technology companies are far more advanced than those of the first computers, however they utilise many of the same types of techniques. Similarly, Daniel Solove argues that the challenges that AI presents for privacy are not new - they have been a long time coming.[98] Solove believes that privacy law, properly conceptualised and implemented, would be a significant step towards addressing the privacy problems with AI.[99] My thesis considers this argument and argues that a social conception of privacy would inform a more effective response to the collective harms of big data.

The relationship between Big Data, Artificial Intelligence, and Privacy

Big data typically refers to very large, diverse sets of information from a variety of sources. Its size and volume mean that big data lends itself to processing by algorithm or artificial intelligence to reveal patterns, associations and trends in human behaviour and other activity.

---

[97] At 5.
[98] Daniel Solove "Artificial Intelligence and Privacy" Florida Law Review (forthcoming Jan 2025) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4713111.
[99] At 6.

Definitions of AI

When people refer to AI they are often making reference to modern machine learning models or generative AI. However, AI is not new - the technology has been in existence for around 80 years.[100] It is the rise of big data, and the rapid developments in computational power, that have facilitated recent and exciting breakthroughs in AI, in the form of machine learning technologies.[101]

The OECD defines an AI system as:[102]

> ... a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.

The OECD definition encompasses algorithms that input personal information to make decisions about people. The definition of an AI system in Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) also encompasses algorithms used to make predictions, recommendations, or decisions about people.[103] I discuss the Artificial Intelligence Act in detail in Chapter Six.

Algorithms

An algorithm is simply a sequence of steps used to solve a problem.[104] Predictive algorithms or predictive models (I use the terms interchangeably) harness the processing

---

[100] At 11.
[101] At 12.
[102] OECD Legal Instruments "Recommendation of the Council on Artificial Intelligence" (amended on 3 May 2024) OECD/LEGAL/0449 https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449
[103] Artificial Intelligence Act, Article 3(1).
[104] John Paul Mueller and Luca Massaron *Algorithms for dummies* (John Wiley & Sons Inc., Hoboken, New Jersey, 2017) at 11.

power of computers to compute thousands of steps at speed to solve a problem or predict an outcome. Predictive algorithms are trained on large datasets of real-world scenarios where the input variables and outcome variables are known. A predictive algorithm can estimate the likelihood of someone doing something based on the past behaviour of other people with whom they share common characteristics (input variables). For example, an algorithm can be used to predict someone's likelihood of reoffending or defaulting on a loan.

Algorithms can be used for social sorting; to categorise people into groups for companies to more effectively market to them, for example. There are concerns associated with the use of algorithms in relation to their accuracy, transparency, the potential for discrimination, as well as the assumption that they are objective decision-making tools. These concerns will be discussed in detail in Chapter Three.

The privacy problems of AI

AI includes the use of algorithms that input personal information and produce outputs that can impact on people's lives. Solove explains that there are privacy problems associated with both the inputs and outputs of AI.[105] AI exacerbates existing privacy problems and shortcomings of contemporary data privacy regimes.[106] Solove identifies the problems with AI's inputs as the problem with data collection, including the scraping of the websites on the internet to train generative AI models. He adds that there are also problems with consensual forms of data collection.[107] I discuss issues with consent in further detail in Chapter Five.

Solove articulates the problems with AI outputs as involving the generation of information about people (predictions, inferences and associations). He states that "[d]ata generation blurs the line between data collection and data processing, allowing end-runs around many privacy law protections".[108] Contemporary data privacy regimes offer a right of correction as a means for data subjects to correct inaccurate information held

[105] Solove, above n 98, at 6.
[106] At 6-7.
[107] At 6.
[108] At 6.

about them. However, it is very difficult, if not impossible, to establish that an inference or a prediction about future behaviour is inaccurate.[109] Additionally, to be able to exercise a right of correction, we need to know what inferences or predictions have been made about us, and many of these algorithms operate without our knowledge.

The ever-increasing demand for AI to generate personal information, and its vast capacity to do so, is incompatible with the privacy principle of collection limitation.[110] Solove concludes that:[111]

> AI raises vexing challenges for regulatory oversight, stakeholder participation, and accountability. AI severely complicates transparency, as AI algorithms are dynamic and often inscrutable. AI presents challenges for individual due process. The development of AI technologies often excludes many affected stakeholder groups, especially underrepresented and marginalised groups.
>
> ...
>
> Overall, AI is not an unexpected upheaval for privacy; it is, in many ways, the future that has long been predicted. But AI starkly exposes the longstanding shortcomings, infirmities, and wrong approaches of existing privacy laws.

This thesis examines the extent to which privacy law's focus on individual rights and remedies contributes to the "wrong approaches" of New Zealand's privacy law.

Surveillance

Critical to a central argument of this thesis is the ability of AI to exacerbate the privacy harms of surveillance. I discuss how mass surveillance is a social privacy harm, inadequately addressed by the current regulatory regime, in Chapter Six.

Surveillance Capitalism

The rise of big data has facilitated surveillance capitalism: the monetisation of data generated by the monitoring of people's behaviour and movements online. Concerns

[109] At 41.
[110] At 23.
[111] At 7.

about surveillance are not limited to the actions of governments, such as the widespread generalised surveillance by U.S. intelligence services that motivated Schrems to challenge Facebook's transatlantic transfers of EU data. Equally concerning is surveillance capitalism: a phenomenon coined by academic and social scientist, Shoshana Zuboff, in her influential book, *The Age of Surveillance Capitalism.*[112]

Zuboff describes the concept as a "new economic order that claims human experience as free raw material for hidden commercial practices of extraction, prediction, and sales".[113] She provides multiple descriptions of surveillance capitalism, including "a rogue mutation of capitalism marked by concentrations of wealth, knowledge, and power, unprecedented in human history".[114] Zuboff argues that surveillance capitalism exploits the immense asymmetries in knowledge, and consequently power, between the big technology companies that operate as surveillance capitalists,[115] and the populations that use their services and purchase their products. Zuboff states:[116]

> Surveillance capitalists know everything *about us,* whereas their operations are designed to be unknowable *to us.* They predict our futures for the sake of others' gain but not ours. As long as surveillance capitalism and its behavioral futures markets are allowed to thrive, ownership of the new means of behavioral modification eclipses ownership of the means of production as the fountain head of capitalist wealth and power in the twenty-first century.

---

[112] Shoshana Zuboff *The Age of Surveillance Capitalism: The Fight For A Human Future At The New Frontier of Power* (Profile Books Ltd., London, 2019).
[113] At i.
[114] Zuboff also describes surveillance capitalism as: "...2. A parasitic economic logic in which the production of goods and service is subordinated to a new global architecture of behavioral modification; ...
4. The foundational framework of a surveillance economy;
5. As significant threat to human nature in the twenty-first century as industrial capitalism was to the natural world in the nineteenth and twentieth;
6. The origin of a new instrumentarian power that asserts dominance over society and presents startling challenges to market democracy;
7. A movement that aims to impose a new collective order based on total certainty;
8. An expropriation of critical human rights that is best understood as a coup from above: an overthrow of the people's sovereignty" (Zuboff, above n 112, at i).
[115] Zuboff states that at the time of writing, in 2019, not all of the 5 big tech companies are pure surveillance capitalists, for example Apple has pledged to abstain from surveillance capitalist practices. Amazon also does, to a lesser degree. Apple and Amazon gain revenue from the physical and digital products they sell, so therefore face less financial pressure to act as surveillance capitalists, unlike Google and Facebook. Zuboff's focus is on surveillance capitalists Google, Facebook, and Microsoft (Zuboff, above n 112, at 22-24).
[116] Zuboff, above n 112, at 11.

Zuboff distinguishes surveillance capitalism from traditional capitalism. She argues that when a company collects personal information with consent and solely for the purposes of improving a product or service, it is committing capitalism.[117] Surveillance capitalism still follows capitalist norms of profit maximisation and competitive production, but these dynamics "operate in the context of a new logic of accumulation".[118] Surveillance capitalism is driven by the economic imperatives of extraction and prediction: the constant need to find new sources of personal information to fuel its predictive analytics.[119] Zuboff believes that the requirements of behavioural modification steer surveillance capitalists towards the ultimate objective of total information and control.[120]

Surveillance capitalism highlights the shortcomings of data privacy regimes focussed on the individual rights of data subjects; premised on the basic information rights of the late 1970s-1980s. The value of big data to tech companies lies in the sheer volume of information about people and how they relate to each other. The harms of a few big tech companies, or governments, holding excessive amounts of information about populations are collective, public harms, not simply individualistic ones.

Similarly, the collective harms of mass surveillance enabled by networks of CCTV cameras cannot be addressed by privacy regimes that rely on the individual to exercise control over their data and assert their own privacy rights. Problematically, the contemporary individual control model of data privacy works in favour of those that benefit from increased surveillance. Solove argues that privacy law has never been well-equipped to address the harms of surveillance, and concerningly, the rise of big data enables mass surveillance at unprecedented levels.[121]

[117] At 22.
[118] At 66-67.
[119] At 87; 128; 338.
[120] At 67.
[121] Solove, above n 98, at 50.

United States of America

For as long as the world's biggest technology companies are based in the United States, the U.S. will continue to leverage significant influence over international data privacy practices, norms, and laws. The U.S. lobbied aggressively for its interests in the EU legislative processes in negotiations leading up to the Data Protection Directive and the GDPR and managed to exert influence over negotiation outcomes. Bygrave states that:[122]

> US policy preferences continue to shape real levels of data privacy in online worlds. They are embodied in the culture of the US-based corporations that maintain a powerful grip over much of Internet development and set many of the default standards for the routine processing of data on Internet end-users. For a huge proportion of the latter, the Internet has become largely a Google-, Facebook-, and Apple-mediated experience. The data privacy standards governing this experience remain primarily rooted in US law, even if EU law increasingly holds influence.

The overall data privacy framework in the U.S. is piecemeal. It was not until 1996 that the Health Insurance Portability and Accountability Act was passed, ensuring privacy protection for medical records. In 1998 the Children's Online Privacy Protection Act was passed providing online privacy protections for children, but only those under the age of thirteen. There is no omnibus federal privacy law,[123] yet there is a statute prohibiting the wrongful disclosure of video tape rentals or sales records.[124] There is no national data protection agency in the United States. The agency that comes closest to this is the Federal Trade Commission (FTC) which can uphold data privacy standards in contexts involving misleading or deceptive business practices.[125]

[122] Bygrave, above n 14, at 107.
[123] Note that there is a federal Privacy Act 1974 that establishes a code of fair information practices governing the collection, maintenance, use, and dissemination of information about individuals that is maintained in systems of records by federal agencies.
[124] United States of America Video Privacy Protection Act 1988.
[125] Bygrave, above n 14, at 110-111.

Federal Trade Commission

Section 5 of the Federal Trade Commission Act 1934 prohibits "unfair or deceptive acts or practices in or affecting commerce".[126] Gehan Gunasekara and Jingyi Xiong state that, in the absence of a privacy regulator in the U.S, "...the FTC has assumed the de facto role of one".[127] In their research, Gunasekara and Xiong analyse the FTC's publicised settlements involving unfair or deceptive business practices against the approach of conventional privacy principles such as those in New Zealand's Privacy Act.[128] They conclude that conventional information privacy principles can, for the most part, address the problems encountered in the FTC settlements.[129] While the case-by-case approach of the FTC allows for flexibility in the face of technological change, Gunasekara and Xiong identify deficiencies in the U.S. approach.[130] They include a lack of transparency, comparative lack of rights for individuals to access and seek correction of their information, and the reactive approach of FTC settlements in contrast to the principles-based approach of New Zealand's Privacy Act.[131]

California Consumer Privacy Act 2018

California became the first state to introduce its own data privacy law in 2018 with the California Consumer Privacy Act (CCPA), described both as "the absolute toughest"[132] and "most ambitious and comprehensive piece of privacy legislation in the history of the United States".[133] The CCPA states that it is the intention of the Legislature to further Californians' right to privacy by providing them with the right to know what data is being collected about them[134] and to be able to access that data;[135] to know if their personal data

[126] The Federal Trade Commission Act of 1934 15 USC § 45.
[127] Gehan Gunasekara and Jingyi Xiong "Lost in Translation? Privacy and Unfair or Deceptive Acts or Practices in Commerce in the United States" (2016) 22 New Zealand Business Law Quarterly at 4.
[128] The Privacy Act 1993 was in force at the time of publication, however the Privacy Act 2020 applies the same principles-based approach of the 1993 Act but with an additional information privacy principle addressing disclosures of information overseas.
[129] Gunasekara and Xiong, above n 127, at 30.
[130] At 30.
[131] At 30-31.
[132] John Stephens "California Consumer Privacy Act" (American Bar Association, 2019) https://www.americanbar.org/groups/business_law/publications/committee_newsletters/bcl/2019/201902/fa_9/
[133] Jordan M. Blanke "Protection for 'Inferences Drawn': A Comparison Between the General Data Protection Regulation and the California Consumer Privacy Act" (2020) Global Privacy Law Review I 2 81-92 at 88.
[134] The California Consumer Privacy Act of 2018 [CCPA], s 1.
[135] Section 4.

is sold or disclosed and to whom;[136] and to be able to object to the sale of their personal data.[137] Californians have the right to equal service and price even if they exercise their privacy rights.[138] The CCPA is targeted at big business and data brokers.[139] Civil litigants started filing class actions for statutory breaches of the CCPA shortly after it went into effect on 1 January 2020.[140]

The California Privacy Rights Act of 2020 amends the CCPA, enshrining additional data protection provisions and creating the California Privacy Protection Agency to investigate breaches of, and enforce, state privacy laws. The California Privacy Rights Act gives consumers greater control over the information businesses collect about them, including a right of correction and a right to limit the use and disclosure of sensitive information.[141]

Colorado, Connecticut, Utah, and Virginia began enforcing similar statutes to California's CCPA in 2023.[142] A number of other states passed privacy laws in 2023.[143] Maryland is the latest of five states to pass privacy laws in the first half of 2024.[144] Commentators point to the growing number of state privacy laws as both a cause and effect of the lack of comprehensive federal privacy legislation.[145] However, two omnibus consumer privacy bills were introduced to Congress in 2023: the Data Care Act of 2023 and the Online Privacy Act of 2023.[146] The Data Care Act would impose a duty of care on online

[136] Section 2.
[137] Section 3.
[138] Section 5.
[139] The CCPA applies to for-profit businesses that do business in California, and have a gross annual revenue of over 25 million dollars; or buy, receive or sell the personal information of 50,000 or more Californian residents, households, or devices; or derive 50% or more of their annual revenue from selling Californian residents' personal information (State of California Department of Justice, Office of Attorney-General https://oag.ca.gov/privacy/ccpa)
[140] *Barnes v Hanna Andersson and Salesforce* was filed in the Northern District of California on 3 February 2020 and *Burke v Clearview AI* was filed in the Southern District of California on 27 February 2020.
[141] State of California Department of Justice, Office of Attorney-General "California Consumer Privacy Act (CCPA)" https://oag.ca.gov/privacy/ccpa
[142] Frederic D. Bellamy "U.S. data privacy laws to enter new era in 2023" (13 January 2023) Westlaw Thompson Reuters https://www.reuters.com/legal/legalindustry/us-data-privacy-laws-enter-new-era-2023-2023-01-12/
[143] Iowa, Indiana, Tennessee, Montana, Florida, Oregon, Delaware, Florida, and Texas passed privacy laws in 2023 (Sheppard Mullin "U.S. State Comprehensive Privacy Laws" https://discover.sheppardmullin.com/us-state-comprehensive-privacy-laws/p/1).
[144] Sheppard Mullin Eye on Privacy "Maryland, the Old Line State, Creates New Lines with Consumer Privacy Law" (20 May 2024) https://www.eyeonprivacy.com/2024/05/maryland-the-old-line-state-creates-new-lines-with-consumer-privacy-law/
[145] Müge Fazlioglu "U.S. privacy legislation in 2023: Something old, something new?" (26 July 2023) IAPP https://iapp.org/news/a/u-s-federal-privacy-legislation-in-2023-something-old-something-new/
[146] Fazlioglu, above n 145.

service providers to reasonably secure personal data from unauthorised access. The Online Privacy Act is an amended version of legislation previously introduced in 2019 and then again, in 2021. It contains basic information rights of access, correction, deletion, portability, and human oversight of automated decision-making, as well as provisions on disclosure and consent. Notably, it would establish a new federal entity, the Digital Privacy Agency.[147]

Google, along with other large technology companies based in California, lobbied extensively to dilute the provisions of the CCPA.[148] It is typical of lobbyists representing business interests to exert significant pressure on legislators who attempt to pass bills giving consumers greater data privacy rights. Commercial factors are a significant reason for the United States' lack of comprehensive privacy protections. The right to free speech has also been used to overturn privacy legislation in the United States.[149] Carissa Véliz, Associate Professor at the Institute for Ethics in AI at the University of Oxford, argues that big technology companies hold too much personal information about us. Véliz describes how after 9/11 U.S intelligence agencies saw an opportunity to expand their surveillance by obtaining a copy of all the personal data that corporations were collecting. Hence there was no incentive for the U.S. government to regulate for privacy. The opposite was true - government surveillance becomes more powerful with the more data that businesses collect.[150]

The vested interests of big tech are not the sole explanation behind the absence of comprehensive privacy protection in the United States. Bygrave argues that the reasons for the transatlantic divergence are nuanced and complex.[151] He believes that Americans view the primary importance of privacy as protecting individual freedoms from government intrusion. Whereas in Europe, the protection of privacy is linked to respect

[147] Fazlioglu, above n 145.
[148] Kartikay Mehrotra, Laura Mahoney and Daniel Stoller "Google and other tech firms seeks to weaken landmark California data-privacy law" (4.9.19) Los Angeles Times https://www.latimes.com/business/story/2019-09-04/google-and-other-tech-companies-attempt-to-water-down-privacy-law; Zack Whittaker "Silicon Valley is terrified of California's Privacy law. Good" (20.9.19) TechCrunch https://techcrunch.com/2019/09/19/silicon-valley-terrified-california-privacy-law/
[149] Bygrave, above n 14, at 111.
[150] Véliz, above n 2, at 36-37; Michael Birnhack and Niva Elkin-Koren "The Invisible Handshake: The Re-emergence of the State in the Digital Environment" (2022) 8 2 Virginia Journal of Law and Technology 1.
[151] Bygrave, above n 14, at 107-116.

and dignity of the individual, as well as being critical for maintaining societal values of civility, pluralism, and democracy.[152] The last few years have seen lawmakers in both the Democrat and Republican parties growing increasingly concerned about the behaviour, size, and dominance of the big technology companies.[153] This concern, accompanied by the international focus on AI governance in 2023 and the increasing legislative activity of State lawmakers, could indicate that the passing of a comprehensive federal privacy law in the United States is more likely in 2024 than ever before.

New Zealand's privacy legislation

OECD Guidelines: The foundation of New Zealand's Privacy Act 1993

New Zealand's privacy legislation has been influenced by international developments in data privacy law, in particular the OECD Guidelines 1980.[154] As discussed earlier in this Chapter, the 1970s and 1980s saw a number of countries pass data privacy legislation in response to the risks posed by the development of large, computerised databases. New Zealand shared similar concerns about the potential for misuse of personal information held in large databases by government agencies.[155] However, where New Zealand's Privacy Act 1993 departed significantly from equivalent legislation in other jurisdictions was in its application to private sector agencies in addition to public sector ones.[156]

The OECD Guidelines formed the foundation of the Privacy Act 1993. The long title of the Privacy Act 1993 states that it is an Act "to promote and protect individual privacy in general accordance with [the OECD Guidelines]". The long title also states that, in particular, one of its purposes is:

---

[152] At 112.
[153] Joe Biden "Republicans and Democrats Unite Against Big Tech Abuses" (11 January 2023) Wall Street Journal https://www.wsj.com/articles/unite-against-big-tech-abuses-social-media-privacy-competition-antitrust-children-algorithm-11673439411; Cecilia Kang "Democratic Congress Prepares to Take On Big Tech" (26 January 2021) New York Times https://www.nytimes.com/2021/01/26/technology/congress-antitrust-tech.html; Theodore F. Claypoole "U.S Senate Takes Aim at Big Tech" (2021) National Law Review XI:40; Makena Kelly "All The Ways Congress Is Taking On The Tech Industry" (3 March 2020) The Verge https://www.theverge.com/2020/3/3/21153117/congress-tech-regulation-privacy-bill-coppa-ads-laws-legislators
[154] Stephen Penk "The Privacy Acts 1993 and 2020" in Nikki Chamberlain and Stephen Penk (eds) *Privacy Law in New Zealand* (3rd ed) (Thompson Reuters, Wellington, 2023) at 144.
[155] At 143-144.
[156] At 142.

> ... to establish certain principles with respect to ... the collection, use, and disclosure ... of information relating to individuals; and ... access by each individual to information relating to that individual and held by public and private sector agencies ...

Another purpose set out in the long title of the 1993 Act is "to provide for the appointment of a Privacy Commissioner to investigate complaints about interferences with individual privacy". The Privacy Act 1993 established the Office of the Privacy Commissioner[157] and provided a framework for the investigation and resolution of privacy complaints.[158]

Background to the Privacy Act 1993

Access to information

Prior to the passing of New Zealand's first privacy legislation, the Privacy Act 1993, the Official Information Act 1982 (OIA) and the Local Government Official Information and Meetings Act 1987 (LGOIMA) constituted the access to information regime in New Zealand.[159] Both the OIA and the LGOIMA aim to provide for a balance between access to information, and its legitimate protection.[160] The enactment of the Privacy Act in 1993 broadened access rights to include personal information held by private agencies.[161]

Wanganui Computer Centre Act 1976

The Privacy Act 1993 repealed the Wanganui Computer Centre Act 1976.[162] The Wanganui Computer Centre Act established a national police computer system.[163] Under this Act individuals had the right to apply to the Wanganui Computer Centre Privacy Commissioner for a copy of the information held about them on the police computer system[164] and could make a complaint about any information they believed to be inaccurate, incomplete, or likely to create a misleading impression.[165]

---

[157] Privacy Act 1993, Part 3.
[158] Part 8.
[159] Graham Taylor and Paul Roth *Access to Information* (LexisNexis NZ Limited, Wellington, 2011) at 1.
[160] At 9.
[161] At 2.
[162] Privacy Act 1993, s 129(2).
[163] Wanganui Computer Centre Act 1976, s 3.
[164] Section 14.
[165] Section 15.

Privacy Commissioner Act 1991

The now repealed Privacy Commissioner Act 1991 established the office of the Privacy Commissioner and authorised information matching by specified government departments.[166] The Privacy Commissioner had a range of functions including reviewing information matching programmes and publishing an annual report on those programmes. However, the Privacy Commissioner Act did not confer on the Commissioner any power to investigate privacy complaints.[167]

Impetus behind the Privacy Act 1993

The Privacy Act 1993 was passed in response to several factors. An important motivation was the anticipated EU Data Protection Directive[168] and the advantages to be gained from making New Zealand's data privacy law "adequate" under the Data Protection Directive.[169]

Article 25 of the EU Data Protection Directive stated that member states and the Commission shall inform each other of third countries that do not ensure an adequate level of data privacy protection. The principles for assessing a third country's data privacy regime were set out in the Article 29 Data Protection Working Party in *Transfers of personal data to third countries: Applying Articles 25 and 26 of the EU data protection directive*.[170] The Working Party's "Opinion 11/2011 on the level of protection of personal data in New Zealand" (4 April 2011) made a recommendation of adequacy.[171]

---

[166] Penk, above n 154, at 147.
[167] At 147.
[168] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, OJ 1995 L 281/31.
[169] New Zealand formally achieved adequacy status in 2015.
[170] Article 29 Data Protection Working Party "Transfers of personal data to third countries: Applying Articles 25 and 26 of the EU data protection directive" (WP12, DG XV D/5025/98, adopted on 24 July 1998). D/5025/98, adopted on 24 July 1998) https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/1998/wp12_en.pdf
[171] Article 29 Data Protection Working Party "Opinion 11/2011 on the level of protection of personal data in New Zealand" (4 April 2011) 00665/11/EN WP 182. https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2011/wp182_en.pdf

In their criticism of this adequacy finding, Graham Greenleaf and Lee Bygrave reiterate the point that adequacy does not equate to equivalence and argue that the adequacy recommendation was given "...despite considerable shortcomings in the NZ data protection regime".[172] Greenleaf and Bygrave state that despite the Working Party finding seven instances of New Zealand's privacy regime not being fully adequate, they were held not to be sufficiently serious to prevent a finding of adequacy. They argue that this was a pragmatic decision, premised on New Zealand's relative geographical and economic isolation; the unlikelihood of EU data being transferred to New Zealand and the lack of direct marketing from New Zealand into the EU.[173] Greenleaf and Bygrave also make the critical point that "...it is the effect of a third party's laws on EU citizens that counts toward adequacy, not the effect on the country's own citizens".[174]

A significant domestic motivation behind the enactment of the Privacy Act 1993 was to gain public acceptance of the government's Information Matching Agreements.[175] There are 55 Authorised Information Matching Programmes in place in New Zealand.[176] Information matching is the comparison of one set of government records with another set of government records that relate to the same person. When it is done by computer it is referred to as 'data matching'. The process is predominantly used to detect fraud in public assistance programs, one example being the comparison of people on an unemployment benefit with those incarcerated. Sometimes the data matching process can be used to identify individuals entitled to assistance or entitlements. For example, new citizens are sent reminders to enrol to vote. Without sufficient scrutiny, the information- or data- matching process is vulnerable to misuse and breaches of privacy.[177] The

[172] Graham Greenleaf and Lee Bygrave "Not entirely adequate but far away: Lessons from how Europe sees New Zealand data protection" (2011) Privacy Laws and Business International Report 111 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1964065
[173] Greenleaf and Bygrave, above n 172.
[174] Greenleaf and Bygrave, above n 172.
[175] Office of the Privacy Commissioner "Information matching / Data Matching" https://privacy.org.nz/privacy-act-2020/information-sharing/information-matching/
[176] Office of the Privacy Commissioner "Information matching provisions" https://www.privacy.org.nz/privacy-act-2020/information-sharing/information-matching-provisions/
[177] The Office of the Privacy Commissioner outlines the following privacy risks of information matching: "using information obtained for one purpose for an unrelated purpose; 'fishing' in government records with the hope of finding wrongdoing by someone; automating decisions affecting individuals and removing human judgment; presuming people guilty simply through their being listed in a computer or requiring people to prove their innocence; multiplying the effects on individuals of errors in some government databases; and undermining personal information by dispersing information obtained by one agency in confidence onto a variety of other agencies' databases" (https://www.privacy.org.nz/privacy-act-2020/information-sharing/information-matching-overview-2/).

government aimed to minimise these risks, and perhaps equally importantly, the public perception of these risks, by regulating the use of information matching agreements by the public sector under the Privacy Act 1993.

Another impetus behind the enactment of the Privacy Act 1993 was the health sector reforms of the early 1990s. The reforms were accountability-driven with a focus on increasing efficiencies, reducing waiting lists, and containing expenditure.[178] It is not coincidental that the Health Information Privacy Code 1994 was the first code of practice to be issued under the Privacy Act 1993.

The Privacy Act 1993 recognised the individual's interest in protecting and safeguarding their personal information. Rather than setting out to recognise privacy as a fundamental human right, the Privacy Act 1993 was passed primarily to facilitate economic and strategic objectives.

New Zealand's privacy reforms

Unlike the United States, New Zealand is unable to address the issues of big tech dominance in a meaningful or comprehensive way. However, that does not mean there are no measures that could be taken to limit the harms of surveillance capitalism and improve New Zealanders' privacy. The core argument of this thesis is that we need to think about privacy as a social concept. Overly individualistic data privacy regulation, such as New Zealand's Privacy Act 2020, falls short in addressing the challenges of big data. The period from 1993, when New Zealand's first Privacy Act came into force, to its repeal in 2020, saw fundamental developments in computing technology. In particular, the exponential increase in capacity for the storage of data, and corresponding erosions of privacy. More of people's day-to-day activity moved online, facilitating the big technology companies' ability to generate increasing amounts of personal data for their own gain. Unfortunately, these developments were not adequately addressed in the reforms to New Zealand's Privacy Act in 2020.

[178] Nancy Devlin, Alan Maynard, Nicholas Mays "New Zealand's new health sector reforms: back to the future?" (2001) BMJ 322 1171-1174 at 1171.

Privacy Act 2020

The repeal and replacement of the Privacy Act 1993 with the Privacy Act 2020 saw the introduction of several significant provisions, including mandatory data breach notifications, compliance notices, and tighter rules around the transfer of personal information overseas. These changes were recommended by the New Zealand Law Commission and supported by the Privacy Commissioner. The provisions of the Privacy Act 2020 will be discussed in detail in Chapter Five. The fundamental framework of the Privacy Act remains the same. It continues to be a principles-based Act, premised on an individual consent-based model.

New Zealand Law Commission recommendations

An extensive review of privacy law in New Zealand was undertaken by the New Zealand Law Commission from 2008 - 2011. The length of time that elapsed between the publication of the Law Commission's report in 2011 and the passing of the amended Act in 2020 saw significant developments in computing technology as well as developments in data privacy law, such as the GDPR coming into force. Many of the Law Commission's recommendations were implemented in the Privacy Act 2020, although the strength of some of the recommendations was ultimately watered down in the Act. Two recommendations of the Law Commission were dismissed by the government in the early stages of its review. They were the creation of a power for the Privacy Commissioner to report on surveillance activities, and a review of the handling of health information. The government's response to these two recommendations was that it believed "these recommendations do not appear to address any identifiable problem or issue, are likely to cause confusion, or add administrative burden without clear benefits".[179]

Another recommendation of the Law Commission, ultimately dismissed by the government, was that the role of the Director of Human Rights Proceedings in privacy cases be discontinued. The Law Commission recommended that the Privacy Commissioner take cases to the Human Rights Review Tribunal, streamlining the

[179] New Zealand Government "Government Response to Law Commission Report on *Review of Privacy Act 1993* (2011)" at 6.

complaints process; a recommendation also supported by the Privacy Commissioner at the time, John Edwards.[180]

Changes that were made to the Privacy Act, as recommended by the Law Commission, included the Privacy Commissioner having the power to require an audit of an agency's information handling practices and the ability to issue enforcement notices.[181] The Law Commission also recommended that there should be a new framework under the Act that allowed for the sharing of information between government agencies where it is in the public interest to do so, but with appropriate safeguards.[182]

Office of the Privacy Commissioner recommendations

The Privacy Commissioner welcomed reform of the Privacy Act 1993 and recommended that all of the Law Commission's proposed recommendations be implemented in the Privacy Bill. The Privacy Commissioner also recommended strengthening protections for individuals from the privacy risks of de-identification and providing safeguards for re-identification events;[183] the addition of rights to information portability and the right to erasure of personal information;[184] as well as adding a new Information Privacy Principle to address automated decision-making and algorithmic transparency.[185] The recommendations of the Privacy Commissioner came the closest to acknowledging the potential threats posed by new information technologies and the rise of surveillance capitalism. It is unfortunate that not all of his recommendations were incorporated into the Act.

Overall, the reforms to the Privacy Act were underwhelming. The opportunity to address the privacy concerns raised by big data was missed. Nevertheless, some useful provisions were added to the Act. In Chapter Five, I assess their effectiveness and conclude that the

---

[180] John Edwards, Privacy Commissioner "Privacy Commissioner's submission on the Privacy Bill to the Justice and Electoral Select Committee" (31 May 2018) at 1.
[181] Law Commission *Review of the Privacy Act 1993: Review of the Law of Privacy Stage 4* (March 2010, Wellington) at 166.
[182] At 276.
[183] John Edwards, Privacy Commissioner, above n 180, at 22-25.
[184] At 25-30.
[185] At 30-34.

Privacy Act's continued focus on the individual limits its efficacy in addressing the social privacy harms of big data.

Privacy Amendment Bill

The Privacy Amendment Bill 292-1 (2023) will introduce a new information privacy principle 3A (IPP 3A), addressing a current gap in the Privacy Act[186] and ensuring New Zealand continues to maintain its adequacy status under the GDPR.[187] IPP 3A is similar to IPP 3 which deals with the direct collection of personal information. Currently, there is no requirement for agencies to notify individuals when they collect their information from a third party. The new indirect notification obligation requires an agency to notify an individual of a range of matters when collecting their information indirectly, including the name and address of the agency, the purposes for which the information is being collected, and the rights of access to, and correction of, the information. Nonetheless, the requirement will be subject to a wide range of exceptions to ensure the efficient administration of certain public functions and to protect against other unintended consequences.[188]

New Zealand's tort of privacy

Tort law was the precursor to New Zealand's privacy legislation. The Privacy Act 2020 does not cover all areas of information privacy. Data protection legislation routinely excludes the news media and information held in connection with one's domestic or personal affairs, and this is also the case in New Zealand.[189] Tort law is useful for filling in the gaps of the Privacy Act.

---

[186] The Privacy Amendment Bill also introduces other minor changes to the Privacy Act 2020.
[187] European Commission "Report From The Commission To The European Parliament And The Council on the first review of the functioning of the adequacy decisions adopted pursuant to Article 25(6) of Directive 95/46/EU (15.1.2024) at 13. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52024DC0007
[188] Privacy Amendment Bill 292-1 (2023) https://www.legislation.govt.nz/bill/government/2023/0292/15.0/whole.html#d17500865e2
[189] Section 8(b)(x) of the Privacy Act 2020 states that a New Zealand agency (to which the Privacy Act applies) does not include a news entity, to the extent that it is carrying on news activities; s 27 sets out a restricted application of the information privacy principles to personal information collected or held for personal or domestic affairs.

New Zealand is the only country outside of the United States to have been receptive to the idea of a tort of privacy from as early as the 1980s.[190] The New Zealand Court of Appeal declared a separate common law tort of privacy in 2004.[191] Other common law countries have been reluctant to recognise a privacy tort.[192] New Zealand's common law actions for breach of privacy are similar to the British tort of equitable breach of confidence which covers the wrongful disclosure of confidential information. In 1849 in the precedent-setting case of *Prince Albert v Strange*[193] the court recognised that Prince Albert's privacy was the right invaded by the prospective exhibition, and publication of a catalogue, of personal family etchings. The court issued an injunction to prevent their publication, holding that Strange's possession of the etchings must have originated in a breach of trust or confidence of an employee of the printer in taking more impressions than were ordered. The action was based on the actual or constructive knowledge of the confidential relationship between Prince Albert and the printer to whom he had entrusted the plates.

In the UK there are two versions of an equitable tort of breach of confidence. The first, the classic breach of confidence, requires a confidential relationship between the parties and information disclosed in circumstances where the information is confidential and there is some loss or harm.[194] (Breach of confidence is an equitable cause of action in New Zealand and can be a useful remedy in a data breach.) The second version of a breach of confidence grew out of the classic breach of confidence action and overlaps with misuse of private information. It requires publication of information where there is a reasonable expectation of privacy (irrespective of a confidential relationship) and the recipient of the information knows, or ought to know, that it must be treated as confidential.[195]

[190] *Tucker v News Media Ownership Ltd.* [1986] 2 NZLR 716.
[191] *Hosking v Runting* [2005] I NZLR 1.
[192] There were indications that a privacy tort might be introduced in Australian law but this has not developed into an independent right to privacy (*Hosking v Runting* [2005] I NZLR 1 at para 54).
[193] *Prince Albert v Strange* (1849) 41 ER 1302.
[194] *Campbell v MGN Ltd* [2004] 2 AC 457.
[195] *Weller v Associated Newspapers Ltd* [2015] EWCA Civ 1176.

Tort of public disclosure of private facts

The tort of public disclosure of private facts is one of two privacy torts recognised in New Zealand[196] and it corresponds in most respects to the second version of the British equitable tort of breach of confidence. The New Zealand privacy tort of public disclosure of private facts drew upon the US tort, although did not replace it.[197] In *Hosking v Runting,* the New Zealand Court of Appeal recognised the influence of the American privacy torts on British jurisprudence but also acknowledged how the first amendment right to freedom of speech had diluted the right to privacy since the days of Warren's and Brandeis's article expounding the right 'to be let alone'.[198] The Court of Appeal held that the elements of a breach of confidence action are well-established in New Zealand, but maintained that many privacy cases do not fit within this action "yet undoubtedly justify legal remedies".[199] Privacy and confidence are different concepts and the Court of Appeal was reluctant to cause confusion by merging the two.[200]

The New Zealand privacy tort of public disclosure of private facts requires the existence of facts in respect of which there is a 'reasonable expectation of privacy', and publicity given to those facts that would be considered 'highly offensive' to an objective, reasonable person.[201] It is a defence if the publication of the private facts is justified by a legitimate public concern (as opposed to being "merely interesting").[202] This tort could potentially be used to seek a remedy for the publication of sensitive personal information, such as health information.

---

[196] The second privacy tort is intrusion on an individual's seclusion or solitude. In 2012 the High Court in *C v Holland* [2012] 3 NZLR 672 held that an intentional invasion of privacy of the type that occurred in this case (the non-consensual filming of the plaintiff showering) even without publicity, or the prospect of publicity, is an actionable tort in New Zealand. It involved the infringement of a reasonable expectation of privacy that would be highly offensive to a reasonable person.
[197] Rosemary Tobin, "The Common Law Tort of Invasion of Privacy in New Zealand" in Nikki Chamberlain and Stephen Penk (eds) *Privacy Law in New Zealand* (3rd ed.) (Thomson Reuters, Wellington, 2023) at 195-197.
[198] *Hosking v Runting*, above n 191, at paras 67-75.
[199] At para 46.
[200] At para 48.
[201] At para 117.
[202] At paras 129-134.

Tort of breach of confidence

The tort of breach of confidence can be used to seek an injunction to prohibit the disclosure, or cease the ongoing disclosure, of confidential information. There are three elements required to establish a breach of confidence. They are firstly, that 'the information has the necessary quality of confidence about it'; secondly, the information has been 'imparted in circumstances' that import 'an obligation of confidence'; and thirdly, there must be an unauthorised use or disclosure of the information.[203]

It is quicker to bring a common law action as opposed to filing a complaint with the Privacy Commissioner for a privacy infringement. However, civil litigation is expensive. It is free to make a complaint under the Privacy Act, but it may take up to two or three years for a resolution if the complaint makes it all the way to the Human Rights Review Tribunal for consideration.

Conclusion

New Zealand's regulatory framework for data privacy has been influenced by multiple factors, including international developments in data privacy law, commercial and governmental interests in collecting and using personal information, the rise of big data and the assorted predictive and surveillance technologies it enables, as well as political and cultural factors unique to New Zealand. International developments in privacy law, and the regulation of AI in other countries, will arguably have increasing influence on New Zealand's regulatory framework, given the cross-jurisdictional nature of information practices.

The Privacy Act 2020, like its equivalent in many other jurisdictions, is coming under increasing strain from its inherently contradictory objectives: to facilitate the commercial interests in the free flow of personal information, while at the same time protecting individuals' privacy interests. This tension has always been present but is exacerbated in an age of big data. Personal information is collected and repurposed to an extent that the creators of the first data privacy regimes could not have envisaged. The privacy principles

[203] *Coco v AN Clarke (Engineers) Ltd* [1969] RPC 41 (Ch) applied in *EQC v Krieger* [2013] NZHC 3140 at 10-11.

of the OECD Guidelines, which reflect basic information rights recognised around the world, and which form the basis of New Zealand's Privacy Act, fall short in addressing the social privacy harms of big data.

History shows us that, with the political will, privacy law is adaptable and flexible enough to respond to challenging new uses of technology. In the following chapters I will describe the benefits and harms of big data and argue that a shift in attention from the rights and responsibilities of the *individual* data subject, to focus more on the *social context* of the information exchange, is necessary if data privacy regimes are to meet the challenges of big data.

Chapter Two - The Promise of Big Data: An overview of big data in health

Introduction

Big data delivers greater speed and precision across many sectors of society: directing major retailers' marketing and pricing strategies;[1] determining the news we see on social media;[2] facilitating improved driving experiences through GPS systems and traffic prediction tools.[3] Governments around the world harness big data to prevent fraud and increase efficiencies.[4] Law enforcement agencies deploy big data analytics in risk assessments and crime prevention.[5] Big data is everywhere, but health is where big data promises the most: to improve both the quality and the longevity of our lives. This chapter focusses on health in assessing the advantages of big data because this is an area where there are big expectations and the potential for the greatest benefits to society.[6]

There has been considerable enthusiasm for the potential of big data in health research. In November 2012, many of the world's leading experts in medical research and biomedical informatics came together, for the first time, to discuss the potential of big data in health at the *Oxford-Stanford Conference on Big Data*[7], a conference organised

[1] Amazon, Marriott Hotels and Starbucks are examples of companies that utuilise big data applications to personalise their marketing and / or apply dynamic pricing (Nina Tudor, "7 real world examples of how brands are using big data analytics" (2021) https://www.bornfight.com/blog/7-real-world-examples-of-how-brands-are-using-big-data-analytics/

[2] Akos Lada "Facebook: How Does News Feed Predict What You Want To See?" (2021) Meta https://about.fb.com/news/2021/01/how-does-news-feed-predict-what-you-want-to-see/

[3] Joao Laranjeira "What is traffic prediction and how does it work?" (2020) TomTom https://www.tomtom.com/newsroom/behind-the-map/road-traffic-prediction/

[4] In New Zealand, authorised information matching agreements under Part 7, subpart 4 of the Privacy Act 2020, assist in preventing fraud by allowing information matching by government agencies to identify people who are receiving allowances or benefits they are not entitled to because they are domiciled overseas or incarcerated, for example. (Office of the Privacy Commissioner https://privacy.org.nz/privacy-act-2020/information-sharing/information-matching/).

[5] Location based predictive policing algorithm, Predpol, shows police officers a digital map where, when, and whay type of crime it predicts will occur (www.predpol.com).

[6] John Bell, Regius Professor of Medicine, University of Oxford, believes that the impact of big data in medicine will be as big or bigger than in many other disciplines. The potential of big data to significantly improve health outcomes was recognised by the participants of the Oxford-Stanford Conference on Big Data (https://www.bdi.ox.ac.uk/news/newsitem-3).

[7] University of Oxford Big Data Institute "Big Data: challenges and opportunities for human health" Oxford-Stanford Conference on Big Data, November 2012 https://www.bdi.ox.ac.uk/news/newsitem-3

by both universities and supported by the Li Ka Shing Foundation.[8] The conference organisers concluded that "we are poised for a revolution in the way society understands disease and treats patients in the twenty first century".[9] The general consensus was that big data had the potential to significantly improve patient outcomes through better disease surveillance, earlier and more accurate detection of diseases, and a greater understanding of the causes of disease, *if* the challenges of how to effectively analyse big data could be overcome.[10] Concerns about the challenges inherent in applying big data to healthcare have been raised from the outset[11] and, over a decade later, are still considered problematic. Crucially, the promise of big data's potential to dramatically change the delivery of healthcare is yet to be realised.

In the first part of this chapter, I provide an overview of big data in healthcare with examples of algorithms matching, or outperforming, humans at medical pattern recognition and risk prediction, and big data assisting in the surveillance of infectious diseases. I also examine some of the methodological hurdles of applying big data to a healthcare context that have limited its progress to date. Despite these shortcomings, the government of the United Kingdom (UK) and the National Health Service (NHS) have been eager to partner with technology companies and artificial intelligence (AI) firms in pursuit of the increased efficiencies and improved health outcomes that big data promises. In the second part of this chapter, I consider the ethical issues raised by the relationships between government and technology companies and the lessons that New Zealand can learn from the UK.

This Chapter examines the question of whether diminished privacy is a price worth paying for the potential of big data. It is important to distinguish the potential of big data from the big technology companies that profit from it. One can dislike the business

[8] The La Ki Shing Foundation is a philanthropic organisation with a focus on supporting education and healthcare initiatives (www.lksf.org).
[9] University of Oxford Big Data Institute, above n 7.
[10] University of Oxford Big Data Institute, above n 7.
[11] Richard Hillestad, James Bigelow, Anthony Bower, Frederico Girosi, Robin Meili, Richard Scoville, Roger Taylor "Can Electronic Medical Record Systems Transform Health Care? Potential Health Benefits, Savings, And Costs" (2005) 24 Health Affairs 5 1103-1117; James S. Kahn, Veenu Aulakh, Adam Bosworth "What It Takes: Characteristics Of The Ideal Personal Health Record" (2009) 28 Health Affairs 2 369-376; Clifford Lynch "How do your data grow?" (2008) 455 Nature 28-29; Alexander Szalay and Jim Gray "Science in an Exponential World" (2006) 440 Nature 413-414; Doug Howe and Seung Yon Rhee et al "The Future of Biocuration" (2008) 455 Nature 47-50.

practices of Amazon and Google, for example, while still embracing big data research and development. Responsible, socially beneficial, big data initiatives are dependent on the effective regulation of personal information. The rules that govern the disclosure and use of health information should respect health values. An important value in healthcare is patient autonomy. Therefore, in many health contexts, the right of the individual to give, or withhold, consent to certain uses of their data should be upheld. An understanding of social context, in particular, social norms of information exchange, is critical to the effective regulation of personal information.

The three Vs of big data

There is no agreed definition of "big data", however, big data is typically characterised by the three Vs - Volume (very large amounts of data), Variety (data from different and diverse sources), and Velocity (high speed at which data is generated and analysed).[12] The concept of big data encompasses more than just the source or type of data; it includes big data analytics - the processing of large volumes of data through machine learning techniques, AI, or algorithm. In this thesis, my consideration of big data includes the technologies that it enables.

Advantages of big data in health

Algorithms perform certain tasks better than humans

An algorithm, as defined in Chapter One, is simply a sequence of steps used to solve a problem.[13] However, the 3 Vs of big data mean that big data naturally lends itself to processing by algorithm. Predictive algorithms can harness the processing power of computers to compute thousands of steps at speed to solve a problem or predict an outcome. Therefore, it is not surprising that algorithms can perform better than humans

[12] Doug Laney is credited with introducing the 3 V's concept in 2001 (TechTarget "3V's (volume, velocity, and variety)" https://www.techtarget.com/whatis/definition/3Vs#:~:text=The%203%20V's%20(volume%2C%20velocity,number%20of%20types%20of%20data.) Some authors include a fourth 'V' of veracity. However, the accuracy of many big data sources is questionable.
[13] John Paul Mueller and Luca Massaron *Algorithms for dummies* (John Wiley & Sons Inc., Hoboken, New Jersey, 2017) at 11.

at certain tasks, such as calculation, pattern recognition, risk prediction, and surveillance.[14]

In her book, *Hello World: How to Be Human in the Age of the Machine,*[15] Professor in the Mathematics of Cities, Hannah Fry, argues that algorithms are not inherently good or bad; "it's how they're used that matters".[16] Fry is optimistic about the use of algorithms, particularly in medicine where their positive contribution to society is far more clear cut than in other social contexts.[17] In a similar vein, I argue that we can benefit from allowing algorithms to carry out the tasks that they perform better than humans while recognising their inherent limitations in other contexts.

Medical pattern recognition

Fry explains how our scientific understanding of medicine has progressed through our ability to spot patterns, classify symptoms, and then use those patterns to predict ill health.[18] Fry identifies pattern recognition, symptom classification, and prediction as skills that algorithms perform very well. Algorithms can be particularly useful in medical pattern recognition, particularly pathology.[19]

Breast cancer screening algorithms

Breast tissue biopsy samples can reveal normal cells at one end of the spectrum and invasive carcinomas at the other. In between the two extremes of normal and horribly malignant are several other more ambiguous categories – atypical cells or pre-cancerous growths. Fry explains that distinguishing between the ambiguous categories is very difficult.[20] She states that:[21]

---

[14] However, the accuracy of the outcomes of these tasks will always be dependent upon the accuracy and completeness of the data the algorithm relies on.
[15] Hannah Fry *Hello World: How to Be Human in the Age of the Machine* (Transworld Publishers, Penguin, London, 2018).
[16] At 3.
[17] At 78.
[18] At 80.
[19] At 80-82.
[20] At 83.
[21] At 88.

> There are two things you want from a good breast cancer screening algorithm. You want it to be **sensitive** enough to pick up on the abnormalities present in all the breasts that have tumours, without skipping over the pixels in the image and announcing them as clear. But you also want it to be **specific** enough not to flag perfectly normal breast tissue as suspicious.
> …
> The problem is that refining an algorithm often means making a choice between sensitivity and specificity. If you focus on improving one, it often means a loss in the other. If, for instance, you decide to prioritize the complete elimination of false negatives, your algorithm could flag every single breast it saw as suspicious.

Recent studies reveal advanced deep learning algorithms show comparable performance with radiologists in interpreting mammograms.[22] Human radiologists are better than algorithms at avoiding false positives - not labelling normal cells as cancerous - but can miss tiny tumours, whereas algorithms are particularly good at avoiding false negatives - spotting abnormal cells - even if they do often identify normal cells as cancerous also.[23] This is because algorithms / AI and humans perceive slightly different image features as suspicious for cancer.[24] Fry believes that the best outcome is to combine the strengths of human and machine:[25]

> The algorithm does the donkey-work of searching the enormous amount of information in the slides, highlighting a few key areas of interest. Then the pathologist takes over. It doesn't matter if the machine is flagging cells that aren't cancerous; the human expert can quickly check through and eliminate anything that's normal. This kind of algorithmic pre-screening partnership not only saves time but bumps up the overall accuracy of a diagnosis to a stunning 99.5 per cent.

Combining the strengths of algorithm and practitioner can result in greater accuracy and efficiency. While the notion of the full screening process being performed by algorithm

---

[22] Karin Dembrower, Alessio Cripper, Eugenia Colôn, Martin Eklund, Fredrik Strand "Artificial Intelligence for breast cancer detection in screening mammography in Sweden: a prospective, population-based, paired-reader, non-inferiority study" (2023) Lancet 5; Ioannis Sechopoulos and Ritse M. Mann "Stand-alone artificial intelligence - The future of breast cancer screening?" (2021) Special issue: Artificial Intelligence in Breast Cancer Care Science Direct 56; Mattie Salim, Erilk Wahlin, Karen Dembrower, Edward Azavedo, Theodoros Foukakis, Yue Liu, Kevin Smith, Martin Eklund and Fredrik Strand "External Evalution of 3 Commercial Artificial Intelligence Algorithms for Independent Assessment of Screening Mammograms" (2020) JAMA Oncology 6(10).
[23] Fry, above n 15, at 88.
[24] Dembrower et al, above n 22, at 708.
[25] Fry, above n 15, at 90.

alone is some way off, what is likely in the not-too-distant future is automated identification of 'normal' cases.[26] A study evaluating three commercial algorithms for independent screening of mammograms found that the combination of the best algorithm in the study with a radiologist identified more positive cancer cases than two human radiologists.[27]

A 2023 population-based study of AI breast cancer detection in Sweden found that a double reading by one radiologist plus AI, resulted in a four per cent increase in screen detected cancers.[28] There was a 21 per cent increase in the number of examinations with an abnormal interpretation with the combined AI and radiologist double reading due to the differences between AI and human interpretations of suspicious tissue.[29] However, this did not translate into an increased recall rate because the consensus discussions that review mammograms, medical history, and the AI result, made more informed determinations. The total number of recalls decreased by four per cent, resulting in a significant workload reduction.[30] Advanced breast cancer screening algorithms look set to deliver efficiencies to overburdened screening programmes, freeing up practitioners' time and resources.

Risk prediction models for somatic illnesses

DeepMind and Moorfields Eye Hospital

The use of big data analytics to screen for non-communicable diseases and chronic conditions is particularly promising in its potential to extend access to preventative care to underserviced populations. The collaboration between Google's DeepMind and Moorfields Eye Hospital, London, has achieved a significant milestone in artificial intelligence reaching, and even exceeding, the performance of experts in referral recommendations on a range of retinal diseases.[31] This has significant implications given

[26] Dembrower et al, above n 22, at 709; Sechopoulos et al, above n 22.
[27] Salim et al, above n 22.
[28] Dembrower et al, above n 22, at 708.
[29] At 708.
[30] At 708-709.
[31] Jeffrey De Fauw et al. "Clinically applicable deep learning for diagnosis and referral in retinal disease" (2018) 24 Nature Medicine 1342-1350; Nima John Ghadiri "AI in Opthalmology" in Michael F. Byrne (ed.) *AI in Clinical Medicine: A Practical Guide for Healthcare Professionals* (2023, John Wiley & Sons Ltd., Hoboken, USA) at 260.

the shortage of human expertise in interpreting an ever-increasing quantity of medical imaging across many clinical specialties. The widespread availability of Optical Coherence Tomography has not been met by the availability of experts to interpret the scans and refer patients for the appropriate clinical care.[32] The progress made by DeepMind paves the way for AI technology to be used in a clinical setting,[33] which could have major implications for under-resourced countries.

Additionally, DeepMind's work in modelling the proteome (the entire set of proteins expressed by the eye at a certain time) is expected to lead to new ways of developing more accurate drugs to treat eye diseases. Administering treatments to the eye has historically been a challenge due to "various anatomical barriers in the eye preventing the drug from reaching its target tissue".[34] Big data and AI may play an important role in understanding drug constituents, and consequently lead to therapeutic breakthroughs. It is argued that DeepMind's work in modelling the proteome "will open new gateways for more accurate structure-based drug design, enabling the optimization of small-molecule drugs".[35]

Airdoc

Airdoc, a Chinese start-up founded in 2014, utilised Microsoft's machine learning capabilities to create a cloud-based algorithm trained on the data analysed from thousands of retinal scans. The AI system takes and analyses photos of the retina and identifies signs of dozens of chronic illnesses, including macular degeneration and diabetes. Airdoc states that it has trained its AI on 10 million fundus (interior surface of the eye) images to date.[36] Microsoft reports that the system has a higher accuracy rate than conventional diagnoses by doctors and is much faster.[37] In 2018, Microsoft, and the founder of Airdoc, Ray Zhang, believed the system was groundbreaking because of its potential to make preventative screening available to millions of people in China, and around the world,

[32] De Fauw et al, above n 31, at 1342.
[33] At 1349.
[34] Ghadiri, above n 31, at 262.
[35] At 262.
[36] Airdoc https://world.airdoc.com/#/
[37] Geoff Spencer "AI and preventative healthcare: Diagnosis in the blink of an eye" (17 September 2018) Microsoft Asia News Centre https://news.microsoft.com/apac/features/ai-and-preventative-healthcare-diagnosis-in-the-blink-of-an-eye/

who might not otherwise have had access to scarce medical resources.[38] In 2020, Airdoc was approved as a medical device in China for the auxiliary diagnosis of diabetic retinopathy and is currently deployed in more than ten countries.[39] Arguably we should not underestimate the potential of such programmes – assuming people will have access to appropriate treatment once conditions are detected.

While the accuracy of AI in ophthalmology, evidenced in the work of Airdoc and Deepmind, is encouraging, it is important not to over-rely on technology in clinical contexts:[40]

> The excitement that accompanies AI can include the temptation to side-step important parts of the eye clinic evaluation. AI is an adjunctive tool rather than a clinical surrogate; there are many nuances to ophthalmic evaluation that AI is unable to synthesize, particularly the social and psychological aspects.

A study of Airdoc's retinal AI system in detecting *multiple* fundus diseases in primary healthcare settings found that while AI has great potential, "the effectiveness of AI systems in screening multiple retinal abnormalities in primary healthcare settings remains unclear".[41] Airdoc's AI "performed satisfactorily in identifying multiple retinal diseases during the initial development and validation".[42] However, the poor quality of images and photographic conditions in real-world settings as well as the low proportion of fundus diseases in primary healthcare might have contributed to the "low predictive value and low sensitivity in the study".[43] The authors concluded that Airdoc's algorithm needs improving to perform better in real-world, primary healthcare settings.[44]

---

[38] Spencer, above n 37.
[39] Airdoc, above n 36.
[40] Ghadiri, above n 31, at 263.
[41] Chufeng Gu, Yujie Wang, Yan Jiang, Feiping Xu, Shasha Wang, Rui Liu, Wen Yuan, Nurbiyimu Abudureyimu, Ying Wang, Yulan Lu, Xiaolong Li,Tao Wu, Li Dong, Yuzhong Chen, Bin Wang, Yuncheng Zhang, Wen Bin Wei, Qinghua Qiu, Zhi Zheng , Deng Liu, Jili Chen, "Application of artificial intelligence system for screening multiple fundus diseases in Chinese primary healthcare settings: a real-world, multicentre and cross-sectional study of 4795 cases" (2023) Br J Ophthalmol at 1.
[42] Gu et al, above n 41, at 4.
[43] At 4-5.
[44] At 6.

Google and Stanford Healthcare

Advances in machine learning and deep learning techniques offer many exciting opportunities for medicine: greater accuracy, earlier detection of disease, as well as the financial savings to be gained from greater efficiencies. The collaboration between Google and Stanford Health Care and Palo Alto Veterans Affairs, Northwestern Medicine, Chicago and New York University-Langone Medical Center, has resulted in the creation of a deep learning algorithm that can predict the risk of lung cancer at a performance level on par with radiologists.[45] Researchers believe that this could provide an opportunity for more patients to be screened for lung cancer and highlight the "potential for deep learning models to increase the accuracy, consistency and adoption of lung cancer screening worldwide".[46] This could lead to fewer unnecessary follow-up procedures and fewer missed cancers.[47] The deep learning techniques applied in this study are also relevant to other types of 3D imaging data[48] and the results of this research bring automated image evaluation for lung cancer risk malignancy estimation a step closer to being implemented clinically.[49]

Bias in datasets

The use of big data analytics to screen for diseases and predict the risk of individuals developing certain conditions, is promising. However, shortcomings include algorithms trained on non-representative datasets and the under-representation of minority groups in clinical trials. Machine learning algorithms trained predominantly on lighter skin tones for example, may not perform as well at identifying melanoma on darker skin.[50] Researchers need to be mindful of this and take steps to compensate for ethnicity (or ancestry), gender, and other biases in their data to avoid biased results.[51]

---

[45] Diego Ardila, Atilla P. Kiraly, Sujeeth Bharadwaj, Bokyung Choi, Joshua J. Reicher, Lily Pera, Daniel Tse, Mozziyar Etemadi, Wenxing Ye, Greg Corrado, David P. Naidich, Shravya Shetty “End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography” (2019) Nature Medicine 25 954-961.
[46] At 954.
[47] At 958.
[48] At 959.
[49] At 960.
[50] Ademole S. Adamson and Avery Smith "Machine Learning and Health Care Disparities in Dermatology" (2018) JAMA Dermatology 154 11 at 1247.
[51] Benjamin Harris “AI’s future in healthcare is not entirely rosy” Healthcare IT News (23 August 2019) Healthcare IT News https://www.healthcareitnews.com/news/ai-s-future-healthcare-not-entirely-rosy

Having a representative testing (or training) dataset will not always remedy bias. Commercially available pulse oximeters overestimate oxygen saturation measurements for Black patients, reducing their likelihood of receiving necessary supplemental oxygen.[52] The U.S. Food and Drug Administration requires that pulse oximeters are calibrated on a cohort that is representative of the U.S. population.[53] If they were calibrated against a predominantly dark-skin cohort, "the reverse bias would occur".[54] In this instance, the problem cannot be fixed by changing the sampling requirements for calibration. Promisingly, a recent study has shown that developing a pulse oximeter that uses a "narrow spectral bandwidth light source" could virtually eliminate the existing skin pigment bias.[55] Forehead thermometers, and tools for locating veins, may also be inaccurate for darker-skin toned patients.[56] Both of these devices were tested on cohorts consisting primarily of men of European descent.[57]

Risk prediction models for psychiatric illness

From 2015, big data analytics were gaining momentum in psychiatric research. It was hoped that big data could provide predictive models for use in clinical practice and public health.[58] Ronald Kessler, Professor of Health Care Policy, Harvard Medical School, researches the social determinants of mental health and illness. In *Personalized Psychiatry* published in 2019, Kessler et al review the limitations of machine learning models to predict suicidality.[59] Various ways of improving the accuracy of the machine learning models are suggested, however, Kessler et al acknowledge that even if all the improvements are adopted "it is almost certain that prediction accuracy would be

[52] Marie V. Plaisime "Invited Commentary: Undiagnosed and Undertreated - the Suffocating Consequences of the Use of Racially Biased Medical Devices During the COVID-19 Pandemic" (2023) American Journal of Epidemiology 192 5 at 715.
[53] Mark S. Rea and Andrew Bierman "Light source spectra are the likely cause of systemic bias in pulse oximeter readings for individuals with darker skin pigmentation" (2023) British Journal of Anaesthesia" 131 4 at 102.
[54] At 102.
[55] At 102.
[56] Plaisime, above n 52, at 717.
[57] At 717.
[58] Ives Cavalcante Passos, Benson Mwangi, Flavio Kapczinski “Big data analytics and machine learning: 2015 and beyond” (2016) 3 Lancet 13-15.
[59] Ronald Kessler, Samantha L. Bernecker, Robert M. Bossarte, Alex R. Luedtke, John F. McCarthy, Matthew K. Nock, Wilfred R. Pigeon, Maria V. Petukhova, Ekaterina Sadikova, Tyler J. VanderWede, Kelly L. Zuromski, Alan M. Zaslavsky “The Role of Big Data Analytics in Predicting Suicide” in Passos I, Mwangi B, Kapczinski F (eds) *Personalized Psychiatry* (2019) Springer, Cham 77-98.

insufficient to allow treatment planning to be based on such a model".[60] Despite this, the authors are critical of those clinicians who dismiss the potential of machine learning models on that basis. There is an ongoing debate in psychiatry about the use of structured risk assessment tools altogether (whether machine learning or not) with some clinicians believing that in-depth clinical evaluation of needs is what is needed for effective treatment planning.[61] Kessler et al maintain that "it makes much more sense to see this phase of machine learning analysis as a useful first step in a multi-step process of need and risk evaluation".[62] They believe that, with successive refinements, machine learning models combined with clinical evaluations, are the answer to improving the detection and treatment of suicidal patients.[63]

Machine learning models have proven helpful in identifying risk factors for suicidality in U.S. soldiers transitioning back to civilian life.[64] Findings from complex machine learning models have been analysed to produce a simple risk calculator based on data inputs known prior to a soldier leaving active service.[65] Consequently, intensive case management intervention can be directed to transitioning soldiers predicted to be at high-risk of suicide. A pilot study is underway to evaluate the effects of targeting high-risk soldiers with intensive intervention and future work will examine the accuracy of the model and look at ways to refine it.[66]

Surveillance of infectious diseases: mapping malaria distribution

Big data is also contributing to the area of infectious diseases – a major target of public health measures. In the case of malaria, for example, understanding its geographical distribution is critical to combating the disease. Clinical incidence data can be combined with other data sets to map patterns of *Plasmodium vivax* and *Plasmodium falciparum*

[60] At 91.
[61] At 80.
[62] At 91.
[63] At 92.
[64] Jaclyn C. Kearns, Emily R. Edwards, Erin P. Finley, Joseph C. Geraci, Sarah M. Gildea, Marianne Goodman, Irving Hwang, Chris J. Kennedy, Andrew J. King, Alex Luedtke, Brian P. Marx, Maria V. Petukhova, Nancy A. Sampson, Richard W. Seim, Ian H. Stanley, Murray B. Stein, Robert J. Ursano and Ronald C. Kessler "A practical risk calculator for suicidal behavior among transitioning U.S. Army soldiers: results from the Study to Assess Risk and Resilience in Servicemembers-Longitudinal Study (STARRS-LS)" (2023) Psychological Medicine, Cambridge University Press.
[65] At 7098.
[66] At 7104.

parasites, which cause the bulk of the world's malaria burden.[67] The maps provide a valuable evidence-based resource for informing global policy on malaria control, enabling more efficient targeting of high-needs areas with malaria control programmes.[68] Although substantial progress has been made in reducing global rates of malaria since 2000, these maps suggest that there are still some areas where the malaria burden has plateaued or increased over the last five years, indicating that efforts towards elimination of malaria need to be strengthened in those areas.[69]

Significantly, the movements of individuals infected with malaria parasite *Plasmodium falciparum* can increase the spread of the parasite further than mosquitoes alone.[70] Data derived from mobile phones has been useful in tracking and analysing the impact of human movement on malaria transmission.[71] Malaria control programmes need to account for the spread of malaria parasites through human travel to avoid undermining local control and elimination strategies. A Harvard study used mobile phone data to map the travel patterns of nearly 15 million individuals in Kenya over a one-year period.[72] This data was combined with a transmission model of malaria infection prevalence data to map routes of parasite dispersal.[73] Networks of parasite movements within Kenya were used to identify those areas that were primarily "source" areas of parasites or "sink" areas – i.e. those areas that were net receivers of parasites.[74] Mobile phone data was a useful resource for informing control programmes targeting large volumes of human traffic between regions, incorporating measures such as advice to travellers, restricting travel, and conducting surveillance in high-risk areas.[75]

[67] Simon I Hay and Peter W Gething et al. "Mapping the global prevalence, incidence and mortality of *Plasmodium falciparum,* 2000-17: a special and temporal modelling study" (2019) 394 Lancet 322-331.
[68] At 329.
[69] At 329-330.
[70] Amy Wesolowski, Nathan Eagle, Andrew J. Tatem, David L. Smith, Abdisalan M. Noor, Robert W. Snow and Caroline O. Buckee "Quantifying the impact of human mobility on malaria" (2012) 338 Science 267.
[71] At 268.
[72] At 268.
[73] At 268.
[74] At 268-269.
[75] At 270.

Limitations of big data in understanding and accommodating social context

West African Ebola epidemic

Despite some positive outcomes, not all big data initiatives in epidemiology have been lauded a success. The use of unconventional or 'non-health' sources of data to infer health information is not without its pitfalls; one example being the use of big data during the West African Ebola epidemic of 2014-2016. Professor of Global Health at Simon Fraser University, Canada, Susan Erikson, argues that big data's utility was over-hyped during this outbreak.[76] In early 2014, big data was credited with detecting the outbreak of Ebola before public health authorities were aware of the problem.[77] As the epidemic developed it was believed that big data was able to help stop the spread of the disease by tracking the movements of contagious individuals by utilising the data from millions of mobile phones in the region; applying the same model used in the Harvard malaria study[78] (discussed above). However, Erikson argues that the Harvard model, while useful in finding links between human mobility and the spread of malaria parasites, was unsuitable for Ebola detection:[79]

> Estimated location and prevalence – rather than specific people and incidence – inform the Harvard malaria mobility study. But estimations are not good enough to contain Ebola. The approximate-able ecologies of the malaria findings in the Harvard study model do not translate well to Ebola containment. With Ebola, the contact tracing that brought an end to Ebola relies on the **exact** person in an **exact** location with the disease. Contact tracing is the door-to-door public health process to identify and then diagnose people who potentially came into contact with somebody already sick with a contagious disease.

---

[76] Susan L. Erikson "Cell Phones as an Anticipatory Technology: Behind the Hype of Big Data for Ebola Detection and Containment" (2018) Working Papers of the Priority Programme 1448 of the German Research Foundation Adaption and Creativity in Africa: technologies and significations in the making of order and disorder Nr.24 at 5.
[77] Associated Press "Healthmap software flagged Ebola 9 days before outbreak announced" (10 August 2014) https://www.cbc.ca/news/health/healthmap-software-flagged-ebola-9-days-before-outbreak-announced-1.2732464; Associated Press "Online Tool Nailed Ebola Epidemic" (9 August 2014) https://www.politico.com/story/2014/08/healthmap-ebola-outbreak-109881 as cited in Erikson, above n 40, at 10.
[78] Erikson, above n 76, at 3.
[79] At 8.

This is a salient reminder that an algorithm or model designed for a particular purpose should not be used for another, without sufficient consideration of the differences in context.[80] Ebola and malaria are fundamentally different diseases. Ebola is a virus that is spread from human to human, whereas malaria parasites are predominantly spread by mosquito (but can also be passed from human to human). Most people with malaria do not die from the infection, but more than half of those who contract Ebola die.[81] Erikson explains how big data's failure to contain the West African Ebola outbreak in 2014 was also due to epidemiologists' lack of understanding of social context, more specifically, how mobile phones are used in Sierra Leone.[82] Mobile phones did not act as accurate indicators of an individual's movements, enabling tracking and surveillance of infected individuals, because people in Sierra Leone often have multiple phones and sim cards that are frequently exchanged and passed from individual to individual.[83]

Not only did this use of big data fail to deliver public health gains, Erikson believes that the excitement of big data distracted "from the urgent need to develop care infrastructures in the short term, and a functioning health care system in the longer term".[84] Erikson argues that big data 'fixes' have too much influence on global public health decision-making. She states that "there is a prevailing belief that technology will provide answers to complex postcolonial postconflict health care challenges hundreds of years in the making".[85]

---

[80] An example of the problems that can arise when algorithms are used for a purpose other than that for which they were designed is the controversial use of risk assessment models (designed for use in correctional facilities) by judges in sentencing decisions. See Danielle Kehl, Priscilla Guo and Samuel Kessler *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing* (2017) at 13-14. https://dash.harvard.edu/bitstream/handle/1/33746041/2017-07_responsivecommunities_2.pdf?sequence=1&isAllowed=y

[81] Erikson, above n 76, at 8; The World Health Organisation estimates the fatality rate to be around 50% with fatality rates varying from 25% to 90% https://www.who.int/health-topics/ebola/#tab=tab_1

[82] Note that the practice of sharing phones and/or having multiple SIM cards is not limited to people in Sierra Leone. People in Kenya also share phones and/or have multiple SIM cards (see Amy Wesolowski, Nathan Eagle, Abdisalan Noor, Robert Snow, Caroline Buckee "Heterogenous Mobile Phone Ownership and Usage Patterns. in Kenya" (2012) PLoS ONE 7(4)). The pertinent factor is the difference in transmissibility and lethality of the two diseases. With Malaria, approximating the location of individuals through mobile phone records was sufficient, but this was not the case with Ebola which demands much more precision in contact tracing.

[83] Erikson, above n 76, at 3-4.

[84] At 12.

[85] At 12.

Additionally, Erikson argues that during the Ebola outbreak public health officials had a better understanding of the situation without big data. She cites evidence showing that anthropological scholarship was generated promptly and provided much needed insights for containing Ebola.[86] What was needed was resources to implement established and proven strategies for containing the outbreak[87], not the assumption that a big data solution that worked in one context could be applied to another, quite different, one and produce the same outcome.

Google Flu Trends

In addition to the Ebola experience, the failure of the Google Flu Trends algorithm is a stark reminder that infectious disease tracking techniques that utilise big data may complement traditional epidemiological surveillance networks, but are not an adequate substitute for them.[88] Traditionally, flu surveillance data is derived from reports of doctors on the number of cases of patients with flu-like symptoms. This estimate is then calibrated by testing a subset of people with the same symptoms to determine how many of them have the flu (and not some other flu-like illness). The Google Flu Trends algorithm, launched in the U.S. in November 2008, was designed to estimate the distribution and extent of the flu from flu-related internet searches. The idea was that when people were sick, they would search for 'flu' or 'influenza' or google their symptoms. The increase in these types of searches would imply a corresponding increase in cases of the flu. However, this inference proved false.

In addition to missing the onset of the H1N1 pandemic in April 2009, Google Flu Trends failed to predict the early outbreak of influenza in early 2013. An article in *Nature* published shortly after the influenza outbreak reported that the "glitch" was described by

---

[86] At 13.

[87] Isolation of suspected cases of Ebola, contact-tracing with monitoring and travel restrictions, sanitary funeral and burial practices, and raising awareness in communities, have a proven track record of containing Ebola outbreaks. See World Health Organisation "Statement on the 1st Meeting of the IHR Emergency Committee on the 2014 Ebola Outbreak in West Africa" (8 August 2014). https://www.who.int/mediacentre/news/statements/2014/ebola-20140808/en/; Annette Rid and Ezekiel J. Emmanuel "Why Should High-Income Countries Help Combat Ebola?" (2014) 312 JAMA 13 1297-1298; Abhishek Pandey and Katherine E. Atkins et al "Strategies for Containing Ebola in West Africa" (2014) 346 Science 6212 991-995.

[88] Declan Butler "When Google got flu wrong" (2013) 494 Nature 155-156 at 155; David Lazer, Ryan Kennedy, Gary King, Alessandro Vespignani "The Parable of Google Flu: Traps in Big Data Analysis" (2014) 343 Science 1203-1205.

experts as a temporary setback, and that Google was sure to tweak its algorithms in response.[89] However, Google later abandoned its algorithm in 2015.[90] The algorithm had been trained against seasonal fluctuations of the flu. Ordinarily, most people searching for flu symptoms and related topics are unwell, but in the wake of widespread media reporting, the flu became a newsworthy topic in itself. As a result, many more people started searching for flu-related information, throwing the algorithm off.[91] The variability, and sometimes unpredictable nature, of human behaviour is difficult to account for in designing a surveillance network that incorporates big data from non-traditional sources such as search engines and social media, resulting in potentially inaccurate predictions.

Technical and methodological challenges of big data in health

The technical and methodological challenges associated with big data raise questions about its utility in the context of health. The use of a variety of datasets from different sources (including those outside of a healthcare context) to glean meaningful information, is problematic for a variety of reasons. Whether the issues raised can be solved by more advanced big data analytics or a more discerning use of data sources, is uncertain.

Noise and inaccuracies

Data that is not collected in the context of providing a health or disability service can lack accuracy and structure.[92] Dr Wendy Lipworth, a bioethicist at Macquarie University, Sydney, sets out many of the methodological concerns of big data research in her co-authored article *Ethics and Epistemology in Big Data Research.*[93] One particular concern that Lipworth et al raise is that of "noise" or vast amounts of irrelevant information. They state that "minor noise due to errors or low quality information can easily be translated into false signals. Analysing rubbish or incommensurable datapoints may not yield useful inferences".[94]

[89] Butler, above n 88.
[90] Michael Eisenstein "Infection forecasts powered by big data" (2018) 555 Nature.
[91] Butler, above n 88; Eisenstein, above n 90.
[92] Wendy Lipworth, Paul H. Mason, Ian Kerridge, John P. A. Ioannidis "Ethics and Epistomology in Big Data Research" (2017) 14 Journal of Bioethical Inquiry 489-500.
[93] At 493.
[94] At 494.

Large data sets from online sources are often unreliable: the errors and omissions they contain are compounded when multiple data sets are used together.[95] Even health data that is collected in the context of providing a health or disability service can be a mess.[96] The data in electronic medical records is inevitably incomplete.[97] Hannah Fry notes that inaccuracies, handwritten and indecipherable medical records, nuances of language like figure of speech and hyperbole for example, mean that this data is very difficult to analyse digitally. Fry states that "medicine is really, really complicated, and every single layer of complexity makes the data a little less penetrable for a machine".[98]

Interoperability

There are inherent difficulties in comparing and combining different datasets from different sources, and consequently in identifying and linking individuals across those datasets.[99] Data interoperability is a considerable hurdle to overcome in the context of big data health applications. It can be very difficult to link an individual's health information across different datasets that vary in structure. For example, linking an individual's A&E records held by a hospital to those held by their GP and pharmacy, all of which run on different systems, and store and record data in different formats.

Confounding by indication

Confounding by indication[100] is particularly problematic for big data researchers who do not start out with a clear hypothesis, but instead rely on inferences drawn from the data. This method is the opposite of traditional scientific inquiry which starts with a research question, and then collects the necessary data. Causal inference studies utilising big data, such as electronic health records, can be prone to the same confounding biases as in

[95] Danah Boyd and Kate Crawford “Critical Questions for Big Data” (2012) Information, Communication & Society 662-679 at 668.
[96] Fry, above n 15, at 103.
[97] Choong Ho Lee and Hyung-Jin Yoon "Medical big data: promise and challenges" (2017) Kidney Research and Clinical Practice 36:3-11 at 7.
[98] Fry, above n 15, at 103-104.
[99] Lipworth et al, above n 92, at 493.
[100] Confounding by indication is “a distortion that modifies an association between an exposure and an outcome, caused by the presence of an indication for the exposure that is the true cause of the outcome” (University of Oxford https://catalogofbias.org/biases/confounding-by-indication/).

observational studies.[101] Confounding occurs when outcomes and treatment decisions share common causes. For example, when mild cases of a condition are treated differently to severe cases and have better outcomes, irrespective of the treatment received.[102] Additionally, associations between risk factors and disease identified in observational studies are not always replicated in randomized trials.[103] Confounding is not unique to big data, but it is important to acknowledge that the sheer size of big data does not compensate for this limitation.[104]

Replication and peer review

Big data analytics and machine learning can be very difficult, or impossible, to replicate unlike traditional scientific inquiry which allows for peer review. The complexity of big data analytics limits the ability of the public, and experts, to understand or challenge research findings.[105]

Curse of dimensionality

The 'curse of dimensionality'[106] refers to the various issues that arise when analysing high dimensional data. As dimensionality (characteristics of the data) increases, the volume of space increases significantly, and consequently the available data becomes sparse. This makes it difficult to draw correlations and to obtain a statistically sound result. Reducing dimensionality may cause a loss of key information. Hence there becomes a tradeoff between having lower false positive rates and identifying novel insights.[107]

---

[101] Kathryn Rough and John T. Thompson "When Does Size Matter? Promises, Pitfalls, and Appropriate Interpretation of "Big" Medical Records Data" (2018) American Academy of Opthalmology at 1137.
[102] At 1137.
[103] Michael V. Boland "Big Data, Big Challenges" (2016) American Academy of Opthalmology 123 1 at 7.
[104] Rough and Thompson, above n 101, at 1138.
[105] Lipworth et al, above n 92, at 494.
[106] The term "curse of dimensionality" was coined by Richard Ernest Bellman, an American applied mathematician famous for dynamic programming.
[107] Lee and Yoon, above n 97, at 8.

Challenging traditional conceptions of knowledge

The advent of big data has challenged traditional conceptions of knowledge. Some view the advent of big data as the 'end of theory'.[108] A feature of big data analytics is finding associations in vast and complex data. Big data analytics is focused on finding patterns or correlations not causal inference.[109] The idea is that the data speak for themselves – the meaning is in the data, just waiting to be discovered.[110] Chris Anderson, editor-in-chief of *Wired* states that "faced with massive data, [the traditional] approach to science — hypothesize, model, test — is becoming obsolete".[111]

The misconception that big data is objective is problematic and is exacerbated by the fact that big data research is largely observational, as opposed to experimental. Lipworth et al argue that:[112]

> Crucially, despite hopes to the contrary, the deficiencies of observational studies (e.g. confounding by indication) do not get eliminated with big data, and in fact they may be compounded by the volume and often suboptimal quality of the information.

Social scientists and media studies scholars, Danah Boyd and Kate Crawford, describe how all researchers, in seeking to understand what data means, are involved in a process of interpretation:[113]

> For example, in the case of social media data, there is a 'data cleaning' process: making decisions about what attributes and variables will be counted, and which will be ignored. This process is inherently subjective.

Boyd and Crawford explain that bigger does not always make for better, more objective, data:[114]

---

[108] Chris Anderson "The end of theory: the data deluge makes the scientific method obsolete" (2008) www.wired.com/2008/06/pb-theory
[109] Lee and Yoon, above n 97, at 6.
[110] Brent Daniel Mittelstadt and Luciano Floridi "The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts" (2016) 22 Sci Eng Ethics 303-341 at 320.
[111] Anderson, above n 108.
[112] Lipworth et al, above n 92, at 494.
[113] Boyd and Crawford, above n 95, at 667.
[114] At 668.

> A data set may have many millions of pieces of data, but this does not mean that it is random or representative. To make statistical claims about a data set, we need to know where data is coming from; it is similarly important to know and account for the weaknesses in that data. Furthermore, researchers must be able to account for the interpretation of the data. To do so requires recognising that one's identity and perspective informs one's analysis.

There are methodological limits to the potential of big data to revolutionise healthcare. As I have highlighted with the examples of big data health research in this Chapter, the literature predominantly describes big data's potential and the knowledge to be gained from its use. However, for the most part, the final translation of big data discoveries into actual clinical practice is limited. It is possible that many of the technical challenges of big data in health research could be resolved through advances in technology. However, Lipworth et al believe that the "issues are more fundamental – revealing epistemological misconceptions and raising questions about the very possibility of big data research achieving its goals".[115]

IBM Watson

IBM's promise that its supercomputer, Watson, would become an omniscient AI doctor is a prime example of big data failing to live up to its hype. In 2011, Watson thoroughly defeated two human champions at the American game show *Jeopardy!* Shortly afterwards, IBM announced that Watson would become an AI doctor, its first services to be available commercially within 18 to 24 months.[116] However, its achievements have fallen well short of its promise: "IBM has discovered that its powerful technology is no match for the messy reality of today's healthcare system".[117] Unstructured medical data such as doctors' notes and hospital discharge records are not easily understood by AI. While Watson can sift through massive amounts of data and medical journals, "it proved impossible to teach Watson to read articles the way a doctor would".[118]

---

[115] Lipworth et al, above n 92, at 494.
[116] Eliza Strickland "How IBM Watson Overpromised and Underdelivered on AI Healthcare" (2019) IEEE Spectrum https://spectrum.ieee.org/how-ibm-watson-overpromised-and-underdelivered-on-ai-health-care#toggle-gdpr
[117] Strickland, above n 116.
[118] Strickland, above n 116.

Medical information is complex, and it is very difficult to translate a medical practitioner's expertise into code. Lab results and measurements of vital statistics are the type of data easily understood by Watson, however unstructured written information incorporating acronyms and shorthand is not so easily understood.[119] The structured and unambiguous nature of genomic information led to Watson showing some promise in this field, identifying potentially important mutations missed by doctors for 32 per cent of patients in an unpublished study at the University of North Carolina in 2017.[120] However, due to enduring methodological issues, Watson for Genomics was discontinued in 2020.[121]

IBM's biggest failure to deliver has been in cancer care. IBM claimed that Watson for Oncology could process massive amounts of data to generate new insights and approaches to the treatment of cancer.[122] However, after investigating this claim and interviewing dozens of doctors, IBM executives, and AI experts, Casey Ross and Ike Swetlitz of STAT news found that "the system doesn't create new knowledge and is artificially intelligent only in the most rudimentary sense of the term".[123] It doesn't come up with its own insights - doctors input their own recommendations into Watson.

While it is theoretically possible to build a machine that can diagnose disease, it is exceptionally difficult. Cancer-spotting algorithms have the advantage of being able to view the actual cells that might be causing a problem whereas a machine such as Watson is several steps removed from the underlying issue.[124] Fry states that:[125]

> ... an all-seeing diagnostic machine like Watson would need to understand virtually every possible disease. That would require an army of incredibly highly qualified human handlers prepared to feed it information about different patients and their specific characteristics for a very long time.

[119] Strickland, above n 116.
[120] Strickland, above n 116.
[121] Advisory Board, Daily Briefing "10 years ago, IBM's Watson threatened to disrupt healthcare. What happened?" (July 2021) advisory.com https://www.advisory.com/daily-briefing/2021/07/21/ibm-watson
[122] Advisory Board, above n 121.
[123] Casey Ross and Ike Swetlitz "IBM pitched its Watson supercomputer as a revolution in cancer care. It's nowhere close" (2017) STAT news https://www.statnews.com/2017/09/05/watson-ibm-cancer/
[124] Fry, above n 15, at 100.
[125] At 101.

Ross and Swetlitz found that doctors' reliance on Watson varies.[126] Those hospitals that have few oncology specialists rely heavily on Watson's recommendations, while others with greater human expertise relegate it to a background role.[127] IBM's Watson has not been critically and independently reviewed by scientists outside of IBM. Its own studies, and those of paying customers, show positive results - Watson saves doctors time and has high levels of concurrence with practitioners' own treatment recommendations.[128] However, hospitals in Denmark and the Netherlands have not adopted Watson as a diagnostic tool because Watson is fed diagnostic and treatment decisions from a select group of American doctors. It is slanted towards their preferences with associated race, gender, and class biases, and possibly also the medico-legal environment in which United States' clinicians operate. Watson has been criticised for placing a lot of emphasis on American studies at the expense of other, equally significant, international studies.[129]

Watson's failure to live up to its hype may be obscuring its actual achievements in healthcare. The Advisory Board, a group of experts that identify the top challenges that healthcare leaders face, stated in July 2021 that IBM was "promoting Watson primarily as a collection of software tools that streamline and automate tasks such as accounting, payments, technology operations, marketing, and customer service"[130] and achieving commercial success in that field. IBM sold Watson Health in January 2022 to a global investment firm.[131] Watson's contribution as an AI assistant would be commendable had IBM not promised a revolution in cancer care.

Precision medicine

Precision medicine has also been criticised for failing to deliver on its hype.[132] Precision medicine promises to drastically improve human health by taking into account individual differences in genes, environment, and lifestyle in the prediction, diagnosis, and treatment

[126] Ross and Swetlitz, above n 123.
[127] Ross and Swetlitz, above n 123.
[128] Ross and Swetlitz, above n 123.
[129] Ross and Swetlitz, above n 123.
[130] Advisory Board, above n 121.
[131] IBM newsroom "Francisco Partners to Acquire IBM's Healthcare Data and Analytics Assets" (21 January 2022) https://newsroom.ibm.com/2022-01-21-Francisco-Partners-to-Acquire-IBMs-Healthcare-Data-and-Analytics-Assets?mhsrc=ibmsearch_a&mhq=Watson%20health
[132] Michael J. Joyner and Nigel Paneth “Promises, promises, and precision medicine” (2019) 129 Journal of Clinical Investigation 3 946-948.

of disease.[133] Precision medicine in the form of DNA sequencing has improved the clinical evaluation of many patients with rare diseases, however there are limited clinical interventions making use of this new information.[134] One successful intervention has been the impact of the drug Ivacaftor on cystic fibrosis patients. It is the first drug that treats the underlying cause, rather than the symptoms, of the disease caused by specific mutations and is heralded as a shining example of the potential of precision medicine.[135]

However, the treatment of cancer is, arguably, where precision medicine has had the most impact. For example, the association between breast cancer and germline mutations in BRCA1 and BRCA2 tumour suppressor genes has enabled the lifetime risk of developing cancer to be calculated.[136] Individuals identified with these mutations can elect to have increased cancer screening or bilateral prophylactic mastectomy, thus decreasing their chances of developing cancer. Nevertheless, only a minority of cancer patients currently benefit from targeted cancer therapies.[137] The task of translating the knowledge gained from genomics into long-term health benefits continues to be challenging.[138]

Precision medicine comes under criticism from those who believe that the focus on individually targeted treatments comes at the expense of other health initiatives that could have a greater impact on overall public health, such as measures addressing the causes of obesity and cancer.[139] Anaesthesiologist and physiologist Michael Joyner, and epidemiologist, Nigel Paneth state that:[140]

---

[133] U.S. Department of Health and Human Services, National Institutes of Health "The Promise of Precision Medicine" https://www.nih.gov/about-nih/what-we-do/nih-turning-discovery-into-health/promise-precision-medicine#:~:text=Precision%20medicine%20is%20an%20innovative,grown%20directly%20from%20biomedical%20research.
[134] Joyner and Paneth, above n 132, at 946.
[135] Lisa B. Feng, Scott D. Grosse, Ridgely Fisk Green, Aliza K. Fink, Gregory S. Sawiki "Precision Medicine in Action: The Impact of Ivacaftor on Cystic Fibrosis-Related Hospitalizations" (2018) 37 Health Affairs 5 773-779.
[136] Ramya Ramaswami, Ronald Bayer, Sandro Galea "Precision Medicine From a Public Health Perspective" (2018) 39 Annual Review of Public Health 153-168 at 157-158.
[137] Hannah Wise and David B. Solit "Precision Oncology: Three Small Steps Forward" (2019) 35 Cancer Cell 825-826.
[138] Bruce M. Psaty "Comparison of 2 Treatment Models: Precision Medicine and Preventative Medicine" (2019) 320 JAMA 8 751-752.
[139] Ramaswami et al, above n 136.
[140] Joyner and Paneth, above n 132, at 947.

> ... nearly two decades after the first predictions of dramatic success, we find no impact of the human genome project on the population's life expectancy or any other public health measure, notwithstanding the vast resources that have been directed at genomics.

Precision medicine's failure to live up to its hype, and the underwhelming utility of Watson, provide salutary caution against having unrealistic expectations of big data applications in health. Interestingly, Associate Professor at the Institute of Ethics in AI at the University of Oxford, Carissa Véliz, argues that artificial intelligence offers the most promise in finding new antibiotics, anti-viral and anti-fungal drugs, and vaccines - areas of research that do not require the use of any personal information.[141]

Privacy versus innovation

Véliz argues that 'privacy versus innovation' is a false dichotomy.[142] She argues that it is not about choosing 'more' or 'less' privacy but the choice to adopt responsible data privacy practices. Arguably, adapting for privacy considerations can result in better systems and uses of technology, however, there may be new uses of technology that unavoidably compromise individual privacy. The question we should be asking is whether the benefits of an innovative technology outweigh the disruption to established norms of information use. How can we facilitate big data research and initiatives while at the same time promote civic, or more specifically, health, values?

The answer begins with our conception of privacy. A regulatory framework informed by a social conception of privacy would support rules that govern the disclosure and use of health information in accordance with health values. In Chapter Four, I outline Helen Nissenbaum's Framework of Contextual Integrity as a useful response to new technologies that challenge established norms of information flow.[143] Nissenbaum argues that the values of a particular social context should govern the uses of information in that context.[144]

---

[141] Carissa Véliz *Privacy Is Power: Why And How You Should Take Back Control Of Your Data* (Penguin Random House, London, 2020) at 165-167.
[142] At 158-160.
[143] Helen Nissenbaum *Privacy In Context: Technology, Policy, and the Integrity of Social Life* (2010, Stanford Law Books, California) at 3.
[144] At 165.

Of fundamental importance to healthcare is the trust between medical practitioner and patient. Data practices that seriously undermine that trust should be prohibited. New technologies, particularly those that rely on big data, can challenge social norms and, arguably, contribute to the creation of new norms of information exchange. This is one reason why we need a regulatory framework for data privacy that is flexible enough to adapt and evolve, while remaining true to enduring social values. Society's values should inform the development of new technologies, not the other way around.

In the field of healthcare there is the very real prospect that medical imaging and risk prediction technologies will open up preventative screening to underserviced populations and also result in earlier detection of disease, and consequently, more effective treatment. Although the big data revolution in healthcare remains an unfulfilled promise, the very real potential of big data applications in certain healthcare contexts should not be ignored. We need to pave the path ahead with responsible data privacy practices that respect health values and facilitate socially beneficial innovation. The hype of big data must not blind us to the very real risks of harm that I set out in detail in Chapter Three.

'The Folly of Technological Solutionism'

The hype of big data in health, and of big data generally, is problematic when it leads to politicians, public servants, and citizens buying into the idea that the progression of technological development and its adoption is inevitable; that the future is tech and will happen with or without our collective permission or input. It is risky because it can lead to decision-makers applying technology to problems without thinking through the nature of the problem and the ramifications of applying a technological solution. "Technological solutionism" - a term coined by Evgeny Morozov - describes an ideology that legitimises and sanctions the idea that all complex social situations can be neatly defined as problems with definite technological solutions.[145] Proponents of this ideology, when presented with the availability of a new technology, ask 'what can we apply it to?' Rather than thinking through the nature of the problem (if indeed there is a problem that needs solving) and then looking for a solution from a range of available options and choosing the most

[145] Evgeny Morozov *To Save Everything Click Here: The Folly of Technological Solutionism* (Perseus Books, USA, 2013) at 5-6.

appropriate (which may very well be a technological response, but may on the other hand, be political or social). Getting the order of this process right is fundamentally important if we are to avoid technological 'solutions' that create more problems than they solve.

Governments around the world are keen to embrace technological solutions because adopting technology is viewed as progressive; it promises cost savings, accuracy, and objectivity. And when the right technology is applied in a considered manner it can deliver the promised benefits. However, as Morozov highlights, proponents of technological solutionism often reach for the answer before the question has been fully asked. How problems are conceptualised is just as important as their resolution. Many problems are not suited to a quick and easy solution and some of the 'problems' in need of solving are not actually problems at all.[146] The theme of technological solutionism continues through the second half of this Chapter in which I use the UK government's relationship with technology companies to illustrate some of the pitfalls of adopting technology too quickly, without thinking through the longer-term implications of its use.

The UK government and its relationships with technology companies

The UK government has been quick to partner with technology companies in pursuit of the efficiencies and health benefits that technology promises in healthcare. In this section I examine the arrangement between the Royal Free London NHS Foundation Trust (the Royal Free) and Google's DeepMind Technologies Ltd (DeepMind)[147] and the creation of a smartphone application (app) to detect Acute Kidney Injury (AKI). I also look at other arrangements the UK government has forged with technology companies, consider the ethical issues that these relationships raise, and ask what New Zealand might learn from the UK's experience.

[146] At 5-6.
[147] DeepMind Technologies Limited [DeepMind] is owned by Alphabet Inc., a holding company of which Google is also a subsidiary.

DeepMind / Royal Free

Acute Kidney Injury (AKI)

AKI is described by Kidney Care UK as a sudden drop in kidney function, often as a complication of another serious illness.[148] AKI affects up to one in five hospitalised patients in the United Kingdom and the United States.[149] It is difficult to detect and can have serious, and even life-threatening, consequences if not treated early enough. Approximately 100,000 deaths per year are associated with AKI in the United Kingdom, 30 per cent of which could be prevented with the right care and treatment.[150] Improvements in the detection of AKI could potentially save tens of thousands of lives.

Royal Free transfers personal health information to DeepMind without patient consent

The collaboration between DeepMind and the Royal Free to develop a smartphone app to detect AKI resulted in the Royal Free transferring vast amounts of identifiable personal health data to DeepMind without its patients' consent.[151] The data consisted of every patient admission, discharge, and transfer within the constituent hospitals of the Royal Free for over a period of more than five years.[152] The patient data included health information unrelated to AKI and its causes, such as personal information about broken bones, drug overdoses, and abortions, as well as routine administrative data on patients who presented at the Royal Free hospitals.

The disclosure of very large amounts of personal health information to a conglomerate that includes the world's leading data miner with a history of engaging in unfair trade practices[153] gives cause for unease. There are two main concerns. First, if Google obtains

[148] Kidney Care UK https://www.kidneycareuk.org/news-and-campaigns/facts-and-stats/
[149] Mustafa Suleyman and Dominic King "Using AI to give doctors a 48-hour head start on life-threatening illness" (31 July 2019) DeepMind https://deepmind.com/blog/article/predicting-patient-deterioration
[150] NHS England "Factsheet: Implementation of NICE guideline on Acute Kidney Injury (AKI)" https://www.england.nhs.uk/wp-content/uploads/2014/02/rm-fs-10-4.pdf
[151] Julia Powles and Hal Hodson "Google DeepMind and healthcare in an age of algorithms" (2017) 7 Health Technology 351-357.
[152] At 353.
[153] In 2017 the European Union fined Google / Alphabet 2.42 billion euros for positioning and displaying its own comparison shopping service more favourably than competing shopping services. See European Commission *Prohibition Decision (Art.102 Ex 82)* (27 June 2017) https://ec.europa.eu/competition/elojade/isef/case_details.cfm?proc_code=1_39740; New rules agreed by

this information no-one will really know what it's doing with it. Can DeepMind / Google be trusted to use this data ethically, or even lawfully? The second concern is more immediate and pragmatic – if Google gains control over the health analytics industry, how can it be prevented from leveraging its dominant market position over the NHS and other national health systems?

Google and the Royal Free – contrasting histories and objectives

Google

Google is the world's largest search engine, dominating the market with over 90 per cent of the search engine market share.[154] Founded in 1996 by Larry Page and Sergei Brin, it began as a research project when both of its founders were PhD students at Stanford University, California. Initially called 'BackRub' the search engine was designed to measure the importance of a website by counting its backlinks[155] – the link one website gets from another website. Its creators never intended to start a company but their brilliant innovation, in the form of Google, has grown to become one of the world's most valuable brands.[156] Google states that its mission is "to organize the world's information and to make it universally accessible and useful".[157] Google is the quintessential surveillance capitalist. Its market model depends on obtaining access to new sources of personal information in order to offer increasingly personalised and targeted advertising spaces to its customers. Over the last 20 years Google has spent tens of billions of dollars on

EU negotiators in February 2019 mean that technology companies will have to disclose how they rank their own and rival products on their platforms. See Foo Yun Chee "Google, Amazon among those targeted in EU unfair practices digital rules" (14 February 2019) https://www.reuters.com/article/us-eu-tech/google-amazon-among-those-targeted-in-eu-unfair-practices-digital-rules-idUSKCN1Q30ZY. In April 2019, the Federal Trade Commission announced an antitrust investigation of Google. See Consumer Watchdog for more information on how Google abuses its dominant position in the marketplace in the US: https://www.consumerwatchdog.org/blog/ftc-targets-google-antitrust-investigation

[154] StatCounter Global Stats (August 2019) https://gs.statcounter.com/search-engine-market-share

[155] Newton Lee *Google It: Total Information Awareness* (Springer, Burbank USA, 2016) at 3.

[156] In 2024, Google is the world's third most valuable brand. https://www.visualcapitalist.com/most-valuable-brands-in-2024/. In 2018, Google was the world's most valuable brand. For the twelve years prior to 2018, it held the title jointly with Apple. Lucy Handley "Amazon beats Apple and Google to become the world's most valuable brand" (11 June 2019) CNBC https://www.cnbc.com/2019/06/11/amazon-beats-apple-and-google-to-become-the-worlds-most-valuable-brand.html

[157] Google https://about.google

purchasing new start-ups and innovative product ideas.[158] One significant acquisition was Google's purchase of DeepMind in 2014 for $500 million. DeepMind is an artificial intelligence company with an inter-disciplinary approach to AI and was founded in 2010, before there was much hype around AI's potential.[159] In 2015, Google became the main subsidiary of Alphabet Inc., a holding company created to facilitate a restructuring of Google. In 2016, DeepMind Health was launched and the Streams smartphone application was developed in collaboration with the Royal Free. Alphabet is one of the so-called 'big four' global technology companies developing and extending its business in the lucrative healthcare industry.[160] I argue that a significant motivation behind this is to obtain access to personal health data.

The Royal Free London NHS Foundation Trust

The UK's National Health Service (NHS) provides a comprehensive health service free of charge to all individuals ordinarily resident in the UK and is funded through general taxation.[161] The Royal Free is one of the UK's biggest NHS trusts with 10,000 staff and three hospitals.[162] It has a long history of caring for patients.[163] The Royal Free Hospital was one of the first hospitals in England to provide free healthcare: its founding principle is the commitment to take in anyone needing treatment regardless of their ability to pay.[164] The Royal Free states that its mission "is to be world class in terms of our innovative healthcare services, clinical research and teaching excellence".[165] Presumably, the Royal

---

[158] Matt Reynolds "If you can't build it, buy it: Google's biggest acquisitions mapped" (25 November 2017) Wired https://www.wired.co.uk/article/google-acquisitions-data-visualisation-infoporn-waze-youtube-android

[159] DeepMind https://deepmind.com

[160] Insider Intelligence "BIG TECH IN HEALTHCARE: Here's who wins and loses as Alphabet, Amazon, Apple, and Microsoft target niche sectors of healthcare" (24 January 2023) https://www.insiderintelligence.com/insights/big-tech-in-healthcare-report/; Nancy Huynh "How the 'Big 4' Tech Companies Are Leading Healthcare Innovation" (27 February 2019) Healthcare Weekly https://healthcareweekly.com/how-the-big-4-tech-companies-are-leading-healthcare-innovation/

[161] A small proportion of funding also comes from national insurance (a payroll tax). International Healthcare System Profiles https://international.commonwealthfund.org/countries/england/

[162] Barnet Hospital, Chase Farm Hospital and Royal Free Hospital. Royal Free London NHS Foundation Trust *About Us* http://s3-eu-west-1.amazonaws.com/files.royalfree.nhs.uk/AboutUs/A5_About_us_booklet_general_recruitment.pdf at 2.

[163] The Royal Free Hospital was established in 1828 by William Marsden. Marsden had recently qualified as a surgeon and was appalled when he could not find anyone willing to treat an indigent young woman. Royal Free London NHS Foundation Trust https://www.royalfree.nhs.uk/about-us/

[164] Royal Free London NHS Foundation Trust https://www.royalfree.nhs.uk/about-us/our-history/

[165] Royal Free London NHS Foundation Trust https://www.royalfree.nhs.uk/about-us/world-class-expertise-local-care/

Free's commitment to innovation motivated its clinicians to approach DeepMind in July 2015 to create an application to assist with detecting AKI.[166]

The arrangement between the Royal Free and DeepMind - a bigger goal in mind?

The Royal Free's decision to approach DeepMind, an AI company that had no experience in providing healthcare services, to build a relatively straightforward app, raised questions among many people concerned about patient privacy and the integrity of the NHS.[167] Julia Powles, a legal researcher at the University of Cambridge, and Hal Hodson, a reporter for *The Economist* argue that DeepMind's publicly stated purpose for holding this sensitive data (the management and direct care of AKI) was a much narrower purpose than that which contractually constrained its use of the data.[168] The first of several agreements between DeepMind and the Royal Free, an Information Sharing Agreement, was signed on 29 September 2015. It states that the information is being shared for the purpose of "Patient Rescue" – "a Proof of Concept Technology Platform that enables Analytics as a Service for NHS Hospital Trusts".[169] The analytical tools to be developed:

> ... will consist of: (i) Patient Safety Alerts for Acute Kidney Injury, and (ii) Real time clinical analytics, detection, diagnosis and decision support to support treatment and avert clinical deterioration **across a range of diagnoses and organ systems**.[170] (emphasis added)

The purpose stated in the agreement was much broader than the creation of an application to facilitate the management and treatment of AKI. It envisioned the creation of tools to manage and support the treatment of a range of conditions.

---

[166] There has been no explanation from the Royal Free as to why DeepMind (as opposed to other technology companies) was approached to build the app.
[167] Powles and Hodson, above n 151, at 352; Philip Hunter "The big health data sale" (2016) 17 Science and Society 8 1103-1105; Christof Stache "If Google has nothing to hide about NHS data, why so secretive?" (4 May 2016) New Scientist https://www.newscientist.com/article/mg23030722-900-big-data-if-theres-nothing-to-hide-why-be-secretive/
[168] Powles and Hodson, above n 151, at 352.
[169] Royal Free London NHS Foundation Trust and Google UK Limited "Information Sharing Agreement" (29 September 2015) at 3.
[170] At 3.

Lack of consultation and transparency

The Royal Free's decision to transfer patient data to DeepMind was made without consultation with the relevant public bodies such as the Information Commissioner's Office, responsible for enforcing the UK's Data Protection Act[171] and the Health Research Authority, which provides a framework for releasing confidential health information without consent (through the Confidentiality Advisory Group). Nor was the Medicines and Healthcare Products Regulatory Agency, responsible for regulating medical devices, contacted. It was only after *New Scientist* published an exposé by Hal Hodson[172] on 29 April 2016 that the public became aware of the extent of the volume of personal health data processed by DeepMind.

Information Commissioner's Office Investigation

The Royal Free was investigated by the UK Information Commissioner's Office (ICO) to determine whether it had complied with its legal responsibilities as a data controller under the Data Protection Act 1998 (UK). A 'data controller' is a person or agency that, alone or jointly with others, determines the purposes and means of the processing of personal data. A 'data processor' is a person or agency (other than an employee of the data controller) that processes personal data *on behalf of* the controller.[173] On 3 July 2017, the ICO concluded that the processing of 1.6 million patients' data by DeepMind for the purpose of the clinical safety testing of the Streams app did not comply with the requirements of the Data Protection Act.[174] The ICO also held that the relationship between the Royal Free and DeepMind was one of data controller to data processor. Thus, it did not investigate the actions of DeepMind, because as data processor it was just processing the personal data on behalf of the Royal Free. The ICO held that there were a

[171] Data Protection Act 1998 (UK), repealed by the Data Protection Act 2018 (UK).
[172] Hal Hodson "Revealed Google AI has access to huge haul of NHS patient data" (29 April 2016) New Scientist https://www.newscientist.com/article/2086454-revealed-google-ai-has-access-to-huge-haul-of-nhs-patient-data/#ixzz5xrJRhk3a
[173] The definitions of 'data controller' and 'data processor' vary slightly in wording across the original DPA 1998 (UK), its current iteration, and the General Data Protection Regulation, however the meanings remain consistent.
[174] Information Commissioner's Office "Letter to Sir David Sloman, Chief Executive, Royal Free NHS Foundation Trust" (3 July 2017).

number of shortcomings which amounted to non-compliance with the following data protection principles by the Royal Free. They included:

Principle One: personal data shall be processed fairly and lawfully

The Royal Free did not inform patients of the processing of their sensitive personal data by DeepMind, nor did it fulfil a condition of processing that would remove the need to obtain the informed consent of patients for that processing. The "data subjects were not adequately informed that the processing was taking place … as a result, the processing was neither fair nor transparent".[175] The ICO held that the information was not processed for the purpose of direct patient care (which the Royal Free had maintained was the justification for transferring the data to DeepMind). The Royal Free did not have the implied consent of patients to process their data in that way, therefore the processing was not lawful under the Data Protection Act.

Principle Three: personal data should be adequate, relevant and not excessive

The ICO was not convinced that it was necessary or proportionate to process 1.6 million partial patient records in order to test the clinical safety of the app. The processing of these records was held to be excessive and in breach of Principle Three.

Principle Six: Personal data shall be processed in accordance with the rights of data subjects

Data subjects were unaware that their data was being processed by DeepMind therefore they could not exercise their right to opt-out or prevent processing. The ICO held that the Royal Free failed to comply with Principle Six.

Principle Seven: Appropriate technical and organisational controls shall be taken – this includes the need to ensure that appropriate contractual controls are in place when a data processor is used

The ICO held that the initial Information Sharing Agreement dated 30 September 2015 did not go far enough to ensure that the processing was undertaken in accordance with the Data Protection Act. It did not ensure that only minimal data was available to

[175] At 5.

DeepMind and that processing would only be conducted for limited means. The ICO states:[176]

> I am also concerned to note that the processing of such a large volume of records containing sensitive health data was not subject to a full privacy impact assessment ahead of the project's commencement.

The ICO held that the Royal Free failed to demonstrate compliance with Principle Seven in its initial agreement with DeepMind, but notes that the Royal Free has since improved the documentation in place between it and DeepMind, and has increased the transparency of patient data use for the Streams app.

DeepMind partners with other NHS Trusts

Despite the data privacy controversy and ongoing investigation of the Royal Free by the ICO, several other NHS Foundation Trusts also partnered with DeepMind, some embarking on a five-year rollout of the Streams app.[177] Before data was transferred to DeepMind the Trusts de-identified it and obtained ethics research approval - avoiding the compliance issues raised by the Royal Free / DeepMind agreement.[178]

Advantages of the Streams app

The Streams app uses the pre-existing NHS algorithm to predict the risk of AKI in hospitalised patients at the Royal Free hospitals. It works by sending a text message to clinicians to notify them of likely cases of AKI, which affects one in five hospitalised patients and is very difficult to detect.[179] The Streams app has been credited with saving clinicians' time, reducing the average cost of admission of patients with AKI by 17 per cent, and fewer cases being missed (3.3 per cent compared to 12.4 per cent before the app was in use).[180] While the Streams app is a useful and efficient digital tool, DeepMind

[176] At 8.
[177] Imperial College London NHS Foundation Trust; University College London NHS Foundation Trust; Taunton and Somerset NHS Foundation Trust; and Yeovil NHS Foundation Trust (although Yeovil NHS Foundation Trust never actually implemented the Streams app for use in a clinical setting).
[178] Powles and Hodson, above n 151, at 356.
[179] C.Wu, Y.Zhang and S. Nie et al "Predicting in-hospital outcomes of patients with acute kidney injury" (2023) Nature Communications 14 at 1.
[180] Suleyman and King, above n 149.

discovered that it couldn't use the Royal Free patient data to develop a new algorithm for AKI prediction because of the poor quality of the data.[181]

Deep learning promised to identify at-risk patients sooner

Nevertheless, DeepMind aims to upgrade or replace the existing algorithm with one based on a database from the U.S. Department of Veterans Affairs. The algorithm will need training on much more diverse datasets (only 6.38 per cent of the data subjects are women) before it can be rolled out for use in a clinical setting.[182] The hope is that DeepMind's deep learning approach will provide opportunities to identify patients at risk of AKI within a timeframe that allows for earlier treatment.[183]

DeepMind is subsumed by Google Health

On 18 September 2019, DeepMind became part of Google Health,[184] validating concerns about DeepMind's ability to keep Royal Free patient data separate from other sources of identifiable personal data held by Google. NHS Trusts that wanted to continue to use the Streams app had to form new contracts directly with Google.

Streams app is decommissioned

Two years later, in August 2021, Google made the decision to decommission its Streams app altogether. Techcrunch reporter Natasha Lomas states:[185]

> Google is infamous for spinning up products and killing them off, often in very short order. It's an annoying enough habit when it's stuff like messaging apps and games. But the tech giant's ambitions stretch into many domains that touch human lives these days. Including, most directly, healthcare. And — it turns out — so does

[181] Fry, above n 15, at 104.
[182] Trevor Back, Christopher Nielsen, Joseph R. Ledsam, Shakir Mohamed et al "A clinically applicable approach to continuous prediction of future acute kidney injury" (2019) 572 Nature 116-119 at 118.
[183] University of Michigan Institute for Healthcare Policy and Innovation "An AI model predicting acute kidney injury works, but not without some tweaking" (20 January 2023) https://ihpi.umich.edu/news/ai-model-predicting-acute-kidney-injury-works-not-without-some-tweaking; Back et al, above n 136.
[184] Natasha Lomas "Google completes controversial takeover of DeepMind Health" (20 September 2019) Techcrunch https://techcrunch.com/2019/09/19/google-completes-controversial-takeover-of-deepmind-health/
[185] Lomas, above n 184.

> Google's tendency to kill off products that its PR has previously touted as "life saving".

Lomas argues that once DeepMind/Google realised that the Streams app would not facilitate an endless flow of patient data that could be used to create new predictive AI health tools, it lost interest in the project. She goes on to state that:[186]

> Google's decision to decommission Streams may be factoring in a lack of enthusiasm from involved Trusts to continue using the software - although if that's the case, it may, in turn, be a reflection of Trusts' perceptions of Google's weak commitment to the project.

When Google confirmed the decommission of the Streams app, the Royal Free was the only NHS Trust still using it. Its decommission meant there would no longer be security updates for the Streams app, exposing it to data security concerns if it were to continue to be used after the decommissioning.

Royal Free's failings

The Royal Free transferred the personal health data of 1.6 million patients to DeepMind, in breach of several data protection principles. It did not consult with the ICO's office or the Health Research Authority. I argue that there was a lack of integrity and transparency in the Royal Free's arrangement with DeepMind. Ultimately, very little was gained from the Streams app. It used the NHS algorithm that pre-existed its creation. It led to some modest savings of time and resources and made it easier to treat patients at risk of kidney injury. However, it did not deliver significant improvements in kidney health outcomes.[187] It was, arguably, a case of senior clinicians' eagerness to apply a technological 'solution' to a health problem without first thinking through the implications of transferring patient health data to a conglomerate that includes the world's largest data miner.

---

[186] Lomas, above n 184.

[187] This is because the app, relying on the NHS algorithm, estimates the risk of creatinine in the blood, and creatinine levels can rise hours after AKI has already set in. (Matt Reynolds "DeepMind's new AI predicts kidney injury two days before it happens" (2019) Wired https://www.wired.com/story/deepmind-streams-ai-algorithm-kidney-injury/).

We will we probably never know with any certainty how the NHS patient data has been used; if, or when, it will be deleted, or how it could potentially be used in the future. Powles and Hodson argue that once data makes its way onto Google's servers the ability to track that data and to understand how it is being used ceases.[188] Google is contractually required to delete all NHS patient data after the termination of each contract with a NHS Trust. However, Powles and Hodson make the critical point that Google "benefits from relying on commercial secrets and the absence of public law obligations and remedies against it. This leaves it with few incentives for accountability".[189]

DeepMind and Google claim to be working in the public interest, striving to improve health outcomes and save lives. But the Royal Free has been left in the lurch now that the Streams app has been decommissioned. Big technology companies have the necessary resources and expertise required to create medical AI and other technology tools, making them a tempting partner for the NHS. However, these deals must be transparent - the public needs to be a part of the debate about public/private healthcare collaborations. The costs and benefits of such arrangements over the longer term need to be fully considered. These can only be properly calculated when the public servants entering into these deals understand, not only the potential for personal information to be misused but also, the underlying power dynamic associated with tech companies holding vast amounts of information about us.

NHS COVID-19 Data Store

The UK government continues to pursue its relationships with big tech. During the Covid-19 outbreak of 2020, the UK government transferred the personal health information of millions of NHS users to Amazon, Microsoft, Google, and AI firms Faculty and Palantir to facilitate the creation of the NHS COVID-19 Data Store (Data Store). The aim of the Data Store is to provide a comprehensive dataset of information related to Covid-19. The NHS states that information from a range of sources is brought together to "provide a

[188] Powles and Hodson, above n 151, at 360.
[189] At 360.

single version of the truth" to support Covid-19 decision-making.[190] This diverse range of data sources includes, but is not limited to, aggregated 111 calls, hospital admissions, GP appointment and contract services data, information on diabetes patients, cancer treatment waiting times, deprivation index data, maternity services data, mental health and learning disabilities data, census data, as well as Google mobility reports and Apple mobility trends.[191]

At the time of its creation there were concerns that the Data Store - rushed through without public consultation - would outlast the coronavirus.[192] The British public was assured that the project was only temporary, and after the pandemic the Data Store would be closed.[193] However, initial concerns proved to be valid with the UK government signing another two-year contract with Palantir in December 2020 that went beyond the scope of Covid-19 to include Brexit and general business planning. In February 2021, openDemocracy and tech-justice firm, Foxglove, filed for judicial review of the UK government's decision to award this new contract to Palantir.[194] One month later, in March 2021, the government backed down, conceding that it cannot offer firms like Palantir a long-term NHS role without consulting the public; that it would not expand Palantir's work on the Data Store beyond the pandemic without notifying the public; and also agreeing to engage citizens about Palantir's role in the NHS through the use of patient juries.[195]

The UK government credited the Data Store with helping to monitor the spread of Covid-19 and ensuring that resources, such as ventilators, are directed to where they are most needed.[196] It is also used to direct patients to the services that are best able to care for

[190] Ming Tang, NHS blog "Data integration - driving improvements in patient care" (18 December 2020) https://www.england.nhs.uk/blog/data-integration-driving-improvements-in-patient-care/
[191] For a full list of data sources see NHS COVID-19 Data Store Reference Library: https://data.england.nhs.uk/covid-19/
[192] Paul Lewis, David Comm and David Pegg "UK government using confidential patient data in coronavirus response" (12.4.20) The Guardian https://www.theguardian.com/world/2020/apr/12/uk-government-using-confidential-patient-data-in-coronavirus-response
[193] Lewis et al, above n 192.
[194] Foxglove "Success! UK government concedes lawsuit over £23m NHS 'data deal' with controversial US tech corporation Palantir" (2021) https://www.foxglove.org.uk/2021/04/01/success-uk-government-concedes-lawsuit-over-23m-nhs-data-deal-with-controversial-us-tech-corporation-palantir/
[195] Lindsay Clarke "Under threat of judicial review, UK.gov agrees to consultation before extending Palantir's NHS role beyond pandemic" (30 March 2021) The Register https://www.theregister.com/2021/03/30/ukgov_caves_over_nhs_palantir_lawsuit/
[196] NHS COVID-19 Data Store privacy notice https://www.england.nhs.uk/contact-us/privacy-notice/how-we-use-your-information/covid-19-response/nhs-covid-19-data-store/

them, as well as identify particularly vulnerable patients.[197] NHSX - the technology arm of the NHS - states:[198]

> In a crisis, knowledge is power - and a lack of it can cost lives. Different parts of the NHS were already capturing much of the data we needed to inform our COVID-19 response. However, without a single place to gather and analyse this data, decision-makers were unable to see a complete picture of the situation or move as quickly as the response to the pandemic demanded.

Whether the Data Store was a proportionate and measured response to the data requirements of the government and NHS during the pandemic remains to be seen. The Data Store has been criticised for its lack of transparency[199] and many of the companies involved in the Data Store have a history of poor data privacy practices.[200] Questions about the efficacy and necessity of the Data Store are still to be answered. Nevertheless, there is a strong desire on the part of some politicians and public servants to keep building on the momentum of the NHS Covid-19 Data Store beyond the pandemic.[201]

General Practice Data for Planning and Research

The desire to put 'data at the centre of decision making' was evident in the UK government's push during 2021 to implement the General Practice Data for Planning and

---

[197] NHSX "The COVID-19 Data Store: putting data at the centre of decision making" https://www.nhsx.nhs.uk/key-tools-and-info/data-saves-lives/improving-health-and-care-services-for-everyone/the-nhs-covid-19-data-store-putting-data-at-the-centre-of-decision-making/

[198] NHSX, above n 197.

[199] Audrey Guinchard "Our digital footprint under Covid-19: should we fear the UK digital contact tracing app?" (15 July 2021) International Review of Law, Computers & Technology 35:1; M.J. Mourby "'Leading by Science' through Covid-19: the NHS Data Store and Automated Decision-Making" (24 February 2021) International Journal of Population Data Science 5:2:03.

[200] Prior to its involvement with the NHS COVID-19 Data Store, Google had been fined €50 million in France for GDPR violations https://www.irishtimes.com/business/technology/google-hit-with-50m-fine-for-data-privacy-breach-1.3765575; Amazon, Apple and Google all faced data privacy complaints in 2019 https://www.bbc.com/news/technology-46944694; Amazon was investigated by both the Austrian Data Protection Authority and Germany's Antitrust Watchdog in 2019.

[201] Ming Tang, National Director for Data Analytics NHS England, refers to the NHS COVID-19 Data Store and states that we must "retain this progress and continue to improve the way data is managed..." https://www.england.nhs.uk/blog/data-integration-driving-improvements-in-patient-care/; Rt Hon MP Matt Hancock, Secretary of State for Health and Social Care, states that we must "unleash the unlimited potential of data in health" and that "our shared experience of the pandemic has shown us the prize that's on offer. Now we must ...use the full power of data to deliver real solutions..." (Ministerial forward to the UK Department of Health and Social Care Policy Paper: *Data saves lives: reshaping health and social care with data (draft)* updated 8 September 2021).

Research (which has the very confusing acronym for data privacy scholars of G**P**DPR). The aim of the GPDPR is to replace the 300 collections of GP data with one collection; creating one massive database of identifiable patient health information for research and the more efficient allocation of resources.[202] Initially, it also allowed private corporations access to data where there was "a benefit to health and social care".[203]

In July 2021, the UK government was pressured to back down in relation to several key aspects of the GPDPR after openDemocracy, Foxglove, and other campaigners[204] threatened legal action over what they described as "the great NHS data grab".[205] Once again, openDemocracy teamed up with Foxglove, to protect the rights of patients to have some say about their personal health information being shared with private companies.

In 2014, a similar, though less ambitious project of the UK government, to upload and share GP patient records - care.data - was stalled over ethical concerns and finally abandoned in 2016.[206] In 2021, with the GPDPR, there was no public consultation. The health secretary Matt Hancock issued a direction to every GP in England, instructing them to upload their patient records to a central database. Patients were to be given just a few weeks to find out about the plans. Both the British Medical Association and the Royal College of GPs expressed concern about this, in particular, the lack of transparency and failure to engage with the public; the potential for private companies to profit from the data; and being left with the burden of doing the government's job of explaining the plan to patients.[207]

---

[202] NHS Digital "About the GPDPR programme" https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-collections/general-practice-data-for-planning-and-research/about-the-gpdpr-programme

[203] NHS "How NHS Digital Makes Decisions About Data Access" https://digital.nhs.uk/services/data-access-request-service-dars/how-nhs-digital-makes-decisions-about-data-access

[204] Just Treatment, the Doctors' Association UK, The Citizens, the National Pensioners convention, and David Davis MP.

[205] Foxglove "The government has scrapped the deadline for the NHS data grab" (22 July 2021) https://www.foxglove.org.uk/2021/07/22/the-government-has-scrapped-the-deadline-for-the-nhs-data-grab/

[206] Paraskevas Vezyridis and Stephen Timmons "Understanding the care.data conundrum: New information flows for economic growth" (2017) Big Data & Society at 2.

[207] Digitalhealth.net "BMA and RCGP issue joint letter to NHS Digital over GPDPR programme" (7 June 2021) https://www.digitalhealth.net/2021/06/bma-and-rcgp-issue-joint-letter-to-nhs-digital-over-gpdpr-programme/

The UK government's concessions regarding the GPDPR included scrapping the deadline for patients to opt out and amending the system whereby an opt-out only applied to your data being taken in the future, but your life-long health record remained in the database.[208] While these changes significantly reduce the threat of patient data being taken without consent, big concerns remain, including the issue of private companies being able to access data in the GPDPR. The project is currently on hold, with no start date for the collection of GP patient data.[209]

NHS Federated Data Platform

The concern about private companies being able to access patient data continues with Palantir being awarded the contract for the NHS Federated Data Platform in November 2023. NHS England states that the Federated Data Platform will improve and connect health information, resulting in better provision of health services.[210] The UK government has said the Federated Data Platform will only share data with private companies if it is in a de-identified form.[211] Foxglove argues that this is not good enough given how easily de-identified data can be re-identified.[212] Polling suggests that 48 per cent of NHS patients in England will opt-out of the NHS Federated Data Platform if it is opened up to private companies.[213] This is a significant proportion of NHS patients and indicates the level of the UK public's aversion to private corporations' access to NHS patient health data. Foxglove, Doctors' Association UK, National Pensioners Convention and patient organisation, Just Treatment, will challenge the legal basis for the Federated Data Platform on the grounds that it lacks the public's consent to expand access to confidential

[208] Foxglove, above n 205.
[209] NHS Digital "About the GPDPR programme" https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-collections/general-practice-data-for-planning-and-research/about-the-gpdpr-programme
[210] NHS England "Federated data platform - improving and connecting our health information" (20 November 2023) https://www.england.nhs.uk/long-read/federated-data-platform-improving-and-connecting-our-health-information/
[211] Foxglove "Legal action launched: no legal basis for the £330 Palantir NHS Federated Data Platform" (30 November 2023) https://www.foxglove.org.uk/2023/11/30/legal-action-palantir-nhs-federated-data-platform/
[212] Foxglove, above n 211.
[213] YouGov.co.uk https://d3nkl3psvxxpe9.cloudfront.net/documents/Foxglove_NHS_Data_Sharing_230522.pdf

patient data.[214] Foxglove and Doctors' Association UK agree that the NHS is in need of a data system upgrade, which could help to get the most out of NHS data, leading to greater efficiencies and improved patient care.[215] However, they point to repeated attempts by the government to implement change that have failed because of public concerns about private sector access to patient health data.[216]

Ethical concerns raised by the arrangements between government and big tech

The UK government's relationships with technology companies raises significant ethical issues. Firstly, can technology companies be trusted with very large datasets of health information, particularly if their business model relies on generating or collecting increasing amounts of personal information? The "extraction imperative"[217] that drives some technology companies to continually look for new sources of personal information makes them high-risk partners in collaborations where they become the recipients of large, population-wide datasets of personal information. There may be contractual clauses that stipulate when data must be deleted, but as Powles and Hodson point out, once data makes its way onto Google's servers, the ability to keep track of how it is being used, ceases. It can become impossible to keep track of how technology companies use personal information.

Secondly, the longer-term implications of these arrangements must be considered. There is a lack of effective safeguards built into these projects to protect the NHS from a scenario whereby the processing of this data helps to create products and services that are then sold back to the NHS at an exorbitant rate. This raises the question of *who should be* allowed to profit from NHS, or other public health, data. Additionally, these arrangements between the NHS and technology companies involve the transfer of, and/or access to, population-wide datasets, contributing to information- and power- asymmetries between

[214] Dennis Campbell "NHS England faces lawsuit over patient privacy fears linked to new data platform" The Guardian (30 November 2023) https://www.theguardian.com/society/2023/nov/30/nhs-england-faces-lawsuit-patient-privacy-fears-new-data-fdp-platform?CMP=Share_iOSApp_Other
[215] Cori Crider "Why Palantir's latest NHS land-grab is such bad news for patients" openDemocracy (17 March 2023) https://www.opendemocracy.net/en/palantir-foundry-faster-data-flows-nhs-cori-crider-foxglove/
[216] Crider, above n 215.
[217] Shoshana Zuboff *The Age of Surveillance Capitalism: The Fight For A Human Future At The New Frontier Of Power* (Profile Books Ltd., London, 2019) at 87.

citizens and the technology companies that hold increasing amounts of information about them.

New Zealanders are not protected from the kinds of arrangements that the UK has pursued. Our health information is vulnerable to government agencies in New Zealand entering into contracts with overseas technology companies offering health-related tech services in exchange for access to datasets of health information.

New Zealand government and big tech

The New Zealand government has memorandums of understanding with both Microsoft and Amazon Web Services (AWS).[218] AWS has made significant headway in working itself into the New Zealand healthcare system; holding a key role in shifting health data to cloud computing platforms.[219] The government's cloud-first policy results in data being shifted from within agencies to hyper-scale data centres. Experts warn that when big tech companies hold your data, they then get to charge you for using it.[220] Rupert Taylor-Price, head of Australian-owned, Vault Cloud has warned:[221]

> When you look at the economics of this, when you give up your sovereignty of your data, you know, data is the new gold, you're giving up the value of your economy.
> ...
> I don't criticise the multinationals for trying to take advantage of weak leadership in a government. They should be trying to further their commercial interests.
> ...
> The sad truth is that your local industry rarely has the financial capability to make those big lobbying plays that large multinationals do. And that's not in the interest of the country.

[218] Phil Pennington "'Data is the new gold' - Warning New Zealand at risk with reliance on foreign firms" (4 July 2023) RadioNZ https://www.rnz.co.nz/news/business/493161/data-is-the-new-gold-warning-nz-at-risk-with-reliance-on-foreign-firms
[219] Phil Pennington "Revealed: Amazon's efforts to work its way into the NZ healthcare system" (6 September 2023) RadioNZ https://www.rnz.co.nz/news/national/497352/revealed-amazon-s-efforts-to-work-its-way-into-the-nz-healthcare-system
[220] Pennington, above n 218.
[221] Pennington, above n 218.

New Zealand does not need to rely on overseas technology companies for solutions to its technology and data system woes. We have local technology companies that are proving successful and capable, ironically selling most of their products and services to overseas customers.[222] The transition to locally grown and managed data solutions for government might not be easy, but I argue that we must look to the longer term when the control of our information management systems is at stake.

Lessons to be learned from the UK

There are many lessons New Zealand can learn from the UK's experience; most importantly, to be transparent and open with the public about any intention to use personal health information, and to fully consider the implications of engaging with companies that profit from the processing of personal information. There is a strong argument to be made against partnering with companies that operate as surveillance capitalists. Governments should exercise due diligence; considering the backgrounds of the companies that they contract with and thinking through the practicalities of auditing compliance with, or enforcing the terms of, a contract with an overseas-based company. If governments fail to understand the nature and value of personal information, including the ways in which it can be repurposed, they may not be able to assess whether the proposal makes business sense in the longer term. Governments should consider the future ramifications of any arrangement that involves the transfer of personal information and who stands to profit. Could a local or national company provide the same service? Does the arrangement comply with data privacy laws, and perhaps most importantly, is it ethically, socially, and culturally appropriate in the context of New Zealand?

Conclusion

Significant amounts of time, money and expertise have been invested in big data health initiatives. Yet there are few examples of big data breakthroughs being translated into actual clinical practice. The question of when, or if, it is realistic to expect to see big data transform the delivery of healthcare, is beyond the remit of this thesis. There is real

[222] NZ Trade and Enterprise "New Zealand's best tech companies are taking on the world" (25 June 2023) https://www.nzte.govt.nz/blog/new-zealands-best-tech-companies-are-taking-on-the-world

potential, particularly in medical image analysis and risk prediction tools, but the extent to which we will ultimately benefit from big data in healthcare is uncertain.

What is certain is that the unequivocal beneficiaries of big data initiatives in healthcare, to date, are the technology companies that have gained access to unprecedented amounts of personal data - data that are also increasingly valuable outside of a healthcare context. The possible future uses of this data are unknown and the potential harms of its misuse can only be surmised. While the privacy risks of big data may be speculative and uncertain, so too are the overall health benefits for those populations whose personal data is being disclosed.

Ultimately, we need new responses to the challenges of increasing commercialisation of personal data. How do we limit big tech's exploitation of personal information without relinquishing the potential benefits of big data research? I argue that we start by shifting our focus from the individual's narrowly defined right of control over their data, to focus more on the context of information exchanges and uses. The rules that determine how, and to whom, health information is disclosed, and the purposes for which it can be used, must reflect and promote health values. Trust is critical in a healthcare context - not just in the relationship between doctor and patient, but also the trust citizens must place in their government to hold their data responsibly. If we are to learn anything from the UK government's relationship with big tech, it is how easily this can be broken.

Chapter Three - Social Privacy Harms of Big Data

Introduction

In Chapter Two I argued that the clear beneficiaries of big data initiatives in healthcare are the technology companies that have gained access to unprecedented amounts of personal information. In this Chapter I will address the question of *why* this is problematic. The processing of big data is prolific and extends across many sectors of society - while the contexts vary, the rationale remains the same: large amounts of personal data are valuable, arguably, a new form of capital.[1] The momentum is self-perpetuating: to remain competitive, many companies generate ever-increasing amounts of data from which to make increasingly detailed profiles and predictions about us.

For the big technology companies, the primary purpose of big data is to place people into groups based on their shared characteristics, preferences, and beliefs. It is the relational aspect of big data that gives it its power and reach. Big tech isn't interested in the individual's data in isolation; value lies in the sheer volume of information about people and how they relate to each other.[2]

There is nothing inherently wrong with big data nor is there anything necessarily problematic about big data analytics: the predictive models, algorithms, and artificial intelligence that process big data. However, when used carelessly or for ill-conceived objectives, big data has the potential to cause significant harm, predominantly to those already struggling the most.

I use the term 'social privacy harms' to describe the collective harms of big data. The harms that ripple beyond the individual data subject in a particular information

[1] Jathan Sadowski argues that data is a new form of capital for many industries who are constantly looking to exploit new ways of accumulating data: "Just as we expect corporations to be profit driven, we should now expect them to be data driven. This is why so much of smart tech is built to suck up data". (Jathan Sadowski, "How 'Smart Tech' Masks an Emerging Era of Corporate Control" (9 March 2020) OneZero https://onezero.medium.com/how-smart-tech-masks-an-emerging-era-of-corporate-control-779c96b05f85)

[2] Salomé Viljoen "A Relational Theory of Data Governance" (2021) Yale Law Journal 131 at 573-578.

transaction. The harms that our current regulatory framework for data privacy, with its narrow focus on the individual, is ill-equipped to address.

I draw on Joel Feinberg's concept of accumulative public harms in my assessment of the social privacy harms of big data. I argue that uses of big data that contribute to increasing inequality set back the public interest in living in economically successful, and socially thriving, societies. The cumulative effects of discriminatory and unfair big data applications harm society by making it more unequal. Additionally, big data has facilitated the generation and collection of vast amounts of personal information held by a handful of powerful corporations. Very large online platforms, such as Facebook, and very large online search engines, such as Google, exercise significant control over the information we view online; know a great deal about us, and thus, can influence the decisions we make. Too much personal information in the hands of a few powerful corporations creates power asymmetries that are harmful to democracies. Societies have grappled with the issue of power being concentrated in in the hands of a few for centuries - it is not a problem created by the rise of big data; however, I argue that it is made worse by the predictive power and reach of big data analytics.

There will be winners and losers of social sorting, automated decision-making, and recommender systems. Individuals stand to lose and gain from big data. The strongest justification for changing our conception of data privacy from a narrowly-defined focus on individual data rights, to one that acknowledges the social context of information use, is to mitigate social privacy harms.

In New Zealand there is limited information on private sector uses of big data, however public sector uses of big data and artificial intelligence (AI) are better documented.[3] There are some benign and useful applications of algorithms by government agencies in New Zealand, but we cannot be complacent about potential future uses of big data. Some questionable uses of big data have been abandoned after public condemnation.[4] Whether

[3] Colin Gavaghan, Alistair Knott, James Maclaurin, John Zerilli, Joy Liddicoat *Government Use of Artificial Intelligence in NZ: Final Report on Phase One of the New Zealand Law Foundation's Artificial Intelligence and Law in New Zealand Project* (New Zealand Law Foundation, Wellington, 2019); Stats NZ *Algorithm Assessment Report* (2018) https://www.data.govt.nz/assets/Uploads/Algorithm-Assessment-Report-Oct-2018.pdf.

[4] For example, an algorithm commissioned by the Minister of Social Development and Employment to predict at-risk children before birth was scrapped in 2015 before any testing of its accuracy was

this shows that the public sector is responsive to the mood of the nation or reveals the apparent eagerness of public agencies to adopt unproven technology,[5] can only be surmised.

Surveillance capitalism – with its significant power imbalances and market model based on the monetisation of personal information - demands a new approach to the regulation of personal information. Privacy harms in an age of big data affect more than just the individual data subject and are not adequately addressed by New Zealand's current regulatory framework. Big data, if not properly regulated, has the potential to exacerbate and entrench existing inequalities in society and harm democracies.

Social Privacy Harms

Harm

In *Harm To Others: The Moral Limits of the Criminal Law*, Joel Feinberg explores the moral constraints that limit a law-maker's options.[6] He does not set out to answer the question of what would be good to legislate for, but rather what *may* be legislated for; the kinds of conduct that a legislature may make criminal without infringing on the moral autonomy of its citizens.[7] Feinberg states that:[8]

> ... we can assert tentatively that it is legitimate for the state to prohibit conduct that causes serious private harm, or the unreasonable risk of such harm, or harm to important public institutions and practices.

---

undertaken; In 2020 NZ Police trialled Clearview AI's facial recognition technology without the signoff of the Police Commissioner or the input of the Privacy Commissioner. The technology was later abandoned by Police due to it being "ineffective in New Zealand" https://www.rnz.co.nz/news/national/416913/police-stocktake-surveillance-tech-after-clearview-ai-facial-recognition-trial

[5] Another example of a public agency considering questionable technology is that of NCEA looking into automated exam essay marking. One fool proof way to inhibit original thought and creativity is to employ an algorithm that can only assess the work before it by comparing its likeness to what has scored well in previous years (Lauren Wiltshire "Qualifications authority running trials into automated NCEA essay marking" (24 February 2021) https://www.stuff.co.nz/national/education/124328937/qualifications-authority-running-trials-into-automated-ncea-essay-marking).

[6] Joel Feinberg *Harm To Others: The Moral Limits of the Criminal Law* (Oxford University Press, Oxford, 1984) at 3-4.

[7] At 3-11.

[8] At 11.

Feinberg believes that John Stuart Mill's harm principle[9] is a valid principle for determining legitimate invasions of individual liberty.[10] He defines harms as setbacks to interests that are also wrongful.[11] Interests are described as all things in which one has a stake; we flourish or languish as our interests flourish or languish.[12] A person harms another by setting back their interest; leaving that interest in a worse condition than it otherwise would have been had the setback not occurred at all.[13] Feinberg argues that something undesirable or unpleasant is not necessarily a harm unless its presence is sufficient to impede an interest.[14]

Feinberg believes that in order for A to have harmed B -

(1) A must act (or omit to act) in a manner that intentionally produces the consequences for B that follow (or similarly adverse ones) or act with negligence or recklessness in respect of those consequences; and
(2) A acts in a manner that is neither excusable nor justifiable; and
(3) A's action causes a setback to B's interests, which is also a violation of B's right[15] (B's setback interest is one he has a right to have respected).[16]

The harmer's act must be blameworthy or morally indefensible[17] because only people who voluntarily break the law without good reason or excuse should be punished.[18] Feinberg explains that justified wrongdoing is not wrongdoing at all and without wrongdoing there is no harming, however severe the harm that results.[19]

---

[9] Mill asserts the principle that "the sole end for which mankind are warranted, individually or collectively, in interfering with the liberty of action of any of their number, is self-protection. That the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others." (John Stuart Mill *On Liberty* (Cambridge University Press, Cambridge, 2011) at 21-22.)
[10] Feinberg, above n 6, at 12-13.
[11] At 105.
[12] At 33-34.
[13] At 34.
[14] At 46-47.
[15] At 105-106.
[16] At 108.
[17] At 108.
[18] At 108-109.
[19] At 109.

Feinberg considers the degree of harm required in order for the harm principle to warrant legal coercion to prevent it and concludes that it must cause more than mere annoyance, hurt, inconvenience, or offense:[20]

> No would seriously suggest, for example, that repeated rude and disrespectful remarks to parents, spouses, teachers, and others who have a right to better treatment should be forbidden by criminal statutes, even though such discourtesies not only "wound feelings," but indirectly harm the interest in personal efficiency by causing depression and anger sufficiently great to distract and debilitate.

Feinberg's rationale for excluding minor harms is because to punish them could cause more harm than it prevents.[21] Also, because it would cause harm to wrongdoers out of all proportion to their guilt and to the harm that they would otherwise cause.[22] Feinberg argues that 'risk' should be a significant concern to the legislator guided by the harm principle.[23] Consideration should be given to whether or not the private interests of agencies engaging in certain activities are valuable, and if they are backed up by public interests.[24]

Feinberg describes problematic cases where a type of activity tends to cause harm to people who are affected by it but prohibiting that activity would thwart the substantial interests of those who engage in it, harming them.[25] Consequently, the legislator must decide whose interest is more or less important.[26] When weighing interests, Feinberg maintains that legislators must consider the extent to which each interest is thwarted; their importance; the degree to which they are backed up by other interests, and also their "inherent moral quality".[27]

[20] At 188-189.
[21] At 189.
[22] At 190.
[23] At 191.
[24] At 191.
[25] At 203.
[26] At 203.
[27] At 205.

Public interest

Feinberg describes the second of two categories of "public interest" as a commonly shared, specific interest - an interest that all, or most, members of a community share. For example, economic prosperity, or the interest in avoiding natural disasters, crime waves, riots and the "spreading of distrust and incivility".[28] Not all people will share the common interest to the same degree - some people benefit financially from war and times of economic depression, and these events are not equally hard on all people.[29] Similarly, I argue that some people also benefit from the spreading of disinformation online and the erosion of democracy.

Feinberg argues that public interests should be given a priority over other interests because the most commonly shared public interests happen to have the "greatest degree of vitality in the interest networks of those who share them".[30] Public interests are also given additional weight through social reinforcement:[31]

> An act that sets back to a small extent a widely shared interest may do little harm to each person who has the interest, but it does some harm to almost everyone, a consideration that multiplies its significance.

In designing legislation to reduce or prevent harms to public interests, Feinberg identifies the difficulties in attempting to attribute harm as a problem best addressed by civil law or regulation, and not the criminal law in the first instance (although the criminal law can be used as a backup sanction). Resorting to the criminal law in the first instance is too blunt an instrument argues Feinberg[32], who states:[33]

> Rather, the question a legislature must ask, in the spirit of the harm principle, is this: In the effort to minimize public harms generally, within the limits of efficiency, equity, and fair play, what sort of regulative scheme should be devised?

---

[28] At 223.
[29] At 223-224.
[30] At 225.
[31] At 225.
[32] At 227.
[33] At 227-228.

Accumulative public harms

Feinberg categorises air and water pollution as accumulative public harms.[34] He states that they "provide challenging hard cases for the concept of *harming*".[35] Feinberg uses the example of one polluting car not causing sufficient harm to the environment to be worth pursuing but at some point, between one polluting driver and most people driving polluting cars, the threshold of harm would be met.[36]

Analogies can be drawn between the public harms of pollution and the public harms of degraded personal information ecosystems. There are information practices or uses of big data analytics where an isolated instance may not reach the threshold of harm to an individual, but when repeated at scale, constitute a public harm. Feinberg argues that it is only against the background of a regulatory system that individual imputations of public accumulative harm make sense.[37] He states that:[38]

> For *A* to harm the public interest is for *A's* wrongful conduct to cause a setback to that interest. In the context of industrial polluting, "wrongful" must mean unlawful as judged by a regulative agency applying rules for allocating permits in accordance with specified requirements of fairness and efficiency. In these contexts, no prior standard for wrongfulness exists. There is nothing inherently wrongful or right-violating in the activity of driving an automobile, generating electricity, or refining copper. These activities can be meaningfully condemned only as violations of an authoritative scheme as allocative priorities.

Just as there is nothing inherently wrongful in many activities that generate pollution, there is also nothing inherently wrongful about deploying big data-enabled technologies. Feinberg argues that the harm principle lends legitimacy to legislative efforts to address pollution, "but in its bare formulation without supplement, it offers no guide to policy".[39] Feinberg's notion of accumulative public harms is helpful in the assessment of the social

[34] At 227.
[35] At 227.
[36] At 228.
[37] At 230.
[38] At 230.
[39] At 232.

privacy harms of big data. However, it does not purport to address how a regulatory framework should address those harms.

New Zealand legislative definitions of 'harm'

The concept of privacy harm *to the individual* has been considered by the New Zealand legislature in both the Privacy Act 2020 (the relevant wording remaining unchanged from the original 1993 Act) and in the Harmful Digital Communications Act 2015.

Under the Privacy Act, an action of an agency is an interference with privacy if the action breaches, in relation to an individual, an information privacy principle[40] **and** the action:[41]

(i) has caused, or may cause, loss, detriment, damage, or injury to the individual; or

(ii) has adversely affected, or may adversely affect, the rights, benefits, privileges, obligations, or interests of the individual; or

(iii) has resulted in, or may result in, significant humiliation, significant loss of dignity, or significant injury to the feelings of the individual.

In the Harmful Digital Communications Act (HDCA), an Act designed to deter, prevent, and mitigate harm caused to individuals by digital communications,[42] "harm" is defined as "serious emotional distress".[43] In both Acts, emotional harm is recognised as a harm that can be caused by the misuse of personal information. The threshold is set high in both Acts; it must be *serious* emotional distress under the HDCA and *significant* humiliation, *significant* loss of dignity, or *significant* injury to the feelings of the individual under the Privacy Act. The scope of both Acts is limited to harms to the individual, although the Privacy Act does state that a complaint may be made on behalf of one or more aggrieved individuals.[44] Neither Act addresses the concept of public harms, but both recognise intangible emotional harms caused by the misuse of personal information.

---

[40] Or the provisions of an approved information sharing agreement; or the provisions of an information matching agreement, or s 115 of the Act which requires an agency to give notice to affected individuals of a notifiable privacy breach (Privacy Act 2020, s 69(2)(a)).
[41] Privacy Act 2020, s 69(2).
[42] Harmful Digital Communications Act 2015, s 3(a). Another purpose is to provide victims of harmful digital communications with a quick and efficient means of redress (s 3(b)).
[43] Harmful Digital Communications Act 2015, s 4.
[44] Privacy Act 2020, s 71(2).

Harm is often considered in the context of harm to the individual and privacy harms have traditionally been conceptualised in this way. However, with the development of big data, our conception of privacy harm must extend beyond the individual to include social privacy harms. Drawing on Feinberg's notion of accumulative public harms, I will describe two broad categories of social privacy harm caused by big data.

Harm of increasing inequality

If not properly regulated, big data-enabled technologies can perpetuate unjustified discrimination and exacerbate inequalities in society. Predictive technologies can discriminate on many grounds, in particular, race, gender, and socioeconomic status. While not every instance of unjustified discrimination or unfairness will cause harm to an individual by setting back or thwarting their interests, in the following paragraphs I make the argument that, cumulatively, unfair or unethical information practices set back the public interest in living in successful, healthy, and harmonious communities.

The notion of what constitutes fairness is contested. My concept of fairness is influenced by John Rawls' theory of justice. Rawls states that:[45]

> ... the principles for the basic structure of society are the object of the original agreement. They are the principles that free and rational persons concerned to further their own interests would accept in an initial position of equality as defining the fundamental terms of their association. These principles are to regulate all further agreements; they specify the kinds of social cooperation that can be entered into and the forms of government that can be established. This way of regarding the principles of justice I shall call justice as fairness.

Two principles of justice that Rawls believes would be agreed to in the original position are:

(1) each person to have an equal right to basic liberties; and

[45] John Rawls *A Theory of Justice* revised edition (Oxford, Oxford University Press, 1999) at 10.

(2) social and economic inequalities are to be arranged to everyone's advantage and attached to positions and offices open to all.[46]

The first principle has priority over the second.[47] Rawls argues that:[48]

> All social values - liberty and opportunity, income and wealth, and the social bases of self-respect - are to be distributed equally unless an unequal distribution of any, or all of these values is to everyone's advantage.
>
> Injustice then, is simply inequalities that are not to the benefit of all.

Rawls' theory of justice draws on the ideal of equality of opportunity - people should not be penalised on the basis of circumstances they cannot control.[49] Individuals' successes and failings in life do not stem solely their own skill and effort (or lack of) but are influenced by the circumstances they are born into and the environment that they live in. A fair society recognises that those who have succeeded in life have not achieved that success in a vacuum, but as individuals within a society and therefore should contribute some of the economic reward of those successes through taxation, to be redistributed to those at the bottom of society. Rawls believes that "undeserved inequalities call for redress".[50] Inequalities of birth and natural ability are undeserved and should be compensated for in a fair society.

In *Inequality: an assessment,*[51] Ken Mayhew and Samuel Willis note that the idea of equality of opportunity, or the notion of a level playing field, is where there is most widespread agreement amongst economists in their assessment of inequality.[52] However, they recognise that, in practice, implementing a level playing field can be very difficult. They point to the employment context where measures designed to promote equality of opportunity have been hampered by the problem of unconscious discrimination in hiring

[46] At 52-53.
[47] At 53-54.
[48] At 54.
[49] At 62-63.
[50] At 86.
[51] Ken Mayhew and Samuel Willis "Inequality: an assessment" (2019) Oxford Review of Economic Policy 35 3.
[52] At 354.

decisions.[53] Mayhew and Willis argue that too much inequality may have harmful effects on the economy overall but also that too little inequality may also be harmful, for example, a more progressive tax system may discourage entrepreneurism or hard work and ambition.[54]

Like Rawls, and Mayhew and Willis, Venkat Ventkatasubramanian believes that income inequality *in itself* is not necessarily bad.[55] He points to the varied talents, skills, and capabilities of individuals who make differing contributions, arguing that it is only fair that those who contribute more, earn more.[56] Some income inequality is needed to incentivise people to contribute more to society, benefitting everyone overall. A certain level of inequality can benefit society.[57] However, Ventkatasubramanian qualifies this by stating that:[58]

> ... when a large section of the society is not benefiting from society's growth (e.g., in GDP), with its real incomes stagnating for decades and its access to health care, education, and opportunities for upward mobility declining, then the system is not working as it should.

Ventkatasubramanian discusses the moral issues caused by extreme inequality in the U.S. and the social justice issues it raises.[59] Economic inequality gives wealthy people too much power over the lives of others and undermines the fairness of political institutions and the economic system.[60] Economic inequality can erode the cooperation and social cohesion required for society to function effectively.[61] Ventakasubramanian also highlights concerns about the adverse impacts of extreme inequality on educational outcomes, health, the opportunity for social mobility, happiness and wellbeing, in addition to the "immorality and illegitimacy of extreme income inequality".[62]

---

[53] At 354.
[54] At 357.
[55] Venkat Venkatasubramanian "Extreme Inequality in Income and Wealth" in Venkat Venkatasubramanian (ed.) *How Much Inequality Is Fair? Mathematical Principles of a Moral, Optimal, and Stable Capitalist Society* (Columbia University Press, New York, 2017) at 12.
[56] At 12.
[57] At 12.
[58] At 13.
[59] At 13.
[60] Venkatasubramanian draws on the argument of Harvard philosopher Tim Scanlon, at 13.
[61] At 13-15.
[62] At 15.

In their psychological assessment of income inequality, Nicholas Buttrick and Shigehiro Oishi discuss the association between decreased levels of trust, and increased levels of competition and status anxiety, in societies with greater inequality.[63] They point to surveys that reveal people are less likely to trust each other in areas with greater income inequality.[64] Low trust in other people is associated with a range of negative outcomes, including a greater likelihood of engaging in unethical behaviour, less participation in civic life and lower participation in social group membership.[65] Buttrick and Oishi argue that countries with greater income inequality have greater political instability, worse institutions, and more corruption.[66] Interestingly, they note that status anxiety is higher amongst all members of unequal societies, including the wealthy, in comparison to citizens of equivalent socioeconomic status in other, more equal, societies.[67]

Societies prosper when their citizens are treated fairly.[68] The cumulative effects of unjustified discrimination harm society by making it more unequal. In *The Spirit Level: Why Equality is Better for Everyone*, Richard Wilkinson and Kate Pickett draw on decades of research in health inequalities and official statistics across numerous countries to conclude that not only are health outcomes for *everyone* better is a more equal society, but many other determinants of social wellbeing are improved in more equal societies.[69] Their research shows that more unequal societies have more social problems overall.[70] Reducing inequality improves the quality of life for everyone, including the wealthiest, and reducing inequality is key to how a country performs in lots of different fields.[71] A much higher proportion of the population suffer from mental illness in more unequal

[63] Nicholas R. Buttrick and Shigehiro Oishi "The psychological consequences of income inequality" (2017) Social and Personality Psychology Compass 11 at 2.
[64] At 2.
[65] At 3.
[66] At 3.
[67] At 4.
[68] Deloitte "The economic benefits of improving social inclusion" (August 2019) https://www2.deloitte.com/content/dam/Deloitte/my/Documents/risk/my-risk-sdg10-economic-benefits-of-improving-social-inclusion.pdf at 6-7; The Legatum Prosperity Index 2023 "The Foundational Elements of Prosperity" https://www.prosperity.com/feed/foundational-elements-prosperity; Mayhew and Willis, above n 51; Venkatasubramanian, above n 55; Buttrick and Oishi, above n 63; Richard Wilkinson and Kate Pickett *The Spirit Level: Why Equality is Better for Everyone* (Penguin Books, London, 2009).
[69] Wilkinson and Pickett, above n 68.
[70] At 25.
[71] At 29.

countries.[72] The use of illicit drugs is more common in unequal societies[73] and there is a correlation between greater inequality and higher rates of murder.[74] Wilkinson and Pickett persuasively argue that "the vast majority of the population is harmed by greater inequality".[75] Big data is harmful when it exacerbates and entrenches existing inequality in society through unjustified discrimination on the grounds of race, gender, disability, or socioeconomic status. Not every act of unjustified discrimination causes harm by setting back or thwarting the interest of the individual discriminated against. However, the cumulative effects of big data that contribute to increasing inequality are harmful to the public interest in living in a healthy and prosperous society.

Harm to democracy

Big data has facilitated the generation of vast amounts of personal information that is held and processed by the big technology companies. Too much information in the hands of a few harms society's interest in living in a well-functioning democracy. Without adequate transparency and oversight, big tech's control over what information is presented to us and how it is prioritised through search engine rankings and newsfeeds poses a threat to the public interest in being able to access reliable and accurate information online. Information- and power- asymmetries between citizens and big technology companies gives those companies an over-sized influence that erodes democratic societies. In Chapter One I described the extensive lobbying by big tech to dilute the provisions of the European Union's General Data Protection Regulation and the California Consumer Privacy Act; the *Schrems'* cases highlight the difficulty in holding Facebook to account for breaching the law. The size and influence of the big technology companies, fuelled in a large part by their ability to generate big data, is inherently harmful to democracy.

Relational Aspect of Big Data

Contemporary 'notice and consent' models of data privacy regulation are ripe for reform. Salomé Viljoen argues persuasively that a central feature of data processing is to place people into "population-based relations with one another" and that this priority of big tech

[72] At 66-67.
[73] At 70.
[74] At 135.
[75] At 176.

is overlooked by data privacy reforms that aim to either redistribute the excessive wealth created by data processing, or reassert the individual data subject's control over their personal information and its use.[76] Viljoen states that "this relational aspect of data production drives much of the social value and harm of data collection and use in a digital economy".[77]

Viljoen provides a compelling theoretical account of data as social relations and explains how data relations result in supraindividual legal interests. Individual data subject rights cannot account for, or address, population-level effects of data processing. Additionally, Viljoen states that:[78]

> ... both the status quo and reform proposals suffer from a common conceptual flaw: they attempt to reduce legal interests in information to individualistic claims subject to individualistic remedies, which are structurally incapable of representing the interests and effects of data production's population-level aims.

The crux of Viljoen's argument is that big data's power to convey information about others and their relationship to one another makes big data particularly adept at both benefiting and harming those other than the data subject.[79] Viljoen rightly asserts that there is a broader social concern that is left unprotected by the status quo and its focus on the individual data subject and data controller. She states that "digital-surveillance technologies used to enhance user experience for the rich simultaneously provide methods of discipline and punishment for the poor".[80] A constant theme in the analysis presented in this chapter is how the burdens of big data fall heaviest on the poor and marginalised.

Unfairness

Big data, processed by algorithms, to aid decision-making in contexts such as criminal justice, finance, housing, and employment, raises a range of fairness and justice issues.

[76] Viljoen, above n 2, at 573.
[77] At 573.
[78] At 578.
[79] At 583.
[80] At 580-581.

There is a lack of transparency when an algorithm operates as a black box and no-one actually knows how or why it reaches its outcomes. When the workings of proprietary algorithms are not disclosed for commercial reasons, it becomes difficult to challenge a decision that is influenced by the outcome of such an algorithmic assessment.

Mathematician Cathy O'Neill describes many of these algorithmic harms in her book *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*.[81] O'Neill believes that mathematical models or algorithms can be used with the best of intentions, for example, to be fair and impartial by removing human bias. However, the human bias is often not eliminated, but merely camouflaged by technology.[82] O'Neill states that there will always be mistakes with models because they are by their very nature, simplifications. She states that "no model can include all of the real world's complexity or the nuance of human communication".[83] In making a model the designers make decisions about what is important to include: "a model's blind spots reflect the judgments and priorities of its creators".[84] With some models this is unproblematic: e.g. Google Maps models the world in terms of roads, tunnels, and bridges, but ignores the buildings.[85] However, with other models, particularly in social contexts, judgements about what is, or is not, important to include are more significant.

One example is a value-added algorithm used by the Washington D.C. school system, the objective of which was to raise the performance of the city's underperforming schools.[86] However, the underlying theory was that students were underachieving because teachers were not doing a good enough job.[87] An algorithmic model was used to evaluate teachers' performance, largely on the basis of their students' test scores from one year to the next. By focussing on students' test results it ignored other important measures of what it takes to be a good teacher, for example how well teachers engage and motivate their students, or help them with family and personal problems.[88] O'Neill describes how trying to calculate the impact that a teacher may have on a student over a course of a school year

[81] Cathy O'Neill *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (Penguin Random House, United Kingdom, 2016).
[82] At 24-25.
[83] At 20.
[84] At 21.
[85] At 21.
[86] At 3.
[87] At 4.
[88] At 21.

is much more complex than using big data to target advertising, for example.[89] The algorithmic model was used to inform decisions to fire "underperforming teachers" at the risk of also firing some high-performing teachers who had received positive reviews from principals and parents.[90] The high-stakes outcomes of student test results as well as the high number of erasures and unusually high test scores raised the suspicion that some teachers had corrected their students' answers.[91] O'Neill states that algorithms or mathematical models "despite their reputation for impartiality, reflect goals and ideology".[92] She believes that "whether or not a model works is also a matter of opinion".[93] We must ask who designed the model and what they were trying to achieve.

When used at scale, a predictive algorithm can potentially gain an entrenched and inflexible hold over people's employment, housing, or credit prospects. Nevertheless, it is important to acknowledge that human decision makers can also be discriminatory and are, at best, imperfect decision-makers. However, it is the scope of predictive algorithms; their ability to reach further than one discriminatory employer, landlord, or bank manager that I believe makes them potentially very harmful.

Big data, *if not appropriately regulated*, is likely to exacerbate existing inequalities in society. Algorithms that lack mechanisms to check the accuracy of their predictions against real world outcomes, can contribute to the outcomes that they predict. If employers, for example, use predictions of future ill health as a reason not to hire people, those people remain unemployed and consequently, are likely to endure a poorer diet, colder homes, and reduced access to health care in comparison with employed citizens. Unemployment propels people into poverty which contributes to poor health outcomes which could make them even more unemployable. This potential for harm highlights the importance of understanding how, and for whom, algorithms work.

[89] At 6.
[90] At 4 & 21.
[91] At 9.
[92] At 21.
[93] At 21.

Opacity

Unfortunately, the Internet and finance companies that collect vast amounts of information about people often fail to be transparent when it comes to the workings of their companies and the algorithms they use to make important decisions about us. In his book *The Black Box Society: The Secret Algorithms That Control Money and Information*[94] Frank Pasquale states:[95]

> Reputation. Search. Finance. These are the areas in which Big Data looms largest in our lives. But too often it looms invisibly, undermining the openness of our society and the fairness of our markets.

Pasquale describes a repeatedly occurring and alarming dynamic of corporate secrecy expanding as people's privacy contracts.[96] He argues that automated scoring systems used to classify and rate individuals invariably operate in secret while claiming to treat everyone in the same way and thus avoiding the discrimination of human decision makers. However, like O'Neill, Pasquale believes that automation does not eliminate discrimination, it merely drives it upstream to the software engineers whom, he states:[97]

> ... construct the datasets mined by scoring systems ... define the parameters of data mining analyses ... create the clusters, links, and decision trees applied ... [and] generate the predictive models applied. Human biases and values are embedded into each and every step of development.

While marginalised groups in society are likely to be hit hardest by algorithmic decision-making, Pasquale states that anyone can be unfavourably labelled in a database as "unreliable", "high medical cost", or "declining income", for example.[98] Reputational algorithms that operate like a black box make it impossible to determine whether they operate discriminatorily. Pasquale states that:[99]

---

[94] Frank Pasquale *The Black Box Society: The Secret Algorithms That Control Money and Information* (Harvard University Press, London, 2015).
[95] At 5.
[96] At 26.
[97] At 35.
[98] At 38.
[99] At 38.

> Reputational systems are creating new (and largely invisible) minorities, disfavored due to error or unfairness. Algorithms are not immune from the fundamental problem of discrimination, in which negative and baseless assumptions congeal into prejudice. They are programmed by human beings, whose values are embedded into their software. And they must often use data laced with all-too-human prejudice.

I argue that there is a clear need for more openness from the agencies and companies that process personal information and a greater level of accountability. Pasquale argues that the financial crisis of 2007-2008, sparked by the crash of mortgage-backed securities, was "the natural consequence of a black box financial system".[100] He believes that the misuse of algorithms is not confined to that period and that algorithms "have supercharged classic forms of self-dealing".[101] Pasquale argues that there is a pressing need for more transparency and accountability from financial companies because "far too much of contemporary finance is premised on hiding information: from borrowers, lenders, clients, regulators, and the public".[102]

Big data is not intrinsically bad, nor is there anything necessarily problematic about big data analytics. However, big data must be used knowledgeably and responsibly to harness its benefits and limit social privacy harms. Algorithms designed for use in one context should not be transplanted to another without considering the implications of that change in context. Responsible deployment of big data requires an awareness of the limitations of AI, and the potential for bias in datasets and in how software is programmed. Predictive systems require feedback - they need a mechanism by which they can learn from their mistakes. AI requires accountability and oversight, which can be time-consuming and expensive. When big data is deployed as a cost-saving measure, these safeguards can be overlooked. However, to use big data ethically and to gain the greatest insights from it, it must be used in an informed and responsible manner. In Chapter Six, I conclude that relying on agencies to do this voluntarily has failed; and now demands a hard law approach. The following part of this Chapter considers how the social privacy harms of big data fall heaviest on those relying on government assistance, the marginalised, and the indigent.

---

[100] At 134.
[101] At 103.
[102] At 135.

The Burden of Big Data Falls Heaviest on the Poor

In the following section, I review Virginia Eubanks' book, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*[103], in which Eubanks sets out several case studies of big data being deployed in harmful ways in the provision of social services in the United States. Eubanks clearly articulates the heart of the issue when she states that "we all inhabit this new regime of digital data but we don't all experience it in the same way". [104] Marginalised groups face higher levels of data collection when they access public benefits, financial assistance, and health care, and by virtue of living in more heavily policed neighbourhoods. Eubanks states that:[105]

> ... data acts to enforce their marginality when it is used to target them for suspicion or extra scrutiny. These groups seen as undeserving are singled out for punitive public policy and more intense surveillance. It is a kind of collective red-flagging, a feedback loop of injustice.

Eubanks states that in the United States automated systems are being integrated into human and social services "at a breathtaking pace with little or no political discussion about their impacts".[106] Eubanks describes a long held American mindset towards the poor that holds them individually accountable for their own misfortune; an attitude that breeds intolerance, contempt, and hatred for the poor in the United States, which is arguably not the case in New Zealand, at least not to the same level of intensity. The social, economic, cultural, and political contexts in the United States and New Zealand are different. Nevertheless, the lessons to be learned from these algorithmic experiments are just as pertinent to the New Zealand context. The algorithmic sorting and monitoring systems discussed in the following paragraphs could have similarly harmful consequences for people if implemented in New Zealand.

---

[103] Virginia Eubanks *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (St. Martin's Press, New York, 2017).
[104] At 5.
[105] At 7.
[106] At 11-12.

The "digital poorhouse"

The array of digital databases and algorithmic tools that exacerbate discrimination is described by Eubanks as the "digital poorhouse". These digital systems are opaque and difficult to understand; massively scalable; hard to decommission; and enduring - digital data lasts a long time.[107] Eubanks argues that "when automated decision-making tools are not built to explicitly dismantle structural inequities, their speed and scale intensify them".[108] The personalisation of automation might help some people, but it also leaves the vulnerable open to persecution.[109] I argue that in the interests of fairness, automated decision-making tools that run the risk of further marginalising vulnerable groups should not be deployed without suitable safeguards in place.

Automated welfare eligibility in Indiana

In the mid-2000s the state of Indiana contracted with IBM to create an automated model to determine welfare eligibility. Eubanks argues that this resulted in devastating impacts on the poor and working class in Indiana. She states that:[110]

> Between 2006 and 2008, the state of Indiana denied more than a million applications for food stamps, Medicaid, and cash benefits, a 54% increase compared to the three years prior to automation.

Automation meant that applicants no longer dealt with one case worker from the start to finish of an application for welfare. Employees responded to tasks in a queue and applicants spoke to a different case worker each time they called. This was designed to sever the caseworker-client relationship and prevent caseworkers from seeking the maximum entitlements for their clients. Applications were declined for 'failing to cooperate' if a document was missing or incorrectly filled out.[111] But the notifications sent to applicants provided no detail and applicants lacked a case worker to call to ask for an explanation or help.[112] Eubanks describes how IBM's automated welfare eligibility

[107] At 190.
[108] At 190.
[109] At 199-200.
[110] At 51.
[111] At 51.
[112] At 51.

system was rife with service issues "including excessive wait times, lost documents, inaccurate data, interview scheduling problems, slow application processing, and incorrect instructions to clients".[113]

After months of defending the automation experiment, Governor Daniels admitted that it had failed and cancelled the contract with IBM, describing it as a "flawed concept that simply did not work out in practice".[114] In May 2010, the state of Indiana sued IBM for $437 million, claiming breach of contract because of the incorrect benefit denials. IBM countersued for $100 million for the server, hardware, software, and algorithms that the county was still using. IBM won the case and was awarded $52 million. Marion Superior Court Judge David Dreyer stated that neither party deserved to win, and both were to blame, referring to "misguided government policy and overzealous corporate ambition".[115]

Judge Dreyer held that there had been no material breach of the contract by IBM. IBM, as a contractor, was only liable to its employer and shareholders. It was not its job to measure the automation experiment's impact on the poor and working-class. The human costs of the system had not been anticipated or addressed.[116]

I argue that in some circumstances, justice requires flexibility. When human discretion is taken away from frontline public servants and decisions are automated according to the inflexible discretion of software engineers and commercial contractors, discrimination is exacerbated.[117] Admittedly, human discretion also allows for bias to creep into decision-making, but this is arguably less harmful than when an entire system is set up with biased assumptions ‘baked in’. Eubanks states that in Indiana:[118]

> The "social specs" for the automation were based on time-worn, race- and class-motivated assumptions about welfare recipients that were encoded into performance metrics and programmed into business processes: they are lazy and must be

[113] At 71.
[114] At 72.
[115] At 72 [*State of Indiana v International Business Machines Corporation* No.49S02-1408-PL-00513 (Indiana Supreme Court 2006)].
[116] At 75.
[117] At 81.
[118] At 81.

> "prodded" into contributing to their own support, they are sneaky and prone to fraudulent claims, and their burdensome use of public resources must be repeatedly discouraged.

Eubanks is persuasive in her assessment that automated decision-making in Indiana's welfare system inflamed the harmful effects of pre-existing punitive attitudes towards the poor. Rather than being a new high-tech tool that allowed for greater accuracy in ensuring that every applicant got all the benefits that they were entitled to, it worked to filter and divert applications; acting as "a gatekeeper, not a facilitator".[119]

In 2006, when the Governor of Indiana signed the contract with IBM, 38 per cent of poor families with children were receiving cash benefits, but by 2014 it was eight per cent.[120] Eubanks concludes that "in the end, the Indiana automation experiment was a form of digital diversion for poor and working Americans. It denied them benefits, due process, dignity, and life itself".[121]

Automated tools, implemented to improve efficiencies and better allocate resources, can fall short in other equally important social justice and fairness measures. Automated tools should not be deployed in important decision-making contexts that have significant impacts on people for the primary purpose of saving resources without proper consideration for other important objectives, such as fairness, accuracy, and avoiding harm.

Electronic registry of homeless in Los Angeles

Another failure outlined by Eubanks involved Los Angeles' co-ordinated entry system that was intended to match the county's most vulnerable unhoused people with appropriate housing and social services.[122] It was designed to reduce waste and prevent people obtaining more than what they were entitled to.[123] It originated in 2013, when the charitable organisation, United Way of Greater Los Angeles, and the Los Angeles Area

[119] At 82.
[120] At 82.
[121] At 83.
[122] At 84.
[123] At 85.

Chamber of Commerce collaborated on a project called *Home for Good*. The project pledged to house one hundred of the most vulnerable people on Skid Row in one hundred days. To do this, a list of Skid Row's homeless was created, ranked in order of need, and matched to housing opportunities. A digital registry was built to store the data.[124]

The co-ordinated entry process begins whenever a homeless personal engages with a social service worker or volunteer during a shelter admission or other social service program. The *Vulnerability Index – Service Prioritisation Decision Assessment Tool (VI-SPDAT)* includes intimate questions about the subject's health including use of mental health crisis services, sexual activity, risk-taking behaviours such as prostitution, drug running and drug use, and suicidality, as well as full name, date of birth, and social security number.[125] This information can be shared with up to 168 other organisations. All the information is fed into the federal *Homeless Management Information System (HMIS).* The federal government requires all organisations receiving homeless assistance funds to collect a set of universal data points, which provides aggregated information about America's homeless.[126]

Data from the *VI-SPDAT* is entered into the Los Angeles *HMIS*. A ranking algorithm then tallies up a score between 1 – 17: '1' indicating low risk – i.e. a relatively small chance of dying or ending up in a mental hospital; and '17' for the most vulnerable. Those scoring 0-3 get no housing intervention; people with scores of 4-7 get limited-term rental subsidies and case management services; those scoring 8-17 are assessed for permanent supportive housing.[127] A second algorithm (the matching algorithm) identifies a person who is in the greatest need (by virtue of the *VI-SPDAT* score) and who also meets the eligibility criteria for the particular housing, and then if the match is approved the person receives the housing. If not, the match disappears, and the algorithm is re-run to find a new match with another applicant.[128]

The 'catch 22' with the *VI-SPDAT* survey is that an applicant needs a high score to qualify for housing, but if the score is too high it raises the question of whether the applicant is

[124] At 93.
[125] At 93-94.
[126] At 94.
[127] At 94.
[128] At 95.

capable of living on their own safely.[129] Even with a 'section 8 voucher' for housing, which provides a subsidy to be paid directly to the landlord by a Public Housing Agency, private landlords may be unwilling to lease property to those most downtrodden and in the most desperate need. Section 8 vouchers expire after six months, and the applicant will have to re-apply all over again if they have been unable to find housing.[130] Even for those homeless who have had no assistance through the co-ordinated entry process, their personal information remains in the database.[131]

Eubanks argues that the real need is for greater resources. Co-ordinated entry might increase efficiencies, but it does not solve homelessness. Eubanks makes a critical argument when she states that:[132]

> ... the pattern of increased data collection, sharing, and surveillance reinforces the criminalization of the unhoused, if only because so many of the basic conditions of being homeless - having nowhere to sleep, nowhere to put your stuff, and nowhere to go to the bathroom - are also officially crimes.

If you can get a ticket for sleeping in a public park or leaving your possessions on the sidewalk and you have no way of paying those fines, those tickets turn into warrants and law enforcement then has reason to search the databases:[133]

> Thus, data collection, storage, and sharing in homeless service programs are often starting points in a process that criminalizes the poor.
> ...
> Coordinated entry is not just a system for managing information or matching demand to supply. It is a surveillance system for sorting and criminalizing the poor.

Eubanks argues that "in new data-based surveillance, *the target often emerges from the data.* The targeting comes after the data collection, not before".[134] The co-ordinated entry

---

[129] At 107.
[130] A Section 8 voucher allows for a subsidy to be paid directly to the landlord by a Public Housing Agency https://www.benefits.gov/benefit/710.
[131] Eubanks, above n 103, at 108.
[132] At 117.
[133] At 117-121.
[134] At 122.

system harms those most vulnerable who need to disclose their risky behaviours in order to obtain housing assistance. Eubanks critiques the reasoning of those in favour of the co-ordinated entry system and their enthusiasm to apply technical 'fixes' to complex social problems:[135]

> These perspectives assume that complex controversies can be solved by getting correct information where it needs to go as efficiently as possible. In this model, political conflict arises primarily from a lack of information. If we just gather all the facts, systems engineers assume, the correct answers to intractable policy problems like homelessness will be simple, uncontroversial and widely shared.
>
> But, for better or worse, this is not how politics work. Political contests are more than informational; they are about values, group membership, and balancing conflicting interests.

The co-ordinated entry system described by Eubanks as a response to homelessness reveals a lack of understanding of the complexities that lead to people ending up homeless, as well as a failure to acknowledge the conflicting interests of the citizens of Los Angeles and their reluctance to spend the large amounts of money necessary to properly solve the housing crisis,[136] a scenario that is not unique to Los Angeles but repeated in cities around the world. Eubanks references the work of Gary Blasi, a homeless advocate and emeritus professor of law at University of California, Los Angeles, who believes that algorithms cannot handle the nuance and complexity that human beings present.[137] Blasi's argument is that co-ordinated entry systems draw resources and attention from other aspects of the problem of homelessness.[138]

In certain contexts, such as the allocation of housing or other scarce social resource, an algorithm cannot capture all the unknowable variables that are at play for the individuals that are subject to their outputs. An element of human discretion is needed (even if this does unavoidably allow for human bias to creep in). The alternative is a system that lacks flexibility to respond in a way that treats people with dignity; as individuals whose

[135] At 124.
[136] At 124.
[137] At 125.
[138] At 125.

circumstances cannot always be understood and assessed by algorithm. Also problematic is where the deployment of algorithmic decision-making tools distracts from the need for adequate resourcing, or is used to justify reduced expenditure, in underfunded social services.

Predictive Risk Modelling in Child Protection Services

Allegheny County, United States

Eubanks provides a third case study critiquing the use of predictive risk modelling in child protection services. In 2012, the Department of Human Services in Allegheny County had its funding cut by 10 per cent, aggravating an already critical situation in which demand for services had risen following a recession.[139] In response to the need to reduce expenditure and increase efficiencies, it was decided that decision support tools and predictive analytics would be implemented to direct resources to where they could do the most good. The chosen proposal was one submitted by a team from Auckland University of Technology, led by economist Rhema Vaithianathan, and Emily Putnam-Hornstein, director of the Children's Data Network at the University of Southern California. Their proposal aimed to put in place a decision-making tool that would mine the Allegheny County data warehouse to make predictions about which children might be at greatest risk of abuse or neglect.[140] The data warehouse holds more than one billion electronic records - an average of 800 records for each person in Allegheny County. Data is regularly sent from social agencies including the police, probation, drug and alcohol services, housing authority, jail, office of income maintenance, office of mental health and substance abuse services, and schools.[141]

A Similar Risk Model is Commissioned in New Zealand

Also in 2012, New Zealand's Minister for Social Development and Employment, Paula Bennett, commissioned Vaithianathan's team to create an algorithm to predict which children were most likely to be abused or neglected, as part of a broader package of

[139] At 136.
[140] At 136-137.
[141] At 134.

welfare reforms.[142] The algorithm was designed to target voluntary early intervention programmes to families of high-risk children whose parents or guardians were welfare recipients.[143] The predictive model draws on 132 variables, including length of time on benefits, past involvement with the child welfare system, mother's age, if the child was born to a single parent, mental health, and correctional history.[144] Vaithianathan's team found their core algorithm had a "fair, approaching good" level of accuracy in predicting the risk of substantiated findings of maltreatment of children.[145]

In 2015, Social Development Minister, Anne Tolley, who had replaced Paula Bennett the year before, halted the plan to launch the two-year observational experiment to risk-rate 60,000 newborn babies to test the accuracy of the tool. Tolley notably annotated the proposal with: "not on my watch! These children are not lab rats".[146] The main issues with the use of predictive risk modelling (PRM) in child protection services are the stigma that attaches to those families and children with a high-risk rating; the lack of evidence that these tools improve real-world outcomes for children; privacy concerns, and the risk of bias.[147]

Vaithianathan's team recognised significant ethical issues associated with the implementation of their PRM, but did not believe these posed an "insurmountable barrier" to its implementation.[148] The report, titled *Vulnerable Children: Can Administrative Data be Used to Identify Children at Risk of Adverse Outcomes?* states that there is a need for confidentiality and sensitivity, and to "ensure that the benefits are large enough to warrant the stigmatisation and the false-positives that might be inevitable from a risk assessment".[149]

---

[142] Ministry of Social Development "White Paper for Vulnerable Children" (2012) https://www.beehive.govt.nz/feature/white-paper-vulnerable-children
[143] Ministry of Social Development, above n 142.
[144] The full list of variables has not been published.
[145] Rhema Vaithianathan, Tim Maloney, Nan Jiang, Irene De Haan, Clare Dale, Emily Putnam-Hornstein, Tim Dare "Vulnerable Children: Can Administrative Data be Used to Identify Children at Risk of Adverse Outcomes?" (Centre for Applied Research in Economics, University of Auckland, 2012) at 15.
[146] Stacey Kirk "Children - 'not lab rats'. Anne Tolley intervenes in child abuse experiment" (30 July 2015) Stuff https://www.stuff.co.nz/national/health/70647353/children-not-lab-rats---anne-tolley-intervenes-in-child-abuse-experiment
[147] New Zealand Herald "Predicting Trouble: Child abuse database raises eyebrows" (20 October 2012) https://www.nzherald.co.nz/nz/predicting-trouble-child-abuse-database-raises-eyebrows/UXLTJHIEKWAF5JDSZA7MPAHKQM/?c_id=1&objectid=10841709
[148] Vaithianathan et al, above n 145, at 32.
[149] At 33.

However, with no child protection services anywhere in the world using PRM at the time the report was written,[150] it would be very difficult to ensure that the benefits of the tool would be sufficient to outweigh its harms. Britain created a national database of all children in 2004 and disbanded it in 2010 due to concerns about accuracy and security of data.[151] The PRM tool would have been revolutionary if implemented, but there was no evidence of its efficacy. Dr Patrick Kelly, head of the child abuse unit at Starship Hospital, was part of an expert forum looking at the high rates of child abuse in New Zealand.[152] Dr Kelly told the New Zealand Herald that the danger of a revolution is that if it is not successful, it can undo a lot of what is working in the process. He stated his full support for more information sharing between relevant government agencies, but believed that could be done without expenditure on new IT.[153]

Vaithianathan and two members of her team, Tim Dare and Irene De Haan, argue in favour of implementing PRM in child protection services, particularly in the absence of multi-party consensus on the necessary measures to tackle child poverty.[154] They refer to "a powerful ethical imperative to find a way to protect the most vulnerable members of our community".[155] They believe that significant ethical concerns associated with the use of PRM are either outweighed by the potential benefits or can be mitigated by appropriate implementation strategies.[156] These include addressing the stigma of being identified as high-risk by ensuring that interventions are at the minimum level necessary; information is disseminated as narrowly as possible; and for interventions to be preventative and supportive, not punitive.[157]

Vaithaianathan et al argue that PRM should only be used as an opportunity to offer additional services to high-risk children, not to reduce existing universal services.[158] They question why critics of PRM believe that it poses greater risk of under-resourcing than

[150] At 30.
[151] New Zealand Herald, above n 147.
[152] New Zealand Herald, above n 147.
[153] New Zealand Herald, above n 147.
[154] Tim Dare, Rhema Vaithianathan, Ireene De Haan "Addressing Child Maltreatment in New Zealand: Is Poverty Reduction Enough?" (2014) Educational Philosophy and Theory 46:9.
[155] At 991.
[156] At 991.
[157] At 991-992.
[158] At 992.

the status quo. Eubanks' discussion of the use of PRM in the United States reveals that these tools are often implemented as cost saving strategies.[159] I question why the PRM was commissioned *before* the resources and staffing to implement any new support services were in place.[160] Was it because the commissioning of the tool was primarily motivated by the desire to reduce existing expenditure? Or perhaps the predictive risk model was being considered for purposes other than the targeting of additional voluntary services? A large part of the problem lies with the rationale behind implementing such a tool. However, the question of whether predictive risk modelling tools perform a comparable, or better, job than human case workers has not been comprehensively addressed. I question whether such tools will ever be agile enough to take into consideration all the varied and unknowable factors that will impact on such critical decisions for human beings. Even if we had a tool that was more accurate than the human decision-makers it replaces, do you allow case worker overrides in order to provide the flexibility required in human contexts, but in doing so allow the opportunity for human bias to creep in?

Associate Professor of Social Work at the University of Otago, Emily Keddell, warns that the individualisation of risk "clearly represents a neo-liberal concern with personal responsibility and a limited role of the nation state".[161] The focus on the risk factors of the individual is at the expense of the societal factors that contribute to child abuse.[162] Keddell argues that when too much information sharing happens too quickly, it risks leading to overly intrusive interventions that shut down the potential for other more supportive options for families under stress, damaging the critical relationship between social worker and family.[163]

---

[159] See paragraphs on Automated welfare eligibility in Indiana; Electronic registry of homeless in Los Angeles; and Predictive risk modelling in child protective services: Allegheny County, United States where I discuss Eubanks' critique of high tech tools used to criminalise the poor.
[160] The *Vulnerable Children* report uses an American initiative called the Nurse Family Partnership Program in its risk modelling calculations (at 19). The initiative appears very similar to Plunket which already offers support to all families of preschool children in New Zealand, arguably calling into question the validity of the calculations in the report.
[161] Emily Keddell "The ethics of predictive risk modelling in Aotearoa/New Zealand child welfare context: Child abuse prevention or neo-liberal tool?" (2004) Critical Social Policy at 82.
[162] At 76.
[163] At 74.

Restricting the scope of the tool to children whose families are receiving benefits reinforces the association between welfare and child abuse, and contributes to surveillance of the poor, highlighted as particularly problematic by Eubanks. While Keddell acknowledges that the tool's focus on prevention is commendable, "the nature of the data informing the model must be carefully considered".[164] The variables included in the model would result in a disproportionate number of Māori tamariki (children) being identified as at high-risk, contributing to existing structural inequities.[165] Keddell states there is a need to go beyond the welfare and child protection data used by the tool[166] (an issue acknowledged by Vanaithanan's team in their report).[167]

Keddell argues that recognition of poverty as a significant contributor to a range of poor outcomes is fundamental. Also, implementing the proven best practice of providing universal services which can then be used as a platform from which to offer targeted support services to families. This approach lessens stigma, particularly if the targeted support comes from the same service provider that offers the universal services.[168]

No database is ever 100 per cent secure. However, public agencies in New Zealand have a particularly poor track record of securely storing personal information.[169] I believe the individuals subject to the PRM (with sensitive personal information stored in its database) would be justified in having concerns about its security and integrity. Even if concerns over storage and security could be mitigated, the issue of enduring digital data remains.

---

[164] At 78.
[165] At 76.
[166] At 78.
[167] Vaithianathan et al, above n 145, at 31.
[168] Keddell, above n 161, at 83.
[169] Multiple government agencies have experienced data privacy scandals over the last decade. Some of which include, an ACC employee breaches client privacy in March 2022: https://www.rnz.co.nz/news/national/463044/acc-staffer-breached-client-privacy-after-snooping-in-sensitive-claim-review-authority; In 2021 ACC employees share personal information of clients in snapchat group https://www.acc.co.nz/newsroom/stories/investigation-into-sharing-client-information-has-been-completed/; The Ministry of Health Tū Ora data breach in August 2019: https://www.health.govt.nz/about-ministry/information-releases/general-information-releases/tu-ora-compass-health-cyber-security-incident-further-information; IRD breached clients' privacy in 2018 and 2012: https://www.nzherald.co.nz/business/ird-says-sorry-for-privacy-breach/54MDLADNVTX2HHIEKPI3URJBEQ/; https://www.rnz.co.nz/national/programmes/ninetonoon/audio/2018668093/ird-privacy-breach-raises-data-handling-concerns; Work and Income NZ disclose contact details of more than 100 clients in 2021: https://www.newshub.co.nz/home/new-zealand/2021/08/work-and-income-accidentally-reveals-more-than-100-clients-personal-info-to-each-other.html; Work and Income NZ employee emails personal information of 34 beneficiaries to another claimant: https://www.databreaches.net/winz-privacy-breach-major-stuff-up/

There is the possibility that risk scores could pop up to haunt individuals in the years to come in unknown and harmful ways. There is no guarantee that this data would not be used in different contexts or for a different purpose in the future.

While it is counterintuitive to argue against implementing a tool that is designed to prevent child abuse and neglect, it is easier to do so if there is no evidence that the PRM tool improves real-world outcomes for children. Its implementation could result in the harm of stigmatising families and children, as well as the associated stress on parents that accompanies this type of government surveillance. Additionally, there is the potential for PRM scores to be used for other purposes and the risk of sensitive personal information being compromised.

A recurring theme of this thesis is, in the words of Evgeny Morozov, the "folly of technological solutionism",[170] evidenced in the desire of public agencies to adopt technical solutions as the sole response to complex, multifaceted social problems. A range of factors contribute to child maltreatment, including economic deprivation, drug and alcohol abuse by parents or caregivers, lack of education, and intergenerational family violence.[171] There is a growing body of research in New Zealand and overseas that shows that high rates of child abuse and neglect correspond with poverty and socioeconomic inequality.[172] Ultimately, the advice of Dr Kelly, and his expert forum to tackle high rates of child abuse in New Zealand, was followed with an Information Sharing Agreement for Improving Public Services for Vulnerable Children signed on 25 June 2015, to facilitate information sharing between public agencies to better identify vulnerable children;

---

[170] Evgeny Morozov *To Save Everything, Click Here: The Folly of Technological Solutionism* (Perseus Books Group, United States of America, 2013).
[171] Denise Wilson and Melinda Webber "The People's Report: The People's Inquiry into Addressing Child Abuse and Domestic Violence" (2014) The Glenn Inquiry.
[172] The Child Poverty Action Group website states that: "Although child abuse is the result of a range of interacting factors, there is an ever-expanding body of evidence showing poverty and inequality significantly increase the risk of child maltreatment and neglect. The Avon Longitudinal Study of Parents and Children (2001) findings suggest that the greater a family's material deprivation, the greater the risk of child maltreatment. A 2014 study found a strong association between higher income inequality among United States counties and higher levels of child abuse and neglect in those counties". https://www.cpag.org.nz/child-abuse-and-neglect-the-links-with-inequality/;
The following report's findings also contribute to the growing literature linking income inequality with child maltreatment and other poor health and wellbeing outcomes: John Eckenrode, Elliott G. Smith, Margaret E. McCarthy, Michael Dineen "Income Inequality and Child Maltreatment in the United States" (2014) Pediatrics 133 3.

protect them from harm; and to promote the wellbeing of vulnerable children and their families.[173]

Allegheny Family Screening Tool

Unsurprisingly, any plan to trial the PRM tool in New Zealand was quickly abandoned after the media furore over its proposed use. However, by this time, Vaithianathan's team had won the contract to create a risk prediction model for Allegheny County.[174] The Allegheny Family Screening Tool (AFST) produces a risk score for each notification of suspected child maltreatment received by the Department of Human Services hotline. The AFST risk score assists hotline workers to determine if referrals should be screened in or out for investigation.[175] Eubanks states that the AFST has inherent flaws that reduce its accuracy:[176]

> It predicts referrals to the child abuse and neglect hotline and removal of children from their families - hypothetical proxies for child harm - not actual child maltreatment. The data set it utilizes contains only information about families who access public services, so it may be missing key factors that influence abuse and neglect. Finally, its accuracy is only average. It is guaranteed to produce thousands of false negatives and positives annually.

Eubanks highlights how the outcome variables of the model are subjective. She states:[177]

> Was a parent re-referred to the hotline because she neglects her children? Or because someone in the neighbourhood was mad that she had a party last week? Did caseworkers and judges put a child in foster care because his life was in danger? Or because they held culturally specific ideas about what a good parent looks like, or feared the consequences if they didn't play it safe.

---

[173] The Information Sharing Agreement was authorised under Part 9A of the Privacy Act 1993 https://orangatamariki.govt.nz/assets/Uploads/Working-with-children/Information-sharing/Information-Sharing-Agreement-for-Improving-Public-Services-for-Vulnerable-Children.pdf
[174] Eubanks, above n 103, at 138.
[175] Department of Human Services, Allegheny County "Allegheny Family Screening Tool: Predictive Risk Modelling in Child Welfare in Allegheny County" https://www.alleghenycounty.us/Human-Services/News-Events/Accomplishments/Allegheny-Family-Screening-Tool.aspx
[176] Eubanks, above n 103, at 145-146.
[177] At 146.

There is a lot of data on people who use *public* services in the Allegheny data warehouse, but the County has no access to data about parents who access *private* drug treatment services or mental health counselling.[178] Referral bias in the community results in more calls to child abuse and neglect hotlines about Black and biracial families than white.[179] Nuisance calls introduce contaminated data into the model and compromise its accuracy.[180] The AFST views the use of public services as a risk factor for child maltreatment and a quarter of the predictive variables are directly linked to poverty.[181] Eubanks states that the "overwhelming majority of child welfare investigations in the United States involve neglect not abuse".[182] Eubanks questions where the line is drawn between poverty and neglect. Is it a lack of food, cold and damp housing, homelessness, a lack of healthcare..?[183]

It is important to avoid the trap of comparing a flawed algorithm to a hypothetical (and non-existent) perfect human decision-maker. However, Eubanks makes the critical point that predictive risk models incorporate the discretion and bias of a few software engineers whereas the discretion and bias of social workers in the field is varied and many. She states that "the automated discretion of predictive models is the discretion of the few. Human discretion is the discretion of the many. Flawed and fallible, yes. But also flexible".[184] Social workers, like all humans, are at risk of making biased decisions. However, they are also capable to responding to changing cultural norms and expectations and can adapt their viewpoints as society evolves. A few social workers may hold outdated and unacceptable views on minority groups or make unsubstantiated judgements based on gender, race, or socioeconomic status, but their decisions will be reviewed by other human decision-makers (colleagues and team leaders) and not set in a predictive model that impacts on all decision-making in that context.

A concern associated with targeting high-risk families is that it might see them withdraw from services offering support, particularly when predictive risk models view parents who access such services as a danger to their kids. Eubanks believes it would be horribly ironic

[178] At 146-147.
[179] At 153.
[180] At 154.
[181] At 156.
[182] At 156.
[183] At 156-157.
[184] At 168.

if the AFST contributed to the abuse it was designed to prevent.[185] The risk factors for child abuse and neglect all increase when parents feel the stress of being watched all the time. Eubanks states that:[186]

> It is difficult to say if a predictive model works if it produces the outcomes it is trying to measure. A family scored as high-risk by the AFST will undergo more scrutiny than other families. Ordinary behaviors that might raise no eyebrows before a high AFST score become confirmation of a decision to screen them in for investigation.

However, Eubanks concedes that the AFST may be the "best-case scenario" for predictive risk modelling because it is designed to support human decision-making. Similar models have been introduced in other U.S. states, but this one is at least open in its design; it is participatory and transparent. Eubanks states that in other states similar systems have been designed by private companies and deployed without public consultation.[187] Eubanks raises the concern of the AFST and other predictive risk models being misused in times of fiscal austerity or in response to child neglect scandals: she asks whether a very high-risk score could, one day, be used to justify pre-emptively taking a child from their home?[188]

Eubanks formulates two critical questions that should be asked of automated tools in any humane society. First, 'does the tool increase the self-determination and agency of the poor?' Second, 'would the tool be tolerated if it was targeted at non-poor people?'[189] I believe that these are pertinent questions that should be asked of every algorithm used to assist in making significant decisions about people's lives.

Race and Gender Discrimination

Race and gender discrimination can be exacerbated by big data-enabled technology. Safiya Noble explains how search engines facilitate, and embed, discrimination and asks how Google can fail to provide reliable or credible information about women and people

[185] At 168.
[186] At 168.
[187] At 171.
[188] At 172.
[189] At 211-212.

of colour and face no repercussions for this.[190] Noble believes an algorithmic web search that "offers up racism and sexism as the first results ... reflects a corporate logic of either wilful neglect or a profit imperative that makes money from racism and sexism".[191]

Noble believes that the way in which commercial search engines currently operate does not present "appropriate social, historical, and contextual meaning to already overracialized and hypersexualized people".[192] The problematic ways that women have been represented in print advertising are reproduced in online search advertising, which Noble says is unsurprising. She states that:[193]

> Of course, this makes sense, because Google search is in fact an advertising platform, not intended to solely serve as a public information resource in the way that, say, a library might. Google creates advertising algorithms, not information algorithms.

In her book *Race After Technology: Abolitionist Tools for the New Jim Code*[194] Ruha Benjamin describes the "New Jim Code" (a play on words that refers to the Jim Crow laws that enforced racial segregation in the U.S. in the late nineteenth and early twentieth centuries). The New Jim Code is "the employment of new technologies that reflect and reproduce existing inequities but that are promoted and perceived as more objective or progressive that the discriminatory systems of a previous era".[195]

Benjamin provides the shocking example of an audit of California's gang database that revealed that not only were 87 per cent of those listed Blacks and Latinas, but many of the names also belonged to babies, some of whom were listed as "self-described gang members".[196] Benjamin states that:[197]

---

[190] Safiya Umoja Noble *Algorithms of Oppression: How Search Engines Reinforce Racism* (NYU Press, New York, 2018) at 3-4.
[191] At 5.
[192] At 36.
[193] At 38.
[194] Ruha Benjamin *Race After Technology: Abolitionist Tools for the New Jim Code* (Polity Press, Cambridge, 2019) at 197.
[195] At 5-6.
[196] At 6.
[197] At 6.

> So far, no one ventures to explain how this could have happened, except by saying that some combination of zip codes and racially coded names constitute a risk. Once someone is added to the database, whether they know they are listed or not, they undergo even more surveillance and lose a number of rights.

Benjamin highlights this as an example of where racism can be facilitated and exacerbated by technologies that appear to be objective and neutral. Technology can be racist without a conscious intention for it to be so. Machine learning for example, relies on large real-world datasets that are steeped in racial bias. Robots use this data to learn, consuming "deeply ingrained cultural prejudices and structural hierarchies".[198] Robots learn to speak the coded language of the software programmers who create them, as well as the language of everyone online who contributes to the datasets from which they learn.[199] Benjamin explains how the more "intelligent" a machine becomes the more likely it is to become racist because of the data it feeds on: "robots, designed in a world drenched in racism will find it nearly impossible to stay dry".[200]

However, Benjamin believes that this is not inevitable if the racism that structures the social and technical elements of the design of technology is addressed and taken seriously;[201] technology companies simply employing individuals from minority groups will not change anything.[202] Benjamin examines how 'whiteness' is treated as the norm and everything else as 'other', such as the automated soap dispenser that only recognises white skin. The concern is that if technology is viewed as value-free then it is less likely to be questioned.[203] And if it is not questioned, unjustified discrimination can go unchecked.

Big Data in Policing and Criminal Justice

The problem of racial discrimination is particularly pertinent in the context of big data in policing and criminal justice. Hannah Fry states that the use of algorithms by Police can help to identify violent criminals - there are recognisable patterns in where crimes are

[198] At 59.
[199] At 62.
[200] At 62.
[201] At 59.
[202] At 62.
[203] At 66-69.

committed.[204] Fry argues that "crime is not random. People are predictable".[205] For example, the concept of "distance decay" relates to the chances of finding an offender's home decreasing the further investigators move away from a crime scene.[206] Also, a "buffer zone" describes the concept of offenders being unlikely to target victims who live very near them.[207] Algorithms that take this kind of geographical information into account can be used to prioritise existing lists of suspects. Fry states that they are not relied on too heavily but can make an investigation more efficient.[208] I argue that when used in this way, these types of algorithms does not raise issues of racial discrimination or exacerbate existing social inequities, unlike other algorithms used in the criminal justice system.

Predpol

Location-based predictive policing algorithm, Predpol, shows police on a digital map where, when, and what type of crime is likely to occur.[209] Police can then focus their attention on those neighbourhoods that Predpol identifies as high crime areas. It uses historical crime data to direct patrol operations to areas where it predicts crime will occur, based on what the data reveals has happened in the past. However, this algorithm has been accused of contributing to the outcomes it predicts.[210] For example, in places where you have police on the ground you are more likely to detect crime. Even if there are two geographical areas where the same amount of crime is occurring, the police will detect more crime in the area that they are in as opposed to the area where they are not present. This can create a feedback loop. Fry states that:[211]

> There are ways, theoretically at least, to ensure that the algorithm doesn't disproportionately target particular neighbourhoods – like randomly sending police to medium-risk areas as well as high-risk ones. But, unfortunately, there's no way to know for sure whether PredPol is managing to avoid these feedback loops entirely, or indeed whether it is operating more fairly generally, because Predpol is a

[204] Hannah Fry *Hello World: How to Be Human in the Age of the Machine* (Transworld Publishers, Penguin, London, 2018) at 141-173.
[205] At 144.
[206] At 144.
[207] At 144.
[208] At 146.
[209] Predpol https://www.predpol.com
[210] Fry, above n 204, at 156-157.
[211] At 157.

> proprietary algorithm, so the code isn't available to the public and no one knows exactly how it works.

There is also the concern that predictive policing algorithms like Predpol target high crime areas which are predominantly populated by Black people. These algorithms are trained on biased data - Black men are more likely to be stopped by Police than white men, resulting in Black men being more likely to be found carrying contraband and therefore more likely to be arrested. These predictive algorithms run the risk of entrenching existing inequities through the over-policing of Black and poor neighbourhoods. Fry believes that:[212]

> Predictive policing algorithms undoubtedly show promise … But the concerns around bias and discrimination are legitimate. And for me these questions are too fundamental for a just society for us simply to accept assurances that law enforcement agencies will use them in a fair way. It is one of many examples of how badly we need independent experts and a regulatory body to ensure that the good an algorithm does outweighs the harm.

This raises the issues of fairness, best practice, audit, and oversight of predictive systems. I look at the use of independent experts and regulatory bodies in Chapter Six and consider their effectiveness as a safeguard against unjustified bias and discrimination in algorithmic tools.

Algorithmic Recidivism Models

Algorithmic recidivism models have been introduced into the criminal justice system in the U.S. in an attempt to reduce bias and inconsistency and to make more efficient use of limited resources by improving accuracy in predicting recidivism. However, these models have raised issues of fairness and transparency, particularly when used outside of correctional facilities (for which they were originally designed) but in the courtroom as an aid to judges' decision-making in sentencing.

[212] At 158-159.

Cathy O'Neill explains how these models use data from lengthy questionnaires that the defendant or prisoner completes. Some questions are relevant to the risk of recidivism such as – *how many prior convictions have you had?* Or *what part did others play in the offence? What part did drugs and alcohol play?* But O'Neill states that as the questions continue, digging deeper into the person's life, you can start to see how people from privileged backgrounds will answer one way and those from poorer neighbourhoods another.[213] Questions such as – *when was the first time you were ever involved with the police?* And *do your friends and family members have criminal records?* In the U.S., Black and Latino men are much more likely to have been stopped by the police for stop-and-frisk checks than white men.[214] If early contact with police signals recidivism, then it follows that minority groups and the indigent look much riskier.[215]

COMPAS

Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) is an example of an algorithmic risk assessment tool originally designed as a decision support for correctional agencies in determining placement decisions, offender management and treatment planning.[216] COMPAS was adopted for use in sentencing to address increasing levels of incarceration and overcrowding in U.S. jails. COMPAS promises to assist in reducing prison populations while at the same time ensuring that offenders most likely to commit crimes are incarcerated.

The focus of COMPAS is assessing risk to reduce recidivism,[217] which is an appropriate focus for a tool used by a corrections department. However, it raises questions of fairness when used as a decision-making aid in sentencing. When sentencing a defendant for the crime he has committed, is it fair to consider a prediction of the likelihood of *future offending*? Particularly, when the prediction is based not on the *individual's* likelihood of reoffending, but on a group likelihood of re-offending, of which the individual is a member due to shared characteristics with other members of the group. O'Neill argues

[213] O'Neill, above n 81, at 25.
[214] At 25-26.
[215] At 26.
[216] equivant Northpointe Inc. "Practitioner's Guide to COMPAS Core" (2019). https://www.equivant.com/wp-content/uploads/Practitioners-Guide-to-COMPAS-Core-040419.pdf
[217] equivant Northpointe Inc., above n 216.

that a defendant should be judged on what he has done, not on who he is, and therefore, what he is expected to do in the future.[218]

State v. Loomis

In the Supreme Court of Wisconsin in *State v. Loomis,*[219] Eric Loomis challenged the use of the risk assessment portion of the COMPAS report at his sentencing. He argued that the consideration of the COMPAS risk assessment violated his constitutional right to due process for three reasons:

(1) the inability to assess COMPAS's accuracy;

(2) it violated his right to an individualised sentence;

(3) it improperly used gendered assessments in sentencing.

The court recognised that sentencing decisions are guided by due process protections that may not necessarily apply to other correctional decisions, therefore care needs to be taken when importing a risk assessment tool designed for use by Corrections.[220] However, the court held that Loomis failed to show that his due process protections were in fact violated. It held that, if used properly, with an awareness of the need for cautions and inherent limitations, a circuit court's consideration of a COMPAS risk assessment at sentencing does not violate a defendant's right to due process. The court held that:[221]

> Although it cannot be determinative, a sentencing court may use a COMPAS risk assessment as a relevant factor for such matters as: (1) diverting low-risk prison-bound offenders to a non-prison alternative; (2) assessing whether an offender can be supervised safely and effectively in the community; and (3) imposing terms and conditions of probation, supervision, and responses to violations.

The use of COMPAS in sentencing is premised on the necessity of a judge exercising discretion. In *Loomis* the court held that consideration of COMPAS risk score in

[218] O'Neill, above n 81, at 26.
[219] *State v. Loomis* 881 N.W.2d 749 (2016).
[220] At para 41.
[221] At para 88.

sentencing is permissible so long as it is not the only factor relied upon by judges in making their decision. However, the fact that a judge also considers other factors does not rule out the possibility of harm from using the tool. Sonja Starr states that "anything treated as a sentencing factor will at least sometimes solely trigger a change in the sentence relative to what it would otherwise have been".[222]

The concept of automation bias suggests that judges may be more likely to place significance on risk predictions that are drawn from data and perceived to be scientific.[223] It is difficult for human decision-makers to override recommendations suggested by an algorithm. Starr believes that judges are likely to follow the predictions suggested by risk assessment tools. An algorithmic assessment seems more reliable and objective.[224] There is the potential for judges to be reluctant to disregard a COMPAS high-risk score even when other factors suggest that the classification is incorrect. A 2006 report on juvenile detention risk assessment revealed that "detain overrides" are more frequent than "release overrides".[225] There is a lack of empirical research on how risk prediction instruments affect judges' weighing of recidivism risk versus other factors. Nevertheless, in other contexts, it has been shown that decision-makers defer to scientific or technical models.[226] However, there is also the contrasting notion of algorithmic aversion where humans distrust an algorithm (even where proven to be more accurate than human decision-makers). People are not good at judging the accuracy of algorithms and deciding when to overrule them; this can lead human decision-makers to override algorithmic outputs in detrimental ways.[227]

The accuracy of COMPAS has been called into question. ProPublica's extensive investigation into COMPAS's predictions of recidivism found that only 20 per cent of people predicted to commit violent crimes ended up doing so, and when taking into

[222] Sonja Starr "Evidence-Based Sentencing and the Scientific Rationalization of Discrimination". (2014) Stanford Law Review 66 4 at 863.
[223] At 866.
[224] At 866-867.
[225] The Annie E. Casey Foundation "Juvenile Detention Risk Assessment: A Practice Guide to Juvenile Detention Reform" (2006) https://assets.aecf.org/m/resourcedoc/aecf-juveniledetentionriskassessment1-2006.pdf
[226] John Zerilli, Alistair Knott, John Maclaurin, Colin Gavaghan "Algorithmic Decision-Making and the Control Problem" (2019) Minds and Machines 29 at 556-557.
[227] Ben Green "The flaws of policies requiring human oversight of government algorithms" (2022) Computer Law and Security Review 45 at 7.

consideration all crimes, only 61 per cent of those were arrested within two years.[228] Another study concluded that "COMPAS is no more accurate or fair than predictions made by people with little or no criminal justice experience" and comparable accuracy could be achieved by considering two features as opposed to the 137 features that COMPAS considers.[229] Sonja Starr states that "when it comes to predicting individual behaviour, the model offers fairly modest improvements over chance".[230] Starr argues that the wrong question is being asked by recidivism models: "what judges need to know is not just how "risky" the defendant is in some absolute sense, but rather how the sentencing decision will **affect** his recidivism risk".[231]

O'Neill makes a similar point that recidivism models are logically flawed. She states that:[232]

> The unquestioned assumption is that locking away "high risk" prisoners for more time makes society safer. It is true, of course, that prisoners don't commit crimes against society while behind bars. But is it possible that their time in prison has an effect on their behavior once they step out? Is there a chance that years in a brutal environment surrounded by felons might make them more likely, and not less, to commit another crime?

Defendants have the right not to be sentenced on the basis of inaccurate information. Precedents in Wisconsin case law emphasise the importance of accuracy in sentencing, and more importantly, the defendant's right to be able to assess that accuracy himself.[233] In *Loomis* the court held that the score and the report itself were not hidden from Loomis;

---

[228] Julia Angwin, Jeff Larson, Surya Mattu, Lauren Kirchner "Machine Bias: There's software used across the country to predict future criminals and its's biased against blacks" (2016) Propublica https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing
[229] Julia Dressel and Hany Farid "The accuracy, fairness, and limits of predicting recidivism" (2018) Science Advances https://www.science.org/doi/epdf/10.1126/sciadv.aao5580
[230] Starr, above n 222, at 806.
[231] At 807.
[232] O'Neill, above n 81, at 97.
[233] In *Gardner v. Florida* 430 U.S. 349, 97 S Ct. 1197 (1977) it was held that withholding information in the pre-sentence investigation report from the defendant was a violation of his due process rights. In *State v Skaff* 152 Wis. 2d 48 (Wis. Ct. App. 1989) the pre-sentencing investigation report was provided to the defendant's counsel with the instruction that he was not to allow the defendant to read it. The court held that to deny the defendant access was to deny him due process. The defendant plays an important role in assessing the accuracy of information relied on in sentencing. - Katherine Freeman "Algorithmic Injustice: How the Wisconsin Supreme Court Failed To Protect Due Process Rights In State v. Loomis" (2016) North Carolina Journal of Law & Technology 75 at 86-87.

Loomis provided the answers to the COMPAS questionnaire himself and that the other inputs determining the COMPAS score are a matter of public record. However, this reasoning fails to acknowledge that Loomis was denied the means to ascertain if there was any misinformation. Northpointe/equivant does not disclose the weight various factors are given by COMPAS, therefore the defendant is also denied the opportunity to question the significance placed on each input. Due to its opacity, COMPAS's source code cannot be scrutinised by independent organisations or academics, and its accuracy cannot be independently audited. Even the Supreme Court was not allowed access to its workings. A defendant cannot appeal a COMPAS score, nor hold its designers accountable for the harm it causes. While the exercise of judicial discretion without the aid of algorithms is not perfect, the decisions of judges can at least be analysed and are open to public criticism. Judicial rulings can be appealed, but there is no way of appealing a COMPAS score when its inputs are kept secret.

Bias and Competing Notions of Fairness

ProPublica, a group of investigative journalists in the U.S., argues that COMPAS is biased against Blacks because among those defendants who did *not* reoffend, Blacks were more than twice as likely as whites to be classified as medium or high risk. Northpointe/equivant argues that COMPAS is not biased because the scores mean the same thing regardless of race. Within each risk category, the proportion of defendants who reoffend is approximately the same, regardless of race.

Both arguments are correct due to competing notions of fairness. Corbett-Davies et al describe competing notions of fairness in the following way:[234]

- Within each risk category, the proportion of defendants who reoffend is approximately the same regardless of race; this is Northpointe's definition of fairness.

[234] Sam Corbett-Davies, Emma Pierson, Avi Feller, Sharad Goel "A computer program used for bail and sentencing decisions was labeled biased against Blacks. It's actually not that clear" (17 October 2016) The Washington Post https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/

- The overall recidivism rate for black defendants is higher than for white defendants (52 per cent vs. 39 per cent)

- Black defendants are more likely to be classified as medium or high risk (58 per cent vs. 39 per cent) ...

- Black defendants who don't reoffend are predicted to be riskier than white defendants who don't reoffend; thus ProPublica's criticism of the algorithm.

  The key - but often overlooked - point is that the last two disparities in the list above are mathematically guaranteed given the first two observations.

Corbett-Davies et al believe that it would violate the principle of equal treatment if a risk score for Black defendants did not mean the same thing as a risk score for white defendants.[235] The issue of competing notions of fairness is complex and requires considered deliberation of relevant arguments. It is not a question that should be determined for society solely by computer scientists or software engineers.

Due to the opacity of COMPAS, additional biases cannot be ruled out, including bias against indigent people. It is arguable that the use of COMPAS compromises the individual's right to an individualised sentence. COMPAS calculates risk scores based on group data. It asks questions in relation to socioeconomic factors, like education and employment, neighbourhood, family criminality, and friends. People who live in high crime, lower socioeconomic neighbourhoods are penalised because of this.

O'Neill argues that risk assessment tools which judge a defendant on factors that would not be admissible in court are unfair.[236] Many people benefit from the implementation of risk assessment tools that sentence low-risk criminals to shorter spells in jail. Nevertheless, where a risk assessment tool is designed as a decision support for correctional agencies in determining placement decisions, offender management, and treatment planning, it can take into account factors beyond an individual's control. However, when used in sentencing (as opposed to targeted treatment interventions) these

[235] Corbett-Davies et al, above n 234.
[236] O'Neill, above n 81, at 26.

tools can penalise individuals for being Black, living in a poor neighbourhood, or having friends and family members with criminal records. These tools (designed for one context but used in another without appropriate consideration) help to create the environment that justifies their assumptions. O'Neill states that:[237]

> A person who scores as "high risk" is likely to be unemployed and to come from a neighborhood where many of his friends and family have had run-ins with the law. Thanks in part to the resulting high score on the evaluation, he gets a longer sentence, locking him away for more years in a prison where he's surrounded by fellow criminals - which raises the likelihood that he'll return to prison. He's finally released into the same poor neighborhood, this time with a criminal record, which makes it that much harder to find a job. If he commits another crime, the recidivism model can claim another success. But in fact the model itself contributes to a toxic cycle and helps to sustain it.

I argue that this is particularly harmful because it exacerbates and entrenches existing inequalities in society based on race and socioeconomic status. Hannah Fry believes that creating an algorithm for use in the criminal justice system requires collective consideration about the purposes of the criminal justice system. She states that:[238]

> Rather than just closing our eyes and hoping for the best, algorithms require a clear, unambiguous idea of exactly what we want them to achieve and a solid understanding of the human failings they're replacing. It forces a difficult debate about precisely how a decision in a courtroom should be made. That's not going to be simple, but it's the key to establishing whether an algorithm can ever be good enough.

If an algorithm is to be used in a criminal justice context the relative weight a risk assessment algorithm places on various inputs must be monitored to ensure the outcomes of the algorithm are consistent with the values of the criminal justice system. I argue that Northpointe/equivant, a corporation that is not accountable to the public and is driven by profit, should not be exercising this discretion.

---

[237] At 27.
[238] Fry, above n 204, at 77-78.

For these tools to work with the efficacy that is desired of them, they require significant investment in training for the personnel using them; regular auditing;[239] as well as public consultation where their use is contentious - none of which can be undertaken expeditiously or without expense. Predictive algorithms require feedback - some mechanism to tell them when they're issuing faulty and harmful analyses. O'Neill describes how many predictive algorithms, or 'weapons of math destruction' as she calls them, behave in this way because they are unable to learn from their errors: "they define their own reality and use it to justify their results. This type of model is self-perpetuating, highly destructive - and very common".[240]

My research on the use of algorithms in the U.S. criminal justice context raises concerns about the potential for bias and discrimination. Concerns that can be heightened by both the opacity of proprietary algorithms and the use of algorithms outside of the context for which they were designed. However, there can be effective ways of deploying algorithmic decision-making aids. The following section analyses the use of algorithms in the New Zealand criminal justice context.

The New Zealand Criminal Justice Context

The Algorithm Assessment Report published by Stats NZ in 2018 provided an assessment of algorithm use across government agencies in New Zealand. It concluded that algorithms have an essential role in supporting government services and delivering policies that benefit the public.[241] Significantly, the report found that government agencies are applying safeguards in their use of algorithms and made several operational and procedural recommendations.[242] I outline the developments in New Zealand following the Algorithm Assessment Report in Chapter Six.

---

[239] Jack Bandy "Problematic Machine Behavior: A Systematic Literature Review of Algorithm Audits" (2021) Proc. Human Computer Interactions 5 CSCW1, 74; Shea Brown, Jovana Davidovic, Ali Hasan "The algorithm audit: Scoring the algorithms that score us" (2021) Big Data and Society 1-8.
[240] O'Neill, above n 81, at 7.
[241] Stats NZ "Algorithm Assessment Report" (Te Tari Taiwhenua / Internal Affairs, Wellington, 2018).
[242] Recommendations include preserving appropriate human oversight; ensuring the views of stakeholders are taken into consideration, including incorporating a te ao Māori perspective in the development and use of algorithms; safeguards to identify bias, protect privacy, and ensure appropriate levels of transparency; and the importance of seeking expertise from outside government. Stats NZ "Algorithm Assessment Report" (Te Tari Taiwhenua / Internal Affairs, Wellington, 2018) at 4.

*RoC*RoI*

RoC*RoI is a risk assessment algorithm used in New Zealand by the Department of Corrections to assist in the Department's decision-making. Corrections uses RoC*RoI - which stands for 'Risk of ReConviction' multiplied by 'Risk of Imprisonment' - rather than risk of conviction alone because RoC*RoI "gives some indication of serious re-offending, the type of offending which Corrections is attempting to target".[243] RoC*RoI is used by Corrections and in the context of parole decisions. It can be distinguished from COMPAS in that its workings are not opaque and it is deployed for the purposes of predicting risk of reoffending, which is the purpose for which it was designed for.

The guiding principles of the Parole Act 2002 state that "when making decisions about, or in any way relating to, the release of an offender, the paramount consideration for the Board in every case is the safety of the community".[244] However, an inherent limitation of RoC*RoI is that it cannot calculate the risk of reoffending and *not getting convicted*. The RoC*RoI score is a prediction that considers the likelihood that an offender will be both reconvicted in the future and be sentenced to a term of imprisonment for that offence. It is possible that there are certain classes of offenders (categorised by socioeconomic status, race, or the type of crime they commit) who are less likely to get caught by police, and therefore less likely to be convicted of an offence. Corrections NZ recognises that men who commit sexual crimes against children fall into this category and acknowledge that the RoC*RoI score for this group of offenders is not a good indicator of future likelihood of offending.[245]

RoC*RoI has been proven more accurate than experts in predicting likelihood of *reconviction*.[246] However, there is no way of testing to see if RoC*RoI can accurately predict likelihood of *reoffending* just as there is no way of knowing if parole boards accurately predict likelihood of reoffending. The community is clearly at risk from offenders with a high risk of reoffending regardless of whether they have a high risk of

[243] Department of Corrections "Risk of Reconviction" https://www.corrections.govt.nz/resources/research/risk-of-reconviction
[244] Parole Act 2002, s 7(1).
[245] Department of Corrections "Risk of reconviction: The ROC*ROI Measures - Explanatory Note" https://www.corrections.govt.nz/resources/research/risk-of-reconviction#:~:text=Hence%20the%20term%20ROC*ROI,the%20offender's%20Risk%20of%20Imprisonment.
[246] Department of Corrections, above n 245.

conviction as a result of that offending. Therefore, RoC*RoI is one way of measuring the risk to the community of releasing an offender on parole, but it is not the whole picture.

While there is no right or entitlement to be released on parole,[247] offenders do have the right to have their case for parole heard by a parole board as soon as practicable after their parole eligibility date.[248] Offenders before a parole board have already been found guilty, and convicted, of a serious crime. Nevertheless, a parole hearing is a quasi-judicial hearing with significant consequences for the individual concerned. As such, parole hearings arguably merit at least the minimum application of natural justice; the right to a fair hearing and the right to impartial decision-makers.

An offender's RoC*RoI score is considered by a parole board as part of the relevant information before it, including the offender's own submission. Decisions to release an offender on parole must be made on the basis of all the relevant information that is available to the board at the time.[249] It is important that a parole board considers all the relevant information before it because it includes information that cannot be factored into RoC*RoI's calculation, such as:

- the motive for the crime, and any mitigating or aggravating factors;
- the impact on the victim(s);
- the offender's behaviour in prison;
- the offender's repentance and acceptance of responsibility for the crime, and any attempts at rehabilitation.

A parole hearing is designed to be inquisitorial in nature. Section 49(1) of the Parole Act 2002 states that:

> A hearing must be run in the manner of an inquiry, and in an atmosphere that encourages persons appearing before the Board to speak for themselves, and as freely and frankly as possible.

[247] Parole Act 2002, s 28(1AA).
[248] Section 21.
[249] Section 7(2)(c).

However, there is an element of predetermination in assigning a RoC*RoI score to an offender that could compromise his right to a fair hearing before a parole board. The risk is that parole boards will not make their decisions based on all the relevant information before them. Where an algorithm has been proven to be more accurate than experts in determining likelihood of reconviction it could be very difficult for members of a parole board to justify a decision that is inconsistent with the algorithm's prediction. By virtue of its accuracy, an algorithm encourages compliance with the outcome that its prediction directs. Automation bias describes the tendency of humans to minimise or ignore contradictory information during the decision-making process when provided with an algorithmic prediction that is relied on as accurate.[250] Automation bias can result in decisions that are not based on a thorough analysis of all the pertinent facts, but are strongly swayed in favour of an algorithmic prediction.[251] Reliance on an algorithm could have implications for the exercise of professional judgement by members of a parole board. They may find it difficult to identify and question inconsistent outcomes or anomalies produced by the algorithm, particularly if they do not have a comprehensive understanding of how it works.

The use of an algorithm like RoC*RoI also has implications for the transparency of the decision-making process. A parole board is legally required to provide reasons for its decisions. It is a fundamental requirement of fairness that the offender (and the victim) know why a particular decision has been made. Admittedly, the reasons given by a parole board may not necessarily convey the decision-making process or the processes that occurred in the decision-makers' heads. However, the duty to provide reasons for decisions promotes transparency, thus encouraging public confidence in the parole process. Reasons are particularly valuable in the context of a parole hearing where the offender has a right of appeal. The provision of reasons will allow the offender to determine whether he has grounds for appeal and will inform him of the case he has to meet if he does appeal. RoC*RoI scores are often provided in Parole Board decisions as a factor in declining or approving parole,[252] but should more be required? Should

[250] Mary L. Cummings "Automation and Accountability in Decision Support System Interface Design" (2006) The Journal of Technology Studies 32 1 at 25.
[251] Raja Parasuraman and Dietrich H. Manzey "Complacency and Bias in Human Use of Automation: An Attentional Integration" (2010) Human Factors 52 3 at 391-392.
[252] New Zealand Parole Board "Decisions" https://www.paroleboard.govt.nz/decisions

members of a parole board be obliged to understand, and explain, the process an algorithm uses in order to be able to rely on it as a reason for approving or declining parole?

The use of algorithms, like RoC*RoI, offer the temptation of increased accuracy and the perception of objectivity, but RoC*RoI is not an infallible predictor of risk to the community. It is a useful tool to aid decision-making, however, its limitations need to be understood by its users, and its outcomes continually monitored and assessed.

While the source code of RoC*RoI algorithm is not published on the Corrections website, it has been released under an Official Information Act request and is available on the Internet.[253] RoC*RoI is used for a defined purpose and its workings are reasonably well understood. Its use is not hidden from the public or from those who are the subject of its predictions, unlike some algorithms used in criminal justice and law enforcement. Therefore, I argue that it is a benign and uncontentious algorithmic decision-making aid, provided those who use it understand its limitations.

Clearview AI - Facial Recognition Technology

In February and March of 2020, New Zealand Police commenced a trial of Clearview AI's facial recognition technology without the knowledge of the Police Commissioner or the Privacy Commissioner. It was only when an exposé by Radio NZ aired that the public became aware that Police had secretly trialled the controversial technology. This prompted an internal stocktake by Police of its surveillance technology,[254] and an external review of facial recognition technology in policing.[255] An independent Expert Panel on Emergent Technologies was established in 2021 to advise Police on the use of new technology.[256]

---

[253] fyi.org.nz "RoC*RoI formula and briefings https://fyi.org.nz/request/501/response/4002/attach/html/4/Attachment%20C59531.PDF.pdf.html
[254] Mackenzie Smith "Police 'stocktake' surveillance tech after Clearview AI facial recognition trial" (18 May 2020) RadioNZ https://www.rnz.co.nz/news/national/416913/police-stocktake-surveillance-tech-after-clearview-ai-facial-recognition-trial; Phil Pennington "Audit reveals new facial recognition tech tools in police's digital armoury" (5 November 2020) NZ Herald https://www.nzherald.co.nz/nz/audit-reveals-new-facial-recognition-tech-tools-in-polices-digital-armoury/FR7VXHHGE4QUBFQKJ5IRXYJDJU/
[255] Nessa Lynch and Andrew Chen "Facial Recognition Technology: Considerations for Use in Policing" (2021) commissioned by NZ Police https://www.police.govt.nz/sites/default/files/publications/facial-recognition-technology-considerations-for-use-policing.pdf
[256] Radio NZ "Independent panel to advise on police use of new technology" (11 March 2021) https://www.rnz.co.nz/news/national/438147/independent-panel-to-advise-police-on-use-of-emerging-

Clearview AI works like a search engine by comparing an image of a suspect or victim to a database of facial images to find a match. Clearview AI's FRT draws on a database of over 20 billion publicly available facial images sourced from the Internet.[257] When a user[258] uploads an image, it "returns links to publicly available images that contain faces similar to the person pictured in the uploaded image".[259] Clearview's facial recognition algorithm takes into account the effects of aging, changes in facial hair and other visual differences as well as variations in position and pose. Clearview AI states that its FRT performs at 99 per cent or better across all demographics,[260] meaning that its true positives are accurate 99 per cent of the time, not that it returns a positive match 99 per cent of the time.

The New Zealand Police trial found that Clearview AI's FRT had difficulty in identifying Māori and Polynesian people and Police acknowledged that it was ineffective in the New Zealand context.[261] There remain unanswered questions in relation to the use of surveillance technologies by law enforcement in New Zealand, which I discuss in more detail in Chapter Six.

Surveillance Capitalism

The more information the big technology companies have about us, the greater their powers of persuasion. Shoshana Zuboff argues that surveillance capitalism is profoundly antidemocratic due to exclusive concentrations of knowledge and power that sustain privileged influence over the division of learning in society.[262] Whereas other areas of society (such as employment, banking, and the environment) are regulated to protect

technology; https://www.police.govt.nz/about-us/programmes-and-initiatives/police-use-emergent-technologies/advisory-panel-emergent

[257] Note that while the images may be publicly available, Clearview AI has been issued with multiple cease and desist notices from the major social media platforms for scouring images off their sites in breach of the platforms' terms of use. See Zach Whittaker "Clearview AI ruled 'illegal' by Canadian privacy authorities" (4 February 2021) TechCrunch https://techcrunch.com/2021/02/03/clearview-ai-ruled-illegal-by-canadian-privacy-authorities/

[258] Clearview AI restricts the use of its FRT to government agencies.

[259] Clearview AI, "Principles" https://www.clearview.ai/principles

[260] Clearview AI, above n 259.

[261] Smith, above n 254.

[262] Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (Public Affairs, 2019, New York) at 513.

society from the excesses of capitalism, surveillance capitalism's exploitation of personal information has not faced similar restrictions.[263] Big data gives surveillance capitalists (and other very large technology companies) too much power and influence.

Zuboff describes a "democratic recession" - a weakening attachment to democracy as a result.[264] She warns that if democracy is to be strengthened in the coming years, individuals and societies must take back the right to answer the following questions:[265]

> **Who knows? Who Decides? Who decides who decides?** That surveillance capitalism has usurped so many of our rights in these domains is a scandalous abuse of digital capabilities and their once grand promise to democratize knowledge ...

On Zuboff's analysis, surveillance capitalism's threat to democracy is not immediate and dramatic but involves a gradual erosion of the fundamental principles and institutions that democracy is premised on. Informed decision-making by citizens is a necessary component of a healthy, functioning democracy. This requires access to reliable news and information online, not information curated on the basis of secret algorithms that operate to the advantage of big technology companies and their customers who seek to influence our beliefs and actions.

Access to Reliable Information

Access to reliable news and information is fundamental to democracy. People cannot make informed choices and decisions without reliable and accurate information. Democracy is compromised when a few very large players predominantly control the information and news that people can access online. The problem of a few very wealthy individuals controlling news empires is not a new phenomenon, however, it is a problem exacerbated by big data.

Professor of Public Policy at the University of California, Robert Reich, argues that when multi-billionaires take control of vital communication platforms "it's not a win for free

[263] At 514.
[264] At 516-517.
[265] At 521-522.

speech. It's a win for oligarchy".[266] Reich and others critical of Elon Musk's purchase of Twitter in 2022 are concerned about the world's richest men owning influential news corporations and social media platforms.[267] Safiya Umoja Noble agrees that the "monopoly on the information sector is a threat to democracy".[268] Online search engines are advertising companies; their primary purpose is not the provision of reliable information but prioritising their own commercial interests to make a profit.[269] Noble states that:[270]

> Search happens in a highly commercial environment, and a variety of processes shape what can be found; these results are then normalized as believeable and often presented as factual.
> ...
> The assumption (held by many) that search engines such as Google provide access to neutral, credible and accurate information is concerning.

Google is an advertising company. Its search results are biased in favour of its own economic interests: to increase its profit and dominate the market. In 2017, the European Commission fined Google €2.42 billion for breaching European Union anti-trust laws.[271] In 2023, the European Competition Authority, in its preliminary phase of an anti-trust investigation, found that Google had favoured its own ad exchange tool to the disadvantage of its competitors in the advertising market.[272] The information Google provides in its search results appears credible and objective, but in reality reflects its own advertising interests and those of its customers.

Noble argues that we need to pay more attention to our over-reliance on commercial search. A search using Google does not produce the most relevant and therefore the most

---

[266] Robert Reich, X (formerly Twitter) 27 May 2022.
[267] Mark Zuckerberg owns Facebook, Instagram and WhatsApp; Jeff Bezos owns the Washington Post; Elon Musk owns Twitter; Rupert Murdoch owns Fox News and the Wall Street Journal.
[268] Noble, above n 190, at 3.
[269] At 5.
[270] At 24-25.
[271] European Commission "Commission fines Google €2.42 billion for abusing dominance as search engine by giving illegal advantage to own comparison shopping service" (27.6.17) https://ec.europa.eu/newsroom/comp/items/104946/#:~:text=The%20European%20Commission%20has%20fined,Statement%20by%20Commissioner%20Vestager.
[272] Euractiv "EU probe: Google may have abused advertising market dominance" (15.6.23) https://www.euractiv.com/section/digital/news/eu-probe-google-may-have-abused-advertising-market-dominance/

useful or reliable information: "there are commercial interests at play that are embedded in algorithms and implemented by programming code".[273] Noble argues that because "search results play a powerful role in providing fact and authority to those who see them ... they must be examined carefully".[274] Noble believes that:[275]

> The public generally trusts information found in search engines. Yet much of the content surfaced in a web search in a commercial search engine is linked to paid advertising, which in part helps to drive it to the top of the page rank ...

Google's creators, Sergey Brin and Larry Page, saw the value in drawing on the citation model to determine what was legitimate, or at least popular, on the Internet in the creation of their PageRank algorithm.[276] They believed that there was a greater likelihood that a document was relevant if many pages pointed to it.[277] However, Brin and Page acknowledge that commercial interests can compromise search quality. Noble quotes both as stating that "advertising funded search engines will be inherently biased towards the advertisers and away from the needs of consumers".[278] Noble notes that with citation analysis in academia, work is peer-reviewed, vetted, and checked before it is published and cited.[279] However, in online search "credibility checking is not a factor in determining what will be hyperlinked".[280]

Google claims that its autosuggestions and search results are determined by objective algorithms. The company claims that it is committed to organising and presenting information impartially and will only remove information in compliance with local laws or site owner requests.[281] However, Frank Pasquale describes how Google has deviated from its 'objective' positive at times and believes that we need to know more about how

---

[273] Noble, above n 190, at 37.
[274] At 36-37.
[275] At 38.
[276] Sergey Brin and Lawrence Page "The anatomy of a large scale hypertextual Web search engine" (1998) Computer Networks and ISDN Systems 30 at 109.
[277] Noble, above n 190, at 110.
[278] Sergey Brin and Larry Page "The Anatomy of a Large-Scale Hypertextual Web Search Engine" (Stanford California, Computer Science Department, Stanford University, 1998) as cited in Noble at 41.
[279] Noble, above n 190, at 41.
[280] At 42.
[281] Google "Our Approach to Search" https://www.google.com/search/howsearchworks/our-approach/

these types of decisions are made, especially considering Google's dominance and power in search.[282]

Pasquale asks at what point do we demand that platforms like Google take responsibility for the workings of its algorithms? These platforms represent us to the world and present the world to us. We should know more about how they work and whose interests they are serving.[283] Noble highlights the failings of "market-driven information portals such as Google" and the implications for society's ability to access trustworthy and reliable information. She believes that "what is needed is a decoupling of advertising and commercial interests from the ability to access high quality information on the Internet".[284] Noble argues that:[285]

> Despite the widespread belief in the Internet as a democratic space where people have the power to dynamically participate as equals, the Internet is in fact organized to the benefit of powerful elites, including corporations that can afford to purchase and redirect searches to their own sites.
>
> ...
>
> We need to make more visible the commercial interests that overdetermine what we can find online.

Personalised search gives users the results they want, to a certain degree, but also compromises those results in terms of what Google thinks will be good for its advertisers.[286]

The role of search engines in the distribution of information to the public is critical because access to quality information is imperative for a healthy democracy. People increasingly rely on information provided by private corporations such as Google, Facebook, or X, however the reliability and quality of that information is compromised by the commercial interests at play within those corporations. Subeditors face pressure to modify headlines and key words within a story to generate more clicks and sharing among

[282] Pasquale, above n 94, at 73-74.
[283] At 77.
[284] Noble, above n 190, at 179.
[285] At 48-50.
[286] At 54.

readers. This implication of big data analytics has "potential to significantly compromise the quality of reporting to the public".[287]

However, it is not just access to reliable and impartial information that is important. Also, imperative is the ability for political information to be critiqued and for the providers of content to be identifiable and held to account if the information they deliver is intentionally misleading. Information is power, and without access to quality information, electorates and individuals cede their voting and decision-making power to surveillance capitalists pulling the algorithmic strings.

Former justice of the Supreme Court of the United States, Louis Brandeis, once famously stated that "we can have democracy in this country, or we can have great wealth concentrated in the hands of a few, but we can't have both".[288] In an age of big data, personal information is wealth. Justice Brandeis's ultimatum demands we take action to address the growing power imbalance between individuals and the big technology companies that operate as surveillance capitalists. Big data facilitates the collection and processing of unprecedented amounts of personal information by online platforms such as Google and Facebook. The accumulation of vast amounts of personal information in the hands of a few is inherently harmful to democratic societies.

Targeted Political Messaging

The same algorithmic tools that enable commercial products and services to be targeted to potential consumers can also be used to target potential voters with individualised political messaging based on their own fears and prejudices. With individualised political messaging, the ability of journalists and academics to fact check and hold providers to account is severely diminished. In the past, when people received their news predominantly through newspapers or mainstream television news broadcasting, there was greater consistency of content. There was an awareness of what the main political parties were saying about a particular issue. A significant proportion of targeted political

[287] At 154.
[288] Louis Brandeis, former Associate Justice of the Supreme Court of the United States from 1916-1939 (Goodreads.com https://www.goodreads.com/quotes/893721-we-can-have-democracy-in-this-country-or-we-can).

messaging online goes unchecked and unchallenged while perpetuating the fears and prejudices of the lowest common denominator. Véliz explains how personalised advertising "fracture[s] the public sphere into individual parallel realities".[289] She states that:[290]

> When politicians have to design one ad for the whole of the population, they tend to be more reasonable, to appeal to arguments that a majority of people are likely to support. Personalized ads are more likely to be extreme.

When political messaging is individualised and more extreme, the potential for mis- and dis- information is rife. If voters are manipulated by the widespread dissemination of inaccurate information designed to make them to think in a particular way, then elections become farcical and people in positions of power cannot be held to account. Carissa Véliz argues that:[291]

> A democracy in which people are not autonomous is a sham. People whose autonomy is thin will be easily influenced into voting in a way that does not reflect their deepest convictions, but rather the ability of powerful actors to manipulate perceptions and beliefs ...

Traditionally, reputable news media companies employ professional journalists who, for the most part, strive to uphold journalistic principles of accuracy, fact checking, and providing all the information that is relevant to a story. Journalists with integrity aim to be impartial and objective. Unfortunately, these professional standards are absent from much of the targeted political messaging online.

Cambridge Analytica

A prime example of targeted political advertising is the Cambridge Analytica scandal. In 2014, Cambridge Analytica used personal information taken without authorisation to

[289] Carissa Véliz *Privacy Is Power: Why And How You Should Take Back Control Of Your Data* (Penguin Random House, London, 2020) at 107.
[290] At 107.
[291] At 75.

build a system to profile voters, targeting them with personalised political advertising.[292] The personal information was taken through an app on Facebook called 'thisisyourdigitallife' which paid users a nominal sum to take a personality test. The answers to the test were used to build a psychological profile of users. The app harvested the personal data of those Facebook users who participated *and the personal data of all the users' friends* - a breach of Facebook's policy. The personal information of 87 million Facebook users was misappropriated by Cambridge Analytica to build a "psychological warfare tool to influence politics around the world - a textbook illustration of how knowledge is power".[293]

Cambridge Analytica used its illegally acquired personal data, in both the Trump 2016 presidential election campaign and in the 'Vote Leave' Brexit campaign, to target potential voters online with advertisements designed for their individual personality traits and political affiliation.[294] Carissa Véliz describes two elements of Cambridge Analytica's campaigns that were particularly dangerous. First, they showed very different content to different people. The content being discussed in the mainstream media was not what voters were viewing online. Second, the campaigns run by Cambridge Analytica did not look like political campaigns, sometimes they appeared as news articles, other times it looked like content created by other social media users.[295] The Cambridge Analytica scandal serves as a warning about the potential for big data to be used to manipulate people into thinking and voting in particular ways.

Behavioural Modification

Shoshanna Zuboff argues that the ultimate goal of surveillance capitalists is behavioural modification. She believes that "this decade-and-a-half trajectory has taken us from automating information flows about you to automating you".[296]

---

[292] Iga Kozlowska "Facebook and Data Privacy in the Age of Cambridge Analytica" (30 April 2018) University of Washington https://jsis.washington.edu/news/facebook-data-privacy-age-cambridge-analytica/
[293] Véliz, above n 289, at 67.
[294] At 67-71.
[295] At 69.
[296] Zuboff, above n 262, at 339-340.

However, Cory Doctorow argues that big tech is not as good at influencing us as it says it is.[297] He argues that there is little evidence that big tech has ways of bypassing our rational faculties and directing our behaviour to create a stream of purchases, votes, and other desired outcomes. Doctorow states that "the impact of dominance far exceeds the impact of manipulation and should be central to our analysis and any remedies we seek".[298]

Doctorow argues that the threat from surveillance capitalism comes from its business model that relies on dominating the market.[299] Google, for example, organises its search results in a way that impacts significantly on public beliefs. Big tech facilitates the supercharging of lies and fraud through targeting advertising, which is divisive. Doctorow critiques Zuboff's focus on surveillance capitalism's purported behavioural modification techniques and argues that the real harm of surveillance capitalism is its control over the world's information: the results to our online search queries. Doctorow states that:[300]

> ... influence campaigns that seek to displace existing, correct beliefs with false ones have an effect that is small and temporary while monopolistic dominance over informational systems has massive, enduring effects. Controlling the results to the world's search queries means controlling access both to arguments and their rebuttals and, thus, control over much of the world's beliefs.

The persuasion techniques of surveillance capitalists are more effective the more data they have. Targeted advertisements are useful for selling repugnant ideas because they are hidden. Facebook can help individuals locate people who have the same unusual or antisocial views as them. Doctorow argues that the effect of mind control techniques such as "machine learning, 'dark patterns,' engagement hacking, and other techniques to get us to do things that run counter to our better judgment" is small. He states that:[301]

---

[297] Cory Doctorow "How to Destroy Surveillance Capitalism" OneZero (26 August 2020) https://onezero.medium.com/how-to-destroy-surveillance-capitalism-8135e6744d59.
[298] Doctorow, above n 297.
[299] Doctorow, above n 297.
[300] Doctorow, above n 297.
[301] Doctorow, above n 297.

> The vulnerability of small segments of the population to dramatic, efficient corporate manipulation is a real concern that's worthy of our attention and energy. But it's not an existential threat to society.

However, Carissa Véliz believes that if even a small proportion of the population is vulnerable to targeted political advertising, then that can have dramatic effects. She states that:[302]

> Targeted ads may not work very well for businesses, but they might work quite well for swaying elections, as we've seen. A 4 per cent effect in selling a product will not be enough to compensate for the cost of the ad, but that same effect in terms of numbers of voters could very well decide an election.

Doctorow believes that Zuboff places a disproportionate weight on behavioural modification when the crucial issue is the lack of regulation that has allowed a few large technology companies to dominate the internet. However, as Véliz highlights, the ability to sway or manipulate even a small percentage of voters through targeted advertising is powerful, particularly when that power is wielded without public scrutiny in hidden and underhand ways. Regardless of the comparative weights one might attach to the harms resulting from behavioural modification versus market dominance, the fundamental problem is that big data results in surveillance capitalists having too much power. Véliz states that:[303]

> The power that privacy grants us collectively as citizens is necessary for democracy - for us to vote according to our beliefs and without undue pressure, for us to protest anonymously without fear of repercussions, to have freedom to associate, speak our minds, read what we are curious about. If we are going to live in a democracy the bulk of the power needs to be with the people. And whoever has the data has the power. If most of the power lies with companies, we will have a plutocracy, a society ruled by the wealthy. If most of the power lies with the state, we will have some kind of authoritarianism.

[302] Véliz, above n 289, at 123-124.
[303] At 82.

Data privacy regimes focussed on the individual are inadequate at addressing many of the potential harms to democracy of big data because they do not take into consideration the relational aspect of big data. Democracy is harmed by the accumulation of vast amounts of personal information by the big technology companies. Big tech wields significant power through its massive stores of personal data and its control over much of the content that we access online. In the following chapters, I will argue that this is a harm that cannot be adequately addressed by regulatory frameworks premised on the narrowly-defined conception of privacy as an individual right of control over access to one's data.

Conclusion

Access to reliable information alone will not be sufficient to ensure a well-functioning democracy if a society is plagued by significant disparities in wealth and opportunity. Increasing inequality is harmful to the public interest in economic growth and a healthy, flourishing society. There is also a compelling, and closely related, public interest in maintaining a strong, well-functioning democracy.

Big data facilitates the promulgation of racist and sexist narratives online. The burden of algorithmic tools falls heaviest on the poor and marginalised. However, not every instance of algorithmic bias, or unfair use of big data-enabled technology, constitutes harm to the individual by setting back that individual's interests. Nevertheless, the cumulative effect of big data applications that exacerbate existing inequalities in society is harmful and forms the first category of social privacy harm described in this chapter.

The second category of social privacy harm is the vast power imbalances between citizens and the big technology companies, enabled by big data. The size, influence, and power of surveillance capitalists is inherently harmful to democracies. Their persuasive influence is more powerful the more information they hold about us. The business model of surveillance capitalists relies on the monetisation of personal information. The more data that is fed into an algorithmic model, the more accurate its predictions are purported to be. Hence, the race to collect increasing amounts of personal information by agencies, governments, and corporations alike. Integral to big data's power is its ability to categorise people into groups and to predict (with varying degrees of accuracy) an individual's behaviour based on the behaviour of others with whom they share similar characteristics.

We need a regulatory framework for data privacy that is better equipped to address the social privacy harms of big data. In the following chapter, I argue that a social conception of data privacy is the first, and necessary, step in harnessing the benefits, and minimising the harms, of big data.

Chapter Four - Conceptions of Privacy

Introduction

New Zealand's data privacy framework, premised on the individual, is fundamentally flawed in an age of big data. Relying predominantly on the individual to manage their personal information in any comprehensive and meaningful way is unrealistic and unworkable. Adequately addressing the harms of big data will require a fundamental shift in how privacy is conceptualised. The social context of information collection and use is critical in determining its governance. In this chapter I argue that privacy is primarily a social concept. Without the possibility, or necessity, of our interactions and relationships with other people and agencies, there would be no need for privacy.

An overly individualistic conception of privacy has resulted in New Zealand's data privacy framework (and those in other jurisdictions) being incapable of adequately addressing the harms of contemporary information use. Prior to the creation of the Internet and the rise of big data, failure to recognise privacy as a social concept was not critical to the efficacy of data privacy regimes, but in today's digital world it is.

Harms resulting from big data are not always easily identifiable or attributable to a particular action, or even to a specific individual or agency. Social privacy harms evade remedy through a regulatory framework that relies predominantly on the individual to identify the discrete harm caused to them by a particular privacy interference. The social privacy harms of big data require intervention at a regulatory level; they cannot be adequately addressed by a complaints-based privacy framework.

In this chapter I outline various approaches to defining privacy and conclude it is not necessary to reach a conclusive definition of privacy. I also provide an overview of the arguments of various privacy scholars whose work has informed my thesis. Many of these academics reside in the United States and their scholarship relates to the cultural, political, and legal landscape in the U.S., which is different to that in New Zealand. But the underlying premise of their arguments - that privacy is a social concept - is universally applicable.

Unlike the U.S., New Zealand has an omnibus data privacy law: the Privacy Act 2020. But the Privacy Act is premised on the individual and is, therefore, ill-equipped to address the social privacy harms of big data. The collective nature of the harms (and benefits) of personal information use is acknowledged by Māori data governance scholars. Advocates for Māori Data Sovereignty address the shortcomings of Western perspectives that view privacy through the narrow lens of individual rights and interests. In the second half of this chapter, I outline Māori conceptions of privacy as a collective, rather than an individual, good.

I conclude that the first step in addressing the social privacy harms of big data, is moving away from the preoccupation with the individual that underlies our current regulatory framework, to recognise privacy as primarily a social concept not an individualistic one. That is not to say that individual consent is no longer relevant or important - it is, and will continue to be. However, determining *when* individual consent is a necessary and meaningful way of authorising information collection and use, will be dependent on context.

What is privacy?

Privacy evades an all-encompassing or conclusive definition. It has different meanings in different contexts and fulfils different but related purposes in those varied contexts. This thesis is concerned solely with data privacy: information about an identified or identifiable individual or individuals. Different accounts of privacy can help us to better understand privacy's value. The various conceptions of privacy are neither 'right' nor 'wrong' - they all offer valuable insights to privacy's purpose but simply fail to capture or define the concept definitively.

The right to be 'let alone'

In Chapter One I described one of most well-known theories of privacy law first articulated by Samuel D. Warren and Louis D. Brandeis in “The Right to Privacy”

published in the Harvard Law Review in 1890.[1] In response to instantaneous photography and an increasingly intrusive press, Warren and Brandeis invoked the right to be 'let alone'. They explained how technological developments pose a threat to privacy and argued that the common law should correspondingly develop to protect interests in private life. Warren and Brandeis described "mental pain and distress, far greater than could be inflicted by mere bodily injury" caused by invasions of privacy facilitated by "modern enterprise and invention".[2] On this account, privacy is conceptualised as the principle of inviolate personality[3] - the right to one's personality[4] and private life; distinguishing privacy from the principle of private property as something much more intimate and personal.

In 1928, as a Supreme Court Justice, Brandeis wrote his famous dissent in *Olmstead v United States*.[5] In this case, the majority of the Court held that wiretapping was not a privacy violation under the Fourth Amendment because it was not a physical trespass into the home. In his dissent, Brandeis dismissed the notion that the *means* of the intrusion was a pertinent factor. He states:[6]

> The makers of our Constitution undertook to secure conditions favourable to the pursuit of happiness ... They conferred, as against the Government, the right to be let alone - the most comprehensive of rights and the right most valued by civilized man. To protect that right, every unjustifiable intrusion by the Government upon the privacy of the individual, whatever the means employed, must be deemed a violation of the Fourth Amendment.

The right to be let alone described in Warren and Brandeis' article and articulated in the minority dissent in *Olmstead v United States* describes one important attribute of privacy. Daniel Solove argues that while Warren and Brandeis' article was ahead of its time (and, in fairness, not written to provide a comprehensive theory of privacy), the conception of privacy as the right to be let alone remains a broad and vague conception of privacy.[7]

---

[1] Samuel D. Warren and Louise D. Brandeis "The Right to Privacy" *Harvard Law Review* (15 December 1890) 4 5.
[2] At 196.
[3] At 205.
[4] At 207.
[5] *Olmstead et al v United States - Green et al v Same - McInnis v Same* 277 U.S. 438 (1928).
[6] At 478.
[7] Daniel Solove "Conceptualizing Privacy" (2002) California Law Review 90 4 at 1102.

Solove describes a similar conception of privacy as limiting access to the self. This conception includes freedom from government interference as well as freedom from intrusions by the press and others. It describes the ability to protect your thoughts, feelings, and private affairs from unwanted access by others.

Privacy as sensitive or secret information

The conception of privacy as the right to protection from disclosure of secret, private, or unfavourable information about oneself fails to recognise that we may wish to keep some perfectly benign information from some people but not others. The value of intimacy encompasses the ability to choose to share information about ourselves with certain people and not others. Privacy enables intimate relationships by allowing individuals to choose what information about themselves they share, and with whom they share it. Also, in an age of big data and predictive analytics, the conception of privacy as secret or sensitive information becomes increasingly irrelevant. As I have outlined in earlier chapters, even innocuous personal information can be used in harmful ways, particularly when combined with other personal information about the data subject and information about others with whom they share similar characteristics.

Public / Private Distinction

A similar concept of privacy draws on the public / private distinction, holding that private information is information about oneself that is not publicly available. However, a lot of private information is publicly available online; available to an audience the scale of which would have been unimaginable before the creation of the Internet. The line between public and private life is becoming increasingly blurred. When should a social media account be considered public or private? Does the data subject still have an interest in controlling access to personal information even after it has entered the public domain? New Zealand courts have held that under the criminal law there is a measured right to privacy in public places.[8] The European Court of Human Rights has held that even in a public place, in certain contexts, there is a reasonable expectation of privacy.[9]

[8] *R v Rowe* [2005] 2 NZLR 833; *Rowe v Police* HC Dunedin CRI 2005-412-000051, 24 November 2005.
[9] *Peck v The United Kingdom* (2003) 36 EHRR 41.

Physical, spatial, and decisional privacy

Physical and spatial conceptions of privacy are concerned with control over the access others have to you physically and to your private places, such as your home. Closely related is the concept of decisional privacy - protection from unjustified interference in personal decision-making and actions. Decisional privacy is about recognising those decisions where it is valuable for the individual to make their own decision according to his or her own values and beliefs.[10]

Control over information about oneself

According to the concept of privacy as control over information about oneself, a breach of privacy occurs when an individual is unable to exercise control over access to their personal information. This conception is inevitably under strain in an age of big data. I argue that it is impossible to exercise meaningful control over all the information about oneself because personal information is held by so many different individuals, agencies, organisations, and companies. Information about oneself can also be information about another. In these instances, the ability to control access to that information may interfere with the other person's interest in controlling that same information about themselves. I argue that control over information is not as relevant or applicable as it was prior to the development of digital technologies. Privacy expert Helen Nissenbaum argues that what is fundamentally important is not limiting and controlling access to personal information, but ensuring that it flows appropriately, according to social norms that vary from context to context.[11] Nissenbaum's argument forms the basis of my approach to data privacy reform, and I discuss her Framework of Contextual Integrity in detail later in this chapter.

"Ownership" as an outdated concept

The concept of "owning" personal information is not a helpful way of thinking about data privacy. It is conceptually flawed for several reasons. Firstly, data can be easily,

[10] Marijn Sax "Privacy from an Ethical Perspective" in B. Van der Sloot and A. De Groot (eds.) The Handbook of Privacy Studies: An Interdisciplinary Introduction (Amsterdam University Press, Amsterdam, 2018) at 19-21.
[11] Helen Nissenbaum *Privacy In Context: Technology, Policy, and the Integrity of Social Life* (Stanford Law Books, California, 2010).

surreptitiously, and inexpensively copied an infinite number of times. Therefore, asserting the property rights that accompany legal ownership can become impossible, particularly when personal information can be processed by big data analytics without the data subject ever being aware of that fact. Personal information about one person may also be information about another individual or individuals.[12] In this situation who does the data belong to and how do they separate their respective property interests in it? Personal information cannot be sold with the rights and interests that attach to it in the same way that a car or a computer can, for example. Even if "sold", the data continues to be the personal information of the data subject to whom it relates. Regardless of the sale, it remains the personal information of the data subject with the rights and interests that attach to information about a person. Often ownership is referred to when what is really meant is the right to profit, or benefit from, that data.[13]

Privacy and the Individual

Privacy is fundamental to intellectual freedom, personal happiness, and the development of relationships. Privacy enables us to lead autonomous and independent lives. These are values that most people hold especially important. The collective privacy of individuals is also critical for healthy democracies, allowing individuals the necessary space to develop their own ideas and arguments or to voice dissent or controversial viewpoints. Proponents of a conception of privacy as an individual right and responsibility draw on values of autonomy, and freedom from state interference. While the traditional notice and consent model is ill-equipped to deal with the harms of big data, there are diverging views as how to best respond to this problem. Some proponents of an individual rights-based approach to data privacy promote various responses that aim to better enable the *individual* to manage their own data more effectively. Marcin Betkier is a New Zealand privacy scholar who proposes a Privacy Management Model and the right to informational self-determination where expert and neutral third party intermediaries manage information transactions between online service providers and individuals on behalf of individuals, supported by a regulatory system which enables data subjects to

[12] Solove, above n 7, at 1113-1114.
[13] At 1115.

exercise their position as market players.[14] Whereas Elizabeth Renieris argues that an individual rights-based approach to privacy should be informed by a broader conception of rights; beyond a narrow focus on data and control over access to it.[15] I consider both arguments in more detail in Chapter Six.

Social conceptions of privacy

Solove's pragmatic approach to conceptualising privacy

For over twenty years Daniel Solove has argued that traditional theories or approaches to conceptualising privacy are either too narrow or too broad. Solove's response to this problem is to draw from Ludwig Wittgenstein's notion of "family resemblances" to argue that certain concepts, such as privacy, are better understood as drawing from a pool of similar characteristics, rather than sharing one common characteristic. Solove promotes a pragmatic approach to conceptualising privacy - a bottom-up approach that starts with the problem itself rather than trying to fit the problem into a category.[16] Solove believes that existing conceptions of privacy are inadequate in an age of big data. He argues that they do not adequately account for the problem of the aggregation of information – a process he describes as "out of control".[17]

Concern over privacy has developed into an issue that is increasingly critical for freedom and democracy in an era of exponential growth in information and artificial intelligence (AI) technologies. The law of privacy has suffered numerous failures and difficulties in responding to contemporary privacy problems.[18] Solove addresses the question of how we establish a robust and effective law of privacy when the ground is constantly shifting because of new technology. He states that:[19]

---

[14] Marcin Betkier *Privacy Online, Law and the Effective Regulation of Online Services* (Insentia Ltd., Cambridge, 2019).
[15] Elizabeth M. Renieris *Beyond Data: Reclaiming Human Rights at the Dawn of the Metaverse* (The MIT Press, Cambridge, Massachusetts, 2023).
[16] Solove, above n 7, at 1088.
[17] At 1154.
[18] At 1089.
[19] At 1090.

> The difficulty in articulating what privacy is and why it is important has often made privacy law ineffective and blind to the larger purposes for which it must serve … Judges, politicians and scholars have often failed to adequately conceptualize the problems that privacy law is asked to redress. Privacy problems are often not well articulated, and as a result, we frequently do not have a compelling account of what is at stake when privacy is threatened and precisely what the law must do to solve these problems. Thus, the need to conceptualize privacy is significant; yet the discourse about conceptualizing privacy remains deeply dissatisfying.

Solove believes that privacy as a concept might not have a single unifying characteristic, instead it draws from a common pool of similar elements. Rather than seeking to locate fixed and inflexible general principles, Solove's approach is pragmatic.[20] He states that:[21]

> Shifting the focus away from finding a common denominator may prove immensely fruitful. The top-down approach of beginning with an overarching conception of privacy designed to apply in all contexts often results in a conception that does not fit well when applied to the multitude of situations and problems involving privacy.

Privacy should be conceptualised contextually as it is implicated in particular problems. Solove explains how privacy problems involve disruptions to certain practices such as writing letters, talking to one's therapist, engaging in sexual intercourse, making certain decisions, for example. Privacy is a dimension of these practices and should be understood as part of these practices rather than as a separate abstract conception.[22] Solove states that:[23]

> When we protect privacy, we protect against disruptions to certain practices. A privacy invasion interferes with the integrity of certain practices and even destroys or inhibits such practices. "Privacy" as a general term refers to the practices we want to protect and to the protections against disruptions to those practices … instead of attempting to locate the common denominator of these practices, we conceptualize privacy by focusing on the specific types of disruption and specific practices disrupted.

[20] At 1091.
[21] At 1099.
[22] At 1129.
[23] At 1093.

Privacy is not simply an empirical and historical question about what society considers to be private. Privacy requires a normative component to help us guide privacy law and policy. Solove states that:[24]

> If we focus simply on people's current expectations of privacy, our conception of privacy would continually shrink given the increasing surveillance of the modern world. Similarly, the government could gradually condition people to accept wiretapping or other privacy intrusions, thus altering society's expectations of privacy. On the other hand, if we merely seek to preserve those activities and matters that have historically been considered private, then we fail to adapt to the changing realities of the modern world.

Determining what the law should protect as private depends upon a normative analysis, requiring us to examine the value of privacy in different contexts. Privacy is about the distribution of power in relationships between people. Questions of privacy and how it should be regulated are fundamental to the social and political structure of society. Under Solove's approach, the value of privacy depends upon the purposes of the practices that are involved:[25]

> Privacy is an issue of power; it affects how people behave, their choices, and their actions. When we seek to protect, create, disrupt, or halt certain practices, we are basing that decision on our view of the importance of the purposes of these practices.

Privacy should be valued instrumentally and contextually

Many theorists believe that privacy has an intrinsic value: valuable in itself, not just for the utility, pleasure, or satisfaction it brings. Privacy, it is argued, is a form of respect that must be provided to all rational beings. However, Solove (and others) believe that privacy has an instrumental value – it is valuable as a means of achieving other ends that are important. I agree that there is no inherent value in privacy: it is entirely dependent on context. There is no value in the privacy one has on a deserted island. The concept of privacy is meaningless when considered outside of our relationships or interactions with

[24] At 1142.
[25] At 1143.

others. I agree that the focus must be specifically on the value of privacy within certain practices.[26] Solove states that:[27]

> Privacy does not have a universal value that is the same across all contexts. The value of privacy in a particular context depends upon the social importance of the practice of which it is a part.

Olmstead v United States is a good example of the need for flexibility in conceptualising privacy. For nearly forty years the Fourth Amendment failed to apply to wiretapping, one of the most significant threats to privacy in the twentieth century. In his dissenting judgment, Brandeis observed that due to technological invention, the landscape of privacy is constantly changing and the law must be flexible in its conceptions of privacy problems in order to keep up with that change. Solove states that "this flexibility is impeded by the use of an overarching conception of privacy".[28]

Solove believes that courts have struggled to respond adequately to the privacy problems of big data because traditional paradigms are not suitable for conceptualising big data privacy problems.[29] He states that "privacy law has fixed itself too firmly to certain conceptions of privacy, and as a result, has lost flexibility in dealing with emerging privacy problems".[30]

The conception of privacy as control over information does not fully capture the problem. Business models that rely on the collection and processing of personal information are highly lucrative and demand increasing amounts of data to remain competitive. This information can be used to make important decisions about people's lives but is not subject to rigorous controls or scrutiny. The general public has little idea about how their information is used. Solove articulated this problem twenty years ago and the need for it to be addressed is even more urgent today:[31]

[26] At 1146.
[27] At 1093.
[28] At 1146.
[29] At 1152.
[30] At 1093.
[31] At 1153.

> ... the aggregation and uncontrolled uses of personal information by private sector bureaucracies … the disruption of the way power is allocated between individuals and large corporations goes to the structure of our society as a whole. This problem differs from the disclosure of a secret, the exposure of a nude body, or the pervasive surveillance of an individual. It is a problem that goes to the heart of what type of society we are constructing as we move headfast into the Information Age.

Solove identified information asymmetries between big tech and citizens as contributing to power imbalances in society. Today there remains a lack of consensus as to how to best regulate privacy in an age of big data. I believe the most persuasive arguments are those that recognise the relational aspect of big data and the need to understand privacy as a social concept, rather than those that propose additional measures to enable the individual to assert their own privacy rights and interests.

Viljoen - the relational aspect of big data

Salomé Viljoen describes how data privacy law is the subject of several proposed legislative reforms. Some reforms aim to maximise the financial gain of the data subject to redistribute the excessive wealth created by data processing, others aim to reassert the individual data subject's control over their personal information and its processing. However, Viljoen argues that these proposals share a common flaw: they overlook the fact that one of the central priorities of data processing is to put people into population-based groups to enable categorisation based on relevant shared features. Viljoen states that "this relational aspect of data production drives much of the social value and harm of data collection and use in a digital economy".[32]

Viljoen provides a theoretical account of data as social relations and explains how data relations result in supraindividual legal interests. Individual data subject rights cannot account for, or address, population-level effects of data processing. She argues that:[33]

[32] Salomé Viljoen "A Relational Theory of Data Governance" (2021) Yale Law Journal 131 2 at 573.
[33] At 578.

> ... the data collection practices of the most powerful technology companies are aimed primarily at deriving (and producing) population-level insights regarding how data subjects relate to others, not individual insights specific to the data subject ...

This has important implications for how data privacy regimes should be structured. It is time to reconsider the paramountcy of individual consent to data use when big tech can draw inferences about people based on the characteristics that they share with others. It is time to question *why* individuals have a legal interest in information about themselves; do those justifications still hold true, and if so, are they justifiable in every context where information is processed?[34]

Proposals for reform of data privacy law that do not recognise the social or collective nature of data privacy suffer from the same conceptual limitations as the status quo. Viljoen states that:[35]

> ... both the status quo and reform proposals suffer from a common conceptual flaw: they attempt to reduce legal interests in information to individualistic claims subject to individualistic remedies, which are structurally incapable of representing the interests and effects of data production's population-level aims.

The core objective of data privacy law should not be to increase individual control over one's own data (even if such a task were even possible) but:[36]

> ... to develop the institutional responses necessary to represent (and adjudicate among) the relevant population-level interests at stake in data production ... [This requires] moving past proposals for individualist data-subject rights and toward theorizing the collective *institutional forms* required for responsible data governance.

Viljoen asserts that societal or group interests in data are not reducible to individual interests:[37]

---

[34] At 578.
[35] At 578.
[36] At 579.
[37] At 583.

> Data's capacity to transmit social and relational meaning renders data production especially capable of benefitting and harming others beyond the data subject from whom the data is collected.

The social privacy harms of big data described in Chapter Three - increasing inequality and the undermining of democracy - are harms that impact on all of society. The *harm to others* is the strongest justification for a social conception of privacy to form the foundation of data privacy law.

Nissenbaum's Framework of Contextual Integrity

Helen Nissenbaum, Professor at Cornell Tech, believes that new technologies and new uses of technology can disrupt social life in ways that threaten moral and political values. She states that:[38]

> ... computer-based, digital electronic technologies … have hugely magnified the power of human beings over information. We are able, as individuals and in groups … to gather, store, communicate, analyse, play with, and use information in historically unprecedented ways.

Nissenbaum believes that traditional conceptions of privacy such as the public versus private distinction, the notion of private information being that which is secret or sensitive, and the right of control over personal information, come under pressure from digital technologies that facilitate the ubiquitous collection and processing of personal information. She describes how this has led academics and social commentators to call for tighter privacy regulation to protect against the erosion of privacy in the face of digital technologies. Many argue that protecting privacy means limiting access to personal information or enhancing individuals' right to control information about themselves. Like Viljoen, Nissenbaum argues that this is not the correct response to the privacy challenges that big data presents. Nissenbaum believes that what people care about is not simply *restricting* the flow of personal information but ensuring that it flows *appropriately.*[39] An

[38] Nissenbaum, above n 11, at 19.
[39] At 2.

account of the appropriate flow of information is articulated in Nissenbaum's Framework of Contextual Integrity.

Contextual Integrity and context-relative informational norms

Nissenbaum explains how social activity occurs in contexts and is governed by context-relative norms.[40] Privacy is not just about the right to control personal information but the right to the appropriate flow of personal information, which varies from context to context. Nissenbaum states that "context-relative informational norms express entrenched expectations governing the flows of personal information".[41] The Framework of Contextual Integrity provides a method of evaluating new information technologies and prescribing legitimate responses to them:[42]

> According to the framework, finely calibrated systems of social norms, or rules, govern the flow of personal information in distinct social contexts (e.g., education, health care, and politics). These norms, which I call context-relative informational norms, define and sustain essential activities and key relationships and interests, protect people and groups against harm, and balance the distribution of power. Responsive to historical, cultural and even geographic contingencies, informational norms evolve over time in distinct patterns from society to society. Informational technologies alarm us when they flout these informational norms – when, in the words of the framework, they violate contextual integrity.

Technologies that diminish our control over information about us are concerning, but Nissenbaum argues that what is more troubling is when these technologies disregard entrenched norms of information flow and threaten to disrupt social life as we know it.[43] However, not all technologies that disrupt information flows are cause for alarm. Some technologies serve societal values better than the status quo – for example, they may promote health and wellbeing or democracy. In these cases, applying the Framework of Contextual Integrity will result in the conclusion that these technologies should be adopted.[44]

---

[40] At 129-130.
[41] At 129.
[42] At 3.
[43] At 3.
[44] At 3-4.

Nissenbaum argues that it is not enough for proponents of a right to privacy to draw a link between privacy and other values like freedom, autonomy, and democracy. It is not enough to say that new technologies diminish individual control over personal information - the conflicts between privacy and competing values must be addressed systematically on a case-by-case basis. Nissenbaum describes higher order values and principles (of which privacy is one) as being abstract and therefore inadequate at resolving real world disputes that occur in the midst of social structures and norms. She states that:[45]

> Countless works, many of them brilliant, have defended privacy as a fundamental human right (not merely a preference or an interest) by linking it to other values with longstanding moral, political, and legal pedigrees. These works have shown privacy to be a form and expression of self-ownership, an aspect of the right to be let alone, a cornerstone of liberty and autonomy, or a necessary condition for trust, friendship, creativity, and moral autonomy. The shortcomings of these works, and this approach, is not that it gets things wrong, generally speaking, but that it leaves a gap. This gap is acutely felt for those who are interested in analyzing controversial systems and in the practical mission of prescribing sound decisions in relation to them.

The Framework of Contextual Integrity (the Framework) is a practical way of analysing the privacy implications of new technologies within the particular social context of their use. The Framework can be used to evaluate disruptive technologies pragmatically. It moves away from the polarising position where privacy is viewed as being in direct competition with technical innovation. Adapting for privacy considerations could lead to better technological systems and better uses of technology by respecting important social values and norms.

Applying the Framework of Contextual Integrity

Context-relative informational norms prescribe the flow of information in a given context. Expectations over the flow of information are related to the characteristics of the background social situation. One must consider:

[45] At 9.

- the type of information;
- the respective roles of the subject, the sender (who may be the subject), the recipient of the information; and
- the principles under which the information is sent.[46]

Next, the novel or disruptive system or practice must be assessed against the entrenched norms it violates. First, identify the prevailing context to establish what norms prevail. For example, a school falls within an educational context; a hospital within a healthcare context. Secondly, establish key actors. Does the new practice bring about changes in who receives the information? (i.e. enlarged sets of recipients?) Who is it about? Who is sending the information? Do the changes affect the *types* of information transmitted? Does it change the principles of transmission? For example, removing a cash option for payment to cross a toll road mandates the disclosure of personal data.[47]

If the new practice generates changes in actors, attributes, or transmission principles, then the practice violates entrenched information norms and is, prima facie, a violation of contextual integrity.[48]

Nissenbaum directs the reader to:[49]

> Consider the moral and political factors affected by the practice in question. What might be the harms? The threats to autonomy and freedom? What might be the effects on power structures, implications for justice, fairness, equality, social hierarchy, democracy, and so on?

One must evaluate how the system or practices directly impinge on values, goals, and ends of the context: "what do harms or threats to autonomy and freedom, or perturbations in power structures and justice mean in relation to this context?"[50] On the basis of this

[46] At 148-150.
[47] At 149-150.
[48] At 148-150.
[49] At 182.
[50] At 182.

analysis, the Framework either recommends in favour of, or against, systems or practices under assessment.[51]

Nissenbaum acknowledges that the Framework is inherently conservative. Any departure from an entrenched practice is flagged as problematic. However, Nissenbaum argues that there is a strong similarity between contextual integrity and the concept of a reasonable expectation: "both concepts attribute moral authority to entrenched social practice".[52] The argument being that social practice deserves to be respected because of the concept of principle or settled rationale and the idea of accumulated wisdom.[53]

To assess whether a new practice is morally or politically superior to an entrenched practice, one must compare how effective each is in supporting or promoting relevant contextual values.[54] The Framework requires that instead of observing and weighing conflicting interests and values, we must assess what contextual values are at stake when we move from an entrenched practice to a novel one.[55] Nissenbaum states that:[56]

> ... flatly ruling out change merely because it violates entrenched normative flows may be morally problematic, unconditionally accepting it is equally so. Instead it needs to be carefully considered and evaluated as prescribed by the CI [Contextual Integrity] decision heuristic.

Nissenbaum illustrates her theory using several case examples. In the context of voting, a highly regulated setting of a polling station during general elections with its enforced privacy signals to the voter that they are alone and free to make their decision.[57] This denies control to voters (they cannot ask for assistance or reveal their choice to others) but the express purpose of doing so is to protect them from having to bend to external pressure. Nissenbaum states that "in an interesting way, this constraint on control indirectly protects autonomy in voting".[58] Electronic voting technologies support

[51] At 182.
[52] At 162.
[53] At 162.
[54] At 165.
[55] At 174-175.
[56] At 205.
[57] At 176.
[58] At 176.

convenience and efficiency but are found wanting in other regards, such as anonymity, accountability and data security.[59] Nissenbaum considers the values of anonymity and autonomy in voting as particularly significant contextual values and any move to alter information flows needs to overcome the threat to these fundamental values.[60]

Nissenbaum applies the Framework to increasing levels of security for travellers at airports. There is a prima facie breach of contextual integrity in relation to full body x-ray scanners, for example. This new information flow increases the flow (and type) of personal data to airport officials. However, on further analysis, the Framework may find that this new practice is preferable to the old because it is more effective at achieving the values of security and safety which are high up on the list of objectives in the context of air transportation.[61] Nissenbaum argues that:[62]

> The aspiration of contextual integrity is similar to that of legal integrity: there is a presumption in favour of entrenched rules rather than strict adherence to the letter that can be overridden if new practices are demonstrably more effective at achieving contextual values, ends, and purposes or the equivalent; in the case of legal integrity, more effectively satisfying or promoting a duly constituted political community's scheme of principles.

The right to privacy is not the right to control personal information nor to have access to personal information restricted. It is a right to live in a world where our expectations about how our personal information is used and who it is shared with are predominantly met. As Nissenbaum describes, these expectations are shaped by the principles of social life within a given context. However, it is acknowledged that social conditions change, and contexts and norms evolve along with them.[63]

The Framework is designed to provide a foundation against which existing or proposed privacy legislation and rules can be tested. Nissenbaum suggests that informational norms may warrant explicit protection through law and public policy where they are being

[59] At 176-177.
[60] At 177.
[61] At 188.
[62] At 179.
[63] At 231.

widespread and systematically violated and the parties perpetrating the violations are overwhelmingly more powerful or wealthy and motivated by self-interest.[64] Nissenbaum also acknowledges that there are other ways to codify and enforce norms, for example, through professional codes. System designers also have a significant role in embodying appropriate norms of information flow in technical design.[65]

Nissenbaum's Framework embodies a social conception of privacy. It acknowledges that the governance of information practices is inextricably linked to social context. I think it is a valuable mechanism for promoting social justice in data governance. Its shortcoming is the need for consensus on the priority given to different values in a particular context. For example, in the context of healthcare there will often be competing, and arguably, equally important values at stake. The value of confidentiality (without which the effectiveness of certain fields of medicine such as psychiatry, are limited), and the safety or health of third parties, are both compelling values in healthcare and the subject of extensive debate in bioethics scholarship. These are not values that lend themselves to consensus in a way that enables the Framework to be applied in a determinative or authoritative manner. However, that does not mean that we cannot use the Framework in other less contentious areas. It also serves as a useful mechanism for drawing out the competing values in the regulation of emerging technologies that may be otherwise overlooked in pursuit of innovation.

Waldman - Privacy As Trust

In his book *Privacy As Trust: Information Privacy for an Information Age,*[66] Ari Ezra Waldman aims to change the way we think about privacy so that the law can protect it more effectively. Contemporary 'notice and consent' models of data privacy law are premised on the notion that each individual has the freedom to choose when and how to disclose personal information. However, Waldman argues that 'notice and consent' data privacy regimes are flawed because they provide little or no assistance in making disclosure decisions and offer even less protection when internet companies use data in

[64] At 237.
[65] At 237.
[66] Ari Ezra Waldman *Privacy As Trust: Information Privacy for an Information Age* (Cambridge University Press, New York, 2018).

unexpected and intrusive ways.[67] In contrast, *privacy as trust* approaches privacy from a social perspective.[68] Waldman argues that trust minimises the inherent vulnerability in sharing personal information. He states that:[69]

> In the context of information sharing, trust gives us the ability to live with and minimize the vulnerability inherent in sharing by relying on expectations of confidentiality and discretion. Indeed, all disclosures create vulnerability and imbalances of power. Elsewhere, as in doctor-patient or attorney-client relationships, where significant disclosures create similar power imbalances, we manage those risks with strong trust norms and legal weapons as well. So, when we share information with others in contexts of trust, this information should be protected as private.

Waldman addresses data privacy in the context of disclosures of information to other individuals and to technology companies. Disclosures of personal information can create power imbalances and vulnerabilities. Waldman believes that "a doctrine of information privacy must mitigate that vulnerability".[70]

Like other proponents of social conceptions of privacy, Waldman argues that traditional rights-based conceptions of privacy struggle to address some modern privacy problems.[71] He believes it is necessary to shift the focus from the individual harms of privacy invasions to the background structure of society that allows the sharing to happen in the first place.[72] Waldman states that "privacy law's traditional focus on injuries to the individual has limited coercive effect: it only works when a person can prove individualized, specific, and often, pecuniary harm ..."[73]

Privacy serves an essential social value by allowing social interactions to occur. It allows information to flow to certain recipients while restricting that flow to others. In this way,

[67] At 8.
[68] At 6.
[69] At 4.
[70] At 11.
[71] At 5.
[72] At 70.
[73] At 70.

privacy facilitates sharing.[74] It is a way of relating to society, not detaching from it.[75] Privacy as trust extends beyond the initial disclosure to cover secondary uses of data.[76] Waldman states that privacy "is, at its core, about the social relationships governing disclosure between and among individuals and between users and the platforms that collect, analyse, and manipulate their information for some purpose".[77]

Waldman's approach to privacy as trust requires unequal relationships between senders and recipients of personal data be readjusted to protect data subjects. Waldman offers a radical alternative to the status quo and argues for data recipients to be considered data fiduciaries, and therefore, not able to profit from personal data at the data subject's expense.

Waldman's critique of social theories of privacy

Waldman critiques other social conceptions of privacy for not going far enough in responding to "the unequal power dynamics at play in a modern world of widespread disclosures"[78] and for remaining theoretical and having no normative application. Through his concept of privacy as trust, Waldman attempts to solve these problems.[79] He identifies the shortcomings in the approaches of other academics who promote social conceptions of privacy. Waldman states that:[80]

> For Solove, there is social value in privacy because privacy protects individual autonomy and freedom. That also sells privacy short. As we will see, privacy is also an element of social structure that allows different kinds of social interaction to happen with different people.

Waldman's criticism of Nissenbaum's Framework is that it is too complicated and vague to apply to an actual privacy dilemma. He states that:[81]

---

[74] At 34.
[75] At 35.
[76] At 9.
[77] At 3.
[78] At 35.
[79] At 35.
[80] At 37.
[81] At 44.

> Although Nissenbaum's is the latest and most profound attempt to bring social theory to our understanding of privacy, the theory is ambiguous and far too complex for judges and privacy professionals on the ground.

I agree that Nissenbaum's Framework may be too complicated for some to navigate. However, it is still a valuable tool and useful for lawmakers in addressing contemporary privacy dilemmas. Alternatively, the simpler one-size-fits-all approach to regulating data privacy fails to accommodate social context, which is inextricable from any meaningful consideration of privacy problems.

Waldman believes that it is the role of trust to mitigate the vulnerabilities and power imbalances that are caused by sharing information.[82] He argues that trust is at the core of our expectations of privacy. The sharing of personal information "is not only inevitable but necessary".[83] Waldman states that "rather than a shield separating individuals and society, privacy is an element of social structure that facilitates sharing and social interaction by constraining the power of information holders".[84] The idea of privacy as an interconnected and social concept finds support in the Indigenous Data Sovereignty movement.

Indigenous Data Sovereignty

Indigenous data sovereignty promotes the control, access, and possession of Indigenous data by the Indigenous population to which it relates. Indigenous data can include data that derives from the Indigenous population or data about its people or culture. Fundamentally, Indigenous data sovereignty is the idea that Indigenous people should be the decision-makers in relation to how data about them is used.[85]

Colonisation and indigenous knowledge

[82] At 45.
[83] At 67.
[84] At 69.
[85] Tahu Kukutai and John Taylor "Data sovereignty for indigenous peoples: current practice and future needs" in Tahu Kukutai and and John Taylor (eds.) *Indigenous Data Sovereignty: Toward an agenda* (Australian National University Press, Canberra, 2016) at 1-4.

As a result of colonisation, Māori customary title to land has been extinguished by political, judicial, and legislative processes.[86] In the late 1840s, under increasing pressure from settlers, "unoccupied land" was classified as "wasteland" - a concept foreign to Māori. On behalf of the Crown, Governor Grey purchased large areas of "wasteland" in the South Island and also in some parts of the North Island, extinguishing native title. The Native Land Purchase Ordinance 1846 further undermined rangatiratanga - Māori governance or control over their land - by restoring the Crown's right of pre-emption, compelling Māori to sell land to the Crown in place of long-term leases to settlers where Māori retained ownership.[87] The New Zealand Wars of the 1840s-1860s, fought between alliances of hapu and iwi against the British army, volunteer settler units and the New Zealand Armed Constabulary,[88] resulted in the confiscation of vast amounts of Māori land, particularly in the North Island.[89]

Colonisation also involved deliberate attempts to stamp out Indigenous knowledge.[90] In colonial New Zealand, Māori children were punished, often beaten, and expelled, for speaking te reo Māori in schools.[91] From the 1840s to 1850s, Māori resistance to sales of land to the Crown had grown significantly. Therefore, the critical challenge for the Crown was to individualise Māori title to land and transform collective rights into individual property rights.[92] The Native Lands Acts of 1862 and 1865 established the Native Land Court.[93] The Native Land Court was tasked with recognising Māori customary title to land, with the objective of removing Māori communal title. During this process, the

---

[86] Danny Keenan *Wars Without End: Ngā Pakanga Whenua O Mua New Zealand Land Wars - A Māori Perspective* (Penguin Random House, New Zealand, 2021) at 236.
[87] John C. Weaver "The construction of property rights on imperial frontiers: the case of the New Zealand Native Land Purchase Ordinance of 1846" in Diane Kirkby and Catherine Coleborne *Law, history, colonialism: The reach of empire* (Manchester University Press, Manchester, 2001) at 221-236.
[88] Keenan, above n 86, at 121.
[89] Te Ara: encylopaedia of New Zealand https://teara.govt.nz/en/new-zealand-wars/page-1
[90] The Tohunga Suppression Act 1907 aimed to suppress "designing persons, commonly known as tohungas" and was passed in response to Pakeha fears that tohunga (itinerant healers) were quacks or charlatans causing harm and, in some instances, deaths. (Raeburn Lange *May the People Live: A History of Māori Health Development 1900-1920* (Auckland University Press, Auckland, 1999) at 246-255.
[91] Michael Neilson "Te Wiki o Te Reo Māori: Beaten for speaking their native tongue, and the generations that suffered" (14 September 2020) NZ Herald https://www.nzherald.co.nz/nz/te-wiki-o-te-reo-maori-beaten-for-speaking-their-native-tongue-and-the-generations-that-suffered/F7G6XCM62QAHTYVSRVOCRKAUYI/; stuff.co.nz "A child should never be punished for speaking the language of their people" (29 September 2021) https://www.stuff.co.nz/national/politics/opinion/300416107/a-child-should-never-be-punished-for-speaking-the-language-of-their-people
[92] Weaver, above n 87, at 235.
[93] Te Ara: the encyclopaedia of New Zealand https://teara.govt.nz/en/te-ture-maori-and-legislation/page-3#:~:text=The%20Native%20Lands%20Acts%201862,it%20individualised%20Māori%20land%20title.

Native Land Court distorted customary information practices: it heard oral testimony from Māori in respect of their claims to territories and resources. Traditionally, whakapapa[94] was tapu[95] and not normally shared with outsiders other than in certain prescribed circumstances. This knowledge would traditionally have been held by a nominated person on behalf of the wider whanau, hapu, or iwi group, and in the interests of the collective.[96] Associate Professor Khylee Quince, Dean of Law at Auckland University of Technology, states that the Native Land Court processes "resulted in the denigration of the traditional laws of tapu, whilst also reifying the status of the individual over the collective".[97]

Colonial states, such as New Zealand, Australia, Canada, and the U.S., have a history of counting and classifying Indigenous populations for their own purposes and continue to collect data on Indigenous populations to form the basis of Indigenous policy. However, Indigenous peoples globally are largely absent from the process of collection and use of data about them, resulting in Indigenous datasets that do not reflect Indigenous worldviews or perspectives, or fulfil Indigenous peoples' data requirements.[98] It is unsurprising that there is low trust on the part of Indigenous populations when it comes to the use of data about them by governments.

Māori conceptions of privacy as a collective, rather than an individual, good

Quince describes how, as a Māori person, she thinks about tapu (the definition of things, people, places, and information as special or restricted) when thinking about privacy. While drawing similarities between Pakeha and Māori concepts of privacy, Quince believes the most significant difference "is the notion of collective privacy that demonstrates the Māori attraction to group identity and living in a manner that maximises

[94] "Whakapapa is the overarching framework of genealogy - that demonstrates the relatedness between people, the natural world and the gods - the three spheres of the Māori universe". [Khylee Quince "Māori Concepts and Privacy" in Steven Penk and Rosemary Tobin (eds) *Privacy Law In New Zealand* (Brookers Ltd, Wellington, 2010) at 30.]
[95] The meaning of tapu includes sacred, prohibited or unclean. Quince believes that tapu provides the closest analogy to Pakeha concepts of privacy. [Khylee Quince "Māori Concepts and Privacy" in Steven Penk and Rosemary Tobin (eds) *Privacy Law In New Zealand* (Brookers Ltd, Wellington, 2010) at 31.]
[96] Khylee Quince "Māori Concepts and Privacy" in Steven Penk and Rosemary Tobin (eds) *Privacy Law In New Zealand* (Brookers Ltd, Wellington, 2010) at 37.
[97] At 37.
[98] The Indigenous World Indigenous Data Sovereignty 2021 https://www.iwgia.org/en/ip-i-iw/4268-iw-2021-indigenous-data-sovereignty.html

the collective good".[99] The Western paradigm of the individual as the starting point from which to form policy and laws informs the Pakeha focus on the individual in the context of privacy law. In contrast, for Māori there is more emphasis on the wellbeing of the collective.[100] Quince states that:[101]

> The individual in Māori thought is really only validated with reference to their membership of broader collectives of whanau, hapu and iwi ... To assert individual rights as taking primacy over obligations owed to the collective turns the Māori paradigm on its head, and has significant implications in the realm of privacy.

Te Mana Raraunga

Te Mana Raraunga is the Māori Data Sovereignty Network. Raraunga means authority over data and data systems. Te Mana Raraunga promotes Māori exercising control over Māori data; jurisdiction over decision-making in relation to Māori data and its physical and virtual storage; as well as Māori rights to data that facilitate self-determination and self-governance. Te Mana Raraunga advocates for Māori rights and interests in data to be protected at a time when increasing amounts of personal information are being shared.[102]

Māori data sovereignty (MDS) draws on the IDS movement and is supported by the United Nations Declaration on the Rights of Indigenous Peoples 2007 (UNDRIP) to which New Zealand is a signatory.[103] Articles 3, 4 and 5 of UNDRIP support the self-determination of Māori, which is inextricably linked to MDS - Māori control and governance of Māori data to promote Māori interests. The right to self-determination is enshrined in article 3 of UNDRIP. Article 4 states that "indigenous peoples have the right to exercise autonomy or self-government in matters relating to their internal and local

---

[99] Quince, above n 96, at 28.
[100] At 29.
[101] At 39.
[102] Te Mana Raraunga: Māori Data Sovereignty Network https://www.temanararaunga.maori.nz
[103] The United Nations Declaration on the Rights of Indigenous Peoples was adopted by the General Assembly on 13 September 2007 by a majority of 144 states in favour, 4 votes against (New Zealand, Australia, Canada, US) and 11 abstentions. By 2011, the 4 states that voted against it had reversed their initial positions to support it. UNDRIP sets out a universal framework of minimum standards for Indigenous Peoples, drawing on existing human rights and fundamental freedoms and applying them specifically to the context of Indigenous Peoples. [United Nations https://www.un.org/development/desa/indigenouspeoples/declaration-on-the-rights-of-indigenous-peoples.html]

affairs". Article 5 states that "indigenous peoples have the right to maintain and strengthen their distinct political, legal, economic, social and cultural institutions".[104] The rights of Indigenous peoples to self-determination, autonomy and their own social and political institutions, support the arguments for MDS.

Te Tiriti o Waitangi / the Treaty of Waitangi 1840 can also be argued in support of MDS. Māori maintain that they did not cede sovereignty in te Tiriti (the Māori version of the Treaty). In article two of te Tiriti it states that Māori retain "tino rangatiratanga" (sovereignty or absolute chiefly authority) and this was confirmed in 2014 in the Waitangi Tribunal Report titled *Te Paparahi o te Raki.*[105] In article two of the English version of the Treaty, it confirms and guarantees to Māori "full, exclusive, and undisturbed possession of their lands, estates, forests, fisheries, and other properties". Arguably 'other properties' encompasses certain personal information and cultural knowledge as taonga. Tino rangatiratanga most definitely incorporates the inherent rights and interests of Māori in Māori data.

Te ao Māori perspectives on data ethics and data governance recognise that personal information is a collective taonga. Information about individuals that belong to a group of people can potentially harm (and benefit) the collective. In Chapter Three I highlighted how information about people can be misused; harming the vulnerable, the marginalised, and minority groups. Indigenous peoples have felt the effects of the misuse of personal information since colonisation. MDS, Māori data ethics, and data governance from te ao Māori perspectives, incorporate the concepts of Māori control and governance of Māori data. That control includes the power to authorise research and data use that is beneficial and empowering for Māori and the use of data that promotes self-determination and governance.[106]

---

[104] United Nations Declaration on the Rights of Indigenous Peoples https://www.un.org/development/desa/indigenouspeoples/wpcontent/uploads/sites/19/2018/11/UNDRIP_E_web.pdf

[105] Tahu Kukutai and Donna Cormack "Pushing the space" Data Sovereignty and self-determination in Aotearoa NZ" in Maggie Walter, Tahu Kukutai, Stephanie Russo Carroll, Desi Rodriguez-Lonebear (eds) *Indigenous Data Sovereignty and Policy* (Routledge, London, 2021) at 23.

[106] Te Mana Raraunga: Māori Data Sovereignty Network "Principles of Māori Data Sovereignty" (2018) https://static1.squarespace.com/static/58e9b10f9de4bb8d1fb5ebbc/t/5bda208b4ae237cd89ee16e9/1541021836126/TMR+Māori+Data+Sovereignty+Principles+Oct+2018.pdf

Supporters of Māori data governance promote taking care to avoid harm through data use. For example, misleading research results because of a lack of contextual awareness on the part of the researchers, and in the context of big data applications, avoiding the downstream effects of secondary data uses.[107] Improving Māori access to data about Māori that is held by government agencies as well as ensuring the quality and integrity of Māori data is imperative.

Asserting Māori interests and rights in relation to data incorporates safeguarding and protecting Māori data as a taonga and preserving cultural knowledge from misappropriation.[108] Māori expertise in data management and digital technologies and business should be promoted and encouraged as should respect for tikanga and whakapapa in the collection and use of Māori data. This includes the return of bodily samples and the physical sovereignty of Māori data: Māori data should be stored in Aotearoa.[109]

MDS focusses on the collective good. There is an emphasis on beneficence over autonomy (individual consent) in certain contexts. Kiri West, Maui Hudson and Tahu Kukutai argue that "a focus on the collective good provides a better foundation for addressing data commons and ethical issues arising from the use of big data".[110] Critically, it addresses the relational nature of big data and the ability for predictions to be made about groups of the population. A focus on the collective good provides a more effective basis from which to address social privacy harms.

Te Mana Raraunga believes that less rigorous consent requirements must be balanced by more rigorous governance requirements. Māori should determine what datasets are noa (open) and what data are tapu (controlled).[111] Māori concepts of privacy are incompatible with a one-size-fits-all approach to data governance. Context is fundamentally important

[107] Te Mana Raraunga, above n 106.
[108] Te Mana Raraunga, above n 106.
[109] See Phil Pennington "Head in the clouds? Call for NZ to take control of data storage" (1.8.22) RNZ https://www.rnz.co.nz/news/national/471967/head-in-the-clouds-call-for-nz-to-take-control-of-data-storage
[110] Kiri West, Maui Hudson and Tahu Kukutai "Data Ethics and Data Governance from a Māori World View" in Lily George, Juan Tauri and Lindsey Te Ata o Tu MacDonald (eds) *Indigenous Research Ethics: Claiming Research Sovereignty Beyond Deficit and the Colonial Legacy* (Emerald Publishing Limited, Bingley UK, 2020) at 74.
[111] Te Mana Raraunga, above n 106.

in determining how data are collected, used, or disclosed. Lastly, appropriate ethical oversight of researchers (particularly, non-Māori) using population or group level datasets, such as Stats NZ's Integrated Data Infrastructure (IDI) for example, must be maintained.[112]

Maggie Walter and Stephanie Russo Carroll highlight how state agencies use data to inform policy decisions about Indigenous populations that draws on Indigenous disadvantage.[113] They state that this disadvantage is revealed through data that shows Indigenous overrepresentation in incarceration levels, socioeconomic deprivation, poor health statistics, and lower educational and employment outcomes. Walter and Carroll do not dispute these data but believe that what is contentious is the purposes for which they serve and the narratives that accompany them.[114] Walter and Carroll argue that these data do not serve the purposes or interests of Indigenous peoples. As a result, they do not provide the policy outcomes that Indigenous peoples need, nor do they provide the data that Indigenous peoples can use to develop their own policies and initiatives.[115]

Tahu Kukutai and Donna Cormack argue that despite the growing interest across government in IDS, for it to be given full effect, the State would have to shift its view of itself as the sole source of sovereignty. They acknowledge that there is little chance of this occurring.[116] Nevertheless, they maintain that a more limited aspiration of Indigenous data governance over government-held data is possible and that this could mitigate against some harms. Kukutai and Cormack state that:[117]

> Indigenous Peoples agree that self-determination is an inherent right and a desired outcome. Self-determination can be defined in various ways but typically requires that nation states recognise Indigenous Peoples distinct forms of social organisation, governance and decision-making, and redistribute power so that Indigenous Peoples are the ones making decisions over the matters that affect them.

---

[112] West, Hudson and Kukutai, above n 110, at 75-76.
[113] Maggie Walter and Stephanie Russo Carroll "Indigenous Data Sovereignty, governance and the link to Indigenous policy" in Maggie Walter, Tahu Kukutai, Stephanie Russo Carroll, Desi Rodriguez-Lonebear (eds) *Indigenous Data Sovereignty and Policy* (Routledge, London, 2021).
[114] At 2.
[115] At 2.
[116] Kukutai and Cormack, above n 105, at 21.
[117] At 22.

Kukutai and Cormack acknowledge the argument that MDS can never be fully realised under the Crown because of colonisation. However, they argue that not advocating for the strongest form of Māori data governance at the very least, could result in further harm and increasing inequities, particularly when state-controlled data systems penetrate so many facets of day-to-day life for Māori. They describe a process of moving from "data dependency to data self-determination".[118] This involves Māori determining their own data ecosystems built on tikanga,[119] mātauranga[120] and Māori priorities. The concept of collective rights and reciprocal obligations is central to all IDS movements. Also key is the idea of passing knowledge down through the generations. In contrast, contemporary western data privacy frameworks focus on the individual. Kukutai and Cormack argue that:[121]

> ... contemporary data environments overwhelmingly reproduce imperial, colonial ideologies in how data is understood and valued. Proposed solutions, such as paying people for their data, do not disrupt extractive, neoliberal understandings of data but rather replicate them in Indigenous settings. Similarly, digital inclusion strategies are promoted as solutions to inequities within the digital economy, but risk co-opting Indigenous Peoples into oppressive structures that do not fundamentally change underlying logics.

If the Crown was to accept Māori data sovereignty in principle, practical issues of scope and authorisation remain. What data will be classified as Māori data? Who is authorised to make decisions on behalf of Māori data subjects or to speak on behalf of Māori? Consideration of these questions is a critical prerequisite to implementation of MDS. While far-reaching and complex, they are not insurmountable. An in-depth consideration of these issues is beyond the scope of this thesis. Nevertheless, I believe that New Zealand's data privacy regime would benefit from adopting the Māori worldview of the

---

[118] At 28.
[119] Quince provides the following description: "Tikanga Māori is the collection of values that regulate Māori life and reflect our collective customs, beliefs and values. It influences how we think, how we act and who we are. From a Pakeha view, tikanga is law, custom and religion rolled into one". [Khylee Quince "Māori Concepts and Privacy" in Steven Penk and Rosemary Tobin (eds) *Privacy Law In New Zealand* (Brookers Ltd, Wellington, 2010) at 30.]
[120] Mātauranga means "Māori knowledge - the body of knowledge originating from Māori ancestors, including the Māori worldview and perspectives, Māori creativity and cultural practices". [Te Aka Māori Dictionary - maoridictionary.co.nz]
[121] Kukutai and Cormack, above n 105, at 31.

interconnectedness of the individual within the wider group. Quince draws from the wisdom of Sir Mason Durie when she states that:[122]

> ... recognising and enforcing individual rights without reference to the broader context actually undermines the individual eventually. The protection and advancement of group interests maximises protection of individual interests, so that they should not be seen as being at odds with one another.

Recognition of privacy as a social concept and a collective good is not only better for society, but for each individual within it.

The right to privacy

International instruments to which New Zealand is a party describe privacy as a 'right' but its status in New Zealand law is not clearly defined. There is no statutory right to privacy. The courts interchangeably refer to privacy "as a right, a value, an interest or a principle".[123] The New Zealand Bill of Rights Act 1990 "does not preclude the general recognition of a general right to privacy".[124] Associate Professor Stephen Penk believes it is unfortunate that the Court of Appeal in New Zealand's leading privacy case *Hosking v Runting*[125] failed to take the opportunity to declare privacy as a right. Despite the fact that the courts have been reluctant to recognise a general right to privacy, Penk argues that "it does not necessarily follow that privacy has not yet acquired the status of a right in this jurisdiction".[126] The right to privacy is frequently referred to by the media and invoked by lay people in their interactions with others. What is important is that the values that a right to privacy protects are upheld, whether through a statutory or common law right to privacy or some other means. Fundamentally, the right to privacy is consistent with the right to appropriate information flows - the right to have one's data flow in accordance with socially acceptable norms of information exchange as described by

[122] Quince, above n 96, at 30 with reference to M Durie *Mauri Ora: The Dynamics of Māori Health* (Auckland, Oxford University Press, 2001) at 186.
[123] Steven Penk "Thinking About Privacy" in Steven Penk and Rosemary Tobin (eds) *Privacy Law in New Zealand* (Brookers Ltd, Wellington, 2010) at 21.
[124] At 23.
[125] *Hosking v Runting* [2005] 1 NZLR 1.
[126] Penk, above n 123, at 23.

Nissenbaum. In some contexts, the social norm of information exchange will be dependent on individual consent to personal information use.

Conclusion

It is not necessary to reach a conclusive definition of privacy. A normative account of data privacy that addresses the harms and benefits of personal information use, including big data applications, is sufficient. The status quo fails to address the concerns of big data because it does not address the relational aspect of big data processing. Data privacy frameworks premised on the individual fall short in mitigating social privacy harms of big data. A conception of data privacy that recognises the social value of privacy and the interconnectedness of personal information use is imperative.

While individual consent can be an appropriate and necessary form of authorisation for some data uses, there exists an 'authorisation vacuum' predominantly, but not exclusively, in big data contexts. Individual consent is purportedly provided as authorisation for the use of personal information, but it is ethically unsound and, in some instances, legally questionable. I will examine the questionable legality of consent in big data contexts in Chapter Five.

There is still a place for 'notice and consent' data privacy frameworks that operate on a complaints-based model. The Privacy Act fulfils a purpose by providing a remedy in situations where an individual has suffered a discrete and identifiable privacy interference. For example, if a GP discloses their patient's sexual health status, causing the data subject significant emotional distress or embarrassment, the data subject can make a complaint to the Privacy Commissioner. This is an example of a scenario where a privacy interference can be adequately addressed by the current regulatory framework.

However, unlike identifiable and discrete privacy harms suffered by an individual, the Privacy Act is ill-equipped to address information practices that contribute to social privacy harms.

Privacy is fundamental to individual wellbeing and social cohesion. Privacy is an amorphous concept that defies definition, and arguably, this has resulted in privacy law

being incapable of responding to the pressing issues that big data presents. New technologies, and new uses of technology, enable processing of personal information in both new and innovative, and disruptive and damaging, ways. An individual rights-based conception of data privacy ignores the relational aspect of big data processing and its collective implications. Social conceptions of privacy recognise that privacy is contextual and that uses of personal information can both benefit, and harm, society. New Zealand's regulatory framework for data privacy would benefit from adopting a Māori worldview of privacy as a collective concept. It is time to cast off our preoccupation with the individual as the focal point for privacy policy and legislation.

Chapter Five - A comparative analysis of the EU's General Data Protection Regulation and New Zealand's Privacy Act and the deficiencies of individual rights-based regimes

Introduction

The Privacy Act 2020 (Privacy Act) embodies an individual rights-based approach to data privacy, which is typical of data privacy legislation. The European Union's (EU) General Data Protection Regulation[1] (GDPR) goes further than New Zealand's Privacy Act in enhancing the rights of individual data subjects. However, both instruments are premised on the individual, which is at the root of their shortcomings in addressing social privacy harms.

Stronger *individual* privacy rights may go some way towards preventing social privacy harms. There are undoubtedly amendments that could be made to both the Privacy Act and the GDPR that would better protect individual privacy rights. Nonetheless, legislation designed to enhance individuals' control over their information can only go so far when what is actually required is consideration of the broader social implications of information practices. There is a place for individual rights-based privacy frameworks - there are contexts in which a complaints-based model, like the Privacy Act, provides an adequate response to interferences with individual privacy. But individual privacy rights are not a complete response to the harm caused by the processing of personal information in an age of big data.

While data privacy legislation is a fundamental component of the overall privacy landscape, also important are social and cultural norms; expectations of what constitutes reasonable uses of personal information; and in particular, the design of technologies that collect and process personal information. It is unrealistic to expect the Privacy Act in isolation to prevent the type of social privacy harms evidenced in the United States as described in Chapter Three. Information privacy practices are shaped by legal

[1] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1.

requirements but also by societal expectations. For example, Chief Executives' attitudes to privacy and its importance influence workplace cultures. Economies, and political and market structures, both in New Zealand and overseas, also impact on information privacy practices. So, with these limitations in mind, I set out an overview of the Privacy Act's strengths and weaknesses in the first part of this chapter.

In the second part of this chapter, I review the EU's GDPR, a regulation that provides greater privacy rights for the individual than the Privacy Act. I consider how effective it is in enhancing the privacy of individuals and also to what extent it offers solutions to the social privacy harms of big data discussed in Chapter Three.

I conclude that while New Zealand could benefit from incorporating some of the provisions of the GDPR in the Privacy Act, the GDPR is not a panacea for the social privacy harms of big data. What is required is a shift in thinking about data privacy from the individual and their rights and responsibilities, to focus on the social context of information collection and use. Big data processing can discriminate against groups of people and harm society by exacerbating existing inequality. Big data facilitates vast information and power asymmetries, contributing to the oversized influence of the big tech companies, which is harmful to democracies. To begin to address social privacy harms, we must think of privacy as a social concept requiring regulatory solutions, not exclusively as an individual right to be protected and upheld by the individual.

Privacy Act 2020

Purpose

The purpose of the Privacy Act is to promote and protect individual privacy. Section 3 of the Act states that:

> The purpose of this Act is to promote and protect individual privacy by-
>
> (a) providing a framework for protecting an individual's right to privacy of personal information, including the right of an individual to access their personal information, while recognising

that other rights and interests may at times also need to be taken into account; and

(b) giving effect to internationally recognised privacy obligations and standards in relation to the privacy of personal information, including the OECD Guidelines and the International Covenant on Civil and Political Rights.

The Privacy Act provides a framework for protecting an individual's right to privacy of their personal information, while recognising that privacy is not an absolute right and will sometimes need to be weighed against other rights and interests. Accordingly, the Privacy Act has two (overlapping) purposes - to provide a framework that allows for the interest in protecting individual privacy to be balanced against other values where they appear to conflict, and to give effect to international privacy obligations and standards.

Overhaul of the original Privacy Act 1993

The Privacy Act 2020 replaced the original Privacy Act 1993. Over this time the use of personal information changed significantly. The Internet and Internet-connected devices became ubiquitous and the use of social media and e-commerce commonplace: developments that present challenges for data privacy law. The implementation of the Act was a long and drawn-out process. The New Zealand Law Commission reviewed the Privacy Act 1993 over a five-year period from 2006 to 2011. Seven years later, the majority of its recommendations were implemented in the Privacy Bill,[2] which was introduced to Parliament in March 2018, and passed its third reading in June 2020. It came into force on 1 December 2020. The impetus for reform was to bring New Zealand into line with other jurisdictions while keeping pace with technological developments. However, in the six years between the Law Commission's completion of its review of the Privacy Act 1993 and the passing of the Privacy Act 2020, several jurisdictions updated their privacy legislation including Australia, Japan, the UK, and notably the EU with the GDPR coming into force in 2018. While implementing some important changes, New Zealand's Privacy Act 2020 falls short of the protections in the privacy laws of some of

[2] Privacy Bill 2018 (34-3).

its key trading partners, and is not fully aligned with international privacy instruments.[3] I will compare some of the provisions of the Privacy Act to the equivalent provisions of the GDPR in the second part of this chapter.

The main changes to the Privacy Act included broader application to overseas agencies[4] and a new information privacy principle relating to the disclosure of information overseas.[5] The introduction of a breach notification system mandates the notification of a privacy breach[6] where it is reasonable to believe that it has caused serious harm to an affected individual or is likely to do so.[7] A new power for the Privacy Commissioner to issue compliance notices to agencies in breach of the Act was introduced[8] as well as two new criminal offences. The first concerns misleading an agency for the purpose of accessing the personal information of another or having that individual's information used, altered or destroyed;[9] the second relates to the destruction of a document containing personal information knowing that a request has been made under the Act in respect of that information.[10] The Privacy Act 2020 also tightens the loose "news medium" exemption of the Privacy Act 1993[11] (which had a very broad application) by narrowly defining a "news entity"[12] to which the Act does not apply.[13]

[3] Yao Dong *Privacy Act 2020* (2020) Auckland University Law Review 26 at 345; Graham Greenleaf and Lee Bygrave "Not entirely adequate but far away: Lessons from how Europe sees New Zealand data protection" (2011) Privacy Laws and Business International Report 111 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1964065
[4] Privacy Act 2020, s 4(1)(b).
[5] Section 22, Information privacy principle 12.
[6] Section 114.
[7] A notifiable privacy breach means a privacy breach that it is reasonable to believe has caused serious harm to an affected individual or individuals or is likely to do so (Privacy Act 2020, s 112). Section 113 sets out the factors that an agency must consider when assessing whether a privacy breach is likely to cause serious harm.
[8] Privacy Act 2020, s 123.
[9] Section 212(2)(c).
[10] Section 212(2)(d).
[11] Under s 2(1) of the Privacy Act 1993 news medium was defined as "any agency whose business, or part of whose business, consists of a news activity; but, in relation to principles 6 and 7, does not include Radio New Zealand or Television New Zealand Limited".
[12] Under s 7 of the Privacy Act 2020 news entity is defined as an entity (including an individual) whose business in whole, or in part, consists of a news activity; and is, or is employed by an employer that is, subject to the oversight of a regulatory body.
[13] The Privacy Act 2020 applies to a New Zealand agency in relation to an action taken in respect of personal information collected or held by it (s 4(1)(a)). New Zealand agency does not include a news entity, to the extent that it is carrying on news activities (s 8(b)(x)).

Complaints-based Act

The original Privacy Act 1993 established a complaints framework for remedying interferences with individual privacy and the Privacy Act 2020 continues the complaints regime. If an individual or individuals believe that their privacy has been breached, they can make a complaint to the Privacy Commissioner.[14] The Privacy Commissioner can investigate the complaint[15] and if satisfied that there has been an interference with the privacy of an individual in accordance with s 69, may refer the complaint to the Director of Proceedings[16] who may decide to refer the matter to the Human Rights Review Tribunal.[17] The role of the Privacy Commissioner is expanded under the new Act with the introduction of compliance notices issued by the Commissioner requiring agencies to comply with the Act.[18] The Commissioner may issue a compliance notice to an agency if the Commissioner considers that a breach of an Information Privacy Principle (IPP) may have occurred[19] - without the requirement of showing harm to an individual.[20] This is significant in that it allows the Commissioner to address information practices that are in breach of the IPPs where there may not be identifiable harm to an individual, but the practice is nevertheless undesirable.

[14] Privacy Act 2020, ss 70-71.
[15] The Privacy Commissioner, as soon as practicable after receiving a complaint, must consider the complaint and take one of the following courses of action: decide, in accordance with s 74, not to investigate the complaint; or decide, in accordance with s 75, to refer the complaint to another person; or decide, in accordance with s 76, to refer the complaint, or part of the complaint, to an overseas privacy enforcement authority; or decide, in accordance with s 77, to explore the possibility of securing a settlement between the complainant and the agency whose action is the subject of the complaint; or decide to investigate the complaint in accordance with subpart 2 (Privacy Act 2020, s 73(1)).
[16] Privacy Act 2020, ss 84, 91(5)(b), 91(6), 94(4)(a), 94(5). The Privacy Commissioner may also refer a complaint to the Director of Proceedings *without conducting an investigation* if the Commissioner has used best endeavours to secure a settlement and assurance without investigating the complaint and is unable to do so under s 77; or it appears that a term of settlement previously secured between the agency and the aggrieved individual or individuals has not been complied with; or it appears that the action that is the subject of the complaint was done in contravention of any term of settlement or assurance previously secured under the Privacy Act 2020 or Privacy Act 1993 (Privacy Act 2020, s 78).
[17] Privacy Act 2020, s 97.
[18] Part 6, Subpart 2.
[19] Section 123(1)(b).
[20] Before issuing a compliance notice, the Privacy Commissioner may, but is not required to, assess whether any person has suffered harm, for example, the types of harm listed in s 69(2)(b). (Privacy Act 2020, s 123(2)(a)).

Compliance notices as a response to big data harms?

The introduction of compliance notices is arguably the most useful amendment to the Privacy Act for addressing social privacy harms because, unlike establishing an interference with privacy under s 69, a compliance notice can be issued without identifying harm to a particular individual.[21] However, before issuing a compliance notice the Privacy Commissioner may, but is not required to, assess whether any person has suffered harm, such as the types of harm listed in s 69(2)(b).[22] Therefore, breaches of the IPPs can be addressed through the issuing of compliance notices without the necessity of identifying harm to an individual or individuals, thus capturing information practices that can contribute to social privacy harm. However, this power has been under-utilised to date. Only two compliance notices have been issued since the 2020 Act came into force. The first compliance notice was issued to the Reserve Bank in September 2021, which has since been fully complied with.[23] The second compliance notice was issued to New Zealand Police in December 2021, which is still to be fully complied with.[24] Police update the Privacy Commissioner on a quarterly basis on its progress. The Privacy Commissioner has not taken enforcement proceedings in relation to either compliance notice.[25]

The small number of compliance notices issued raises the question of whether the Office of the Privacy Commissioner is sufficiently resourced to apply the provision comprehensively. Nonetheless, compliance notices will only be effective deterrent to breaching the IPPs if agencies care about their privacy reputation. This is because the maximum fine for an agency that fails to comply with an order issued by the Tribunal to comply with a compliance notice is only $10,000[26] which, for large companies with significant annual turnover, is minimal.

---

[21] Privacy Act 2020, s 123(1)(b).
[22] Section 123(2)(a).
[23] Letter from Liz MacPherson, Deputy Privacy Commissioner, to Louise Wilsdon in response to Official Information Act request (4 July 2023).
[24] The Office of the Privacy Commissioner has granted an extension to the timeframes to complete the remaining two of fourteen requirements. (New Zealand Police "OPC Compliance Notice Report" (April 2024) https://www.police.govt.nz/about-us/publication/opc-compliance-notice-progress-report-january-march-2024).
[25] McPherson, Deputy Privacy Commissioner, above n 23.
[26] Privacy Act 2020, s 133(3).

Application

The Privacy Act applies to New Zealand agencies,[27] overseas agencies in relation to any action taken in the course of carrying on business in New Zealand,[28] and individuals not ordinarily resident in New Zealand in relation to any action taken by them in respect of personal information collected while in New Zealand,[29] or personal information held by them while present in New Zealand.[30] The meaning of "New Zealand agency" is defined in s 8 and includes an individual who is ordinarily resident in New Zealand,[31] a public sector agency,[32] or New Zealand private sector agency,[33] and a court or tribunal except in relation to its judicial functions.[34] The Privacy Act applies to overseas agencies that collect and use the personal information of New Zealanders and includes companies such as Facebook and Google. The IPPs do not apply to personal information collected or held for personal or domestic affairs,[35] unless the collection, use or disclosure of the personal information would be highly offensive to a reasonable person.[36]

Information Privacy Principles

The most important section in the Privacy Act is s 22 which sets out the thirteen IPPs. The rationale behind having a principles-based Act is to allow for flexibility in responding to new uses of personal information in contrast to a prescriptive, rules-based Act that may quickly become obsolete in the face of new technologies. Much of the Privacy Act is based around the IPPs, which relate to issues of the collection, security, access and correction, accuracy, retention, use, and disclosure of personal information held by public and private sector agencies, as well as the assigning of unique identifiers.

[27] Section 4(1)(a).
[28] Section 4(1)(b).
[29] Section 4(1)(c)(i).
[30] Section 4 (1)(c)(ii).
[31] Section 8(a)(i).
[32] Section 8(a)(ii).
[33] Section 8(a)(iii).
[34] Section 8(a)(iv).
[35] Sections 27(1) and (2).
[36] Section 27.

IPPs based on the OECD Guidelines

The IPPs in the Privacy Act are based on the principles in the OECD Guidelines, which have informed many other data privacy statutes around the world.[37] Due to the digital nature of personal information and the ease and frequency with which it moves between different jurisdictions, it is important that New Zealand's legislation is compatible with that of other countries. The OECD Guidelines assist with the interpretation of the IPPs in New Zealand's Privacy Act.

Legal effect of the Information Privacy Principles

The IPPs are not legally enforceable rights outside of the scope of the Act,[38] except for IPP 6(1) which relates to access to personal information, and in relation to public sector agencies, confers legal rights which are enforceable in a court of law.[39]

An interference with privacy

To establish an interference with privacy, one must show a breach of one or more of the IPPs AND that the action has caused some loss or 'harm' as set out below:

> **69 Interference with privacy of individual**
>
> (1) In this Act, an action of an agency is **an interference with the privacy of an individual** in any of the circumstances set out in subsection (2) or (3).
>
> (2) The action of an agency is an interference with the privacy of an individual if the action breaches,-
>
> (a) in relation to the individual,-
>
> (i) 1 or more of the IPPs; or
>
> (ii) the provisions of an approved information sharing agreement; or
>
> (iii) the provisions of an information matching agreement or section 179 or 181; or

[37] Paul Roth and Blair Stuart *Roth's Companion to the Privacy Act 2020* (LexisNexis NZ Limited, Wellington, 2021) at 220.
[38] Privacy Act 2020, s 31(1).
[39] Section 31(2).

(iv) section 115 (which requires an agency to give notice to affected individuals or the public of a notifiable privacy breach); and

(b) the action-

(i) has caused, or may cause, loss, detriment, damage, or injury to the individual; or

(ii) has adversely affected, or may adversely affect, the rights, benefits, privileges, obligations, or interests of the individual; or

(iii) has resulted in, or may result in, significant humiliation, significant loss of dignity, or significant injury to the feelings of the individual.

...

The requirement to show some form of loss, harm, or detriment to an individual as described in s 69(2)(b) acts as a mechanism to filter out frivolous or vexatious complaints. However, it also makes it difficult to redress breaches of information privacy principles where an individual cannot establish loss, harm, or detriment or identify the agency responsible.

Strengths of the Privacy Act 2020

The Privacy Act is an omnibus piece of legislation with broad application and a wide definition of personal information in line with many other jurisdictions. The Commissioner has the power to issue a code of practice in relation to the IPPs, modifying one or more of them by prescribing more, or less, stringent standards or exempting any action from an IPP.[40] The Health Information Privacy Code 2020 is a code applying to health information and there are also codes relating to credit reporting,[41] civil defence national emergencies,[42] and telecommunications,[43] which set out the IPPs in a manner that applies specifically to those contexts. This approach is very different to the privacy

[40] Section 32.
[41] Credit Reporting Privacy Code 2020.
[42] Civil Defence National Emergencies (Information Sharing) Code 2020.
[43] Telecommunications Information Privacy Code 2020.

legislation in the United States, for example, where there is a patchy, sectoral approach to data privacy legislation that varies from state to state with no overarching federal privacy legislation.[44] New Zealand's regulatory framework for data privacy law is, in contrast, comprehensive, easily accessible, and consistent in its application.

The Privacy Act works best in the contexts in which the original Privacy Act 1993 was designed to address. It provides a framework for individuals who can identify the agency responsible for a privacy interference and can establish the loss or harm that has resulted from that interference. It is designed for the type of scenario where personal information is disclosed, without consent, to the data subject's detriment, and the loss or harm is discrete and contemporaneous. One example could be a school disclosing the residential address of a student to a former abusive partner of the student's parent, resulting in the family having to move house for their safety. Another scenario could be the disclosure of sensitive health information to a neighbour, colleague, or friend, causing the data subject significant humiliation. These are the types of privacy interferences that the Privacy Act is designed to address by providing a framework in which the data subject can make a complaint to the Commissioner.

On receipt of a complaint, the Commissioner can exercise discretion as to the course of action to be taken. The Commissioner can decide to use best endeavours to secure a settlement of the complaint and an assurance from the agency responsible that there will not be a repetition of the action that gave rise to the complaint without commencing an investigation,[45] or at any time during an investigation.[46] If the Commissioner is unable to secure a settlement or a satisfactory assurance without investigating a complaint, the Commissioner may decide not to investigate the complaint if any of the statutory grounds are met,[47] or if the Commissioner considers that any further action is unnecessary or inappropriate.[48] Alternatively, the Commissioner may decide to investigate the

[44] The United States has no comprehensive federal privacy law. However, there are federal laws that apply to certain types of information, such as the Federal Trade Commission Act 1914, Children's Online Privacy Protection Act 1998, United States of America Video Privacy Protection Act 1988, Cable Communications Policy Act 1984, and the Health Insurance Portability and Accountability Act 1996.
[45] Privacy Act 2020, s 77(1).
[46] Section 83.
[47] Section 77(2)(a)(i).
[48] Section 77(2)(a)(ii).

complaint.[49] The Commissioner also has the power to refer the complaint to the Director of Proceedings without conducting an investigation.[50]

If, on the receipt of a complaint, the Commissioner elects to conduct an investigation, the Commissioner may hear and obtain information from any person[51] and make any inquiries.[52] At any time during an investigation the Commissioner may decide to take no further action on a complaint if the Commissioner considers that further action is unnecessary or inappropriate,[53] or if satisfied of any of the matters set out in s 74.[54] Section 74 sets out the reasons for which the Commissioner may decide not to investigate a complaint and includes where, in the Commissioner's opinion, the complainant has not made reasonable efforts to resolve the complaint directly with the agency concerned;[55] the aggrieved individual or individuals knew about the action that is the subject of the complaint for 12 months or more before the complaint was made;[56] the subject of the complaint is trivial;[57] or the complaint is frivolous, vexatious, or not made in good faith.[58]

After completing an investigation, the Commissioner may make a determination that a complaint has substance or does not have substance.[59] If the complaint has substance but has not been resolved despite the Commissioner's best endeavours to secure a settlement and an assurance that there will not be a repetition of the action that gave rise to the complaint or any similar kind of action, the Commissioner may refer the complaint to the Director of Proceedings.[60] The Director of Proceedings can elect to take the matter to the Human Rights Review Tribunal (HRRT).[61] An aggrieved individual who has referred a complaint to the Commissioner may also commence proceedings in the HRRT.[62] The process is long (often up to two years from the complaint being made, to a case appearing

[49] Section 77(2)(b).
[50] Section 78.
[51] Section 81(2)(a).
[52] Section 81(2)(b).
[53] Section 81(3)(b).
[54] Section 81(3)(a).
[55] Section 74(1)(a).
[56] Section 74(1)(e).
[57] Section 74(1)(i).
[58] Section 74(1)(j).
[59] Sections 91(2)(a) and 94(1)(a).
[60] Sections 91(5)(b) and 94(4)(a).
[61] Section 97.
[62] Section 98.

before the HRRT) and is designed to filter out minor breaches of the IPPs. Only very serious privacy interferences, or very persistent complainants, make it before the HRRT.

The notice and consent model, on which the Privacy Act is founded, provides a framework for addressing interferences with privacy where the loss or harm is borne predominantly by the data subject. However, it is not an appropriate framework for information privacy practices that are harmful to society when repeated at scale and potentially deployed by multiple agencies. There may be uses of big data that contribute to the social privacy harm of increasing inequality but, as isolated actions against an individual or individuals, would be insufficient to establish an interference with privacy under s 69 of the Privacy Act.

The Commissioner's obligation to take account of cultural perspectives on privacy

The purpose of the Privacy Act is to "promote and protect individual privacy".[63] However amendments to the Act in 2020 included a new, arguably contradictory, obligation on the Commissioner in performing any statutory function or duty, and in exercising any statutory power, to "take account of cultural perspectives on privacy".[64] At the core of the differences in Pakeha and Māori worldviews on privacy is the Pakeha focus on the individual (which is clearly reflected in the Act) and the Māori conception of privacy as a collective interest. The obligation of the Commissioner to consider Māori perspectives on privacy is challenging because the Privacy Act is designed to promote and protect individual privacy.

The Office of the Privacy Commissioner's *Exposure draft of a biometric processing code of practice: consultation paper* acknowledges that the use of biometric technologies can have a disproportionate impact on Māori and raises concerns in relation to tikanga and mātauranga Māori.[65] It is significant that in the draft Biometric Processing Privacy Code, an agency must take into account the cultural impacts and effects of biometric processing

[63] Section 3.
[64] Section 21(c).
[65] Office of the Privacy Commissioner *Exposure draft of a biometric processing code of practice: consultation paper* (April 2024) at 11-12.

on Māori, and on any other New Zealand demographic group[66] in its assessment of the proportionality of proposed biometric processing.[67]

As argued in Chapter Four, New Zealand's regulatory framework would benefit from adopting Māori data scholars' perspectives on privacy. Shifting the focus from empowering the individual to manage his or her own data to concentrate more on the impact on groups, and on society, of harmful information practices would go some way towards achieving this.

Exceptions to the IPPs - authorisation

IPPs 10 and 11 on the use and disclosure of personal information respectively, contain an individual authorisation exception. IPP 10(1)(c) states that "an agency that holds personal information that was obtained in connection with one purpose may not use that information for any other purpose unless the agency believes, on reasonable grounds, that the use of the information for that other purpose is authorised by the individual concerned". Similarly, IPP 11(1)(c) states that "an agency that holds personal information must not disclose that information to another agency or to any person unless the agency believes, on reasonable grounds, that the disclosure is authorised by the individual concerned".

Surveillance capitalists, Google and Facebook, have privacy policies that form part of their terms of use that users agree to when using either service. Their privacy policies state the broad uses to which users' personal data is put as well as disclosing that users' data is shared with other agencies. Facebook states that it never sells users' personal data[68], but this does not mean that data is only used for functionality or to improve the service provided. The market model of surveillance capitalists is to generate increasing amounts of information about individuals to create more detailed profiles to better predict our behaviour, likes, fears, interests, purchases, and political persuasion, to sell targeted advertising with increasing precision to retailers, lobby groups, or political actors.

---

[66] Office of the Privacy Commissioner *Biometric Processing Privacy Code: Exposure Draft Only: For Comment* Rule 1(2)(e)-(f).
[67] Rule 1(1)(d).
[68] Meta "Privacy Centre" https://www.facebook.com/privacy/guide/ads/

IPP 1 states that:

> (1) Personal information must not be collected by an agency unless-
>
> (a) the information is collected for a lawful purpose connected with a function or an activity of the agency; and
>
> (b) the collection of the information is necessary for that purpose.
>
> (2) If the lawful purpose for which personal information about an individual is collected does not require the collection of an individual's identifying information, the agency may not require the individual's identifying information.

IPP 1 is the most important principle because it is designed to limit the collection of personal information. There must be a lawful purpose and the collection must be necessary for that purpose. However, the sale of targeted advertising space is not unlawful per se. Much of the questionable uses of personal data by Facebook are connected to this very purpose. The market model of surveillance capitalists means that it is necessary for them to collect increasing amounts of personal information from their users in order to remain competitive in a marketplace where the more data points you have on individuals, the more targeted the advertising that you can sell to your customers.

With its focus on the individual, the Privacy Act is ill-equipped to mitigate unfair or discriminatory information practices, which impact on groups of people unfavourably categorised with others with whom they share characteristics. The female engineer whose Facebook newsfeed excludes ads for jobs in her profession because of her gender, or the African American couple unable to afford a mortgage at a higher level of interest, redlined because they live in an impoverished post code. Big data also discriminates on new and unprotected grounds, such as the type and model of device that a person uses to apply to for a job online, typing speed and accuracy, and other inputs that correlate to socioeconomic status or education level. The Privacy Act is quite simply not designed to address social privacy harms; it is based on a model that was designed before big data was imagined, and at a time when individuals could more realistically exercise control over their information in a comprehensive manner. That time has passed. Even if an action

discriminates on a prohibited ground and is covered by anti-discrimination law, the affected individual may not even know that they have been discriminated against.

Exceptions to the IPPs - maintenance of the law

The Privacy Act contains a "maintenance of the law" exception to several of the IPPs. This exception provides an agency with a justification for not complying with IPP 2 (collection of information directly from the individual); IPP 3 (reasonable steps to inform the individual of collection); IPP 10 (use); and IPP 11 (disclosure). An agency can invoke this exception to an applicable IPP if the agency believes, on reasonable grounds that non-compliance with IPP 2 or IPP 3 is necessary, or the use (IPP 10) or disclosure (IPP 11) of the information, is necessary:[69]

(i) to avoid prejudice to the maintenance of the law by any public sector agency, including prejudice to the prevention, detection, investigation, prosecution, and punishment of offences; or
(ii) for the enforcement of a law that imposes a pecuniary penalty; or
(iii) for the protection of public revenue; or
(iv) for the conduct of proceedings before any court or tribunal (being proceedings that have commenced or are reasonably in contemplation)

This exception is broad and provides public agencies with wide scope for lawful non-compliance with the IPPs while carrying out their law enforcement functions. What is concerning, is the lack of transparency around how surveillance and predictive technologies are being used by New Zealand Police and other law enforcement agencies in New Zealand. There is a lack of awareness of the technologies used by these agencies, even within the organisations themselves. A prime example was the Commissioner of Police being blindsided by his Force's trial of facial recognition technology in 2020 (discussed in Chapter Three). Admittedly, this did trigger an internal stocktake by Police of its surveillance technology as well as the establishment of an Expert Panel on Emergent Technologies to advise Police on the use of new technologies. However, the Expert Panel is advisory only and its mandate is limited to the emergent technologies that are referred

[69] Privacy Act 2020, s 22 IPP(2)(e) - note IPP 2(e) also includes subsection (v) 'to prevent or lessen a serious threat to the life or health of the individual concerned or any other individual'; IPP 3(4)(b)(i); IPP 10(1)(e); IPP 11(1)(e).

to it. It does not exercise oversight over all of Police's use of technology. If they are to be effective, oversight bodies need to have oversight of all technology use by law enforcement, not just the technology that gets referred to them for consideration.

A range of explanations are provided by law enforcement to justify their lack of transparency, such as *criminals will game the system if they know how it works* or *the end justifies the means.* However, if we examine these justifications in any depth they cease to be convincing. If a system is capable of being gamed, there may already be people gaming it (those in the know). It raises the question of whether such a system - a system capable of being gamed - should be deployed by law enforcement. It is understandable how operational staff in law enforcement agencies can believe the worthwhile objectives that they are pursuing justifies some questionable means. However, checks, balances, and limitations on state power exist for very good reasons. It can be difficult to remember this when the focus is on catching "villains" and preventing harm to citizens. The grey areas such as protest activity and activism highlight how subjective the policy of policing can be. A person's political views may influence where they make the call about *the ends justifying the means*, which is precisely why it is not a convincing justification for secrecy surrounding policing tools. It raises the question of the degree of operational specificity required to allow policing by consent.

I argue that it is in public interest to know how we are being policed (as well as how other agencies with law enforcement powers are deploying emergent technologies). The era of self-regulation, optional ethics, and voluntary privacy impact assessments, needs to end. There may be a need for the use of some technology to be kept secret, but *if* that is the case, the ethical bodies that oversee its use must be more than advisory and have some power to approve or decline the use of technology that is not disclosed to the public and be guided by transparent policy on this.

The crime prevention rationale that is often purported as justification for the surveillance of New Zealanders - for example, the notion that CCTV prevents crime because criminals know they are being watched - is inconsistent with keeping surveillance tools a secret. The public doesn’t need a working understanding of the logic behind these tools (commercial sensitivity often prevents the agencies themselves from being provided with that level of understanding anyway). However, we should know how we are being

watched, and consequently, be able to question the necessity or desirability of such surveillance. The counterargument to this is that the panopticon effect works best when people are unsure when they are being watched but believe they could be at any time. But should all citizens online be subject to the panopticon effect? Arguably, this can inhibit freedom of association and speech, particularly in relation to protest activity. Ultimately, these are questions best debated in the public domain, and determined democratically, in the public interest.

Ideally, agencies in New Zealand, and overseas agencies using New Zealanders' information, would implement data privacy practices that match New Zealanders' expectations of reasonable uses of personal information. Obviously not all New Zealanders have the same expectations of how their information should, or should not, be used. Arguably, our reasonable expectations of personal information use can degrade as we become accustomed to exploitative information practices and we come to expect that that is how our information will be used, particularly in the context of surveillance. However, while acknowledging this obvious pitfall, the concept is still sound if we consider our reasonable expectations of information use to be a normative concept – *what should we be entitled to expect?* - not one that can be degraded by conflicting corporate interests. In Chapter Four I provided an account of Helen Nissenbaum's privacy as contextual integrity. Nissenbaum provides a framework for assessing whether technologies, systems, and practices that affect the flow of personal information in a society are "morally and politically legitimate".[70] Nissenbaum also suggests that informational norms may warrant explicit protection through law and public policy where they are being systematically violated and the parties perpetuating the violations are overwhelmingly more powerful or wealthy and motivated by self-interest.[71]

Ari Ezra Waldman argues that personal information disclosed for a particular purpose should not be exploited for another purpose, at the expense of the data subjects concerned. Waldman argues that trust is at the core of our expectations of privacy. Trust "is a social norm of interactional propriety based on favourable expectations of others' behaviour … a singularly significant factor in our decision to share personal information with

[70] Helen Nissenbaum *Privacy In Context: Technology, Policy, and the Integrity of Social Life* (Stanford University Press, California, 2009) at 236.
[71] At 237.

anyone".[72] Waldman believes that much of our online life is necessary, not voluntary, and, as such, our related disclosures cannot be seen as the subject of free choice.[73] Sharing personal information is "not only inevitable but necessary"[74] and trust "facilitates disclosures that serve society well".[75] Waldman states that "rather than a shield separating individuals and society, privacy is an element of social structure that facilitates sharing and social interaction by constraining the power of information holders".[76] He argues that where there exists significant vulnerability; imbalance of knowledge, power, or skills; and the companies collecting personal information hold themselves out to be trustworthy, they should be considered information fiduciaries. He states that:[77]

> All fiduciary relationships have two overarching similarities – namely, asymmetry and vulnerability … Companies like Facebook, Google, Uber, and Match.com should be considered information fiduciaries for the same reasons that doctors, estate managers, and investment analysts are considered fiduciaries ...

Social context is key to understanding, and prescribing, how information should be used. Rather than leaving it to the individual to manage their information, we should be considering the cumulative effects on society of certain information practices and how they can be prevented - a pre-emptive, rather than reactive, response to privacy harm. Reasonable expectations of personal information use should be considered a normative concept; guided (but not dictated) by norms of information use to uphold social values in varied contexts and protected by legislation when repeatedly breached.

The Commissioner has the power to inquire generally into any matter where it appears that the privacy of individuals is, or may be, infringed.[78] However, there is no corresponding power to inquire into the potential for harm to be caused by cumulative effects of information privacy practices (legal or otherwise) that, when repeatedly performed, have negative implications for society overall. The increasing use of

---

[72] Ari Ezra Waldman *Privacy As Trust: Information Privacy for an Information Age* (Cambridge University Press, New York, 2018) at 50.
[73] At 66.
[74] At 67.
[75] At 5.
[76] At 69.
[77] At 86.
[78] Privacy Act 2020, s 17(1)(i).

surveillance is a prime example of this. The Commissioner can issue directions and recommendations for agencies to comply with its guidelines and complete privacy impact assessments, but the OPC does not take a high-level consideration of where the country is heading in terms of surveillance. A privacy impact assessment for each application of a tool by an agency is assessed, independently, for *its* potential harm *to individuals*. There is a lack of consideration of the bigger picture of surveillance and the cumulative effects of increasing surveillance by a multitude of government and private agencies on society.

Part 5 Complaints, Investigations and Proceedings

The 2020 Act includes the failure to notify an individual or give public notice of a notifiable privacy breach under s 115 as an interference with the privacy of an individual under s 69, provided the requirement of 'loss, harm or detriment' as described in s 69(2)(b) is also met.[79] What is problematic with this is that in an age of big data, there can be a significant lapse of time between a notifiable privacy breach and any adverse consequences for the affected individual. The notion that a data breach by one agency will have a definable and contemporaneous detrimental impact on the data subject concerned may be applicable in certain scenarios but excludes many big data contexts. It is also very difficult for an agency to determine what types of breaches could cause serious harm. It is impossible to know all the ways that the compromised information could be combined with other data about the individual or individuals concerned, and the likely consequences of that.

Admittedly, mandatory privacy breach reporting can provide the individual with a limited opportunity to protect themselves. However, other than changing passwords or cancelling credit cards, once the information is out, the damage is done. The OPC might gain an insight to the scale and type of serious privacy breaches that are occurring. It can also publish the identity of the agency, if doing so is in the public interest. Arguably, mandatory privacy breach notifications work most effectively as a measure to encourage agencies to take better care of the personal information they hold. However, as is the case with compliance notices, it is only effective where companies or agencies care about their

[79] Section 69(2)(b).

privacy reputation in New Zealand. It also requires that the regulatory body responsible for oversight is properly resourced to fulfil its functions in a comprehensive way.

Fines under the Privacy Act 2020

The maximum fine that can be issued under the Privacy Act is $10,000.[80] This is possibly the starkest contrast to the GDPR, which allows supervisory authorities to impose significant fines of up to €20,000,000 or 4 per cent of worldwide annual turnover.[81] In the second half of this Chapter I will consider the effectiveness of fines as a deterrent to the deliberate misuse of personal information and poor information privacy practices that can result in unintended data privacy breaches.

Remedies and damages in respect of interference with privacy

If the Tribunal is satisfied, on the balance of probabilities, that any action of the defendant is an interference with the privacy of an individual or individuals, the Tribunal may grant one or more of the following remedies:[82]

(a) a declaration that the action of the defendant is an interference with the privacy of 1 or more individuals:
(b) an order restraining the defendant from continuing or repeating the interference, or from engaging in, or causing or permitting others to engage in, conduct of the same kind as that constituting the interference, or conduct of any similar kind specified in the order:
(c) damages in accordance with section 103:
(d) an order that the defendant perform any acts specified in the order with a view to remedying the interference, or redressing any loss or damage suffered by the aggrieved individual or aggrieved individuals as a result of the interference, or both:
(e) any other relief that the Tribunal considers appropriate.

[80] Sections 104(4), 118(1), 133(3), 197 and 212.
[81] GDPR, Art 83(5).
[82] Privacy Act 2020, s 102(2)(a)-(e).

The Tribunal may award damages against the defendant for an interference with privacy of an individual in respect of pecuniary loss,[83] expenses,[84] loss of any benefit,[85] and humiliation, loss of dignity, and injury to the feelings of the aggrieved individual.[86] The Tribunal's statutory limit on the award of damages is $350,000 but that can be exceeded where an action affects a number of individuals and damages are awarded to each individual.[87] Damages awarded under the Privacy Act are typically between $5,000 to $15,000.[88]

Automated decision-making, including profiling

The Privacy Act has no specific provisions on automated decision-making or profiling unlike the GDPR. The Article 29 Data Protection Working Party (A29WP) which has since been replaced by the European Data Protection Board,[89] acknowledged the increasing use of profiling and automated decision-making across a range of sectors, both public and private. It recognised that advances in artificial intelligence (AI), machine learning and big data analytics "have made it easier to create profiles and make automated decisions with the potential to significantly impact individuals' rights and freedoms".[90] Automated decision-making can deliver increased efficiencies and associated cost savings, but these benefits can come with a significant risk to individual rights and freedoms.[91] The GDPR aims to address these risks with appropriate safeguards.

General Data Protection Regulation

The remainder of this Chapter considers what New Zealand could learn from the EU's GDPR. I will consider some of the provisions of the GDPR that go further than the

[83] Section 103(1)(a).
[84] Section 103(1)(b).
[85] Section 103(1)(c).
[86] Section 103(1)(d).
[87] Section 103(2).
[88] With the notable exception of *Hammond v NZCU Baywide* [2015] NZHRRT 6 where $98,000 damages was awarded for emotional harm.
[89] Article 29 Data Protection Working Party was established by Directive 95/46/EC and was composed of the national supervisory authorities and dealt with privacy issues relating to the protection of privacy and personal data until 25 May 2018 (entry into application of the General Data Protection Regulation).
[90] Article 29 Data Protection Working Party "Guidelines on Automated Individual decision-making and Profiling for the purposes of Regulation 2016/679" 17/EN WP251rev.01 (3 October 2017) at 5.
[91] At 5.

equivalent provisions in the Privacy Act, or are missing from the Privacy Act altogether, in particular Article 22, on automated individual decision-making, including profiling. I conclude that while the GDPR goes further than the Privacy Act and has provisions that New Zealand could benefit from adopting, it is not a panacea for the social privacy harms of big data. Its aspirational rights are diluted in their application by ambiguity and uncertainty. The scope and application of Article 22 is limited, and it arguably raises more issues than it addresses.

Diminishing privacy as individual privacy rights get stronger

In his latest book *Industry Unbound,*[92] Ari Ezra Waldman argues that our privacy is disappearing as privacy laws under individual rights-based privacy regimes get stronger. He believes that measures taken purportedly to address privacy concerns by big tech are window dressing.[93] Waldman argues that it is problematic that the big tech companies have significant influence over the interpretation and implementation of the GDPR and other data privacy regulations. Technology companies may look like they are following the law but they are reframing it. Waldman states that:[94]

> Corporate power over law and design processes means that companies can leverage internal organisation, hierarchies and policies to systematically devalue privacy and maximise data extraction in how they interpret the law and how they design new products, making anti-privacy designs more likely.

Waldman also describes a widespread problem of privacy professionals "steeped in corporate-friendly privacy discourses".[95] Waldman argues that privacy is more than data governance; privacy law should be about regulating the power structures of informational capitalism.[96] Data privacy regulation directly impacts the profit margin of technology companies that rely on generating increasing amounts of personal information to remain competitive. Waldman makes a compelling argument that big tech has a vested interest in diluting the efficacy of data privacy regulations.

---

[92] Ari Ezra Waldman *Industry Unbound: the inside story of privacy, data, and corporate power* (Cambridge University Press, Cambridge, 2021).
[93] At 115.
[94] At 8.
[95] At 7.
[96] At 11.

Privacy-as-control works in the interest of big tech

Waldman argues persuasively that the concept of privacy-as-control - the notion that giving individuals more control over their information is an effective way of addressing privacy harm - works in the interests of big tech. Waldman states that:[97]

> Tech companies use the tools of coercive bureaucracies to routinize antiprivacy norms and practices in privacy discourse, compliance and design. Those bureaucracies constrain workers directly by focussing their work on corporate-friendly approaches to privacy. As info-industry workers perform these antiprivacy routines and practices, those practices become habituated, inuring employees to data extraction, even as they earnestly profess to be privacy advocates.

Waldman believes that the dominant privacy discourse in the U.S. centres around choice, consent, and control, which is a very narrow vision of privacy that allows corporations to say that they care about privacy but do little to improve privacy in real terms for their customers.[98] Waldman also argues that:[99]

> Industry leaders seek to influence how we think about privacy not just to erode our interest in and capacity to enact robust privacy laws, but to entrench corporate-friendly ideas as commonsense and mainstream among workers.

The information industry has a vested interest in defining privacy as control: "the discourse of control is the discourse of self-governance. And self-governance is a sham".[100] Waldman argues that privacy-as-control is premised on 3 false assumptions:[101]

1. that we can adequately process corporate privacy notices;
2. that our decision-making is rational; and
3. that consent is the same as making a real choice.

---

[97] At 5.
[98] At 6.
[99] At 6.
[100] At 52.
[101] At 52.

Instead, we suffer from "digital resignation";[102] giving up the expectation that it is even possible to protect our privacy. Waldman argues that notice and consent may be effective for discrete decision-making, but it doesn't scale,[103] for example in online contexts where pop up consent boxes are an irritant rather than privacy enhancing.

The European Union's General Data Protection Regulation: dual purposes

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1 (GDPR) has two purposes: to protect the fundamental rights and freedoms of natural persons in relation to the processing of their personal data and to ensure the free flow of personal data within the EU.[104] The GDPR aims to protect both the data protection rights of individuals[105] and facilitate the free movement of data for business purposes.[106] The GDPR, adopted in April 2016, replaced Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data (DPD) on 25 May 2018.

The GDPR aims to harmonise data privacy legislation across EU Member States, promoting individual rights. It applies to organisations within the EU[107] as well as agencies operating outside of the EU that offer goods or services to, or monitor the behaviour of, individuals within the EU.[108] The GDPR goes further than New Zealand's Privacy Act in enhancing the privacy rights of the individual. It includes the 'right to be forgotten', rights in relation to automated decision-making and imposes significant fines for non-compliance. I will focus on the key provisions of the GDPR that are missing from our Privacy Act or are more demanding than the equivalent provisions in the Act. The GDPR has been widely criticised for its vagueness and the ambiguity of its provisions, which limit its potential efficacy.

---

[102] At 53.
[103] At 53.
[104] GDPR, Art 1.
[105] Article 1(2).
[106] Article 1(3).
[107] GDPR, Art 3(1).
[108] Article 3(2).

Social privacy harms caused by automated decision-making and profiling

In Chapter Three, I argued that when used carelessly or for ill-conceived objectives, big data has the potential to cause significant harm, predominantly to those already struggling the most. Big data facilitates automated decision-making, social sorting, and profiling which can be discriminatory and unfair, and when used at scale, can contribute to inequality. For example, charging people who live in a poorer neighbourhood a higher rate of interest; screening candidates for employment based on predictions of future ill health; or cancelling benefits without human oversight based on an automated assessment of fraud. Cathy O'Neill describes how harmful kinds of mathematical models (predictive algorithms) can contribute to the outcomes that they predict, creating "pernicious ... feedback loops"[109] that can trap people in poverty and perpetuate injustices.[110] I will consider whether the GDPR provides regulatory responses that could be adopted to address existing, or the potential for future, social privacy harms in New Zealand.

GDPR Principles

The GDPR sets out its overarching principles in Article 5 which establishes principles relating to the processing of personal data. Several of these provisions go further than the equivalent provisions in New Zealand's Privacy Act by placing greater obligations on a controller: "the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data";[111] and bestowing greater rights on individual data subjects. All the data protection principles in the GDPR are relevant to profiling and automated decision-making using personal information.

[109] Cathy O'Neill *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (Penguin Books, Great Britain, 2016) at 12.
[110] At 12-15.
[111] GDPR, Art 4(7).

Lawfulness, fairness and transparency

The transparent processing of personal information is a fundamental requirement of the GDPR.[112] Article 5(1)(a) of the GDPR states that "personal data shall be processed lawfully, fairly and in a transparent manner in relation to the data subject ..." Article 5(1)(a) is a significant principle because profiling often occurs without the knowledge of the data subject. Data subjects may find it difficult to understand the complex processes involved in profiling and automated decision-making. Therefore, the controller must take appropriate measures to convey information about the processing of data to the data subject in a "concise, transparent, intelligible and easily accessible form ..."[113] In addition to being transparent, processing must also be fair. The A29WP recognises that:[114]

> Profiling may be unfair and create discrimination, for example by denying people access to employment opportunities, credit or insurance, or targeting them with excessively risky or costly financial products.

The A29WP gives the example of a data broker selling profiles of financially vulnerable consumers to financial companies that will offer those consumers payday loans and other financially risky products.[115] The sale of these profiles, without data subject consent or an understanding of the underlying data, provides an example of an activity that would not meet the requirements of Article 5(1)(a). A29WP states that "unfair profiling can lead to some consumers being offered less attractive deals than others"[116] which is not fair.

Purpose limitation

Article 5(1)(b) of the GDPR states that:

> Personal data shall be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical

[112] Article 29 Data Protection Working Party, above n 88, at 9.
[113] GDPR, Art 12(1).
[114] Article 29 Data Protection Working Party, above n 90, at 10.
[115] At 10.
[116] At 10.

> research purposes or statistical purposes shall ... not be considered to be incompatible with the initial purposes ...

Controllers must maintain a record of processing activities that they are responsible for, and that record shall contain, among other details, the purposes of the processing.[117] This Article is important because automated decision-making and profiling often involves the use of information collected for a different purpose. IPP 10 of the Privacy Act also attempts to limit further processing for unrelated purposes with similar exceptions.[118]

Data minimisation

Article 5(1)(c) states that personal data shall be "adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ..." The Privacy Act covers this principle in IPP 1; stating that personal information must not be collected by an agency unless it is for a lawful purpose connected with a function or activity of the agency; and the collection of the information is necessary for that purpose.[119]

Accuracy

Article 5(1)(d) states that personal data shall be "accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data are accurate, having regard to the purposes for which they are processed, are erased or rectified without delay ..." The equivalent provisions of the Privacy Act are IPPs 7 and 8, dealing with correction and accuracy respectively. Under IPP 7(1), an individual whose information is held by an agency is entitled to request the agency correct that information. IPP 7(2) states that:

> An agency that holds personal information must, on request or on its own initiative, take such steps (if any) that are reasonable in the circumstances to ensure that, having regard to the purposes for which the information may lawfully be used, the information is accurate, up to date, complete, and not misleading.

[117] GDPR, Art 30(1)(b).
[118] Privacy Act 2020, s 22, Information privacy principle 10.
[119] Section 22, Information privacy principle 1(1).

Where a data subject requests that information about them be corrected, but an agency is not willing to correct that information, the data subject may request that a statement of the correction sought be attached to the information in a manner that ensures that it will always be read with that information.[120]

IPP 8 states that:[121]

> An agency that holds personal information must not use or disclose that information without taking any steps that are, in the circumstances, reasonable to ensure that the information is accurate, up to date, complete, relevant, and not misleading.

In sum, the GDPR and the Privacy Act have provisions on accuracy and correction that are effectively equivalent.

Storage limitation

Article 5(1)(e) of the GDPR states that personal data shall be "kept in a form which permits identification of data subjects for no longer than is necessary *for the purposes for which the personal data are processed* ..." with exceptions for archiving in the public interest, scientific and historical research purposes, and statistical purposes. IPP 9 of the Privacy Act states that "an agency that holds personal information must not keep that information for longer than is required *for the purposes for which the information may lawfully be used".* Information may be lawfully used for several purposes, including an agency auditing how much inaccurate or duplicate data it holds. The Privacy Act's storage limitation principle is particularly weak and lacks the GDPR's direction that information must not be kept for longer than it is needed to fulfil the purpose for which it is processed. IPP 9 does not require the deletion or destruction of stale or inaccurate information.[122] This is problematic because the longer personal information is kept, the greater the likelihood of it being leaked, becoming inaccurate and misleading, or being reused.

---

[120] Section 22, Information privacy principle 7(4).
[121] Section 22, Information privacy principle 8.

[122] However, a data subject does have the power to request correction or deletion of information about themselves under IPP 7, s 22 Privacy Act 2020.

Integrity and confidentiality

Article 5(1)(f) of the GDPR states that personal data shall be:

> ... processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures ...

Similarly, the Privacy Act states that:[123]

> An agency that holds personal information must ensure—
>
> (a) that the information is protected, by such security safeguards as are reasonable in the circumstances to take, against—
>
> (i) loss; and
>
> (ii) access, use, modification, or disclosure that is not authorised by the agency; and
>
> (iii) other misuse; and
>
> (b) that, if it is necessary for the information to be given to a person in connection with the provision of a service to the agency, everything reasonably within the power of the agency is done to prevent unauthorised use or unauthorised disclosure of the information

The Privacy Act's IPP 5 on storage and security (above) sets out similar requirements to Article 5(1)(f) of the GDPR.

Article 6 Lawfulness of processing

Under the GDPR, data controllers who do not have a specific legal basis for data processing outside of the GDPR must rely on one of the six legal grounds for processing in Article 6. Each ground in Article 6 gives rise to different obligations and rights under the GDPR. Data controllers must determine the lawful basis for processing prior to commencing the processing of personal information and must document it. There is no

[123] Privacy Act 2020, s 22 Information privacy principle 5.

requirement to record the purpose of collection or processing of personal information in New Zealand, however, if a privacy impact assessment is completed, it acts as a record of purpose (albeit one that is not required to be made available to the public or to stakeholders). Under the GDPR, the controller cannot change the legal ground once processing has commenced, and the legal ground must be transparent and made available to the data subject.[124] Establishing a legal ground for processing under the GDPR is a more onerous requirement for controllers than establishing a lawful purpose under IPP 1 of s 22 of the Privacy Act. For many agencies subject to the GDPR, selecting the most appropriate legal basis will be part of their overall data governance strategy.

Other than the legal ground of consent, all the other legal grounds for processing require that processing be "necessary".[125] The question that should be asked is *can the controller achieve the same purpose in a less privacy-invasive way?* The requirement of necessity aims to achieve proportionality between the processing and the purpose. It is important to note that the Privacy Act also requires that the collection of personal information be necessary to fulfil the function or activity of the agency that constitutes its lawful purpose of collection under IPP 1.[126] I will consider two of the more contentious grounds for processing under the GDPR: consent and legitimate interests. - The lawful grounds for processing most relied on by online platforms, search engines, and technology companies.

Consent as a lawful ground

Article 4(11) of the GDPR states that "'consent' of the data subject means any freely given, specific, informed and unambiguous indication of the data subject's wishes ..." The Privacy Act requires that agencies collect personal information directly from the individual concerned unless an exception applies,[127] and that data subjects be given notice of collection, the purpose of collection, intended recipients, and information about the

[124] GDPR, Arts 13(1)(c) and 14(1)(c).
[125] GDPR, Arts 6(1)(b)-(f).
[126] Privacy Act 2020, s 22 Information privacy principle 1(1)(b).
[127] Section 22, Information privacy principle 2. Exceptions include, but are not limited to, where an agency believes on reasonable grounds that compliance would prejudice the purposes of collection (IPP 2(2)(b)); or that non-compliance is necessary to avoid prejudice to the maintenance of the law by any public sector agency (IPP 2(2)(e)(i)); or that compliance is not reasonably practicable in the circumstances of the case (IPP 2(2)(f)).

agency collecting the information when information is collected directly from the data subject concerned.[128] However, there is no explicit requirement of consent in the Privacy Act. The GDPR demands a high level of consent for it to be relied on as a lawful basis for processing. The Privacy Act only requires that an individual be informed of the collection of information about them when an agency collects information directly from the data subject.[129]

Even when consent is relied on as the lawful basis for processing under the GDPR, necessity is still required to a certain extent because a valid consent under the GDPR is given for a specific purpose and the processing must be necessary in relation to that purpose under Article 5(1)(c).[130] Art 6(1)(a) states that:

> 1. Processing shall be lawful only if and to the extent that at least one of the following applies:
>
> (a) the data subject has given consent to the processing of his or her personal data for one or more specific purposes;
> ...

Consent is the foundation of modern 'notice and consent' data privacy regimes. However, reliance on consent is flawed if it is not premised on genuine choice and control, or adequate knowledge: three premises coming under significant strain in an increasingly digitalised world. Consent in the context of big data bears little resemblance to consent in other areas of law, such as consent to medical procedures or sexual encounters. Waldman argues that consent may be suitable for discrete decision-making, but it doesn't scale.[131] Consent is arguably an effective mechanism for authorising a surgical procedure or sexual intercourse, but it does not follow that we give the same consideration to the multitude of requests for consent to the collection and use of our personal information

[128] Section 22, Information privacy principle 3(1).
[129] However, the Government is proposing potential changes to the notification rules for the indirect collection of personal information under the Privacy Act 2020. The proposed changes would broaden the Act's requirements for an individual to be notified when an agency collects their personal information indirectly through a third party in a new IPP 3A. (Ministry of Justice "Key initiatives: Broadening the Privacy Act's notification rules" https://www.justice.govt.nz/justice-sector-policy/key-initiatives/broadening-the-privacy-acts-notification-rules/)
[130] Elena Gil González and Paul de Hert "Understanding the legal provisions that allow processing and profiling of personal data - an analysis of GDPR provisions and principles" (2019) ERA Forum 19 at 600.
[131] Waldman, above n 92, at 53.

online, nor are we as informed of the likely outcomes or possible consequences of doing so. Greater control does not always equate to greater privacy protections for individuals, and especially not online.

Article 4(11) of the GDPR states that:

> ... 'consent' of the data subject means any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her.

Thus, controllers must allow for separate consent to different processing functions and cannot require data subjects give more information than necessary as a condition of a contract.[132] The requirement of informed consent is linked to the principles of fairness and transparency in Article 5(1)(a). Most online privacy notices do not constitute informed or freely given consent as demanded under the GDPR. They are typically long, obscurely worded, and often passed over without consideration. Similarly, the "accept cookies" button that must be clicked to access many websites means very little to most users. Easy-to-understand language is criticised for being overly simplistic. Even if one does read a privacy notice, there is no real option to negotiate the terms. *Accept our terms or don't use our service* is not a real choice when some online platforms are monopolistic players in their field, or where all providers of a similar online service demand the same conditions. Many people cannot understand what uses their data may be put to and they shouldn't have to manage the hundreds, if not thousands, of agencies that hold their information: it is too onerous a burden.

Additionally, in a world where our data privacy choices can have negative implications for others, we should all be able to rely on a regulatory regime that sets boundaries for appropriate, non-exploitative personal information uses. It is not enough to legislate for the rights of the individual, we must also consider the broader social context of information privacy practices. The people most burdened by the harms of big data are those least able to manage their own information and already disadvantaged by

[132] González and de Hert, above n 131, at 601.

discrimination, poverty, or marginalisation. Disadvantages that are exacerbated by unethical information privacy practices.

Data privacy frameworks focussed on the individual are conceptually flawed and unable to address the social privacy harms facilitated by the vast and indiscriminate over-collection of personal information online. Online consent, in the form of ticking a box to access a service, is an inadequate form of authorisation for data collection and use - information that, once collected, can be repurposed and applied in other big data contexts, contributing to social privacy harms. Not only is an individual rights-based model of data privacy conceptually flawed and therefore unable to respond to social privacy harms, it is also ill-equipped to protect *individual* privacy in many big data contexts. The issue of online consent reveals a shortcoming of individual rights-based regimes in protecting individual privacy online.

Legitimate interests as a lawful ground

The concept of a 'legitimate interest' can be a vague and subjective lawful ground for processing. Article 6(1)(f) of the GDPR states that:

> 1. Processing shall be lawful only if and to the extent that at least one of the following applies:
> ...
> (f) processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child.
>
> Point (f) of the first subparagraph shall not apply to processing carried out by public authorities in the performance of their tasks.

This lawful basis excludes the processing carried out by public authorities in the performance of their functions. But public authorities can still use legitimate interests as a basis for their activities other than official functions.[133] When a controller relies on

[133] At 605.

legitimate interests as the basis for processing, it must inform the data subject of those legitimate interests in accordance with the principles of transparency and fairness. There are three parts to using Article 6(1)(f) as the lawful basis for processing: firstly, there must be a legitimate interest on the part of the controller or third party; secondly, necessity for processing; and thirdly, the legitimate interest of the controller or third party must be balanced against the interests and rights of the data subjects.[134] Recital 47 provides the example of where there is a relevant and appropriate relationship between the data subject and the controller, for example where the data subject is a client or in the service of the controller. The Recital goes on to state that:

> At any rate the existence of the legitimate interest would need careful assessment including whether a data subject can reasonably expect at the time and in the context of the collection of the personal data that processing for that purpose may take place. The interests and fundamental rights of the data subject could in particular override the interest of the data controller where personal data are processed in circumstances **where data subjects do not reasonably expect further processing**. [emphasis added]

Recital 49 states that "the processing of personal data to the extent strictly necessary and proportionate for the purposes of ensuring network and information security ... constitutes a legitimate interest of the data controller concerned". Recital 50 states that:

> The processing of personal data for purposes other than those for which the personal data were initially collected should be allowed only where the processing is compatible with the purposes for which the personal data were originally collected ... Further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes should be considered to be compatible lawful processing operations.

Necessity requires an evaluation of whether the processing of data is directly linked to achieving the legitimate interest and whether it could be achieved in a less privacy-invasive way. It would be difficult to rely on legitimate interests in some contexts. The A29WP provides the example of intrusive profiling and tracking for marketing or

[134] At 605.

advertising purposes: "tracking individuals across multiple websites, locations, devices, services or data-brokering" would not be sufficient legitimate interests as a legal basis for processing.[135]

Articles 13 - the right to be informed

Article 13 provides the data subject with the right to be informed of more information about the uses to which their data may be put than the equivalent provision of the Privacy Act.[136] For example, Article 13(2) states:

> 2. In addition to the information referred to in paragraph 1, the controller shall, at the time when personal data are obtained, provide the data subject with the following further information necessary to ensure fair and transparent processing:
>
> (a) the period for which the personal data will be stored, or if that is not possible, the criteria used to determine that period;
> ...
> (c) where the processing is based on point (a) of Article 6(1) or point (a) of Article 9(2) [lawful processing on the basis of consent], the existence of the right to withdraw consent at any time, without affecting the lawfulness of processing based on consent before its withdrawal;
>
> (d) the right to lodge a complaint with a supervisory authority;
> ...
> (f) the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.

The Privacy Act does not stipulate that any of the above points be provided to the data subject at the time when personal data are obtained.

Article 13(3) also states:

---

[135] Article 29 Data Protection Working Party, above n 90, at 15.
[136] Privacy Act 2020, s 22, Information privacy principle 3(1).

> 3. Where the controller intends to further process the personal data for a purpose other than that for which the personal data were collected, the controller shall provide the data subject prior to that further processing with information on that other purpose and with any relevant further information as referred to in paragraph 2.

Authorisation by the individual concerned is an exception to the default rule in IPP 10 of s 22 of the Privacy Act that states that information obtained in connection with one purpose may not be used for any other purpose.[137] However, authorisation by the data subject under this exception does not stipulate the provision of the extent of information as set out in Article 13(2) of the GDPR.

Transparency is an important first step to ensuring fair information privacy practices. Controllers must explain clearly to data subjects how profiling or automated decision-making works. However, in an age of big data, the number of agencies that process our information is extensive; we cannot manage them all and many individuals lack the knowledge, time and expertise to take effective action.

Article 17 - the right to erasure or the *right to be forgotten*

There is no right to erasure or *right to be forgotten* in the Privacy Act. The closest a data subject gets to a right to erasure under the Act is the right in Information privacy principle 7 to request that information about them be corrected. The agency that holds that information must take steps that are reasonable in the circumstances to ensure that information it holds is accurate, up to date and not misleading, having regard to the purposes for which that information may lawfully be used.[138] If an agency chooses not to correct the information as requested and has been provided with a statement of correction by the data subject, it must attach that statement to the information so it will always be read with that information.[139]

Article 17 of the GDPR states that:[140]

---

[137] Section 22, Information privacy principle 10(1)(c).
[138] Section 22, Information privacy principle 7(2).
[139] Section 22, Information privacy principle 7(3).
[140] GDPR, Art 17(1).

> The data subject shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay where one of the following grounds applies:
> ...

Those grounds include where the personal data are no longer necessary for the purposes for which they were collected or otherwise processed;[141] the data subject withdraws consent on which the processing is based and there is no other legal ground for processing;[142] the processing is undertaken on the lawful basis of public interest[143] or the legitimate interest of the controller or a third party[144] and the data subject objects to the processing pursuant to Article 21(1) and the controller cannot demonstrate compelling legitimate grounds for the processing which override the interests, rights and freedoms of the data subject;[145] or the data have been unlawfully processed.[146]

While the right to erasure may be invaluable for individual data subjects whose circumstances meet the requirements of Article 17, it is difficult to see how it could be meaningfully leveraged to address social privacy harms. For a data subject to request erasure, they must first know what agencies hold personal information about them. In a big data environment where thousands of agencies hold data of questionable quality about many of us, process it in variety of ways and/or disclose it to other agencies, the efficacy of the right to erasure as a tool to address social privacy harms is limited. Nonetheless, consideration should be given to having an equivalent concept as an IPP in section 22 of the Privacy Act because it serves as a useful tool for individuals who can identify a publication or other use of information about themselves that is causing a level of harm that is disproportionate to the agency's interest in using it in that way or the public interest in having access to that information.

[141] Article 17(1)(a).
[142] Article 17(1)(b).
[143] Article 6(1)(e).
[144] Article 6(1)(f).
[145] Article 17(1)(c).
[146] Article 17(1)(d).

Article 20 - the right to data portability

Article 20 of the GDPR provides another example of a useful right albeit unsuited, and not designed, to address social privacy harms. Article 20(1) states that:

> The data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to transmit those data to another controller without hindrance ...

Lillian Edwards and Michael Veal argue that data portability is grounded more in consumer protection than data privacy purposes.[147] The objective behind Article 20 is to enable data subjects to switch between providers more easily and therefore seek out more advantageous terms, including potentially, more favourable privacy policies. However, Edwards and Veale identify several problems with this idea. They point to "the well-known inertia of consumers" to switch energy suppliers or banks to save money or get a better service and state that:[148]

> It will take a long time for a competing marketplace of algorithmic model choices to emerge and indeed it is hard to see the current marketplace taking to such voluntarily ... it is quite possible that although the data subject may in theory gain greater control over their personal data, in reality they may not have the knowledge or time to safeguard their data against emerging threats.

Edwards and Veale also point to what they describe as "capricious restrictions" of Article 20. Firstly, it only applies to the data that the subject "provides" and there is no consensus as to whether this includes metadata, or the inferences drawn from that data.[149] The Ministry of Business, Immigration and Employment is currently considering public

[147] Lillian Edwards and Michael Veale "Slave to the Algorithm? Why a 'Right to an Explanation' is Probably Not the Remedy You Are Looking For" (2017-2018) Duke Law and Technology Review 18 at 72.
[148] At 73.
[149] At 74.

feedback on a draft law to establish a consumer data right in New Zealand.[150] While it may be useful, it will not address social privacy harms in a substantive way.

Article 21 - the right to object to processing of personal data

Article 21(1) of the GDPR states that:

> The data subject shall have the right to object, on grounds relating to his or her particular situation, at any time to processing of personal data concerning him or her which is based on point (e) or (f) of Article 6(1), including profiling based on those provisions. The controller shall no longer process the personal data unless the controller demonstrates compelling legitimate grounds for the processing which override the interests, rights and freedoms of the data subject or for the establishment, exercise or defence of legal claims.

Under Article 21, the data subject can object to processing of their personal data where that data is processed in an official capacity or in the legitimate interests of the controller or third party, *unless* the controller can establish that its grounds for processing override the interests, rights and freedoms of the data subject. The data subject can object at any time to the processing of their personal data for direct marketing purposes.[151] There is no equivalent right in the Privacy Act. However, data subjects are often unaware that their personal information is being processed in a particular way or lack an understanding of the harm that may result. More significantly, in the context of social privacy harms, a data subject can only object to the processing of their own data. What is needed is a framework for weighing the harms and benefits of data processing in the public interest, ultimately prohibiting the types of data processing that contribute most prolifically to increasing inequality and the erosion of democracy.

[150] Ministry of Business, Immigration and Employment "Consumer data right" https://www.mbie.govt.nz/business-and-employment/business/competition-regulation-and-policy/consumer-data-right
[151] GDPR, Art 21(2).

Article 22 Automated individual decision making, including profiling

Article 22 arguably contains the most anticipated and debated provisions of the GDPR. The promise of Article 22 has drawn the attention of academics, industry, and data protection authorities in the EU and across the world.

Article 22 states that:

> 1. The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or significantly affects him or her.
>
> 2. Paragraph 1 shall not apply if the decision:
>
> (a) is necessary for entering into, or performance of, a contract between the data subject and a data controller;
>
> (b) is authorised by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or
>
> (c) is based on the data subject's explicit consent.
>
> 3. In the cases referred to in points (a) and (c) of paragraph 2, the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision.
>
> 4. Decisions referred to in paragraph 2 shall not be based on special categories of personal data referred to in Article 9(1), unless point (a) or (g) of Article 9(2) applies and suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place.

Article 22 sets out the right not to be *subject to a decision* based solely on automated processing, including profiling, which produces "legal effects" or "similarly significantly affects" the data subject in contexts other than those described in Article 22(2). There has

been extensive commentary on Article 22, including its basic nature, and much debate around whether the GDPR provides a right to an explanation.

Article 22 - right or prohibition?

Luca Tosoni argues that the key ambiguity in Article 22 is the basic nature of the rule - does it establish a prohibition or is it a right to be exercised by the data subject?[152] The A29WP states that:[153]

> Article 22(1) establishes a general **prohibition** for decision-making based solely on automated processing. This prohibition applies whether or not the data subject takes an action regarding the processing of their personal data. [emphasis added]

The A29WP relies on the principle of the data subject having control over their personal data, consistent with the fundamental principles of the GDPR, as well as Recital 71 in forming its interpretation of Article 22 as a prohibition rather than as a right to be invoked. It points to the wording in Recital 71 below:[154]

> However, decision-making based on such processing, including profiling, **should be allowed** where expressly authorised by Union or State Member law ..., or necessary for the entering or performance of a contract ..., or when the data subject has given his or her explicit consent.

It goes on to state that "this implies that processing under Article 22(1) is not allowed generally".[155] The A29WP holds that Article 22(1) is a general prohibition on fully automated individual decision-making, including profiling that has a legal or similarly significant effect. However, there are exceptions to the rule and where one of those exceptions applies, there must be measures in place to safeguard the data subject's rights and freedoms and legitimate interests.[156]

---

[152] Luca Tosoni "The right to object to automated individual decisions: resolving the ambiguity of Article 22(1) of the General Data Protection Regulation" (2021) International Data Privacy Law 11 2 at 145-146.
[153] Article 29 Data Protection Working Party, above n 88, at 19.
[154] GDPR, Rec 71 in Article 29 Data Protection Working Party, above n 88, at 20.
[155] Article 29 Data Protection Working Party, above n 88, at 20.
[156] At 19.

The two possible interpretations offer different levels of protection to the data subject. If Article 22(1) is interpreted as a prohibition, controllers would not be able to make individual decisions based solely on automated processing, unless an exception in Article 22(2) applies. Alternatively, if Article 22(1) is interpreted as a right to be exercised by the individual, then the use of automated individual decisions is only restricted under the GDPR if the individual has expressly objected to it. The latter interpretation is a much weaker protection, however Tosoni persuasively argues that Article 22(1) should be read, based on a "textual, contextual, systematic, and teleological interpretation ..."[157] Tosoni argues that if the EU legislature had decided to introduce a significant change by transforming a right to be exercised by an individual into a ban, such a decision would have been reflected in a change of wording, or that there would have been an indication of that intention in the legislative history of the provision. He states that:[158]

> In essence, the legislative history of Article 15(1) DPD suggests that the 'right not to be subject to a decision' solely based on automated processing was intended to empower data subjects to decide not to be bound by decisions of this kind, and not to ban automated decision-making in general.

The legislative history of Article 22(1) does not support its interpretation as a general ban against automated decision-making, including profiling. Tosoni's interpretation is compelling: Article 22(1) is not a prohibition but a much weaker provision; offering the data subject *a right to object* to being subject to a decision based solely on automated processing, in certain contexts. Nonetheless, Tosoni believes that the interpretation of Article 22(1) as a general prohibition seems to be the majority view.[159] However, its ambiguous wording still raises confusion, and the issue of interpretation is far from settled.[160]

Tosoni distinguishes between the notion of a "decision" and that of "processing" - the decision constitutes the outcome of the processing and should not be confused with the processing.[161] He states that:[162]

---

[157] Tosoni, above n 152, at 145.
[158] At 150.
[159] At 146.
[160] At 147-148.
[161] At 153.
[162] At 160-161.

> ... it has been convincingly argued that it may be assumed that the central aim of Article 22 is to tackle the same sort of concerns that led to the adoption of Article 15 DPD. These may be summarized in the concern that an extensive use of automated decision-making processes may deprive individuals of the ability to influence decisions that affect them, and lead to the automatic acceptance of decisions reached by a 'machine', as well as to potential inaccuracies and discrimination.

The interpretation of Article 22(1) as *a right to object* to being subject to decision based solely on automated processing is consistent with an interpretation of the GDPR as focussed on the individual and individual rights. The alternative interpretation - that Article 22(1) provides a prohibition on subjecting individuals to decisions based solely on automated decision-making that have legal or similarly significant effects (unless an exception in Article 22(2) applies) - is more consistent with a social conception of privacy. Social conceptions of privacy recognise the social privacy harms that result from the cumulative effects of automated decision-making and profiling. It is likely that banks, insurance companies and big tech would argue that an interpretation of Article 22(1) as a blanket prohibition could hinder the development of innovative products. This may be true, but simply because a product is new or innovative does not necessarily mean that it is beneficial for society. Ultimately, the harms of automated decision-making in certain contexts could outweigh the benefits its innovation provides.

Article 22 and its limited application and scope

'based solely' on automated processing including profiling

The A29WP's *Guidelines on Automated Individual Decision-Making* state that human intervention must be meaningful in order to pull the decision-making out of the realm of being solely automated:[163]

> The controller cannot avoid the Article 22 provisions by fabricating human involvement. For example, if someone routinely applies automatically generated profiles to individuals without any actual influence on the result, this would still be a decision based solely on automated processing.

[163] Article 29 Data Protection Working Party, above n 88, at 21.

If a human being reviews an automated process and takes into account other factors in making the final decision, the decision would not be 'based solely' on automated processing.[164] Oversight should be exercised by someone with "authority and competence to change the decision".[165]

'similarly significantly affects' the data subject

The GDPR does not define "legal" or "similarly significant" effects. The A29WP explains that a legal effect requires that the decision, based solely on automated processing, must affect someone's legal rights. It provides the examples of freedom of association, the right to vote, or take legal action. Additionally, a legal effect could also affect someone's status or rights under a contract, such as the cancelling of a contract, the denial of a benefit granted by law or the refusal of admission to a country or the denial of citizenship.[166]

The addition of "similarly significant affects him or her" in the GDPR means that a decision-making process can fall within the scope of Article 22 without affecting people's legal rights if the effect is similarly significant in impact. Recital 71 provides the examples of "automatic refusal of an online credit application or e-recruiting practices without any human intervention". The A29WP states that "for data processing to significantly affect someone the effects of the processing must be sufficiently great or important to be worthy of attention".[167] It provides examples of decisions that could be considered sufficiently significant, such as decisions that affect someone's financial circumstances; their access to health services; deny someone an employment opportunity or put them at serious disadvantage, or affect their access to education.[168] The A29WP states that solely automated decision-making in the context of online advertising, in many typical instances, will not have a similarly significant effect on individuals. But it could do, depending on the context and including the following considerations:[169]

[164] At 20.
[165] At 21.
[166] At 21.
[167] At 21.
[168] At 22.
[169] At 22.

- the intrusiveness of the profiling process, including the tracking of individuals across different websites, devices and services;
- the expectations and wishes of the individuals concerned;
- the way the advert is delivered; or
- using knowledge of the vulnerabilities of the data subjects targeted.

It is difficult to delineate the effects that a decision based solely on automated processing must cause for it to come within the scope of Article 22. However, it is a high threshold, demanding a serious impact on the individual.

Exceptions in Art 22(2)

The right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her, does not apply if the decision is: necessary for the performance of, or entering into, a contract; authorised by Union or Member State law to which the controller is subject (and also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests); or is based on the data subject's explicit consent.

Member States laws and other "suitable safeguards" in the context of Art 22(2)(b)

Article 22(2) states that Paragraph 1 shall not apply if the decision:

> (b) is authorised by Union or Member State law to which the controller is subject and which also lays down **suitable measures to safeguard the data subject's rights and freedoms and legitimate interests** ... [emphasis added]

Gianclaudio Malgieri analysed the national Member States' laws that have implemented the GDPR in the context of automated decision-making, including profiling. Malgieri describes the approaches of the Member States as "very diverse" - some restrict the scope of the provision to automated decisions producing legal or similarly detrimental effects, whereas other Member States extend it to any decision with a significant impact. The specific safeguards set out in national laws also vary widely. Most Member States just

refer to the three safeguards set out in Article 22(3) - expressing one's point of view, obtaining human intervention, and contesting a decision.[170]

However, some Member States have taken innovative approaches. France and Hungary have a right to legibility or explanations about algorithmic decisions.[171] Malgieri describes how Hungarian law has a broader scope than Article 22, encompassing "all automated decisions *prejudicial to the person* or which have a *significant impact* on the person concerned".[172] Malgieri states that "there is no reference to legal or "similarly significant effects", so that any significant impact on the data subject can be considered relevant ...".[173]

Similarly, Belgian law has a wide scope of protection, including not only legal or similarly significant effects but any "significant effect". Importance is placed on human intervention; it is the only safeguard specifically listed in Belgian law.[174] Austrian law also has a very wide scope, encompassing: "any decision with "detrimental consequences for the data subject or that could significantly affect them ...".[175]

The French law includes a total prohibition on fully, or partially, automated decision-making in the context of judicial decisions where the automated processing "is intended to evaluate aspects of personality".[176] In the years before the GDPR, the French legal system recognised the importance of transparency in relation to algorithms.[177] Malgieri describes French law regulating automated decision-making as 'innovative and complex' for at least three reasons. He states that:[178]

> Firstly, it has a wide scope of application. Secondly, it differentiates among different degree of protection on the basis of the contexts/legal grounds in which an automated

[170] Gianclaudio Malgieri "Automated decision-making in the EU Member States: The right to explanation and other "suitable safeguards" in the national legislations" (2019) Computer Law and Security Review 35 at 1.
[171] At 6.
[172] At 17.
[173] At 17.
[174] At 12.
[175] At 17.
[176] At 13.
[177] At 16.
[178] At 14.

> decision is taken. Thirdly, it is one of the clearest examples of the right to algorithm legibility.
> ... the strictest limitations are provided for judicial decisions evaluating personality aspects of individuals; an intermediate level of limitations is provided for administrative decisions; while fewer limitations are requested for private decisions.

Ireland and the United Kingdom have accountability mechanisms such as notification, and explanations as to why contestation has not been successful. Review, or the requirement for an algorithm impact assessment, are additional procedural safeguards.[179] Interestingly, Slovenia requires a human rights impact assessment for algorithmic decision-making.[180]

Malgieri describes the approach of most Member States as not providing for any specific cases of permitted automated decision-making.[181] However, German law states that the right in Article 22(1) GDPR does not apply if the decision is made in the context of providing services pursuant to an insurance contract. Automated decision making in this context is allowed if the data subject receives a positive decision (their request is approved) but if the outcome is negative (the denial of a request by the data subject) then there are conditions and safeguards in place.[182]

Malgieri raises the question of whether Member States are violating the GDPR by broadening the scope of protection for data subjects. He argues that:[183]

> In principle, any Member State should be free to guarantee a higher level of protection for its citizens if it also respects EU legislation's objectives.
> ...
> In sum, any national divergence in terms of data subjects' right or controllers' obligation needs to be explicitly allowed by the GDPR, otherwise it might be an unjustified restriction of free movement of data within the Union.

[179] At 6.
[180] At 18.
[181] At 6.
[182] At 7.
[183] At 19.

The Member States have different approaches to regulating automated decision-making, for example, using regulation as a tool to exempt specific sectors (i.e. insurance) or to relieve the burdens of "suitable safeguards" for data controllers.[184] Most Member States do not include a right to an explanation of an individual decision based on automated processing. However, in Hungary there are obligations on the data controller to inform the data subject of "the methods and criteria used in the decision-making mechanism"[185] and in France, the explanation should be based on the "rules defining the data processing and the main features of its implementation".[186] Malgieri notes that the Hungarian provision does not make it clear whether the explanation is ex ante or ex poste or whether it needs to be specifically tailored to the individual outcome. Whereas the French law specifies that the 'main features of the implementation' should form part of an ex poste explanation.[187] Malgieri states that:[188]

> In sum, the Hungarian regulation, together with the French one, seems one of the most innovative examples of providing individuals with more transparency about the "black box" algorithms, through specific legibility safeguards.

New Zealand could benefit from consideration of the "suitable safeguards" that Member States have adopted domestically, particularly those of Hungary and France in relation to the legibility of automated decision-making systems.

Article 22 and the right to an explanation

Sandra Wachter, Brent Mittlestadt and Luciano Floridi argue that there are several reasons to doubt both the legal existence and feasibility of a right to an explanation of automated decision-making systems. Firstly, the authors argue that the GDPR only mandates that data subjects receive meaningful, but properly limited, information about the logic involved (Articles 13-15) as well as the significance and envisaged consequences: 'a right

[184] At 20.
[185] At 20.
[186] At 20.
[187] At 22.
[188] At 17.

to be informed'.[189] They suggest the ambiguity and limited scope of Article 22 raises questions about the actual protection afforded to data subjects. They stated, prior to the GDPR's implementation, that "the GDPR lacks precise language as well as explicit and well-defined rights and safeguards against automated decision-making, and therefore runs the risk of being toothless".[190]

Their argument stands six years on. I agree with Wachter et al's argument that a right to explanation is not legally mandated by the requirements set out in Article 22(3)[191] despite Recital 71, which states that:

> In any case, such processing should be subject to suitable safeguards, which should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view, **to obtain an explanation of the decision reached after such assessment** and to challenge the decision ... [emphasis added]

Recitals explain the background to the GDPR and provide guidance on how to interpret the articles. However, the recitals are not legally binding and cannot create legal expectations in themselves. They are important, and useful, in that they provide explanations for the adoption of the Regulation. Also, when the articles require interpretation, the recitals should be considered.[192] Wachter et al argue that, despite Recital 71 stating that a suitable safeguard should include the right to obtain an ex-poste explanation of a decision reached after an automated assessment, this is not stipulated in the articles of the GDPR.

Wachter et al argue that a right to an explanation of automated decision-making would require a corresponding duty on the part of the controller, which is absent from the GDPR. They go further to argue that the omission of a 'right to an explanation' from Art 22 "appears to be intentional".[193] Earlier drafts of the GDPR contained "stricter safeguards

[189] Sandra Wachter, Brent Mittlestadt and Luciano Floridi "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation" (2017) International Data Privacy Law 7 2.
[190] At 76.
[191] At 80.
[192] At 80.
[193] At 80-81.

... including a legally binding right to explanation of specific decisions".[194] Wachter et al make a persuasive argument that the GDPR does not create a right to an explanation, but rather a limited 'right to be informed'.[195] The effect of Article 22 may have been over-hyped by commentators.

Wachter et al examine three possible bases for a right to an explanation in the GDPR:

(1) **safeguards** against automated decision making (Art 22(3) & Recital 71)
(2) **notification duties** of data controllers (Arts 13-14 and Recitals 60-62)
(3) **the right to access** (Art 15 and Recital 63)[196]

They argue that the claim that the GDPR offers a right to an explanation:[197]

> ... conflates (i) legally binding requirements of Article 22 and non-binding provisions of Recital 71 and (ii) notification duties (Articles 13-14) that require data subjects to be provided with information about "the **existence of automated decision-making**, including profiling, referred to Article 22(1) and (4) and, at least in those cases, **meaningful information** about the **logic involved**, as well as the **significance** and the **envisaged consequences** of such processing for the data subject" (emphasis added).

The authors argue that the GDPR's right of access in Article 15 provides a stronger legal basis but not for the type of explanation imagined or purported in public discourse. It establishes a 'right to be informed' about the functionality of automated decision-making systems, not to an explanation of specific decisions. A meaningful right to an explanation is not provided by the GDPR. Even if it was, it would have a very limited scope. It would only apply to decisions based *solely* on automated processing and those decisions with legal or similarly significant effects. The GDPR may be lauded as providing protections against automated decision-making but has a significant transparency and accountability

[194] At 81.
[195] At 77.
[196] At 79.
[197] At 77-78.

gap.[198] I argue it is largely ineffectual at addressing social privacy harms because of its ambiguities and limited scope, and its focus on the individual.

Margot E. Kaminski and Gianclaudio Malgieri disagree with Wachter, Mittlestadt and Floridi that there is no right to an explanation of automated decision-making in the GDPR. They argue that Article 22 creates an *ex post* right to an explanation of an individual decision with significant effects, made by a solely automated system. This is because they believe data subjects have to be given enough information about the decision-making process in order to be able to invoke their other rights under the GDPR, including the right to contest the decision.[199] However, Wachter et al's argument is more persuasive because the articles of the GDPR do not contain a right to an *ex post* explanation of an automated decision unlike earlier drafts of the GDPR, which did.

Transparency unlikely to provide a complete remedy to algorithmic harms

Lilian Edwards and Michael Veale believe that transparency as a response to the issues of unfairness, discrimination and opacity of algorithms, particularly machine learning algorithms, is unlikely to provide a complete remedy to algorithmic harms anyway. They agree that the right to an explanation in the GDPR is uncertain and unclear. In addition to confusion or uncertainty as to when such a right may be activated, types of "explanations" that computer scientists have developed may not meet the legal expectation of "meaningful information about the logic of processing".[200] Edwards and Veale "fear that the search for a "right to an explanation" in the GDPR may be at best distracting, and at worst nurture a new kind of "transparency fallacy".[201]

Edwards and Veale argue that other parts of the GDPR might prove more useful in making the use of algorithms more ethical and explainable: the right to erasure; the right to data portability; privacy by design; data protection impact assessments; certification and privacy seals.[202] They also point out that a very similar right to that purported to be a 'right

[198] At 97.
[199] Margot E. Kaminiski and Gianclaudio Malgieri "Algorithmic impact assessments under the GDPR: producing multi-layered explanations" (2021) International Data Privacy Law 11 2 at 127.
[200] At 18.
[201] At 19.
[202] Edwards and Veale, above n 147, at 19.

to an explanation' in the GDPR existed in the EU's DPD, which preceded the GDPR, in place from 1995.[203] Edwards and Veale ask if "transparency, in the form of explanation rights, is really a useful remedy for taming the algorithm as it intuitively seems to be".[204]

Edwards and Veale acknowledge the conceptual flaw in attempting to address algorithmic harms with an individual rights-based approach to data privacy regulation. They state that:[205]

> ... data protection remedies are fundamentally based around individual rights - since the system itself derives from a human rights paradigm - while algorithmic harms typically arise from how systems classify or stigmatise groups. While the problem is known as a longstanding issue in both privacy and equality law, it remains underexplored in the context of the "right to an explanation" in machine learning systems.

They warn that "we are in danger of creating a "meaningless transparency" paradigm to match the already well known "meaningless consent" trope.[206]

Edwards and Veale argue that instead of "providing individual rights on demand to data subjects", what might be a more effective way of reviewing the effectiveness, fairness and accuracy of machine learning systems is building better machine learning systems and empowering agencies to review them (NGOs, civil society groups, or regulators).[207]

Edwards and Veale ask if there is a way beyond individual rights in the GDPR that provides better outcomes for society as opposed to "merely providing individual users with rights as tools which they might find impossible to wield to any great effect?"[208] They argue that even if the rights in the GDPR were useful tools for addressing algorithmic harm, it is still incumbent upon the individual to exercise them. They state that the difficulties of individual data subjects enforcing their rights is recognised by data privacy regimes that place general oversight in the hands of data protection authorities.

---

[203] At 20.
[204] At 21.
[205] At 22.
[206] At 23.
[207] At 23.
[208] At 75.

However, these authorities are often critically underfunded.[209] A social conception of privacy that considers the effects of information practices holistically and does not place the burden of enforcement on individuals, is necessary to address the social privacy harms of big data.

Multi-stage profiling or decision-making systems raise complications for Article 22

Reuben Binns and Michael Veale emphasise that it is not only the degree of human involvement and the issue of 'significance' of the decision that are contentious issues in relation to Article 22. They argue that "multi-stage automated decision-making scenarios ... raise more fundamental questions about the scope of Article 22 which require further analysis in their own right".[210]

Multi-stage profiling or decision-making systems, where the stages are potentially both manual and automated, raise five distinct complications:[211]

1. the potential for selective automation on subsets of data subjects despite generally adequate human input;
2. ambiguity around where to locate the decision itself;
3. whether significance should be interpreted in terms of any potential effects or only selectively in terms of realised effects;
4. the potential for upstream automation processes to foreclose downstream outcomes despite human input;
5. a focus on the final step may distract from the status and importance of upstream processes

Binns and Veale argue that these challenges will make it difficult for courts or regulators to distil a set of clear, fair and consistent interpretations of Article 22 for many real-world contexts.[212]

---

[209] At 74-75.
[210] Reuben Binns and Michael Veale "Is that your final decision? Multi-stage profiling, selective effects, and Article 22 of the GDPR" (2021) International Data Privacy Law 11 4 at 331.
[211] At 332.
[212] At 332.

Multi-stage profiling systems can create ambiguities around human intervention and significance where:[213]

1. there are multiple outcomes with different significance
2. there are different levels of human involvement between outcomes or segments of the population
3. there are multiple stages at which either significance or human involvement differ

Binns and Veale identify different roles that automation can play in assisting decision-making:

(1) Providing information to a decision-maker (Supporting);[214]
(2) Determining which cases get to a human decision maker or passed to another automated process (Triaging);[215]
(3) Consolidating decisions from one or more human decision maker(s) (Summarizing).[216]

The example of summarising provided is exam marking by humans where the automation involves adding up handwritten scores or calculating an average. With the summarising scenarios, "the main determinant is the collection of human judgements represented in the underlying data ... a human and machine would only differ insofar as they made random errors".[217]

The use of automation in a supporting role, for example where a human being considers an automated profile and then makes a decision, is the least controversial under Article 22. Because no decision is made without human input, none of the decisions are based "solely" on automated processing, and therefore, are not subject to Article 22.[218] There are obvious challenges to this. For example, if a human lacks the authority or competence to overturn a decision, or they never or rarely overturn decisions, then it can be argued

[213] At 321.
[214] At 322.
[215] At 322-323.
[216] At 323-324.
[217] At 324.
[218] At 325.

that these decisions remain based solely on automated processing. This issue seems to have been given the most public consideration to date.[219]

Locating decisions

Binns and Veale consider *where* a decision take place. For example, with security screening - is the decision located at the moment an individual is identified as a potential security risk (through profiling) or when an official decides to select that individual for further screening? They state that:[220]

> ... [if the] officials have the discretion to decide whether or not to select the individual for further screening, then arguably any relevant 'decision' has only been taken at this later point. Even if we can build a counterfactual showing that the decision would not have occurred but for the earlier applied profile, the potential of it being overturned at a later point might lead us to believe the decision is best located there.

If the decision happens upstream (at the automatic assignment of risk) then Article 22 applies **but** the prohibition can be avoided through consent, necessity for contract, or a provision in Member State law, provided additional safeguards are in place. However, if the decision happens downstream at the point where a human intervenes, then Article 22 does not apply because it is not a solely automated decision (unless it is a case of rubber-stamping, for example). Binns and Veale highlight the shortcomings of such an approach. They state that:[221]

> The issue of 'representational harm', related to the perpetuation of stereotypes, cultural denigration and the subordination of certain groups is an important structural consideration in automated systems, although is unlikely to be seen as significant in data protection law given the high barrier of significance as being similarly significant to 'legal' effect, particularly due to the limitations of the wording of Article 22, insofar as harms relate to constructed groups.

[219] At 326.
[220] At 325.
[221] At 326.

The harm borne by marginalised groups whose members find themselves unfairly and repeatedly subject to singling out due to their perceived risk is inadequately addressed by the GDPR.

Binns and Veale state that:[222]

> While it is clear that the initial profiling plays a role in producing the ultimate significant effect (of being subjected to further screening), where downstream discretion exists, it may not be sufficient by itself ... at some point, in some cases, a human will act in such a way as to bring about the significant effect through confirming the profile ... While there seem no easy answers to this quandary in case law or regulatory guidance, the first practical step this entails is that regulators may need to zoom out and seek empirical evidence on a much broader system than they initially thought they were investigating.

Significance

Binns and Veale ask is significance "a condition that is determined where there is a 'potential' of a significant outcome, or not until a significant outcome has been 'realised'?"[223] They state that "confusingly, both the 'realisation' and the 'potential' approach have some backing in the law"[224] but maintain that the 'potential' of a significant outcome is the only sensible approach. They state that a "decision mechanism should be considered 'significant' if it is reasonably foreseeable as such for some individuals who would be subject to it".[225]

Is the final step decisive?

The final step should not be decisive in determining whether a system is solely automated because not all systems ending with an automated step *are* solely automated.[226] Where an automated system summarises human input "the upstream human input could be

[222] At 326.
[223] At 326.
[224] At 327.
[225] At 329.
[226] At 329.

substantial enough to render the process not solely automated ..."[227] Binns and Veale state that "conversely, some processes which 'end' with a human step may 'still' constitute decisions which are solely automated.[228] Binns and Veale identify a possible loophole similar to using a human to rubber-stamp automated decisions. They state that:[229]

> Instead of adding a human step onto the end of an automated process in order to render it non-solely automated, this loophole would be exploited by adding an additional 'automated' step, in order to render any prior automated steps outside of the scope of Article 22 altogether ... **If Article 22 only applies to the final decision step, data controllers would have the discretion to hide contentious automated processing steps 'upstream'.** [emphasis added]
> ...
> Closing this loophole would mean bringing upstream automated processing steps back into the scope of Article 22. Automated steps which 'indirectly' result in significant effects via other automated steps would be within scope.

Binns and Veale acknowledge that relaxing the interpretation of 'producing a legal or similarly significant effect' to include the indirect production of these effects is inconsistent with the consensus that agrees where indirect production of such effects via a further human step is not enough to bring the decision within the scope of Article 22.[230] The logic being:[231]

> If 'automated' steps which indirectly produce qualifying effects are within scope when the further step is also automated, why would they not also be in scope when the further step is taken by a human?

Binns and Veale do not aim to resolve this tension but wish to draw attention to the consequences of using the final step to determine the applicability of Article 22 to the entire decision-making process.[232]

---

[227] At 330.
[228] At 330.
[229] At 330.
[230] At 330.
[231] At 330.
[232] At 331.

Article 22 is one of the few parts of the GDPR that lays out a 'bright line rule regime' as opposed to a risk-based approach or balancing controllers' and data subjects' interests.[233] Binns and Veale consider whether a subjective test rather than measuring the extent of automation would be preferable. However, they point out that controllers are unlikely to voluntarily classify their systems as coming within the scope of Article 22:[234]

> Controllers would likely seize the opportunity to claim that systems they operate are not risky, and given the opacity surrounding their use, and limited regulatory capacity, few claims would be likely to go challenged.

Article 22 raises many questions and complications, beyond those concerning the level of human involvement and the significance of the decision, as Binns and Veale highlight. The ambiguities surrounding the interpretation of Article 22, and its limited scope, unfortunately weaken its application and usefulness as a data protection tool.

Article 25 - data protection by design and by default

Article 25 of the GDPR sets out the requirements of the controller to implement data protection by design and by default measures. Article 25(1) states that:

> Taking into account the state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing, the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organisational measures, such as pseudonymisation, which are designed to implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects.

Article 25(1) is concerned with the security of data and imposes a qualified duty on controllers to implement effective data protection principles and safeguards relative to

[233] At 331.
[234] At 331.

the risks of the processing. Lee Bygrave notes that one of the GDPR's core principles is 'integrity and confidentiality' (Article 5(1)(f)) and that Article 25(1) embraces this concept.[235]

Article 25(2) is also concerned with data minimisation; it states that "[t]he controller shall implement appropriate technical and organisational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed..." This obligation of data minimisation is also applied to "...the extent of their processing, the period of their storage and their accessibility".[236]

Bygrave states that:[237]

> A security remit is also clearly present in the 'data protection by default' requirements of Article 25(2) GDPR. To a large extent, those requirements are essentially concerned with keeping personal data 'lean and locked up', as it were.

The Privacy Act does not specifically refer to privacy by design, however, IPP 1 acts as a collection- or data- minimisation principle.[238] Additionally, IPP 5 stipulates that an agency must ensure personal information is protected by security safeguards, that are reasonable in the circumstances, against loss,[239] unauthorised access, use, modification, or disclosure,[240] and other misuse.[241] Where the Privacy Act fails to match the requirements of Article 25 is in regard to retention of data in IPP 9, which as I have previously stated, does *not* stipulate that personal information be retained "for no longer than is necessary for the purposes for which the personal data are processed ..."[242] Other than in that regard, the fact that the Privacy Act does not have a specific section or IPP on privacy by design does not mean that consideration of data minimisation principles and security safeguards are not a necessary consideration for agencies wanting to comply with the IPPs of s 22 of the Privacy Act.

---

[235] Lee Bygrave "Security by Design: Aspirations and Realities in a Regulatory Context" (2021) Oslo Law Review 8 3 126-177 at 135-136.
[236] GDPR, Art 25(2).
[237] Bygrave, above n 235, at 136.
[238] Privacy Act 2020, s 22, Information privacy principle 1.
[239] Section 22, Information privacy principle 5(a)(i).
[240] Section 22, Information privacy principle 5(a)(ii).
[241] Section 22, Information privacy principle 5(a)(iii).
[242] GDPR, Art 5(1)(e).

Article 32 - security of processing

The text of Article 32(1) is very similar to that of Article 25(1), balancing the context of the processing and available technology against the risks to the rights and freedoms of natural persons. Article 32 requires that both controller and processor "implement appropriate technical and organisational measures to ensure a level of security appropriate to the risk ..."[243] These measures include:[244]

> (a) the pseudonymisation and encryption of personal data;
> (b) the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services;
> (c) the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident;
> (d) a process for regularly testing, assessing and evaluating the effectiveness of technical and organisational measures for ensuring the security of the processing.

Article 32(1) requires, among other considerations, that account be taken of the 'state of the art'. Bygrave states that:[245]

> In very general terms, the criterion means that security measures should be in accordance not only with the current state of technological advancement, but also with contemporary practices, standards and other norms for technical and organisational management that are generally regarded as offering the optimal degree of security at the time.

Article 32(1) prescribes a greater level of detail than IPP 5 of the Privacy Act, for example, the requirement that security risk be considered at the design stage of a system for processing personal data. Bygrave states that Article 32 "demands proactive consideration of security at the outset of establishing a system for processing personal data".[246]

---

[243] Article 32(1).
[244] Article 32(1)(a)-(d).
[245] Bygrave, above n 235, at 140-141.
[246] At 141.

Bygrave argues that the GDPR shows a "pronounced concern" for data protection by design and by default and security by design as means to protect against risks to fundamental rights throughout the entire process from prior to collecting or generating data, to its processing and/or disclosure.[247] However, he also notes that there is a lack of economic incentives for companies to invest in cybersecurity. The penalties (not just pecuniary) are affordable for large companies. Manufacturers view security as expensive and not necessarily profitable and will "often escape having to bear the full economic costs of security breaches arising from flaws in their products".[248] Another contributor to many security breaches is the design of the Internet: at its core, 'simple, flexible and open'.[249]

The consensus that Article 32 requires a reasonable effort to establish a level of security appropriate to the risk, is consistent with the interpretation of IPP 5. A security breach is not necessarily a breach of IPP 5, provided the information was protected by security safeguards that were reasonable in the circumstances. Article 32(1) requires a balancing of the state of the art and the costs of implementation against the nature and purpose of the processing and an assessment of the risks and their likelihood and severity. Bygrave refers to "best *reasonable* effort".[250]

The privacy by design and security measures in the GDPR are important. Data minimisation is a fundamental principle that helps to reduce many of the harms of big data processing. However, Bygrave raises the risk of 'security' being manipulated for purposes of national security or in the case of corporations, "as a pretext for protecting their commercial interests at the expense of the legitimate interests of others".[251] A social conception of data privacy would require consideration of the justifications for processing to be based on its broader social implications.

Articles 35 and 36 - Data protection impact assessment and prior consultation

Article 35(1) states that:

[247] At 152.
[248] At 150.
[249] At 151.
[250] At 168.
[251] At 175.

> Where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data ...

Article 35(3) states that:

> A data protection assessment referred to in paragraph 1 shall in particular be required in the case of:
>
> (a) a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;
>
> (b) processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 10; or
>
> (c) a systematic monitoring of a publicly accessible area on a large scale.

A data protection impact assessment (DPIA) shall be completed by the controller where data processing is likely to result in a high risk to the rights and freedoms of natural persons. Article 36(1) states that:

> The controller shall consult the supervisory authority prior to processing where a data protection impact assessment under Article 35 indicates that the processing would result in a high risk in the absence of measures taken by the controller to mitigate the risk.

Where the supervisory authority is of the opinion that the intended processing would infringe the GDPR, it may impose a temporary or permanent ban on that processing under Article 58(2)(f).

The Privacy Act in contrast, does not demand a DPIA be completed where processing is likely to result in high risk to the rights and freedoms of natural persons or in any other context. Nevertheless, privacy impact assessments (as they are more commonly referred to in New Zealand) are routinely completed by government agencies and are recommended by the Office of the Privacy Commissioner as an essential part of many projects or proposals to identify risks associated with the collection and use of personal information.[252]

Limitations of (data) privacy impact assessments

Margot Kaminiski and Gianclaudio Malgieri identify the biggest shortcoming of the DPIA requirement, which is that there is no mechanism for mandatory disclosure. They argue that public disclosure would enable individual data subjects to avoid companies with bad privacy practices and policies. Kaminiski and Malgieri state that:[253]

> The GDPR puts a lot of faith in the behaviour of companies and in the capacity of regulators ... the GDPR often tasks companies with coming up with the substance of (i) how individual rights will be implemented and (ii) how to address unfairness, biases, and discrimination-related concerns about algorithms. In the absence of public oversight, how can we be sure that this hybrid system of individual rights and collaborative governance is working towards the public good?
> One possible solution is to use heavy regulatory oversight. But the GDPR's enforcers have not, historically, been well-resourced in relation to the companies they regulate.

Kaminiski and Malgieri recommend companies link the DPIA content on processing to the information they disclose to individuals.[254] They outline the various requirements of different proposals for an ideal DPIA but add that:[255]

[252] Office of the Privacy Commissioner https://www.privacy.org.nz/publications/guidance-resources/privacy-impact-assessment/
[253] Kaminiski and Malgieri, above n 199, at 133.
[254] At 134.
[255] At 136.

> ... these proposals articulate the ideal ... in practice, DPIAs even in the European context have tended to focus on data quality and data security, leaving out broader social and legal impact despite aspirational language to the contrary.

The release of information to the public (not necessarily the source code) contained within DPIAs could lead to, and facilitate, external input into how a system is developed and promote its legitimacy. Malgieri and Kaminiski also point to more recent proposals calling for expert input and oversight as a central component of impact assessments.[256]

The GDPR's focus on the individual and individual rights, risks missing the impact of surveillance on groups and society at large. Malgieri and Kaminiski state that:[257]

> ... a systemic approach to risk assessment and risk mitigation requires data controllers to analyse how the system impacts not just individuals but groups. We believe that systemic and group-based explanations uncovered during an [algorithmic impact assessment] can and should be communicated to outside stakeholders, and that a case can be made that such a release is required under the GDPR.

Kaminiski and Malgieri support a move toward a model algorithmic impact assessment that produces multi-layered explanations, promoting "stakeholder input, expert oversight, and public disclosure as essential elements of an effective impact assessment".[258] Consideration should be given to "algorithms as systems embedded in human systems" - those who design it, use it, and the training provided, as well as human freedoms or constraints in using the system. A model algorithm impact assessment should be continuous with ongoing assessments and performance evaluation, especially for algorithms that change or evolve over time, and when deployed in different contexts. This would assist in the consideration of social harms that go beyond impacted individuals.[259]

Article 42 - Certification

[256] At 137.
[257] At 138.
[258] At 139.
[259] At 139.

Article 42 describes the "establishment of data protection certification mechanisms and of data protection seals and marks, for the purpose of demonstrating compliance ..."[260] of data processing operations with the GDPR. Edwards and Veale describe two types of certification:[261]

> 1. certification of the algorithm as software specifying its design specifications or output (performance-based standards); and
>
> 2. certification of the whole person or process using the system to make decisions

Edwards and Veale state that:[262]

> Promising as this may sound, voluntary self- or co-regulation by privacy seal has had a bad track record in privacy, with recurring issues around regulatory and stakeholder capture.

They point to the "demise of the EU-US data agreement SafeHarbor" which was "externally validated for years by trust seals like TrustE".[263] They argue that Europeans are right to be sceptical about the likelihood of "substantive compliance with privacy rights by certification".[264] The Office of the Privacy Commissioner in New Zealand had a Privacy Trust Mark program however it has paused consideration of new privacy trust mark applications. The Commissioner has decided to evaluate the programme to see if it is meeting its objectives of promoting understanding and acceptance of the IPPs. There is no timeframe for the completion of this evaluation.[265]

Edwards and Veale warn that "DPIAs, PbD, certification and the general principle of "accountability" in the GDPR bring with them a real danger of formalistic bureaucratic overkill alongside a lack of substantive change ..."[266] Similarly, Waldman describes the process of "performative privacy" where box ticking, the filling out of a multitude of data

---

[260] GDPR, Art 42(1).
[261] Edwards and Veale, above n 148, at 79.
[262] At 80.
[263] At 80.
[264] At 80.
[265] Office of the Privacy Commissioner https://www.privacy.org.nz/resources-2/applying-for-a-privacy-trust-mark/
[266] Edwards and Veale, above n 148, at 80.

privacy impact assessments and a flurry of compliance-related tasks, creates the illusion that tech companies are adopting increased privacy measures.[267]

Article 82 - Right to compensation and liability

Article 82(1) states that:

> Any person who has suffered material or non-material damage as a result of an infringement of this Regulation shall have the right to receive compensation from the controller or processor for the damage suffered.

A controller is liable for the damage caused by processing that infringes the GDPR. A processor is only liable for damage caused by processing where it has not complied with the obligations of the GDPR that apply specifically to processors or where it has acted outside of, or contrary to, the lawful instructions of the controller.[268] Controllers and processors are not liable for damage caused by processing if they can prove they are not in any way responsible for the event giving rise to the damage.[269]

Article 83 - General conditions for imposing administrative fines

The administrative fines under Article 83 that can be issued by a supervisory authority for infringements of the GDPR are considerable: up to €20,000,000 or 4% of total worldwide annual turnover.[270] Fines can be issued to both the controller and processor. The maximum fines are intentionally hefty to encourage compliance.

There has been an assumption that the significant fines in Article 83 would act as a deterrent and encourage compliance with the GDPR. However, this assumption has not proven true. It is correct that the GDPR has had an impact on the way companies operate, but it is not as the initial assumption suggests. The compliance burden of the GDPR has had a disproportionate impact on start-ups and small to medium sized companies.[271] The

[267] Waldman, above n 92, at 211.
[268] GDPR, Art 82(2).
[269] Article 82(3).
[270] Article 83(5).
[271] Christian Peukert, Stefan Bechtold, Michail Batikaas, Tobias Kretschmer "Regulatory Spillovers and Data Governance: Evidence from the GDPR" (2022) Marketing Science 41 4; Rebecca Janβen, Reinhold

targets of the significant fines[272] – the big tech companies that make their profit from the processing of personal data – have not changed their operating models in any significant way in response to the GDPR's implementation. Mona Naomi Lindvedt argues that:[273]

> If the purpose of the high fines in the GDPR was to rein in the large tech companies, it has not been successful. There are no major changes in behaviour or the business models of the companies, although data protection is duly mentioned and privacy policies are updated.

This is consistent with Waldman's argument that the privacy teams in big tech companies are symbolic structures; providing an aura of compliance but effecting no real change in privacy for individuals who use their products and services.[274] Waldman is critical of vague and ambiguous privacy laws that rely on interpretation by privacy professionals subject to corporate capture. He argues that "coercive bureaucracy" and "normalizing performances" are two social forces operating within the information industry that put privacy at risk.[275]

Lintvedt also describes the shortcomings of Article 83 in its inconsistent application by Member States. There are no provisions in the GDPR to ensure the harmonisation of how fines are calculated or imposed and there is a significant divergence in practice between the Data Protection Authorities.[276] Lindvedt points to a lack of transparency in how the fines are calculated and to whom they are given.[277] Bygrave is also critical of the inconsistencies in application of the GDPR's administrative fines scheme but notes that it is on the EDPB's agenda to agree on a methodology for calculating fines.[278]

For government agencies, bad publicity and damage to reputation may be more effective deterrents than fines,[279] but this is arguably not the case for the big tech companies.

Kesler, Michail E. Kummer, Joel Waldfogel "GDPR and the Lost Generation of Innovative Apps" (2022) National Bureau of Economic Research, Cambridge, Massachusetts.

[272] Mona Naomi Lintvedt "Putting a price on data protection infringement" (2022) International Data Privacy Law 12 1 at 5.

[273] At 12.

[274] Waldman, above n 92, at 130-132.

[275] At 5.

[276] At 5-9.

[277] At 9-11.

[278] Bygrave, above n 230, at 171.

[279] Lindvedt, above n 267, at 5.

Lindvedt argues that other enforcement measures in the GDPR may be more effective in stopping undesirable information practices, such as the power of Data Protection Authorities to ban processing temporarily or permanently, which may also be more punitive for a data controller whose business model relies on the processing of data, than a fine.[280] It may be that the starkest difference between the GDPR and the Privacy Act - the respective level of administrative fines - is not as significant a deterrent as might be expected.

Inherent limitations of 'notice and consent' models of data privacy regulation

The GDPR prescribes greater privacy rights for the individual than the Privacy Act. Enhancing individual privacy rights may also go some way towards reducing social privacy harms. However, increased transparency and greater 'control' do not always translate into greater privacy in real terms, particularly in the context of big data. Knowing how badly Facebook uses my information does not change the fact that I need a Facebook account to know when and where my child's cricket match will be each week. I download classroom apps despite their privacy policies that state how they will share my information with their advertising affiliates or sell it along with other assets as part of a merger or acquisition, because the alternative is missing out on updates from my children's teachers and other critical information about school activities. Extra 'consent boxes' popping up continually when I search the Internet are an irritant, not empowering. I do not have the time or inclination to read the privacy policy of every website I visit; some sites are impossible to navigate without accepting cookies. Rather than helping me to make better privacy choices, increased transparency online contributes to sense of futility when it comes to being able to manage my information.

The market dominance of Facebook, Google, Amazon, and Uber mean there is often no viable alternative to using a particular service, leading to a sense of inevitability about relinquishing personal data on behalf of users.

[280] At 10-11.

The GDPR has not changed the business model of surveillance capitalists. Increased notification, 'control' and transparency requirements do not challenge surveillance capitalism in any meaningful way. Instead, they make navigating the Internet more difficult and contribute to a sense of inevitability on the part of users of online platforms about the loss of control of their information.

The burden of big data processing typically falls heaviest on those already marginalised and discriminated against. An individual-rights based model that places the onus of privacy protection on the individual does little to address this burden: those who might benefit the most from data privacy rights are the least equipped to assert them.

The cross-jurisdictional nature of data privacy limits the effectiveness of any one state's or country's data privacy laws, especially a small country like New Zealand with limited influence on the world stage. The market dominance of companies like Meta (Facebook) and Alphabet (Google) contributes to significant asymmetries in knowledge, wealth, and power between platform and user. The shortcomings of our regulatory framework draw attention to the bigger picture of the dominance and influence of a few surveillance capitalists and the threat they pose to privacy and thus, democracy. I agree with Waldman that privacy law should be about regulating the power structures of informational capitalism.[281]

Unlike identifiable privacy harms suffered by an individual and attributable to an agency, the Privacy Act is ill-equipped to address social privacy harms caused by the cumulative effects of information practices. Where certain applications of big data contribute to the social privacy harms of increased inequality and the undermining of democracy - individual consent cannot be an ethically justifiable form of authorisation for that data processing.

[281] A similar concept to Zuboff's "surveillance capitalism" but includes all companies that profit from the monetisation of personal information, unlike surveillance capitalism that just includes those big tech giants whose entire business operation is premised on processing personal information and generating targeted advertising to sell to marketers (Facebook, Google). Surveillance capitalism also encompasses the idea that these companies that operate as surveillance capitalists want not just to be able to predict peoples' behaviour, but ultimately to be able to direct it as well.

That is not to say that individual consent is no longer relevant or important. Instead, determining *when* individual consent is a necessary and meaningful way of authorising information collection and use, will be dependent on context. In most big data contexts, consent falls short in safeguarding both individual, and collective, interests.

A rights-based privacy framework is an incomplete response to the privacy harms of big data because people won't always know when they have been subject to unfair information practices or be able to identify how they have been harmed by predictive systems or biased algorithms. Privacy harms can be amorphous and difficult to define; privacy harms can be cumulative and one action in itself may not harm an individual, but cumulatively the practices cause harm. Additionally, harm can occur well after the initial information exchange, making it difficult to identify the agency responsible in order to seek redress. Harm can also be caused by multiple agencies. Critically, those most likely to bear the burden of big data - the marginalised and indigent - are ill-equipped to assert their rights; and arguably should not have to do so if there are ways of preventing unfair information practices from occurring in the first place. The threat of possible litigation can encourage compliance with the law. The problem is, if individuals don't know how their data is being used, or by whom, and they can't recognise how they have been harmed, then they will be unable to challenge the agencies that are breaching their privacy rights.

Stronger individual privacy rights may go some way towards preventing social privacy harms. There are undoubtedly amendments that could be made to both the Privacy Act and the GDPR that would better protect individual privacy rights. Neither instrument goes far enough but even if it did, a rights-based framework in isolation, cannot adequately address the harms of big data. Individual privacy rights are not a complete response to the harm caused by the processing of personal information in an age of big data.

Conclusion

Big data processing can harm those outside of an information transaction as well as consenting data subjects in ways far beyond what they could reasonably have contemplated at the time of consent. There are actions we can, and should, take to safeguard New Zealanders' data, which I will discuss in Chapter Six. To address social privacy harms effectively, we must start by thinking of privacy as a social concept. A

social conception of data privacy would better inform amendments to the Privacy Act and/or new laws to address social privacy harms. A social conception of privacy could justify restrictions on harmful information practices such as algorithmic profiling that exacerbates existing inequities or creates new ones, or support a moratorium on the use of facial recognition technology by the private sector, for example.[282]

The Privacy Act has been criticised for not going as far as comparable legislation in other jurisdictions. It does not specifically address automated decision-making or profiling; it does not include a 'right to be forgotten', or contain significant administrative fines. While the amendments to the Privacy Act 2020 could have gone further, even aspiring to GDPR equivalency, I argue that it would still not be an adequate response to the social privacy harms of big data. Although the GDPR provides enhanced privacy rights for the individual, data privacy frameworks that focus on individual rights are conceptually flawed when it comes to addressing the social privacy harms of big data. The GDPR embodies an individual rights-based approach to privacy harm that puts the individual front and centre of its responses. Unsurprisingly, the GDPR fails to adequately address social privacy harms described in Chapter Three.

[282] Foodstuffs is trialling facial recognition technology in its North Island supermarkets (Mildred Armag, "Facial recognition in supermarkets could harm Māori, Pasifika experts fear" stuff.coc.nz (17 February 2024) https://www.stuff.co.nz/nz-news/350177479/experts-fear-foodstuffs-facial-recognition-trial-could-harm-maori-pasifika

Chapter Six – Beyond the Privacy Act 2020: responses to the potential harms & benefits of big data

Introduction

As demonstrated throughout this thesis, there are not only benefits and harms to the individual of big data applications but also social privacy harms. The cumulative effects of unfair and discriminatory big data applications are harmful to societies overall. Social privacy harms demand a top-down approach to law-making. The first step in addressing the social privacy harms of big data is to adopt a social conception of privacy that recognises the data externalities of information disclosures and seeks to address the information- and power- asymmetries in the personal information ecosystem.

In Chapter Five I argued that New Zealand's Privacy Act 2020 is ill-equipped to address the harms of big data. In this chapter I assess New Zealand's responses to the potential harms and benefits of big data beyond the Privacy Act. These responses to date embody a soft law approach, consisting largely of self-regulation through voluntary ethical commitments. In addition, I also consider technological responses, human rights-based responses, and European Union (EU) regulation specifically targeted at artificial intelligence (AI) and big tech.

Admittedly, New Zealand is a small country with limited influence on global AI governance or the regulation of data privacy in other jurisdictions. We import other countries' information practices and the problems inherent with those when we use online services and conduct business with overseas companies. However, there are practical safeguards and measures that we can, and should, adopt, particularly within the public sector. To respond effectively to the harms of big data we must conceptualise privacy as a social concept - therefore this chapter evaluates New Zealand's response through this lens.

It may be that the most effective responses to the harms of big data do not take the form of data privacy regulation, but are instead directed at the business model of surveillance capitalists. For example, introducing measures that remove the incentive for (or

profitability of) the ever-increasing generation and retention of personal information. Regardless, there will continue to be a need for data privacy regulation that effectively addresses the collection and use of personal information. A social conception of privacy must inform amendments to the Privacy Act to make it fit for purpose in an age of big data. Also necessary are government policies that limit the disclosure of New Zealanders' information overseas and promote investment in New Zealand-made, transparent, and accountable data solutions.

New Zealand: voluntary ethical commitments, codes, charters, and self-regulation

Outside of the regulatory framework of the Privacy Act, New Zealand's responses to the potential harms and benefits of big data have been voluntary and non-binding, focussed on best practice, and arguably, have little impact on non-government agencies due to their voluntary nature. The processing of personal information through algorithms and AI, such as automated decision-making systems or aids, and predictive analytics, that results in unfair and discriminatory outcomes is a cause of the social privacy harm of increasing inequality described in Chapter Three. The oversight and accountability of algorithms and AI in New Zealand is relevant to considerations of how to harness the benefits of big data while minimising its potential for harm. This section discusses relevant reports undertaken in this context and the main initiatives currently in use.

Algorithm Assessment Report: Stats NZ

In 2018, Stats NZ published the *Algorithm Assessment Report*[1], which provided an assessment of the use of algorithms by government agencies in New Zealand. Self-reported information from 14 government agencies about the algorithms they used to deliver their functions formed the basis of the Report. The Report found that:[2]

> ... while agencies are applying a range of safeguards and assurances and processes in relation to the use of their algorithms, there are also opportunities for increased collaboration and sharing of good practice across government.

[1] Stats NZ *Algorithm Assessment Report* (2018).
[2] At 4.

The Report also stressed the importance of retaining human oversight and consideration of the views of key stakeholders – particularly the people who are affected by the use of algorithms in the public sector. Also noted as important are incorporating te ao Māori perspectives into the development and use of algorithms as well as looking outside of government for privacy, ethics, and data expertise.[3] The Report made the following five recommendations:

(1) human oversight - policies to consider the balance between human and automated decision-making as well as demonstrating accountability at an organisational level;[4]

(2) development and procurement - algorithms should be developed subject to a process that ensures stakeholder views are considered, including te ao Māori perspectives as well as privacy and ethics. Benefits to be delivered should be considered alongside the limitations or trade-offs and made clear to decision-makers;[5]

(3) information and transparency - there is a need for consistency across government of clear, plain English descriptions of algorithm use. Consideration should be given to publishing more technical detail about how data is collected and stored, algorithmic code, and the role that the algorithm plays in the decision-making process;[6]

(4) review and safeguard - there is an opportunity and need to focus on the ongoing monitoring and review of algorithm use to ensure they are performing as intended and have not had unexpected adverse effects;[7]

(5) sharing best practice - sharing best practice across government agencies would assist in raising the transparency and accountability of algorithm use by government.[8]

[3] At 4.
[4] At 32.
[5] At 32 - 34.
[6] At 34.
[7] At 34 - 35.
[8] At 35

Algorithm Charter for Aotearoa New Zealand: data.govt.nz

The Algorithm Charter was established by the government in response to the recommendations of the *Algorithm Assessment Report.* It formed part of the broader work on algorithm transparency that also included a workforce capacity initiative as well as consideration of the possible regulation of algorithms.[9] The Algorithm Charter aims to address the lack of consistency across government agencies in the application of safeguards and assurance processes.[10] Some government agencies voiced concerns that a voluntary charter did not go far enough or was too little in way of a response, while other agencies believed that it went too far and risked "stifling innovation".[11] Ultimately, the Government Chief Data Steward and data.govt.nz chose a voluntary charter, stating that:[12]

> We believe a voluntary charter strikes the right balance between providing clear, consistent expectations for those agencies who seek them, without being overly prescriptive to agencies who already have well developed approaches to managing algorithm transparency and accountability.

The Algorithm Charter was released in July 2020. Agencies that are signatories to the Charter "commit to making an assessment of the impact of decisions informed by ... algorithms".[13] They also commit to five principles, which can be summarised as follows:

(1) partnership - deliver clear public benefit through Treaty commitments by embedding a te ao Māori perspective in the development and use of algorithms consistent with the principles of the Treaty of Waitangi;

(2) people - focus on people by identifying and actively engaging with people, communities and groups who have an interest in algorithms, and consulting with those impacted by their use;

---

[9] data.govt.nz "Developing the Charter - Responding to the feedback on the Algorithm Charter https://data.govt.nz/docs/charter-feedback-response/
[10] data.govt.nz "Developing the Charter - An Algorithm Charter for Aotearoa New Zealand" https://data.govt.nz/docs/algorithm-charter-draft-sub/
[11] data.govt.nz, above n 10.
[12] data.govt.nz, above n 10.
[13] Stats NZ, *Algorithm Charter for Aotearoa New Zealand* (July 2020) https://data.govt.nz/assets/data-ethics/algorithm/Algorithm-Charter-2020_Final-English-1.pdf

(3) data - make sure data is fit for purpose by understanding its limitations, and identifying and managing bias;

(4) privacy, ethics and human rights - ensure that privacy, ethics and human rights are safeguarded by regularly peer reviewing algorithms to assess for unintended consequences and act on this information;

(5) human oversight - retain human oversight by nominating a point of contact for public inquiries about algorithms; providing a channel for challenging or appealing of decisions informed by algorithms; clearly explaining the role of humans in decisions informed by algorithms.[14]

Three and a half years after its inception there are 29 signatories to the Charter, which is a substantial number but does not include all government agencies. The Algorithm Charter was independently reviewed 12 months after its establishment. The results of the review found that most agencies and subject matter experts see real value in the Charter,[15] however there was confusion as to what should be considered an algorithm and captured by the Charter;[16] some agencies are struggling to find the expertise to measure bias and ensure appropriate human oversight;[17] the requirement for partnership considerations have proved challenging - the pool of experts in Māori data is small and many agencies are unsure of how to implement this commitment;[18] public reporting of algorithm use is fragmented and incomplete, and public awareness of algorithm use by government agencies is limited.[19] Notably, the review states that: "the light regulatory nature of the charter places limitations on its ability to offer public assurance and facilitate public trust".[20]

A key purpose of the Charter - public assurance and accountability - is undermined by its voluntary nature and the fact that not all government agencies have signed up to it.

[14] Stats NZ, above n 13.
[15] Taylor Fry *Algorithm Charter for New Zealand: Year 1 Review* (20 December 2021) at 24.
[16] At 24.
[17] At 26.
[18] At 28.
[19] At 29.
[20] At 30.

Significantly, it offers no mechanism for individuals or public interest groups to challenge algorithm use by government agencies. The decision to make the Algorithm Charter a voluntary charter was premised on the desire to avoid being "overly prescriptive" to agencies that already have well-developed approaches to managing algorithm use.[21] However, the review of the Algorithm Charter reveals that many government agencies were struggling not just to understand the Charter's requirements, but to find the compliance resourcing to implement the principles of the Charter.[22] The review notes that at the one-year point, agencies had made some progress in implementing Charter commitments but states that "capability and capacity limitations will limit the rate of implementation progress".[23]

Principles for the Safe and Effective Use of Data and Analytics: Privacy Commissioner & Chief Data Steward

In 2018, the Privacy Commissioner and Government Chief Data Steward jointly developed six key principles to support safe and effective data analytics. The principles incorporate many of the recommendations of the Algorithm Assessment Report and reflect the principles of the Algorithm Charter. They are as follows:[24]

(1) deliver clear public benefit;
(2) ensure data is fit for purpose;
(3) focus on people;
(4) maintain transparency;
(5) understand the limitations;
(6) retain human oversight

The Principles are designed to form the foundation of guidance to support government agencies' best practice in the use of data analytics in decision-making. However, where principles form best practice guidelines as opposed to enforceable rules, they are

[21] data.govt.nz, above n 10.
[22] Taylor Fry, above n 15, at 14.
[23] At 4.
[24] Office of the Privacy Commissioner [OPC] *Principles for the Safe and Effective Use of Data and Analytics* (218) https://www.privacy.org.nz/assets/New-order/Resources-/Publications/Guidance-resources/Principles-for-the-safe-and-effective-use-of-data-and-analytics-guidance3.pdf

vulnerable to be overlooked in favour of other competing priorities of agencies, such as promoting the use of innovative technology that saves both time and resources.

Trustworthy AI in Aotearoa – AI Principles: AI Forum New Zealand

Founded in 2017, the AI Forum New Zealand is a not-for-profit organisation funded by its members: specifically technology companies, government departments, universities, law and accounting firms. Its website states that its purpose is "to find ways to use AI to help enable a prosperous, inclusive and thriving future for our nation".[25] In 2020, the Law, Society and Ethics working group of the AI Forum published a set of guiding principles for trustworthy AI in Aotearoa New Zealand. A point of difference for these principles is that the principles are:[26]

> ... designed to provide high-level guidance for **anyone** involved in designing, developing and using artificial intelligence in New Zealand (AI stakeholders) with the goal of ensuring New Zealanders have access to trustworthy AI. (emphasis added)

The principles are designed for the public and private sector, and aim to start a conversation about the importance of ethics and law in the design, development and operation of AI. Importantly, the AI Forum states that:[27]

> Government regulation and regulators have an important role to play here. Self-regulation in the form of ethical principles or standards may fill a gap where the law is incomplete or out of date, but they are no substitution for democratically-mandated rules backed up by the force of law.

The AI Forum's AI Principles, are:[28]

> (1) Fairness and Justice ...
> (2) Reliability, security and privacy ...

[25] AI Forum New Zealand "Trustworthy AI in Aotearoa: The AI Principles" https://aiforum.org.nz/reports/trustworthy-ai-in-aotearoa-the-ai-principles/
[26] AI Forum New Zealand, above n 25.
[27] AI Forum New Zealand "Trustworthy AI in Aotearoa New Zealand" https://aiforum.org.nz/wp-content/uploads/2020/03/Trustworthy-AI-in-Aotearoa-March-2020.pdf
[28] AI Forum New Zealand, above n 27.

(3) Transparency ...

(4) Human oversight and accountability ...

(5) Wellbeing - Where appropriate, AI stakeholders should design, develop and use AI systems to promote, as much as possible, the wellbeing of New Zealand's people and environment in areas such as health, education, employment, sustainability, diversity, inclusion and recognition of the unique values of Te Ao Māori.

The principles broadly reflect those of the Algorithm Charter and the Principles for Safe and Effective Use of Data and Analytics however, an interesting addition is the fifth principle of wellbeing - encouraging AI stakeholders to promote the wellbeing of New Zealanders, the environment and the values of te ao Māori. Also, worth noting under the principle of "Human oversight and accountability" is the following sentence: "technologies capable of harming individuals or groups should not be deployed until stakeholders have determined appropriate accountability and liability".[29] This goes further than the recommendations of the Algorithm Assessment Report, the principles of the Algorithm Charter, and the Principles for the Safe and Effective Use of Data and Analytics.

Data Ethics Advisory Group: Stats NZ

The Data Ethics Advisory Group was also established by StatsNZ and the Government Chief Data Steward (GCDS) in response to the recommendations of the Algorithm Assessment Report. The Group is convened by the GCDS and aims to assist government to maximise the benefits and opportunities of emerging uses of data, while at the same time managing risks. The Group provides independent advice to the GCDS and also to state sector agencies, at their request, on new uses of data.[30]

Data Protection and Use Policy: data.govt.nz

The Data Protection and Use Policy (DPUP) Guidelines and Principles are designed to assist frontline workers who collect and use personal information to do so in a way that

[29] AI Forum New Zealand, above n 27.
[30] New Zealand Government, *Data Ethics Advisory Group Terms of Reference* v2 (June 2023) at 4. https://data.govt.nz/assets/Uploads/Data-Ethics-Advisory-Group/Terms-of-Reference-DEAG.pdf

complies with the information privacy principles (IPPs) of the Privacy Act 2020. Of note, DPUP goes beyond compliance with the law to encourage best practice in data use, focussing on the wellbeing of people and communities, to promote trust and transparency in government use of data.[31]

Guidelines issued by the Office of the Privacy Commissioner on GenAI

On 15 June 2023, the Office of the Privacy Commissioner published its expectations around the use of generative artificial intelligence (GenAI) and stated that it is the responsibility of agencies to ensure that their use of GenAI complies with the Privacy Act. In Chapter 5 I highlighted some of the shortcomings of the Privacy Act in mitigating or addressing the potential harms of big data. In light of this, it is concerning that there is currently no plan to regulate the use of artificial intelligence in New Zealand, unlike in many other jurisdictions where AI is a focus of regulatory activity. I will compare the European Union's and the United Kingdom's approaches to AI regulation later in this Chapter.

Interim Centre for Data Ethics and Innovation: StatsNZ

An Interim Centre for Data Ethics and Innovation is to be established as a unit within the function of the GCDS in recognition of the advances in data and data driven technologies. The purpose of the Centre is to "provide strong leadership, consolidate expertise, and help coordinate across agencies and the wider system to tackle these issues efficiently and proactively".[32] There is an inherent tension between data-driven innovation and the principles of data privacy law. Therefore, the need for guidance on best practice as well as consistency of standards across government and other agencies is compelling. Guidelines that promote enhanced standards of information privacy practices will be beneficial if they address the issue of repurposing of data and the risks associated with models. The Interim Centre for Data Ethics and Innovation promises to be a useful resource for government agencies that are committed to ethical uses of data. However, the Interim Centre is to be advisory only and will not set or enforce new regulations

[31] digital.govt.nz "Data Protection and Use Policy" https://www.digital.govt.nz/standards-and-guidance/privacy-security-and-risk/privacy/data-protection-and-use-policy-dpup/
[32] New Zealand Parliament, Economic Development, Science and Innovation Committee *Cabinet Economic Development Committee Minute of Decision* (26 July 2023) DEV-23-MIN-0157 at 1.

itself.[33] As we have seen with the Algorithm Charter, principles and guidelines that are not enforced can fall victim to a lack of appropriate resourcing and failure to be prioritised over other objectives that are measurable and/or enforceable.

Proposed Biometrics Code of Practice: Office of the Privacy Commissioner

In October 2021, the Office of the Privacy Commissioner (OPC) published guidance on the regulation of biometrics,[34] outlining how the Privacy Act regulates biometrics use. On 23 November 2023, the Privacy Commissioner announced that he would be progressing draft privacy rules for biometric information, in recognition of the increasing role of biometrics in people's lives and calls for greater regulation of biometric information.[35] The Privacy Commissioner has the power to issue a Code of Practice under s 33 of the Privacy Act. A Code of Practice modifies the application of one or more of the IPPs of s 22 of the Privacy Act with principles that are directly applicable to a specified class of information, such as biometric information.[36] In early 2024, the OPC publicly consulted on an exposure draft of a biometrics privacy code.[37] The OPC's intended approach to regulating biometrics, as set in this initial discussion document, was promising. It proposed a ban on live facial recognition technology (FRT) and restricting the use of biometrics for direct marketing or to infer someone's health information or mood. The OPC's summary of submissions on the potential biometrics code of practice states that:[38]

> Private sector agencies, and some public sector agencies, opposed the suggested purpose limitations, noting that they could stifle innovation and prevent some beneficial use cases.

---

[33] At 2.
[34] OPC *Office of the Privacy Commissioner Position Paper on the Regulation of Biometrics* https://www.privacy.org.nz/assets/New-order/Resources-/Publications/Guidance-resources/2021-10-07-OPC-position-on-biometrics.pdf
[35] OPC Media Release "Privacy Commissioner to consult on new rules for biometrics" (23 November 2023) https://privacy.org.nz/publications/statements-media-releases/privacy-commissioner-to-consult-on-new-rules-for-biometrics/
[36] Privacy Act 2020, s 32.
[37] OPC "Biometrics" https://www.privacy.org.nz/resources-2/biometrics/
[38] OPC *A potential biometrics code of practice: discussion document - summary of submissions* https://www.privacy.org.nz/assets/New-order/Resources-/Publications/Guidance-resources/Biometrics/Biometrics-November-2023/Summary-of-submissions-on-OPC-discussion-document.pdf

Unfortunately, the initial approach of the OPC was watered down after consultation with the public and stakeholders, including government agencies, many of which deploy biometrics and are keen to expand their use. Consequently, the draft biometrics code removes the prohibition on the use of biometrics for marketing and there is no outright prohibition on the use of live facial recognition technology (FRT) by law enforcement.[39] The consultation period for the draft code closed on 8 May 2024.[40] It remains to be seen what form the final Code will take. However, its current iteration raises questions of legitimacy in relying on agencies that wish to use biometrics to undertake assessments of proportionality in determining risk acceptability,[41] particularly when the only purpose limitations are on uses of biometrics to infer health information, inner or physical state, or a prohibited ground of discrimination.[42]

Important role of Chief Executives and boards of government agencies – ethical use of AI

Chief Executives and senior leaders of both public and private agencies contribute to their agencies' privacy culture. Their understanding of privacy, and the directions they give to staff, will influence where on the spectrum an agency sits, from merely compliance-based to embracing people-centred privacy. When an agency's focus is on compliance, privacy impact assessments are often completed late in the process of developing or adopting new technologies (if completed at all). They are prone to becoming a box-ticking exercise, eliminating any opportunity for implementing privacy-by-design measures and reducing an agency's ability to mitigate privacy risks.

As I have set out in the start of this chapter, there are many resources from which leaders can find high-level guidance on the principles of the ethical use of data and algorithms. The first step would be to become a signatory to the Algorithm Charter. The Review of the Algorithm Charter states that "there is at least one significant example of an agency not signing up to the Charter due to fears about compliance resource requirements".[43] The Review found that, one year following its establishment, a lack of compliance and

[39] OPC "Exposure draft of a biometric processing code of practice: consultation paper" (April 2024) at 10. https://www.privacy.org.nz/news/consultations/biometrics/
[40] OPC "Have your say on biometrics" https://www.privacy.org.nz/news/consultations/biometrics/
[41] OPC "Biometric Privacy Code Exposure Draft Only For Comment" (10 April 2024), Rule 1(d).
[42] Rule 4(2).
[43] Taylor Fry, above n 15, at 14.

capacity resourcing has hindered full compliance with the Algorithm Charter by signatories.[44] One advanced public sector user of algorithms, New Zealand Customs Service, is yet to sign up to the Algorithm Charter.[45] Consideration of the establishment of an oversight body to provide support and oversee aspects of the Charter is a recommendation of the review. Signatory agencies indicated that they would like an oversight body to go to for advice on implementing Charter commitments and for oversight of its principles.[46]

Some government agencies appear to be making greater efforts at implementing ethical standards and principles into their use of data and algorithms than others. The Ministry of Social Development (MSD) has developed a Privacy Human Rights and Ethics (PHRaE) framework for the purpose of identifying where MSD's use of personal information could impinge on people's privacy, human rights, or ethics.[47] It is designed to kick in once a proposal moves beyond a mere idea, and thus avoids the pitfalls of privacy impact assessments being completed once the terms of a contract have already been agreed, or even worse, completed retrospectively when a system or software is already in use making it difficult, if not impossible, to accommodate privacy considerations.

Despite MSD's PHRaE framework, from 20 November 2023 new MSD clients have been able to use an online identity check service when applying for ongoing financial support without it being tested for possible racial bias.[48] Identity Check uses FRT from Irish company Daon. It matches a live image taken by the data subject with their phone to the data subject's drivers' licence or passport photo in government databases held by the Department of Internal Affairs (DIA).[49] This Identity Check service has not been tested

[44] At 14.
[45] data.govt.nz "Algorithm Charter for Aotearoa New Zealand https://data.govt.nz/toolkit/data-ethics/government-algorithm-transparency-and-accountability/algorithm-charter/#signatories
[46] Taylor Fry, above n 15, at 5.
[47] Ministry of Social Development "Privacy Human Rights and Ethics Framework" https://www.data.govt.nz/assets/data-ethics/algorithm/phrae-on-a-page.pdf
[48] Phil Pennington, "Facial recognition: Government rolls out new tech despite racial bias concerns" (15 November 2023) RNZ https://www.rnz.co.nz/news/national/502445/facial-recognition-government-rolls-out-new-tech-despite-racial-bias-concerns#:~:text=The%20government%20begins%20the%20roll,of%20Internal%20Affairs%20(DIA).
[49] Ministry of Social Development "Identity checks and online verification" https://www.msd.govt.nz/about-msd-and-our-work/work-programmes/identity-check/identity-checks-and-online-verification.html

on the New Zealand population, meaning it may not be as accurate at identifying Māori and Polynesian faces, which are not as prevalent in training datasets based on North American or European populations.[50] RNZ reporter Phil Pennington refers to an MSD report that describes the level of racial bias in the technology as '"unknown", "unconfirmed", and "untested"' despite the technology being in development for the last four years by the DIA.[51] Identity Check can also be used to detect and investigate fraud. However, MSD has confirmed that this function will not be utilised yet because of the human rights and ethics risk of racial bias. The technology is 90 per cent accurate. However, if the 10 per cent of inaccuracies are predominantly borne by Māori and Pasifika peoples (the Daon facial recognition algorithm works less well on darker skin tones), then that is problematic. Pennington reports that the government is looking to implement Identity Check as the "go-to tech for proving everyone's identity online ...".[52] Hopefully, potential issues in relation to racial bias will be addressed before this becomes a reality. However, there is currently no legal requirement to ensure testing for bias before FRT is deployed.

Police's Expert Panel on Emerging Technology

The Expert Panel on Emerging Technology (EPET) was established in 2021 to provide external and independent scrutiny of Police's proposed use of emergent technology and to provide non-binding advice.[53] EPET was established after the media revealed that Police had trialled Clearview AI's FRT without the public's knowledge. Police had conducted hundreds of searches using Clearview AI without consulting with the Privacy Commissioner or having the use of the technology approved by the Police Commissioner.[54]

The EPET is appointed by the Commissioner of Police and its membership has a diverse range of skills and expertise. The Terms of Reference for EPET describe its key areas of

[50] Pennington, above n 48.
[51] Pennington, above n 48.
[52] Pennington, above n 48.
[53] New Zealand Police [NZ Police] "New Zealand Police Expert Panel on Emergent Technologies Terms of Reference" (May 2021) at 2.
[54] Mackenzie Smith "Police 'stocktake' surveillance tech after Clearview AI facial recognition trial" (18 May 2020) Radio New Zealand [RNZ] https://www.rnz.co.nz/news/national/416913/police-stocktake-surveillance-tech-after-clearview-ai-facial-recognition-trial

focus as reviewing and advising on proposals for Police to test or trial an emergent technology (or significant new functionality within an existing technology) which are referred to it.[55] Also, to review and advise on the use of algorithms by Police to ensure privacy, human rights and ethics are appropriately safeguarded, in line with the Algorithm Charter.[56] EPET can only review technology or algorithms that are referred to it; not all emergent technology is referred to EPET. Its recommendations are advisory only and are published on the New Zealand Police website. The recommendations and advice of EPET are adopted at Police's discretion.[57]

There have been instances where the advice or recommendations of EPET has resulted in a proposed use of technology being abandoned on ethical or privacy grounds. A road policing algorithm utilising predictive analytics to determine a motorist's likelihood of getting into serious trouble in the next 3 years was abandoned after EPET's advice that the use of this tool could draw on historic biases to influence future policing decisions.[58] A proposed trial by Police of Zavy Social Media Sentiment and Analytics tool was abandoned in response to EPET's advice that the tool did not meet the necessity principle of the New Technology Policy and Framework and it had concerns about its accuracy and effectiveness.[59]

Review of EPET and Terms of Reference

EPET's Terms of Reference states that an independent review of the Panel's operations and its Terms of Reference would be conducted after 12-18 months.[60] The review was completed by a specialist consulting firm in May 2023. It found that EPET is functioning as intended and provides robust advice.[61] It recommended that EPET's advice be made

[55] NZ Police, above n 53, at 2.
[56] At 3.
[57] At 2-3.
[58] Phil Pennington "Police drop technology designed to predict motorists? complete" RNZ (15 July 2021) https://www.rnz.co.nz/news/national/446957/police-drop-technology-designed-to-predict-motorists
[59] NZ Police, *We Asked: Expert Panel on Emergent Technology: Proposal to Trial Zavy https://www.police.govt.nz/sites/default/files/publications/zavy-proposal-we-asked-they-said-we-did.pdf;* 'Necessity' is the first of 10 principles that guide decisions on the trial or adoption of new Police technologies in its New Technology Framework https://www.police.govt.nz/sites/default/files/publications/new-technology-framework.pdf
[60] NZ Police "Mid-term health check of the expert panel" https://www.police.govt.nz/about-us/programmes-and-initiatives/police-use-emergent-technologies/expert-panel-emergent
[61] NZ Police, above n 60.

more accessible to the public on Police's website, which has since been implemented. However, it is noted that EPET's more recent advice needs to be uploaded to the website. The review identifies opportunities to improve the effectiveness of EPET by widening its role to create a strategic agenda.[62] While it is unclear precisely what a strategic agenda might involve, I would argue that providing EPET with a complete stocktake of all algorithms and surveillance technologies in use, or in contemplation of use, by Police, and allowing the EPET to determine which tools or technologies to review would be an excellent start.

Despite the robust, expert advice available to it, Police continue to use technology without appropriate privacy and ethical considerations in place. In September 2023, the Dominion Post reported on the rapid expansion of the use of drones by Police and the failure for this use (including different capabilities of new drones) to be considered by EPET.[63] On 4 October 2023, RNZ reported that Police had not assessed the privacy impacts of new AI visual search tool, SearchX. Police responded by stating that this was because it was not new technology and simply improved and extended existing products, however it was reported that the technology was new to New Zealand Police.[64] There continues to be a lack of transparency surrounding the scope and extent of use of surveillance technologies by Police. However, while Police attracts the most media scrutiny, other government agencies are also deploying predictive analytics and surveillance technologies without appropriate oversight or public awareness. There is a pressing need for transparency around government agencies' use of big data in New Zealand. It is concerning that the New Zealand public is relying on journalists' utilising their powers under Official Information Act 1982 to glean this information to protect against infringements of civil liberties and the potential misuse of Police powers.

[62] NZ Police, above n 60.
[63] Mike White, "Huge increase in police drone use: Sensible solution or spies in the sky?" (6 September 2023) Dominion Post https://www.thepost.co.nz/nz-news/350066875/huge-increase-police-drone-use-sensible-solution-or-spies-sky
[64] Phil Pennington, "New police visual search tech not assessed for privacy impacts" (4 October 2023) RNZ https://www.rnz.co.nz/news/national/499415/new-police-visual-search-tech-not-assessed-for-privacy-impacts

Increasing adoption of surveillance technologies

The focus of the Privacy Act is on harm to the individual. Privacy impact assessments consider the privacy implications of a new use or collection of personal information in the context of that agency's particular use and are therefore, ineffectual at addressing cumulative harms of increasing surveillance by agencies overall. Voluntary codes, guidelines and advisory bodies can be ignored in the pursuit of catching criminals and preventing crime. But at what cost? And who gets to decide?

In New Zealand, government agencies' use of surveillance technologies is increasing without the necessary public debate about the associated harms and benefits and the trade-off between competing interests. Immigration has been criticised for purchasing Cobwebs spyware to detect and prevent mass arrivals by asylum seekers or "boat people" despite New Zealand never facing a real risk of a mass arrival by boat. Police has come under media scrutiny for its use of Automatic Number Plate Recognition (ANPR) technology, with two cases currently before the courts challenging the use of ANPR.[65]

In the private sector, Foodstuffs, one of two companies with 90 per cent of the grocery market share in New Zealand, began a trial of FRT in up to 25 of its North Island stores in February 2024. The FRT is being implemented in response to a rise in retail crime in supermarkets, to enable staff to better identify repeat offenders who account for one third of all incidents.[66] The Privacy Commissioner has stated that he will use his inquiry powers to keep a close eye on the trial and has indicated that he is concerned that there is no evidence that FRT reduces shoplifting and, therefore, questions whether FRT is an effective, necessary and proportionate response.[67] The Privacy Commissioner has also raised concerns about bias and the accuracy of FRT, particularly for women and people of darker skin tones.[68] Foodstuffs has stated that images that do not match an image of a

[65] Phil Pennington "Legal challenges to Police use of automated number plate recognition cameras" (26 October 2023) RNZ https://www.rnz.co.nz/news/national/501012/legal-challenges-to-police-use-of-automated-number-plate-recognition-cameras
[66] Foodstuffs "Foodstuffs North Island begins trialling facial recognition in select stores as part of its commitment to keep teams and customers safe by keeping previous offenders out" (8 February 2024). https://www.foodstuffs.co.nz/news-room/2024/Foodstuffs-North-Island-begins-trialling-facial-recognition-in-select-stores
[67] OPC "Privacy Commissioner to keep a close eye on Foodstuffs' North Island FRT trial" (8 February 2024) https://www.privacy.org.nz/publications/statements-media-releases/privacy-commissioner-to-keep-a-close-eye-on-foodstuffs-north-island-frt-trial/
[68] OPC, above n 67.

known shoplifter will be immediately deleted,[69] but who will oversee that? Will there be sufficient controls over how our images are used? The trial raises a bigger question for New Zealanders - do we want this type of surveillance to become commonplace? Where will it emerge next? Our regulatory framework is currently ill-equipped to address the increasing use of surveillance technologies by public and private sector agencies.

Automatic Number Plate Recognition technology

The Police use of ANPR platforms is not widely known or understood by the public and therefore has not been subject to robust public debate. ANPR involves the use of CCTV cameras to capture still images or videos that are processed by software to automatically recognise number plate information. The information captured by ANPR systems also includes images and video of a vehicle's occupants.

There are two main ANPR providers in New Zealand: Auror[70] and SaferCities.[71] SaferCities captures CCTV footage from local council-owned sites, streets, and motorways. Auror is a retail crime prevention platform. Retailers upload their CCTV footage from cameras positioned in and around their carparks or forecourts, and from within their stores, to the Auror platform. Retailers connected to the platform can upload CCTV videos and images of incidents of alleged crimes as well as personal information about the suspects involved.[72] Radio NZ has reported that Auror has stated that it covers 90 per cent of retail outlets across New Zealand.[73]

Police undertake approximately 10,000-15,000 searches a month on vehicles using ANPR platforms.[74] With each search, Police gain access to 60 days' worth of location data for that vehicle. This information is displayed on a map and provides the date, time, and location.[75] It shows where vehicles of interest have been at specific times and can help to

[69] Foodstuffs, above n 66.
[70] Auror https://www.auror.co
[71] SaferCities https://www.safercities.com
[72] Auror, above n 70.
[73] Phil Pennington, "Police step up surveillance activity, tap into CCTV footage from other businesses" (23 September 2022) Radio NZ https://www.rnz.co.nz/news/national/475342/police-step-up-surveillance-activity-tap-into-cctv-footage-from-other-businesses
[74] NZ Police Assurance Group "Audit of New Zealand Police's use of Automatic Number Plate Recognition Technology (ANPR) Platforms" (December 2022) at 1. https://www.police.govt.nz/sites/default/files/publications/police-use-anpr-platforms-audit-report.pdf
[75] Auror, above n 70.

create a picture of a person's pattern of movement or routine behaviour. Defence lawyers for the Public Defence Service argue that this can constitute the use of a tracking device without a warrant.[76] The Search and Surveillance Act 2012 provides mechanisms for Police to obtain warrants and production orders to access this information.[77] Seeking a warrant or production order through the court system provides the necessary oversight that is lacking from current practice. The Public Defence Service states that the extent of the Police's use of ANPR platforms presents "a real risk of going into the territory of mass surveillance".[78]

Auror's live tracking functionality enables law enforcement to generate an alert within the platform and obtain the real-time location details of a vehicle of interest when the vehicle is detected by a camera on the platform.[79] Auror's search and live tracking functionalities are powerful surveillance tools that save Police time and resources but are also inherently vulnerable to misuse.

Police admitted to misusing ANPR technology in 2021. Police officers falsely reported that a car was stolen in order to track its occupants who had triggered a covid lockdown in Northland.[80] This publicised misuse led to a planned review of ANPR by Police being brought forward. Police's internal review found that the "vast majority of Police" are using ANPR appropriately, however five staff were investigated by its integrity group for misuse and 120 staff were spoken to about searching for their own vehicles.[81] During the period from 2018 to 2022, Police searched ANPR platforms over 350,000 times.[82] Over six thousand police staff can access historical vehicle data and just under one thousand

[76] Pennington, above n 65.
[77] Search and Surveillance Act 2012, ss 6 & 71.
[78] Pennington, above n 65.
[79] Auror, above n 70.
[80] Phil Pennington "Police made false report to use ANPR cameras to track women who triggered Northland lockdown" (28 September 2022) RadioNZ https://www.rnz.co.nz/news/national/475662/police-made-false-report-to-use-anpr-cameras-to-track-women-who-triggered-northland-lockdown#:~:text=28%20Sep%202022-,Police%20made%20false%20report%20to%20use%20ANPR%20cameras,women%20who%20triggered%20Northland%20lockdown&text=Just%20one%20month%20after%20the,so%20they%20could%20track%20it.
[81] NZ Police "Police release ANPR audit findings" (26 April 2023) https://www.police.govt.nz/news/release/police-release-automatic-number-plate-recognition-audit-findings#:~:text="It%20involved%20examining%20and%20cross,to%20use%20ANPR%20data%20responsibly.
[82] NZ Police, above n 81.

staff can track vehicles in real time.[83] The Police review recommended tightening up controls around the use of ANPR platforms.[84] Arguably, tighter controls should have been in place prior to Police's deployment of Auror and SaferCities platforms.

Mass surveillance is a social privacy harm

In Chapter Three I used Joel Feinberg's example of the accumulative public harms of air and water pollution[85] to draw an analogy to the accumulative public harm of information practices that contribute to inequality. I categorised the social privacy harms of big data into two broad categories - the harm of increasing inequality and the harm to democracy. Mass surveillance, enabled by networks of CCTV cameras, also constitutes a social privacy harm. It sets back the public interest in democracy by undermining several key civil liberties. Drawing on Feinberg's polluting car metaphor, I argue that one CCTV camera in the forecourt of a petrol station that records me filling my car with petrol does not set back my interest in moving freely without being surveiled. However, at some point between one petrol station capturing my movements on CCTV and every petrol station, retail store, supermarket, airport, train station, and public amenity I visit, and every road I drive along, capturing my movements and the people I'm with, my interest in freedom of movement and freedom of association is thwarted.

New Zealanders have a public interest in maintaining a degree of anonymity in public spaces, and the ability to come and go as part of their day-to-day lives without their movements being recorded and without the possibility that a Police officer could create an alert to notify them whenever their licence plate (or possibly, in the future, their face) is captured on CCTV camera. Mass surveillance thwarts the public interest in people having the freedom to protest, freedom of movement, and freedom of association - all of which are necessary for a healthy, functioning democracy. Also necessary for a democracy is the maintenance of law and order. Surveillance technologies enabled by big data, like ANPR and FRT make the Police's task of tracking down suspects faster and

[83] NZ Police Assurance Group, above n 74, at 1; RNZ "Police investigating staff misuse of number plate software" (27 April 2023) https://www.rnz.co.nz/news/national/488750/police-investigating-staff-misuse-of-number-plate-software
[84] NZ Police Assurance Group, above n 74, at 2.
[85] Joel Feinberg *Harm to Others: The Moral Limits of the Criminal Law* (Oxford University Press, Oxford, 1984) at 227.

more efficient. However, Police's acknowledgement of their misuse of ANPR and its secret trial of Clearview AI's FRT reveal how surveillance technologies can be used without appropriate safeguards in place or full consideration of the associated privacy and ethical implications. The number of CCTV cameras that support FRT and ANPR is proliferating across New Zealand without the necessary public debate on their merits, efficacy, and threat to fundamental civil liberties. New Zealand has not reached the point of mass surveillance yet. However, with Foodstuff's FRT aspirations and increasing use of surveillance by law enforcement, I believe we are headed in that direction.

Are legislative responses targeted at big tech and AI the answer?

I have described how New Zealand's soft law approach to big data harms is a start, but on balance, is ineffectual. The Algorithm Charter is a voluntary charter, and many of the government agencies that are signatories to it admit to being unable to implement its principles effectively. This highlights a need for a different, or additional, regulatory response to address the information practices and associated harms of big data and artificial intelligence deployed in products and services. Arguably the most effective regulatory responses will be those that are consistent with a social conception of privacy: responses that recognise and address the collective harms of information practices. A social conception of privacy demands a hard law approach to big data harms.

New Zealand cannot effectively address the harms that stem from a few very large tech companies holding vast quantities of information about hundreds of millions of individuals. Very large online platforms and search engines control who has access to information and how that information is presented to users of their services. In Chapter Three I explained how democracies are harmed by vast information- and power-asymmetries between citizens and very large technology companies. There is little New Zealand can do to mitigate this harm (although tighter regulations around lobbying[86] could address some concerns in relation to the influence of big tech on New Zealand politics more broadly).

[86] In April 2023 the government announced several measures aimed at promoting greater transparency around lobbying at Parliament. https://www.justice.govt.nz/justice-sector-policy/key-initiatives/political-lobbying/

Nonetheless, there are big data harms that New Zealand can take measures to actively address. As discussed above, the Algorithm Charter, Principles for the Safe and Effective Use of Data and Analytics, the AI Forum's AI Principles, and the DPUP reveal that there is a wealth of information on *how* to attempt minimise the harms of algorithmic decision-making. What is needed is the legislative will to implement and enforce rules for the use of algorithms that come within the ambit of the Algorithm Charter as well as certain uses of AI. The time for relying on voluntary commitments and self-regulation has passed.

European Union regulatory framework on AI

The European Commission (the Commission) initially adopted a soft law approach to AI, publishing non-binding *Ethics Guidelines for Trustworthy AI* in 2019.[87] It has since shifted towards a legislative approach with an EU regulatory framework on AI. In contrast, the United Kingdom has taken a "pro-innovation" approach to AI and has chosen not to legislate specifically for AI, but instead to rely on existing regulators to govern AI applications that fall within their regulatory remit.[88] This raises the question (which I will return to later in this chapter) of whether a decentralised approach can effectively manage the risks of AI technologies.

EU Digital Services Act

Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (DSA) sets out the rules on the provision of intermediary online services in the internal market of the EU.[89] It aims to facilitate a safe and trusted online environment where fundamental rights are protected.[90] It gives users more control over what they see online as well as the ability to choose an option for recommender systems that is not based on profiling.[91] Protecting users from illegal and harmful content,

[87] European Commission High-Level Expert Group on AI, *Ethics Guidelines for Trustworthy AI* (2019) https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai
[88] UK Government "A pro-innovation approach to AI regulation" (2023) https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper
[89] Regulation (EU) 2022/2065 [Digital Services Act] Art 1(1).
[90] Article 1(2)(b).
[91] Article 29.

including health- and political- disinformation, is an important objective of the DSA.[92] Illegal content is to be removed without undue delay.[93] The DSA enables users of intermediary services to become better informed of how their data is being used (monetised) by online platforms. Online platforms that display advertising must ensure that users can identify that the information presented *is* advertising, and provide, for each advertisement, meaningful information about the parameters used to determine the recipient of the advertisement.[94] In this regard it embodies an individualistic approach to mitigating the potential harms of big data. An individualistic approach perpetuates many of the same shortcomings of contemporary data privacy regimes: the onus is placed on the individual to manage their data, and the assumption is that greater information will lead to improved privacy outcomes.

All online intermediaries offering their services in the EU are subject to the DSA. The DSA also sets out additional obligations for very large online platforms (VLOPs) and very large online search engines (VLOSEs) to manage the particular and systemic risks they pose in relation to illegal content and societal harm. Platforms and search engines that have 45 million average monthly users are categorised as VLOPs / VLOSEs.[95]

Google has 92 per cent of the search engine market.[96] Under the DSA, Google is categorised as a VLOSE and therefore has additional online advertising transparency requirements. It must keep a publicly available repository of information about the advertisements it displays, including the content of the advertisement, the person on whose behalf it is displayed, and the parameters used, if intended to be displayed specifically to one or more groups of recipients.[97] The idea is that by increasing the transparency of search engines and targeted advertising, users will have more idea about how much value is extracted from them. Decarolis and Lis state that:[98]

[92] Recital 63.
[93] Article 8.
[94] Article 24.
[95] Article 25(1).
[96] https://financesonline.com/google-search-statistics/; https://www.tooltester.com/en/blog/search-engine-market-share/
[97] Digital Services Act, Art 30.
[98] Francesco Decarolis and Muxin Li "Regulating online search in the EU: From the android case to the digital markets act and digital services act" (2023) International Journal of Industrial Organization 90 at 3.

> In the short run, this will direct users to search engines with better privacy protection. In the long run, its impact will be much more profound. The change in users' tastes may generate stronger motivations for search engines to invest in privacy protection.

There are compelling reasons to question the above assessment of the implications of the DSA. Firstly, people do not always make rational choices based on calculations that align with their own best interests when it comes to choosing brands and service providers. Secondly, Google, like most other big tech companies, relies on its advertising revenue.[99] Its business model is dependent on the generation and collection of personal information from its users to make a profit. To grow its advertising revenue, Google needs to continually innovate and improve its advertising product.[100] An integral part of its business is selling advertising space that is targeted at very specific audiences. To do this, it needs to collect personal information from its users - the more, the better.

EU Digital Markets Act

The purpose of Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act) (DMA) is to establish rules to ensure fairer and more competitive markets in the digital sector.[101] It sets clear rules for large online platforms that occupy a "gatekeeper" position to large numbers of users, making sure that they do not abuse that position to deny businesses that wish to have access to those users. The DMA targets unfair and anticompetitive behaviour. Articles 5, 6 and 7 set out obligations for gatekeepers. Article 5(2) prohibits unfair processing of personal data. It states that:

> 2. The gatekeeper shall not do any of the following:
>
> (a) process, for the purpose of providing online advertising services, personal data of end users using services of third parties that make use of core platform services of the gatekeeper;

[99] Duncan McCann, *Digital Power Players: Power and Accountability Part 4: The Problem and Power of Tech Monopolies* (2018) New Economics Foundation at 8.
[100] At 8-9.
[101] Digital Markets Act, Art 1(1).

(b) combine personal data from the relevant core platform service with personal data from any further core platform services or from any other services provided by the gatekeeper or with personal data from third-party services;

(c) cross-use personal data from the relevant core platform service in other services provided separately by the gatekeeper, including other core platform services, and vice versa; and

(d) sign in end users to other services of the gatekeeper in order to combine personal data,

unless the end user has been presented with the specific choice and has given consent within the meaning of Article 4, point (11), and Article 7 of Regulation (EU) 2016/679.

...

Article 5 aims to limit the over-collection and repurposing of personal data by gatekeepers by prohibiting the processing, combining, and cross-using of personal data as described above, *unless* the end user has been presented with the choice and has given the required consent.

The DMA relies on self-enforcement by gatekeepers.[102] The Commission clarifies obligations under the DMA[103] and investigates possible systematic non-compliance.[104] The consequences for breaching the rules include non-compliance decisions issued by the Commission, and fines of up to 10% of total gatekeeper turnover in the preceding financial year for intentional or negligent non-compliance.[105] Where a gatekeeper is deemed to have engaged in systematic non-compliance with the obligations in Articles 5, 6, and 7, and the Commission has issued at least three non-compliance decisions within an eight-year period, the Commission may impose "any behavioural or structural remedies which are proportionate and necessary to ensure effective compliance with this

[102] Article 8(1).
[103] Article 8(3).
[104] Article 18(1).
[105] Article 30.

Regulation".[106] However, structural remedies (such as breakups) are anticipated to be used as a last resort. It is argued that breaking up a gatekeeper into smaller entities is unlikely to happen, however, the threat of breakups of gatekeepers under Article 18 of the DMA may promote compliance with the Act.[107]

I argue that while the provisions of the DMA are admirable, it does not address the underlying business model of surveillance capitalists that drives ever-increasing generation of personal data. Until this "extraction imperative"[108] is targeted in a meaningful way, legislation focussed on improving the competitiveness and fairness in the online platform market will only go partway towards addressing the social privacy harms resulting from the significant information- and power- asymmetries between citizens and big tech. One approach to address the extraction imperative driving surveillance capitalists is to prohibit personalised advertising.[109] If very large online search engines and very large online platforms were prohibited from selling personalised advertising space to other agencies, the imperative to generate escalating amounts of personal information in order to make increasingly accurate predictions about us, would diminish.

A prohibition would need to apply to all personalised advertising, not just political advertising or messaging, for two reasons. First, determining what is or is not a political advert is very difficult. A politically motivated advert may not name a politician or political party, but target individuals believed to be susceptible to messaging that will instil fear or distrust with the objective of making a particular candidate more appealing. Second, allowing personalised advertising for non-political messaging does not address the ever-increasing generation of personal information to sell personalised advertising spaces for other products. Consequently, only prohibiting personalised *political*

[106] Article 18(1).
[107] Tone Knapstad, "Breakups of Digital Gatekeepers under the Digital Markets Act: Three Strikes and You're out?" (2023) Journal of European Competition Law and Practice at 16.
[108] Shoshana Zuboff describes the first economic imperative of surveillance capitalism as the "extraction imperative" (Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (2019, Profile Books, London) at 87.
[109] Duncan McCann and Miranda Hall, *Blocking the Data Stalkers: Going Beyond the GDPR to Tackle Power in the Digital Economy* (2018, New Economics Foundation) at 17-18; Carissa Véliz *Privacy Is Power: Why And How You Should Take Back Control Of Your Data* (Penguin Random House, London, 2020) at 119-126.

advertising would not address the issue of the size and influence of surveillance capitalists which is, in itself, harmful to democracy.

Other approaches that utilise competition law to address some of the shortcomings of data privacy regimes look promising but fall outside the scope of this thesis.[110] Also outside of scope are the array of collective governance mechanisms premised on individual data rights, such as data trusts, data co-operatives, and data unions to name a few.[111]

EU Artificial Intelligence Act

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)[112] takes a risk-based approach to the regulation of AI. The Artificial Intelligence Act (AIA) applies to both public and private sector[113] providers and deployers of AI systems in the EU as well as providers and deployers of AI systems in a third country, where the output produced by the system is used in the EU.[114]

The Commission's Overview of the Proposal for Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence recognised the economic and societal benefits of AI across a wide range of sectors, particularly in improving prediction, increasing efficiencies, and personalising services. It also acknowledged the concerns about safety when AI is embedded in products and services, stating that "most notably, AI systems may jeopardise fundamental rights such as the right

[110] For further information on the use of competition law to address privacy law shortcomings see Wolfgang Kerber and Karsten K. Zolna "The German Facebook case: the law and economics of the relationship between competition and data protection law" (2022) European Journal of Law and Economics 54.
[111] For a useful overview of collective governance mechanisms see Jamie Duncan "Data protection beyond data rights: governing data production through collective intermediaries" (2023) Internet Policy Review 12 3.
[112] EU Artificial Intelligence Act https://artificialintelligenceact.eu/the-act/
[113] Article 3(3).
[114] Article 2(1).

to non-discrimination, freedom of expression, human dignity, personal data protection and privacy".[115]

The AIA is a horizontal instrument, which means that it applies to applications of AI across all sectors. It is applicable to all AI systems placed on the market or put into service in the EU.[116] For the purposes of the AIA, an AI system is:[117]

> ... a machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments;

In the AIA, legal intervention is tailored to the level of risk. AI systems that create unacceptable risks are prohibited;[118] high-risk AI systems will have to comply with a range of requirements on risk management,[119] testing,[120] technical robustness,[121] data training and governance,[122] human oversight,[123] and transparency.[124] Limited-risk AI systems attract transparency requirements,[125] and low or minimal risk AI systems are not subject to any legal obligations.

Transparency - an important pre-requisite for the governance of AI

Madalina Busuioc, Deidre Curtin and Marco Almada make an insightful and persuasive argument in their article *Reclaiming Transparency: Contesting the Logics of Secrecy*

[115] European Parliament "Briefing: Artificial intelligence act" (June 2023) https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf at 2.
[116] Artificial Intelligence Act, Art 2(1).
[117] Article 3(1).
[118] Article 5.
[119] Article 9.
[120] Articles 9(5)-(7).
[121] Article 15.
[122] Article 10.
[123] Article 14.
[124] Article 13.
[125] Article 52.

within the AI Act.[126] They believe transparency is an important pre-requisite for the governance of AI and argue that current debate over AI technologies has recast the need for transparency into a need for *explanations* for algorithmic outputs or decisions made by black box systems. They argue that this is a fundamental shift away from the original meaning of transparency, which is to "*render visible* that which is hidden, and in doing so ... open up possibilities for oversight that otherwise would not be there".[127]

Transparency, which is a necessary but incomplete requirement for accountability, becomes vulnerable to manipulation if it is reduced to explanations for decisions curated by providers and users of AI systems.[128] Transparency is arguably remodelled in this way because in its original meaning transparency revealed something that people could understand on seeing it. However, with complex AI systems, this is not the case. Lay people cannot understand the workings of complex AI systems, however experts and public interest groups can. Thus, the relevance and importance of transparency as disclosure, as opposed to explanation, holds true when those who can critique an AI system's inner workings are able do so.[129] The authors argue that 'the transparency-as-explanation' response fails to deliver on the promise of transparent and reasoned decision-making.[130]

Interestingly, the authors attribute the AIA's market focus[131] as being largely responsible for it "insufficiently regulat[ing] the use of AI systems by public authorities, directing instead most of its regulatory attention to the behaviour of providers".[132] Providers of high-risk AI systems bear most of the obligations under the AIA. However, the AIA is also intended to govern the use of AI in the public sector[133] and provide protection for fundamental rights. The authors argue that "its transparency measures, geared towards

[126] Madalina Busuioc, Deidre Curtin and Marco Almada "Reclaiming Transparency: Contesting the Logics of Secrecy within the AI Act" (2023) 2 EUR. L. OPEN 79.
[127] At 81.
[128] At 79.
[129] At 87.
[130] At 89.
[131] The primary legal foundation of the AIA is Article 114 of the Treaty on the Functioning of the European Union, which empowers the EU to adopt unifying measures concerning the functioning of the internal market.
[132] Busuioc et al, above n 126, at 89.
[133] Artificial Intelligence Act, Art 3(4).

product safety risks, offer little remedy for the unique risks the use of AI by public authorities imposes upon individuals affected by such deployments".[134]

The role of transparency has been diminished by the shift from disclosure to explanation, reducing the potential for oversight of public sector decision-making utilising AI systems. Additionally, the requirements of disclosure from the providers of AI systems to users are underspecified and allow for providers to exercise discretion around critical transparency choices.[135] The authors state that:[136]

> The absence of robust and unambiguous legal standards as to what measures are required to afford meaningful interpretation of system outputs is a serious shortcoming that will detract from effective transparency and accountability in this area.

Consequently, the authors argue that there is unlikely to be any real insight as to how AI systems function because of the vague and indeterminate transparency requirements. Users of high-risk AI systems are "subject to perfunctory, token obligations"[137] – the authors blame the product safety and market focus reflected in the AIA's legal basis and warn that this is a "glaring omission" particularly in light of the trend across public sectors to rely on algorithmic systems managed by external providers to carry out critical public tasks.[138]

Where public authorities rely on proprietary algorithmic models to carry out public tasks, the logic of these systems will remain obscured not only to the public bodies that deploy them, but also to the individuals adversely affected by them "exacerbating information asymmetries vis-à-vis private providers and removing opportunities for meaningful oversight".[139] The authors recommend that "rather than purchasing proprietary AI models, public authorities could contractually require providers to forgo proprietary protections to be able to sell high-risk AI systems into the public sector".[140] I agree that this would

[134] Busuioc et al, above n 126, at 98.
[135] At 98.
[136] At 99.
[137] At 99-100.
[138] At 100.
[139] At 100.
[140] At 100.

address the issue of transparency of source code of algorithms and AI used in the public sector and would be particularly important for those algorithms that have a high risk of unintended consequences or will have a significant impact on individuals and/or communities if they do go wrong. (The types of algorithms that New Zealand's Algorithm Charter is concerned with).

There are different methods of testing AI systems, not all of which require full source code disclosure.[141] Nonetheless, complete transparency of AI models through the full disclosure of source code is necessary for thorough and comprehensive testing of AI systems. It gives regulators greater ability to ensure that algorithms and AI systems comply with the law.[142] The argument for protecting source code is justified on the grounds of protecting intellectual property and preventing unfair competition as well as preventing less technologically developed countries from gaining access to the AI technology of other countries that they trade with.[143] It is also argued that protecting source code through international laws and treaties can encourage the trade in AI.[144] Nevertheless, I argue that making source code available to regulators and/or auditors of AI systems is a necessary pre-requisite for the accountable use of AI and algorithms. Measures could be put in place to limit the risk of disclosure of source code beyond those who need access to it for verification and testing purposes.

Market surveillance authorities under the AIA

The AIA provides for significant information disclosure to market surveillance authorities (MSAs) in their enforcement roles. While MSAs will have extensive access to information, they will be bound by confidentiality clauses, meaning that this information cannot be fed back into the public debate.[145] Busuioc et al argue that information and deliberation are removed from the public arena, depleting accountability "of the very

[141] Andrew D. Mitchell, Dominic Let, Lingxi Tang "AI Regulation and the Protection of Source Code" (2023) International Journal of Law and Information Technology 31 4 at 291-293.
[142] At 288.
[143] At 288-289.
[144] At 289.
[145] Artificial Intelligence Act, Art 70.

lifeblood that sustains it: transparency".[146] The authors argue that "to sustain accountability", transparency in its full and original meaning, must be reclaimed.[147]

The authors suggest the following solutions:

(1) clear and unambiguous disclosure obligations must be placed on users of high-risk AI systems, especially government agencies. This will encourage users to demand transparency from providers of AI systems and the models they use;[148] and

(2) intellectual property and trade secrets cannot trump public interests in accountability and maintaining individual rights.[149]

Given the potentially high-stakes contexts in which government agencies can deploy AI systems to inform decision-making, I argue that they should demand private providers of AI systems to forego proprietary protections to sell their AI systems to the public sector. Alternatively, models could be developed in-house. Secrecy needs to be relinquished to enable meaningful control of AI systems and to enable public authorities to understand how AI systems work, and to be able to provide reasons for decision-making assisted by algorithms.[150]

In Chapter Three I discussed the use of COMPAS as a decision-making aid originally designed for correctional agencies in determining placement decisions, offender management and treatment planning. COMPAS was then adopted for use in sentencing in the U.S. and its use in this context has raised questions of fairness. COMPAS works like a black box and computer scientists warn that the lack of transparency in black box systems allows errors to spread, which can be detrimental to society.[151] Being able to "explain" black box models does not resolve this problem – AI models need to be fully transparent. Models still need to be understandable to lay people, but their source code should also be made available for inspection – we must be able to see inside the black

[146] Busuioc et al, above n 126, at 102.
[147] At 102.
[148] At 102.
[149] At 103.
[150] At 103.
[151] At 104.

box in order to fulfil the requirement of transparency in its full sense.[152] Enhanced understanding does not equate to transparency.[153] Busuioc et al state that they:[154]

> ... seek to prevent the idea and reality of transparency from being appropriated by those who give it a narrow, discretionary, and exclusively mediated substance, reinforcing self-serving industry narratives that effectively hollow out transparency.

Transparency in the full sense of the word is a critical prerequisite for public accountability of AI use. Therefore, I argue that it is this notion of transparency should be adopted in any New Zealand legislation designed to limit the harms of big data and/or address the uses of AI.

Accountability measures in the AIA

Oversight: from human 'in the loop' to institutional oversight

Closely linked to accountability is the need for effective oversight of AI systems. The AIA does not provide much guidance on how to achieve effective human oversight of AI systems and it leaves the oversight obligations of AI systems undefined.[155] Human oversight aims to improve the accuracy of AI systems, safeguard important values, and build trust in AI.[156] However, research shows that humans fall short in fulfilling their oversight functions.[157] Empirical research of human oversight of AI systems shows that humans can exhibit both "automation bias" by over-relying on algorithms and "algorithm aversion" through under-utilising algorithmic advice.[158] These shortcomings do not affect every human overseer, nor do they apply in all contexts where human oversight is relied on. However, research shows that humans are not good at judging the accuracy of algorithmic predictions.[159] Therefore, Johann Laux, along with many other experts, believes that "the challenge of human oversight is best addressed at the level of

[152] At 104.
[153] At 105.
[154] At 105.
[155] Johann Laux "Institutionalised distrust and human oversight of artificial intelligence: towards a democratic design of AI governance under the European Union AI Act" (2023) AI & Society at 2.
[156] At 1.
[157] At 1.
[158] At 2.
[159] At 2.

institutional design".[160] Accordingly, when agencies rely on human oversight as an accountability mechanism, they must provide evidence that human oversight is effective at mitigating harmful outcomes in that particular decision-making context. Laux argues convincingly that "demonstrably ineffective oversight is not trustworthy".[161]

The AIA requires effective human oversight measures to be identified, and if possible, built into the AI system before a high-risk AI system is placed on the market or implemented.[162] Deployers of high-risk AI systems are required to implement human oversight measures as directed by the provider.[163] Laux notes that this is significantly more than what is required under Article 22(1) of the GDPR under which data subjects have the right not to be subject to *solely* automated decision-making.[164] Laux argues that the problem with Article 14 as currently drafted, is that it does not provide detail or clarity on what will make human oversight effective or meaningful.[165] Laux states that:[166]

> In sum, the AI Act requires human overseers to be sufficiently competent and authorised to intervene in an AI system. It does not clarify which qualifications (let alone certifications) those individuals must have which are tasked with executing Article 14 AI Act functions. For a horizontal law such as the AI Act, and in the absence of existing standards for effective oversight, this may well be as much as is currently possible.

Laux argues that developers and deployers of high-risk AI systems will have considerable discretion in how they exercise their human oversight obligations under the AIA.[167] This seems problematic when the AIA relies on human oversight as an important accountability measure.

Human oversight is not an effective governance mechanism if humans systematically fail at overseeing AI. This can happen when untrained human overseers fall victim to automation bias and also when expert human overseers interfere with the accuracy of

[160] At 2.
[161] At 11.
[162] Artificial Intelligence Act, Art 14.
[163] Article 29.
[164] Laux, above n 155, at 2.
[165] At 3.
[166] At 4.
[167] At 4.

algorithmic predictions by over-relying on their own judgement versus the algorithmic recommendation.[168] Raja Parasuraman and Dietrich Manzey explain how automated decision aids are often misused by human overseers because people prescribe more importance and authority to automatically generated aids than other sources of advice or information. This automation bias can lead to decisions that are not based on a thorough consideration of all relevant information but are strongly biased by the automatically generated advice.[169] Reuben Binns illustrates how retaining human decision makers in order to preserve individual justice may interfere with the advantages that using an algorithmic system provides. For example, algorithms can ensure consistency of decision-making and limit injustices generated through biased human judgement. However, a human-in-the-loop will exercise their own discretion (in potentially beneficial, or harmful, ways) and arguably, not in conformity with the expectations of the designer of the algorithmic system.[170]

Laux refers to Ben Green's argument that effective oversight of AI systems may be an impossible task for humans. Firstly, because they might lack sufficient understanding of an AI system; and secondly, because they might lack the "cognitive infallibility" to oversee such a system.[171] Green suggests that humans "cannot reliably perform any of the desired oversight functions".[172] Mary Cummings also highlights the possibility of a "moral buffer" or distancing when a human decision-maker relies on an automated decision aid, which can diminish people's sense of responsibility for their actions.[173] Cummings' focus is on military applications of decision aids but the issue of accountability is also pertinent to the use of decision-making aids in the provision of social services, particularly when an algorithm fails to work as intended.

Green is critical of the reliance on human oversight as a mechanism for mitigating potential harms of algorithmic decision-making. He makes a persuasive argument that

[168] At 6.
[169] Raja Parasuraman and Dietrich H. Manzey "Complacency and Bias in Human Use of Automation: An Attentional Integration" (2010) Human Factors 52 3 at 391.
[170] Reuben Binns "Human judgment in algorithmic loops: individual justice and automated decision-making" (2020) Regulation and Governance 16 at 205-206.
[171] At 7.
[172] Ben Green "The flaws of policies requiring human oversight of government algorithms" (2022) Computer Law and Security Review 45 at 2.
[173] Mary L. Cummings "Automation and Accountability in Decision Support System Interface Design" (2006) Journal of Technology Studies 32 1 at 26.

while human oversight is relied on as a safeguard against the risks of government use of algorithms in decision-making, very rarely is empirical evidence available to demonstrate that human oversight is effective in protecting human rights, safety, or dignity.[174] If humans cannot properly fulfil their oversight functions as envisioned, "human oversight policies would have the perverse effect of alleviating scrutiny of government algorithms without actually addressing the underlying concerns".[175] I believe this is particularly concerning. New Zealand should incorporate Green's proposed shift from unquestioning reliance on human oversight with its inherent shortcomings, to adopt institutional oversight measures in a regulatory response to algorithmic decision-making.

Green proposes institutional oversight as a means of achieving greater accountability of government use of algorithms. This is a two-stage process. Firstly, agencies must justify their use of algorithms in high-stakes decisions in a written report. Green believes that agencies need to demonstrate that it is appropriate to use an algorithm and believes that there are "'red lines' that mark unacceptable uses of algorithms" such as facial recognition technology and predictive policing tools.[176] However, I believe there can be acceptable uses of both these types of technology. For example, FRT used in smart gates at a country's border can be managed ethically without significant residual privacy risks. Similarly, predictive policing tools, if used appropriately with safeguards in place to prevent feedback loops and address unjustified biases, may also be acceptable. Green goes on to state that if not prohibited by a 'red line', agencies must consider whether it is appropriate to adopt the algorithm in a particular decision-making context. Part of the justification to use an algorithm must include evidence that human oversight is effective. Without such evidence "governments should not incorporate the algorithm into human decision-making processes".[177] Secondly, agencies must make their reports publicly available and not be allowed to use the algorithm until approved by the public, potentially in the form of a democratically accountable body.[178] The burden should fall on agencies to prove that human oversight of algorithmic decision-making is effective at improving outcomes and mitigating concerns.[179]

[174] Green, above n 172, at 2.
[175] At 2.
[176] At 12.
[177] At 3.
[178] At 14.
[179] At 15.

Audits as an oversight mechanism

The AIA requires that providers of high-risk AI systems use internal controls or third-party audits to check compliance with the requirements for high-risk AI systems in the Act.[180] Laux identifies a pitfall of external audits as an oversight mechanism. He argues that:[181]

> With big tech companies leading the development of AI, auditors risk capture by industry interests to receive repeated auditing commissions.
>
> ...
>
> To be itself trustworthy, human oversight of AI must address the challenges of lacking competence and false incentives.

Laux suggests that if a system of rotation for auditors was implemented then that could promote impartiality on the part of auditors.[182] Impartiality, or the lack of it, may also be an issue for the industry standards setting bodies.

Complaints mechanism and right to an explanation

Natural and legal persons have the right to lodge complaints with the relevant market surveillance authority if they have grounds to consider that there has been an infringement of the Act.[183] Also, any affected individual subject to a decision taken by the deployer made on the basis of an output from a high-risk AI system (with some exceptions) which produces legal effects, or similarly significantly affects him or her in a way that adversely impacts their health, safety or fundamental rights, shall have the right to request clear and meaningful explanations of the role of the AI system in the decision-making process from the deployer of the AI system.[184] These remedies may provide an accountability mechanism in contexts where human oversight has failed. However, it is more desirable to prevent harm in the first instance than to provide a remedy for harm suffered.

---

[180] Artificial Intelligence Act, Art 43.
[181] Laux, above n 155, at 7.
[182] At 9.
[183] Artificial Intelligence Act, Art 85.
[184] Article 86.

What can New Zealand learn from the EU?

New Zealand can benefit from the EU's deliberations over its AIA if it chooses to follow the EU's lead and regulate AI. Independent oversight and accountability are fundamentally important. Institutional oversight will result in greater accountability of government use of algorithms. Green's two-stage process requiring government agencies to justify their use of algorithms would reduce the risks of big data having unintended consequences and causing social privacy harms. This is because many of the social privacy harms discussed in Chapter Three stem from algorithmic decision-making without appropriate oversight mechanisms in place. It would also reveal where the implementation of partial or fully automated decision-making is being introduced as a cost saving measure (which is not problematic in itself, but it may flag the need to pay closer attention to the safeguards in place to prevent harm). As Green outlined, human oversight will not always be an effective safety or accountability mechanism. His two-stage process requires agencies to provide evidence that human oversight of algorithmic decision-making is effective in a particular context rather than just relying on the (flawed) assumption that it is.

Trustworthiness and risk acceptability

New Zealand could also learn from the questionable approach of the EU in correlating the trustworthiness of an AI system with the acceptability of the risks it presents. Johann Laux, Sandra Wachter and Brent Mittelstadt believe that the EU was "overselling its regulatory ambition" in, what was, at the time of publication, its *proposal* for an AIA.[185] They argue that "there remains a threat of misalignment between levels of actual trust and the trustworthiness of applied AI".[186] The AIA has a dual purpose of promoting the increasing use and development of AI in the EU (in both the public and private sectors) and addressing the risks of AI systems. Laux, Wachter and Mittelstadt state that:[187]

[185] Johann Laux, Sandra Wachter and Brent Mittelstadt "Trustworthy artificial intelligence and the European Union AI Act: On conflation of trustworthiness and acceptability of risk" (2023) Regulation and Governance at 1.
[186] At 1.
[187] At 1.

> ... in its proposal the Commission chose to understand "trustworthiness" narrowly in terms of the "acceptability" of AI's risks, with the latter being primarily assessed through conformity assessments carried out by technology experts.

Laux et al argue that the use of AI systems by public agencies has the potential to erode trust in public institutions. The public and private sectors both seek to benefit from the efficiencies of AI, however the public sector also has an obligation to protect citizens from harm.[188] Thus, the use of AI in the public sector raises questions of legitimacy.[189]

Laux et al argue that poorly implemented and secretive policies erode public trust. Governments should be transparent about their implementation of AI systems, including where they have failed.[190] The implementation of AI systems too early risks eroding trust in an entire process, whether justified or not.[191] Laux et al believe that "trust in public institutions is closely linked to their perceived legitimacy".[192] Solely automated decision-making systems are perceived as illegitimate whereas human oversight has been shown to strengthen trust,[193] even if the human overseer has little to no influence on the decisional outcome.[194] I argue that the public trust induced by human oversight further strengthens the argument that government agencies should be made to justify their use of algorithms where those algorithms have a high risk of unintended consequences or will have a significant impact on individuals or communities if they do go wrong.[195] Part of that justification requires evidence that human oversight is effective in achieving its objectives in a given context. If, in New Zealand, human oversight of automated decision-making systems was shown to strengthen public trust in those systems (as is the case in the UK), this provides additional weight to the argument that agencies should be made to provide evidence that human oversight is an effective safeguard, not simply a trust-inducing measure.

---

[188] At 2.
[189] At 3.
[190] At 15.
[191] At 16.
[192] At 16.
[193] At 17.
[194] At 24.
[195] These are the types of algorithms that the Algorithm Charter is concerned with.

Trust in AI systems may be dependent on context. For example, where emotional or human factors matter, the AI system may be less trusted.[196] AI systems need to add some value to citizens in order to be trusted.[197] Also, citizens need to be able to trust the accountability mechanisms in place, especially when the public sector is both regulator of the AI system and user.[198] AI auditors must also be trustworthy. To achieve that they must be impartial and independent. The AIA states that notified bodies shall be independent of the provider of a high-risk AI system,[199] and that their organisation and operation will "safeguard the independence, objectivity and impartiality of their activities".[200] However, Laux et al argue that "prescribing independence and impartiality by law may not be enough in practice ... there is an obvious economic incentive to receive commissions from AI developers".[201]

The AIA relies on AI developers' self-assessment or external assessment by notified bodies or auditors. Laux et al state that:[202]

> At a minimum, establishing the impartiality and independence of intermediaries such as the notified bodies will thus be an essential normative element for the AI Act's prospects of engineering citizens' trust in AI.

This could be an important consideration in a New Zealand context where the pool of AI experts is much smaller, thereby increasing the likelihood of experts and providers being acquainted. However, it is likely that New Zealand would accept approvals from overseas notified bodies.

Proportionality in risk management - how much risk is acceptable?

As outlined, the AIA aims to balance promoting the uptake of AI with a proportionate regulatory burden to reduce risks. Part of this risk management approach requires risks to

[196] Laux et al, above n 185, at 23-24.
[197] At 24.
[198] At 25.
[199] Artificial Intelligence Act, Art 31(4).
[200] Article 31(6).
[201] Laux et al, above n 185, at 25.
[202] At 26.

be eliminated or reduced "as far as technically feasible through adequate design and development of the high-risk AI system".[203]

Fraser and Villarino argue that risk acceptability judgements, inherently qualitative in nature, will not further the AIA's goal of trustworthiness unless made by decision-makers with a level of legitimacy. Even where courts have the constitutional legitimacy to weigh up uncertain values and trade off rights and interests, the weighing is still fraught and open to criticism. How will technologists be able to make such difficult value judgements before a high-risk system has even been implemented?[204] Commercial AI providers have commercial incentives to resolve uncertainties in their own financial interests. Big tech will not make principled decisions about AI risk acceptability in the public interest because they are driven by profit imperatives.[205] Standards bodies are also ill-equipped to make value judgements about risk acceptability - Fraser and Villarino state that:[206]

> They are typically constituted by technology industry insiders, have little experience with human rights and make minimal provision for participation by civil society and other stakeholders. Standards generally govern the engineering of technical features and processes rather than trade-offs between competing rights and interests in complex socio-political contexts.

Fraser and Villarino argue that those tasked with deciding how risk acceptability works must have the "capability and the civic, regulatory legitimacy to do so".[207] They believe there needs to be input from stakeholders and those affected by AI decision-making on the acceptability of risks to human rights and "regulators with political and institutional

[203] Artificial Intelligence Act, Art 9(5)(a).
[204] Henry Fraser and José-Miguel y Villarino "Acceptable Risks in Europe's Proposed AI Act: Reasonableness and Other Principles for Deciding How Much Risk Management Is Enough" (2023) European Journal of Risk Regulation at 12.
[205] At 13; There are several instances where big tech companies have fired ethical oversight teams where they have provided advice that does not align with their interests (Huw Roberts, Alexander Babuta, Jessica Morley, Christopher Thomas, Mariarosaria Taddeo, Luciano Floridi "Artificial intelligence regulation in the United Kingdom: a path to good governance and global leadership" (2023) Internet Policy Review 12 2 at 16); Google and Microsoft have either fired or cut back their ethics teams when they highlight the risks of the systems they are developing (Corporate Europe Observatory "Big tech lobbying is derailing the AI Act" (24 November 2023) https://corporateeurope.org/en/2023/11/big-tech-lobbying-derailing-ai-act)
[206] Fraser and Villarino, above n 204, at 13.
[207] At 13.

legitimacy need to be more involved in making or at least guiding risk-acceptability judgments".[208] I believe this is imperative to ensure trustworthiness.

Any legislative attempt by New Zealand to regulate AI would likely draw heavily on the EU's approach. However, there would be ways in which we could improve on the EU's AIA by placing more significance on transparency in the full sense of the word, implementing institutional oversight mechanisms, and facilitating greater input from stakeholders, public interest and civil liberties groups in risk assessments.

Can a decentralised approach effectively manage the risks of AI technologies?

In the UK government's response to the AI White Paper of March 2023, which proposes a "pro-innovation" approach to AI regulation,[209] it states that "the challenges posed by AI technologies will ultimately require legislative action in every country once understanding of risk has matured".[210] It acknowledges that a set of targeted, binding requirements on developers of highly-capable general purpose AI in particular, will be required in the future.[211] It argues that legislating too early risks hampering the ability to benefit from technological progress. The UK government confirms the establishment of a set of non-statutory principles to be applied by existing regulators to govern AI. Those principles are:[212]

- safety, security and robustness
- appropriate transparency and explainability
- fairness
- accountability and governance
- contestability and redress.

[208] At 14.
[209] UK Department for Science, Innovation and Technology "A pro-innovation response to AI regulation" (March 2023) https://assets.publishing.service.gov.uk/media/64cb71a547915a00142a91c4/a-pro-innovation-approach-to-ai-regulation-amended-web-ready.pdf
[210] UK Department for Science, Innovation and Technology, "Consultation outcome: A pro-innovation approach to AI regulation: government response" (6 February 2024) https://www.gov.uk/government/consultations/ai-regulation-a-pro-innovation-approach-policy-proposals/outcome/a-pro-innovation-approach-to-ai-regulation-government-response
[211] UK Department for Science, Innovation and Technology, above n 210.
[212] UK Department for Science, Innovation and Technology, above n 210.

Regulators are not given new enforcement powers, but the government has allocated a £10 million package to boost regulators' AI capabilities.[213] Arguably, this is an insufficient allocation of funding given that it will need to be spread over a number of regulatory bodies tasked with overseeing the use of AI within their remit.

In their article *Artificial intelligence regulation in the United Kingdom: a path to good governance and global leadership?*[214] Huw Roberts et al argue that the UK's approach "has fostered several contextually appropriate and novel governance initiatives"[215] - they point to the risks of AI systems being highly context specific, which may not lend themselves to a generalised response[216] (such as the AIA). However, they believe that the sectoral approach to the governance of AI may be undermined by the UK government's policy shift towards "a weakening of the powers and independence of a number of regulatory and oversight bodies".[217] Roberts et al argue that "the UK's approach to governance rests on the false assumption that deregulation engenders innovation".[218] They also point to the risk of lax checks and safeguards leading to more scandals and undermining public trust. Roberts et al question whether the existing powers of regulators are adequate for addressing the scope and complexity of risks that AI systems present.[219] They conclude that "given the broader deregulatory direction of the UK, it seems unlikely that a sector-based approach will be sufficient for addressing general-purpose AI".[220]

The UK's pro-innovation approach to regulating AI raises the question of whether New Zealand's regulators can effectively manage the risks of AI technologies. In Chapter Five I explain how New Zealand's regulatory framework for data privacy is ill-equipped to address privacy risks to individuals, and society, of the uses of big data associated with AI systems. With no AI regulation on the horizon, existing regulators will need greater enforcement powers and resourcing to harness the benefits, and manage the wide-ranging risks, of AI systems. Risks that extend beyond privacy risks, including health and safety,

---

[213] UK Department for Science, Innovation and Technology, above n 210.
[214] Huw Roberts, Alexander Babuta, Jessica Morley, Christopher Thomas, Mariarosaria Taddeo, Luciano Floridi "Artificial intelligence regulation in the United Kingdom: a path to good governance and global leadership" (2023) Internet Policy Review 12 2.
[215] At 3.
[216] At 9.
[217] At 14.
[218] At 16.
[219] At 16-17.
[220] At 17.

for example. In light of the new coalition government's plan to reduce public sector expenditure,[221] it is unlikely that regulatory bodies will receive increased funding to manage the risks of AI. Arguably, the target of cutting public sector expenditure by 6.5 per cent[222] could encourage government agencies to deploy big data measures that promise increased efficiencies and cost reductions, regardless of whether a suitable regulatory framework is in place. Significantly, in the round of public sector redundancies announced in early May 2024, the Department of Internal Affairs proposed disestablishing the Government Chief Privacy Officer and the privacy functions that sit under that role.[223] The UK government recently signalled a move to reduce the size of its public sector workforce by replacing workers with AI.[224]

A technological response to harms presented by online service providers

The systemic problem of information asymmetries between online service providers and users of those services has prompted technological responses by academics and tech entrepreneurs. Dr Marcin Betkier of Victoria University of Wellington has developed a Privacy Management Model (PMM) and right to informational self-determination as a response to the need for effective regulation of online services.[225] Out of all the technological responses I have encountered, Betkier provides the most persuasive counterargument to my argument for a social conception of privacy in response to big data harms.

---

[221] Reducing public sector expenditure is a priority of the new coalition government (NZ Government "Coalition Government 100-day plan" https://www.beehive.govt.nz/sites/default/files/2023-12/100%20Day%20Plan%20%281%29.pdf)
[222] Newshub "Why cutting public service roles to save costs normally backfires, creates 'consultocracy' instead" (12 December 2023) https://www.newshub.co.nz/home/politics/2023/12/analysis-why-cutting-public-service-roles-to-save-costs-normally-backfires-creates-consultocracy-instead.html
[223] Anna Whyte "More job cuts at Housing Ministry, StatsNZ, DIA" The Post (2 May 2024) https://www.thepost.co.nz/politics/350265201/more-job-cuts-housing-ministry-stats-nz-dia
[224] Vincent Manacourt "Oliver Dowden's 'hit squad' aims to replace UK civil service jobs with AI" (20.11.23) Politico https://www.politico.eu/article/dowdens-hit-squad-aims-to-replace-civil-service-jobs-with-ai/
[225] Marcin Betkier *Privacy Online, Law and the Effective Regulation of Online Services* (Intersentia Ltd, Cambridge, 2019).

In his book, *Privacy Online, Law and the Effective Regulation of Online Services*[226] Betkier presents a potential solution to the problem of information asymmetries between online service providers and users. He describes the problem in the following way:[227]

> ... the online environment creates a great asymmetry of information to the disadvantage of individuals. The power of service providers is based on controlling the architecture designed to incentivise individuals to leave as many data as possible. Individuals' weakness is a combination of factors including lack of awareness, expertise and vulnerability to cognitive biases. Furthermore, there is also a systemic dissonance between the nature of the long-term contract with multiple risks currently not able to be controlled by the individual, and the expectation of making one-off consent decisions. All of this suggests that information asymmetry cannot be solved by simply providing individuals with more information before concluding online contracts.[228]

I agree that asymmetries of information online are a significant problem and that the answer does not lie in providing individuals with even more information than they already have to navigate. Information- and power- asymmetries are a contributing factor to social privacy harms. However, Betkier and I diverge where Betkier's response is to empower the individual to take back control over their data through his proposed PMM and right to informational self-determination. My approach is to adopt a social conception of privacy to address the harms of big data, which requires less focus on individual control and more emphasis on establishing information ecosystems and regulatory environments where harms are considered collectively. However, Betkier's PMM, if implemented at the right setting, and with a rule against paying for enhanced privacy settings, could embody a social conception of privacy and be a useful tool in addressing social privacy harms.

Betkier's Privacy Management Model and the right to informational self-determination

Betkier argues that data privacy laws based on Fair Information Practice Principles, such as the Privacy Act and its information privacy principles, are ill-equipped to address the

[226] Betkier, above n 225.
[227] At 66.
[228] At 66.

significant asymmetries of information resulting from the Internet and other technologies.[229] This asymmetry of information works to the detriment of individuals and in favour of online service providers who rely on data surveillance as part of their business strategy. Betkier argues that human dignity is harmed by data surveillance and the use of algorithms by companies that do not inform the user of the logic behind them.[230] However, he maintains that "'the right to make bad privacy choices' should be preserved".[231]

Privacy regulation is justified as an important guarantee of individual freedom: "the collective level of privacy of all citizens underpins democratic institutions".[232] The privacy of individuals is important to society as a whole. Betkier proposes his PMM as a solution to current information asymmetries. Under the PMM, data subjects manage their own data through the agency of service providers - Personal Information Administrators. Data subjects set their own plans about how their data will be collected and used (meeting their own subjective expectations about privacy). These individual settings override the system settings provided by the service provider. Betkier argues that this is a better form of authorisation than current consent requirements because under PMM the data subject can view how their data is being used and change the function to authorise (or withdraw) the processing of different types of data.[233]

Betkier believes that individuals should make the decisions about their data collection and use, however the regulatory system needs to be set up to support data subjects to do this with the necessary skills and expertise provided by neutral third parties (Privacy Information Administrators). Privacy Information Administrators would provide expertise and simplify the information provided to data subjects to overcome the imbalance between the parties.[234] Thus, argues Betkier, the position of individuals as market players is reinforced.[235]

[229] At 197-198.
[230] At 70-74.
[231] At 74.
[232] At 93.
[233] At 87-105.
[234] At 136-139.
[235] At 95-96.

Shortcomings of the PMM and principle of informational self-determination

Betkier's proposal still requires input and management by data subjects who may not care, or have the time to be bothered, about data privacy. Will people want to pay for the service of Privacy Information Administrators when there is a general lack of awareness and understanding of privacy harms? It also raises the question of why people would choose an option that involves anything less than the highest level of data privacy from online service providers? The only logical reason would be to save money, which tilts the PMM towards protecting wealthier, more informed people at the expense of those who cannot afford privacy or do not understand the risks involved in over-disclosure. Further, can we trust Privacy Information Administrators to be impartial when we have seen the amount of money and extensive lobbying big tech directs at any legislative initiatives that threaten to limit their ability to operate as surveillance capitalists? Additionally, are there enough suitably qualified people to perform these roles?

Under Betkier's model, companies like Facebook would be allowed to charge those users who want high levels of privacy for their services and offer a free service to those who waive their privacy interests. This is, arguably, an unfair inducement for less well-off people to effectively sell their privacy. If we take the right to exchange one's personal information in exchange for free online services away from people, they will be faced with the option of having to pay for those services or going without. However, under the status quo, when people exchange their data for free services online they risk potentially paying in other ways, particularly those who are indigent or marginalised and disclose enough data to be categorised accordingly. They may be branded by social sorting algorithms in an unfavourable way, such as 'bad credit risk' or 'unemployable' and, consequently, may face higher mortgage interest rates or higher insurance premiums. Whether we prohibit the exchange of personal information for free services or preserve the right for people to make bad privacy choices, people pay (in some form or other) for the use of online services. I argue that poor people will pay a higher price for the freedom to make bad privacy choices than the wealthy.

The argument that people should be allowed to make bad privacy choices distracts from the real issue of the out-of-control collection and use of personal information by big tech companies. The right to make bad privacy choices is, in most contexts, not a right or

interest freely exercised with a full understanding of the consequences of doing so. We make a choice between *take it or leave it* when presented with the option of accepting an online platform's terms and conditions or not using its services. Most people do not have the necessary understanding of what the potential harms of agreeing to a platform's privacy policy might be (they may also be unknown or unknowable). Betkier critiques regulation that bans certain uses of data or mandates the disclosure of data in the public interest as paternalistic.[236] But what do individuals stand to lose by having the right to make bad privacy choices taken away from them in the context of online service provision? The answer is that they may have to pay for online services that they currently receive for 'free'. It is not paternalistic to object to the notion of paying for one's rights. We would be appalled at the idea of paying to vote or paying to not be discriminated against. Why is privacy different? Paternalism is the interference with the fully voluntary, self-regarding choices of competent adults.[237] However, privacy policies are often subject to change without notice, at the discretion of the online service provider. We cannot genuinely consent to future unknown uses of our information.

Indigent people suffer the most from inadequate privacy protections and automated decision-making. Thus, the most effective response to the social privacy harm of increasing inequality will avoid the pitfall of allowing consumers to pay for, or waive, their right to privacy. Betkier makes a strong argument for informational self-determination. The PMM is a clever response to information asymmetries but requires more protection for those that need it the most – the marginalised, the indigent, and disadvantaged. Betkier argues that any individual is potentially vulnerable, but in reality, it is those that are already struggling that will be hit the hardest by the collection and processing of their personal information. For Betkier's PMM to be most effective, the default settings need to be at a level that offers an appropriate degree of protection. If a social conception of privacy was applied to the PMM, it could be a particularly useful tool in addressing online information asymmetries and, in turn, could mitigate the harm to democracy of big tech companies wielding too much power as a consequence of holding vast amounts of information about people.

---

[236] At 95.
[237] Joel Feinberg "Legal Paternalism" in Joel Feinberg *The Moral Limits of the Criminal Law: Harm to Self* (Oxford Scholarship Online, 1989) at 15.

Betkier's argument that the right to make bad privacy choices should be preserved also ignores the data externalities of information exchanges. Where a regulatory framework facilitates poor privacy practices, the consequences of bad choices can have negative implications for all of society, in particular the harm of disinformation. Applying a social conception of privacy to the PMM would require the most privacy-enhancing settings as default and remove the ability to sell information or to trade it for free or less expensive services. When considering the broader human rights and interests at stake with the provision of online services, an individualistic response falls short and fails to account for collective interests in healthy information environments (where unjustified bias and disinformation are limited).

Human Rights-based responses

The Privacy Act is premised on a narrowly defined concept of data protection - the right to access and control one's information. Notably, there is no right to privacy in the New Zealand Bill of Rights Act 1990. The right to privacy was omitted from the Bill of Rights Act because privacy was a developing concept at the time and it was therefore held to be inappropriate to entrench a right that was not fully recognised, the boundaries of which were "uncertain and contentious".[238] As discussed in Chapter Five, the General Data Protection Regulation goes further than New Zealand's Privacy Act and attempts to address big data harms, but is also too individualistic in its approach.

In her book, *Beyond Data: Reclaiming Human Rights at the Dawn of the Metaverse*[239] Elizabeth Renieris argues that a rights-based approach to privacy should be informed by a broader conception of rights. She believes that a focus on data distracts from the real and primary aim of protecting people, stating that laws that focus on data "are extremely vulnerable to manipulation ... they risk creating a perception of safety, while sustaining harmful activities".[240] Renieris believes that "human rights are the only meaningful and

[238] Hon Geoffrey Palmer, Minister of Justice *A Bill of Rights for New Zealand: A White Paper* (Wellington, 1985) at 10.144.
[239] Elizabeth M. Renieris, *Beyond Data: Reclaiming Human Rights at the Dawn of the Metaverse* (The MIT Press, Cambridge, Massachusetts, 2023).
[240] At xiv.

sustainable way forward".[241] The changing and evolving nature of data means it is difficult to define. Renieris states that:[242]

> While we cannot realistically govern something we cannot define, focussing on something as abstract and amorphous as data is arguably easier than talking about power, inequality, exploitation, democracy, racism, and misogyny, among other mounting challenges raised by the ongoing evolution of digital technologies. Data has an air of neutrality that veils the deep structural biases and inequities that give rise to our perceived data-governance related challenges.

Renieris argues that data distracts from the real issues at hand or poorly frames them. The focus should be on the protection and safety of people rather than data.[243] Early data protection laws are woefully inadequate because they focus on data as the organising principle.[244] Renieris states that:[245]

> Unfortunately, this decades-old logic, encoded in our early approaches to data protection, still persists in large measure today, even as it is no longer fit for purpose. Instead, it is hampering the way we interpret and apply existing data-related laws and regulations, while also diminishing the urgency of designing better ones to reflect the substantially different technological realities (and relationships to technology) that we face. Worse yet, it is fundamentally failing to protect people.

Renieris describes the "poverty" of an individualistic approach to privacy that is focussed on data. She draws on Woodrow Hertzog's "control illusion" - making the argument that control does not scale and it also distracts from other important objectives like minimising data collection in the first place. By placing the burden on individuals to manage their data, it "hides the power imbalances in the modern digital economy".[246] Meaningful control over one's data is illusory. Renieris believes that "the technology governance challenges we face today are fundamentally about power".[247]

---

[241] At xvi.
[242] At 8.
[243] At 9.
[244] At 31.
[245] At 31.
[246] At 71.
[247] At 72.

Renieris argues that the framework of data protection or data governance is too narrow, the issues of privacy and human rights are broader and framing it as a data issue, risks reducing the questions to data governance issues. She argues that the first issue to address is limits on the specific processes that result in data in the first place.[248] There must be limits placed on datafication and rendition. Renieris identifies the following shortcomings with data privacy laws:

(1) the principle of purpose limitation does not even come into play until data subjects are identified or identifiable, which is often not the case at the first collection stage of passive data collection; and

(2) privacy frameworks are premised on a relationship between data controller and data subject, "with no appreciation for the complexity or obscurity of the parties that are observing, harvesting, and absorbing data en masse".[249]

Renieris believes that effective governance in a big data context begins with "identifying and imposing limits on things that cannot or should not be turned into data at all".[250] She identifies AI that purports to detect emotions or personality traits as a natural limit on datafication.[251] Similarly, neurotechnology that purport to "read minds" because, regardless of whether they deliver on their promise, they run the risk of being applied as though they really do read minds.[252]

The right to privacy has been adapted as a right to data protection, which falls short in responding to the full fleet of threats to human rights in a post digital world.[253] Renieris states that:[254]

> ... the nature of generalized emotion detection and affect recognition technologies, behavioral and soft biometrics, and other advanced technologies, especially when

[248] At 127.
[249] At 131.
[250] At 134.
[251] At 134-136.
[252] At 137.
[253] At 162-163.
[254] At 163.

> deployed in public or common spaces, requires us to go beyond personal data and reinterpret the underlying right to privacy itself through a more collective lens.

I argue that a broader consideration of human rights and interests (beyond control over one's data) should underlie legislative responses to the potential harms and benefits of emerging technologies that utilise big data. I agree that we need to revisit the right to privacy, but it should not be Betkier's version of informational self-determination, but rather in line with Renieris' vision for a collective concept that embodies other rights and interests threatened by emergent technologies and the information- and power-asymmetries between individuals and the big technology companies.

Conclusion

Beyond the Privacy Act, New Zealand has taken a soft law approach to addressing the potential harms of big data, which has proven inadequate. Multiple sets of guidelines for the ethical use of algorithms reveal that we understand the principles that underscore the responsible use of big data. However, until these principles are developed into enforceable rules, it is too easy for government agencies to lose focus and prioritise the other objectives that their performance is measured by, particularly when facing pressure to reduce costs.

As I concluded in Chapter Five, amendments to the existing regulatory framework for data privacy could improve information privacy practices and result in greater privacy protection for individuals. Improved individual privacy may go some way towards mitigating social privacy harms, but relying solely on individualistic approaches is wholly inadequate. Those who are most affected by big data are the least equipped to manage their own information and challenge the agencies and corporations that misuse it. Additionally, individualistic approaches ignore the data externalities of information exchanges. Regulation must be drafted in a way that recognises big data harms fall disproportionately on the marginalised, the indigent, and the disadvantaged. A social conception of privacy demands a top-down approach to law making to address the social privacy harms of big data.

It is plausible that we may import better AI technologies from Europe if the Brussels effect translates into improved, safer, and more ethical AI systems. Alternatively, leaving it too late to regulate artificial intelligence and big data here in New Zealand may leave us vulnerable to substandard AI applications, with no effective oversight measures. Government agencies should not be asking if they can afford to allocate the resources to ensure the safe and effective use of big data - big data should not be used without first having effective safeguards, and institutional oversight mechanisms, in place.

Practical measures should be taken to keep New Zealanders' data in New Zealand. As I argued in Chapter Two, New Zealand does not need to rely on overseas technology companies for data and AI solutions. A social conception of privacy, and in particular, Māori data sovereignty, supports the argument for investing in local AI expertise, software providers, and cloud services so we are not reliant on overseas technology companies for off-the-shelf AI products and database solutions. New Zealand technology providers need to be incentivised to build systems that embody privacy-by-design from the outset, with privacy and ethical considerations incorporated throughout the lifecycle of the technology. We need training datasets based on the New Zealand population that can be reviewed and audited in New Zealand without sending data overseas.

The need for transparency and a public conversation about law enforcement's use of surveillance, and government use of big data generally, is imperative. Government agencies' use of big data and AI is not sufficiently transparent, and thus, the degree of the likelihood of harm is unknown. Transparency, in the full sense of the word, must be a key theme of regulatory responses to big data and AI.

Conclusion

Over 140 years ago, Warren and Brandeis argued for the right to privacy of the individual from invasion by the press. The principle of inviolate personality was invoked in response to new camera technology that allowed for instantaneous photography, threatening the exposure of one's private life in ways not previously encountered.[1] Privacy law was born in response to a new use of technology that threatened existing social norms of information collection and use. Eight decades later, in the 1970s, data protection laws were drafted in response to the risks to privacy of large government computerised databases. History shows us that privacy law is adaptable. It must continue to evolve to meet the challenges of big data and the technology that it enables.

In this thesis, I have argued that New Zealand's regulatory framework for data privacy is not fit for purpose in an age of big data. An individualistic complaints-based Privacy Act combined with a soft law approach to the governance of algorithms and artificial intelligence (AI), is a wholly inadequate response to the harms, and benefits, of big data. The first step in addressing the shortcomings of New Zealand's regulatory framework is to accurately conceptualise privacy. To respond effectively to the challenges of big data we need a regulatory response informed by a social conception of privacy.

The power of big data lies in its sheer volume and ability to place people into groups based on the characteristics that they share with others. These relation-based groups allow agencies to sort, define, and classify people, drawing a range of inferences about our likes, dislikes, interests, and fears, which can then be used to sell us products, services, and ideas. Predictive analytics are used to estimate the probability of people behaving in a certain way. Predictions of the likelihood of an individual repaying a loan, or reoffending while on parole, are based on the behaviour of other people with whom they share similar characteristics.

[1] Samuel D. Warren and Louis D. Brandeis "The Right To Privacy" (1890) Harvard Law Review IV 5 at 205.

Big data can exacerbate existing inequalities in society and entrench them. Increasing inequality is a public harm that sets back society's interest in living in prosperous, healthy, and harmonious communities. Big data facilitates vast information asymmetries and power imbalances between individual data subjects and the big technology companies, which is harmful to democracies. Very large online platforms and search engines exercise significant influence over the information we access online and how that information is presented to us. The social privacy harms of big data demand a collective response; a top-down regulatory response that acknowledges the harms that can accrue beyond the individual data subjects in particular information transactions.

In this thesis I have argued that privacy is primarily a social concept; the very notion of privacy is dependent on the possibility, or necessity, of our interactions and relationships with other people and agencies. I draw on Helen Nissenbaum's Framework of Contextual Integrity, a counterweight to the argument that privacy is about limiting access to personal information and enhancing the individual's right to control information about themselves. Nissenbaum believes that what people care about is not simply *restricting* the flow of personal information but ensuring that it flows *appropriately.*[2] Appropriate information flows respect social expectations and norms. These are often challenged by new technologies, which should be evaluated in light of how they impact on the values and goals of the particular context that they are deployed in. A social conception of privacy would be more responsive to the social privacy harms of big data.

Adopting a social conception of privacy does not mean that individual consent is no longer meaningful or relevant. Consent is, and will continue to be, important. But determining *when* consent is a necessary and meaningful way of authorising information collection and use should be determined by the social context. Consent to information collection and use does not scale, and in most online contexts is ineffective. There is often no real choice for data subjects, who may not have an adequate understanding of the implications of consenting to a platform's terms and conditions. Critically, individual consent in online contexts disregards the information externalities of the exchange. A

[2] Helen Nissenbaum *Privacy In Context: Technology, Policy, and the Integrity of Social Life* (2010, Stanford Law Books, California) at 2.

social conception of privacy recognises the public harm that can accrue beyond the parties consenting to particular collections and uses of personal information.

There is still a place for an individual complaints-based privacy framework. The Privacy Act 2020 fulfils the purpose that the original 1993 Act set out to achieve. It provides a mechanism by which individuals can seek redress where they can establish that an action of an agency has caused an interference with their privacy.

A new power of the Privacy Commissioner to issue compliance notices was introduced in amendments to the Privacy Act in 2020.[3] Compliance notices can be issued without the requirement of establishing loss, harm, or detriment to an individual as described in s 69(2)(b), which is required (in addition to a breach of an information privacy principle) to establish an interference with privacy.[4] Unfortunately, this power has been underutilised by the Commissioner and may point to the need to adequately resource the Office of the Privacy Commissioner (OPC). Nevertheless, even if the OPC was fully resourced to enforce the Privacy Act, the Act's focus on the individual means it is ill-equipped to address social privacy harms.

The OPC recently published its *Briefing to the Incoming Minister of Justice* recommending a set of specific amendments to the Privacy Act 2020 to make it "fit for purpose in the digital age".[5] The recommendations include a new civil penalties regime for major non-compliance, stronger requirements for agencies deploying automated decision-making and for agencies to demonstrate how they meet privacy requirements, as well as greater resourcing of the OPC.[6] The briefing also refers to "new privacy rights for New Zealanders to better protect themselves".[7]

However, as this thesis illustrates, enhancing individual data rights is an incomplete response to big data harms. Making the Privacy Act *fit for purpose in the digital age* requires recognition of social privacy harms and an understanding of the relational aspect

[3] Privacy Act 2020, s 123.
[4] Section 69(2)(b).
[5] Office of the Privacy Commissioner *Briefing to the Incoming Minister of Justice* (4 December 2023) at 1 https://www.privacy.org.nz/assets/New-order/Resources-/Publications/Reports-to-Parliament-and-Government-/BIMs/BIM.pdf
[6] At 1.
[7] At 1.

of big data. An overly individualistic approach to governing the collection and use of personal information is inhibiting privacy law's ability to mitigate the negative implications of big data applications. Consequently, this puts at risk the development and uptake of beneficial big data uses by potentially undermining public trust.

Individuals do not always know when they have been subject to unlawful information practices or how they have been harmed by predictive systems or biased algorithms. Marginalised and indigent people bear the biggest burden of big data and are least equipped to assert their rights. Arguably, they shouldn't have to, if there are ways of preventing unfair information practices from occurring in the first place.

The threat of complaints, compliance notices, and possible litigation can encourage compliance with data privacy laws by agencies, particularly the public sector and those companies that care about their privacy reputation. However, if individuals don’t know how their data is being used, or by whom, and they can't recognise how they have been harmed, then they will be unable to take action against companies or governments that are breaching data privacy rules. Similarly, if regulators are not adequately resourced, the threat of compliance notices carries less weight as the likelihood of enforcement diminishes.

There are amendments that could be made to the Privacy Act that would better protect individual privacy. Stronger individual privacy rights may go some way towards mitigating social privacy harms. While collective levels of individual privacy in society are undoubtedly important, individual data privacy rights are not a complete response to social privacy harms because they do not address harms to groups of the population and to society overall. Big data facilitates the degradation of information environments and can exacerbate existing inequalities in society. Social privacy harms can be caused by cumulative actions of multiple agencies over extended periods of time and can, therefore, be difficult to assign, define, and quantify. One action in itself may not cause loss, harm, or detriment to an individual, but cumulatively certain information practices can be harmful to society overall.

The contemporary response of legislatures to implement measures designed to enhance the individual's control over their own data are not working to increase privacy in real

terms. While there is still a place for individual privacy rights, regulatory regimes that focus on enabling the individual to manage their own data (an increasingly difficult, if not impossible, task) have their limits. The information practices of big tech cannot be adequately addressed by narrowly defined individual rights-based privacy frameworks. Surveillance capitalism demands ever-increasing information generation and collection, fuelling social privacy harms. The underlying business model of the big tech companies, driving increasing information- and power- asymmetries, needs addressing. This is beyond the power of the New Zealand legislature. But there are measures New Zealand can, and should, take to better protect New Zealanders from social privacy harms.

New Zealand's responses to the potential harms and benefits of big data beyond the Privacy Act embody a soft law approach, which has proven inadequate. The principles and guidelines for the ethical use of algorithms documented in multiple codes / charters / frameworks, none of which are binding or enforceable, reveal that we understand the principles that underscore the responsible use of big data. However, until these principles are developed into enforceable rules, they are largely ineffective at preventing unethical uses of personal information.

Admittedly, New Zealand is a small country with limited influence on global AI governance or the regulation of data privacy in other jurisdictions. We import other countries' information practices and the problems inherent with those when we use online services and conduct business with overseas companies. However, in this thesis I have argued that there are some practical safeguards and measures that we can, and should, adopt, particularly within the public sector. We must invest in New Zealand-made, transparent, and accountable data solutions.

Government agencies should be made to justify their use of algorithms that have a high risk of unintended consequences and/or a significant impact on individuals or groups when they do not work as intended. These are the types of algorithms that the Algorithm Charter is concerned with. The principles of the Charter should be developed into enforceable rules. Government agencies should not deploy high-risk AI, or the kinds of algorithms to which the Charter applies, until they can establish that their use is justified, and that they have effective, proven mitigations in place.

New Zealand should invest in home-grown technology companies and transition the procurement of government technology and software from U.S. tech providers to New Zealand-made solutions. A social conception of privacy is consistent with Māori data sovereignty, which supports keeping New Zealanders' information in New Zealand.

Government agencies' use of big data to enable surveillance and AI technologies is not as transparent or accountable as it should be, which is problematic, particularly when agencies are largely left to self-govern their use of big data. We should not be relying on journalists exercising their powers under the Official Information Act 1982 to find out what government agencies are doing with our information. Transparency and institutional accountability for the use of algorithms and AI would be key themes underlying a regulatory response to big data informed by a social conception of privacy.

Bibliography

A. Primary Materials

New Zealand Table of Statutes and Codes

Credit Reporting Privacy Code 2020.

Civil Defence National Emergencies (Information Sharing) Code 2020.

Harmful Digital Communications Act 2015.

Health Information Privacy Code 2020.

Human Rights Act 1993.

Local Government Official Information and Meetings Act 1987.

Medicines Act 1981.

New Zealand Bill of Rights Act 1990.

Official Information Act 1982.

Parole Act 2002.

Privacy Act 1993 (repealed).

Privacy Act 2020.

Privacy Amendment Bill 292-1 (2023).

Privacy Commissioner Act 1991 (repealed).

Search and Surveillance Act 2012.

Telecommunications Information Privacy Code 2020.

Tohunga Suppression Act 1907.

Wanganui Computer Centre Act 1978 (repealed).

United Kingdom Table of Statutes

Data Protection Act 1998.

Data Protection Act 2018.

United States Table of Statutes

Cable Communications Policy Act 1984.

California Consumer Privacy Act 2018.
California Privacy Rights Act 2020.
Children's Online Privacy Protection Act 1998.
Health Insurance Portability and Accountability Act 1996.
Federal Trade Commission Act 1914.
Privacy Act 1974.
United States of America Video Privacy Protection Act 1988.

European Table of Statutes

Charter of Fundamental Rights of the European Union 2000.
Council of Europe, Modernised Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data 2018 (Convention 108).
European Convention on Human Rights 1950.
European Union, Directive 95/46/EC (repealed).
Proposal for Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts.
Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1.
Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act).
Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act).

International Guidelines, Declarations and Covenants

Asia-Pacific Economic Cooperation, APEC Privacy Framework 2015.
General Assembly of the United Nations, International Covenant on Civil and Political Rights 1966.

General Assembly of the United Nations, Optional Protocol to the International Covenant on Civil and Political Rights 1966.
General Assembly of the United Nations, Universal Declaration of Human Rights 1948.
OECD, Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data 1980, and revised Guidelines 2013.
United Nations Declaration on the Rights of Indigenous Peoples 2007.
United Nations General Assembly Resolution 68/167. The right to privacy in the digital age.

New Zealand Table of Cases

C v Holland [2012] 3 NZLR 672.
EQC v Krieger [2013] NZHC 3140.
Hosking v Runting [2005] 1 NZLR 1.
R v Rowe [2005] 2 NZLR 833.
Rowe v Police HC Dunedin CRI 2005-412-000051, 24 November 2005.
Solicitor-General for New Zealand v Krieger [2014] NZHC 172.
Hammond v NZCU Baywide [2015] NZHRRT 6.
Tucker v News Media Ownership Ltd. [1986] 2 NZLR 716.

United Kingdom Table of Cases

Campbell v MGN Ltd [2004] 2 AC 457.
Coco v AN Clarke (Engineers) Ltd. [1969] RPC 41 (Ch).
Prince Albert v Strange (1849) 41 ER 1302.
Regina v Department of Health, Ex parte Source Informatics Ltd. CA 1999.
Weller v Associated Newspapers Ltd [2015] EWCA Civ 1176.

European Table of Cases

Case C-311/18 *Data Protection Commissioner v Facebook Ireland and Maximillian Schrems* [2020] ECR 297.
Case C-362/14 *Maximillian Schrems v Data Protection Commissioner* [2015] ECR 398.

Decision of the Data Protection Commission made pursuant to Section 111 of the Data Protection Act 2018 and Articles 60 and 65 of the General Data Protection Regulation in the matter of Meta Platforms Ireland Limited (previously known as Facebook Ireland Limited) (Data Protection Commission, Ireland, 12 May 2023).
Peck v the United Kingdom (2003) 36 EHRR 41 (ECHR).

United States table of cases

Gardner v. Florida 430 U.S. 349, 97 S Ct. 1197 (1977).
Olmstead et al v United States - Green et al v Same - McInnis v Same 277 U.S. 438 (1928).
State v. Loomis 881 N.W.2d 749 (2016).
State v Skaff 152 Wis. 2d 48 (Wis. Ct. App. 1989).

B. Secondary Materials

Ademole S. Adamson and Avery Smith "Machine Learning and Health Care Disparities in Dermatology" (2018) JAMA Dermatology 154 11 at 1247.

Diego Ardila, Atilla P. Kiraly, Sujeeth Bharadwaj, Bokyung Choi, Joshua J. Reicher, Lily Pera, Daniel Tse, Mozziyar Etemadi, Wenxing Ye, Greg Corrado, David P. Naidich, Shravya Shetty “End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography” (2019) Nature Medicine 25 954-961.

Article 29 Data Protection Working Party "Guidelines on Automated Individual decision-making and Profiling for the purposes of Regulation 2016/679" 17/EN WP251rev.01.

Article 29 Data Protection Working Party "Opinion 11/2011 on the level of protection of personal data in New Zealand" Adopted on 4 April 2011 00665/11/EN WP 182.

Trevor Back, Christopher Nielsen, Joseph R. Ledsam, Shakir Mohamed et al “A clinically applicable approach to continuous prediction of future acute kidney injury” (2019) 572 Nature 116-119.

Jack Bandy "Problematic Machine Behavior: A Systematic Literature Review of Algorithm Audits" (2021) Proc. Human-Computer Interaction 5 CSCW1 74.

Jessica L. Bell “Governing Commercial Access to Health Data for Public Benefit: Charity Law Solutions” (2019) *Medical Law Review*

Frederic D. Bellamy "U.S. data privacy laws to enter new era in 2023" (13 January 2023) Westlaw Thompson Reuters https://www.reuters.com/legal/legalindustry/us-data-privacy-laws-enter-new-era-2023-2023-01-12/

Ruha Benjamin *Race After Technology: Abolitionist Tools for the New Jim Code* (Polity Press, Cambridge, 2019).

Marcin Betkier *Privacy Online, Law and the Effective Regulation of Online Services* (Intersentia Publishing, Cambridge, 2019).

Deryck Beyleveld and Elise Histed “Case Commentary: Anonymisation is not Exoneration” (1999) Medical Law International 4 69-80.

Reuben Binns "Human judgment in algorithmic loops: individual justice and automated decision-making" (2020) Regulation and Governance 16.

Reuben Binns and Michael Veale "Is that your final decision? Multi-stage profiling, selective effects, and Article 22 of the GDPR" (2021) International Data Privacy Law 11 4.

Michael Birnhack and Niva Elkin-Koren "The Invisible Handshake: The Re-emergence of the State in the Digital Environment" (2022) 8 2 Virginia Journal of Law and Technology 1.

Edwin Black *IBM and the Holocaust: The Strategic Alliance Between Nazi Germany and America's Most Powerful Corporation* (Three Rivers Press, New York, 2001).

Jordan M. Blank “Privacy and Outrage” (2018) Journal of Law, Technology & the Internet 9.

Jordan M. Blanke "Protection for 'Inferences Drawn': A Comparison Between the General Data Protection Regulation and the California Consumer Privacy Act" (2020) Global Privacy Law Review I 2 81-92.

Michael V. Boland "Big Data, Big Challenges" (2016) American Academy of Opthalmology 123 1.

Danah Boyd and Kate Crawford “Critical Questions for Big Data” (2012) 15 Information, Communication and Society 5 662-679.

Sergey Brin and Lawrence Page "The anatomy of a large-scale hypertextual Web search engine" (1998) Computer Networks and ISDN Systems 30.

British Academy and the Royal Society "Data management and use: Governance in the 21st century" (2017).

Shea Brown, Jovana Davidovic, Ali Hasan "The algorithm audit: Scoring the algorithms that score us" (2021) Big Data and Society.

Madalina Busuioc, Deidre Curtin and Marco Almada "Reclaiming Transparency: Contesting the Logics of Secrecy within the AI Act" (2023) 2 EUR. L. OPEN 79.

Declan Butler “When Google got flu wrong” (2013) 494 Nature 155-156 at 155.

Nicholas R. Buttrick and Shigehiro Oishi "The psychological consequences of income inequality" (2017) Social and Personality Psychology Compass 11.

Lee Bygrave *Data Privacy Law: An International Perspective* (Oxford University Press, Oxford, 2014).

Lee Bygrave "Machine Learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions" (2022) Published online by Cambridge University Press https://www.duo.uio.no/handle/10852/94223

Lee Bygrave "Security by Design: Aspirations and Realities in a Regulatory Context" (2021) Oslo Law Review 8 3 126-177.

Michael F. Byrne (ed.) *AI in Clinical Medicine: A Practical Guide for Healthcare Professionals* (2023, John Wiley & Sons Ltd., Hoboken, USA) at 260.

Pam Carter, Graeme T Laurie, Mary Dixon-Woods "The social licence for research: why *care.data* ran into trouble" (2015) Journal of Medical Ethics 41 404-409.

Cristina Blasi Casagran and Mathias Vermeulen "Reflections on the murky legal practices of political micro-targeting from a GDPR perspective" (2021) International Data Privacy Law 11 4.

Nikki Chamberlain and Stephen Penk (eds) *Privacy Law in New Zealand* (3rd ed.) (Thomson Reuters, Wellington, 2023).

Jay Pil Choi, Doh-Shin Jeon and Byung-Cheol Kim "Privacy and personal data collection with information externalities" (2019) Journal of Public Economics 173 113-124.

Danielle Keats Citron and Frank Pasquale "The Scored Society: Due Process For Automated Predictions" (2014) Washington Law Review 89 1.

Danielle Keats Citron and Daniel Solove "Privacy Harms" (2021) The George Washington University Law School 2021-11.

Theodore F. Claypoole "U.S Senate Takes Aim at Big Tech" (2021) National Law Review XI 40.

Nathan Cortez “The Evolving Law and Ethics of Digital Health” in Homero Rivas and Katarzyna Wac (eds) *Digital Health: Scaling Healthcare to the World* (Springer International Publishing, 2018, Switzerland).

Kate Crawford *Atlas of AI: Power, Politics, and the Planetary Cost of Artificial Intelligence* (Yale University Press, New Haven, 2021).

Mary L. Cummings "Automation and Accountability in Decision Support System Interface Design" (2006) The Journal of Technology Studies 32 1.

Angela Daly “The law and ethics of ‘self-quantified’ health information: an Australian perspective” (2015) International Data Privacy Law 5 (2).

Tim Dare, Rhema Vaithianathan, Ireene De Haan "Adressing Child Maltreatment in New Zealand: Is Poverty Reduction Enough?" (2014) Educational Philosophy and Theory 46:9.

Data Protection Commission, Ireland "Decision of the Data Protection Commission made pursuant to Section 111 of the Data Protection Act 2018 and Articles 60 and 65 of the General Data Protection Regulation in the matter of Meta Platforms Ireland Limited (previously known as Facebook Ireland Limited) (12 May 2023).

Francesco Decarolis and Muxin Li "Regulating online search in the EU: From the android case to the digital markets act and digital services act" (2023) International Journal of Industrial Organization 90.

Karin Dembrower, Alessio Cripper, Eugenia Colôn, Martin Eklund, Fredrik Strand "Artificial Intelligence for breast cancer detection in screening mammography in Sweden: a prospective, population-based, paired-reader, non-inferiority study" (2023) Lancet 5.

Nancy Devlin, Alan Maynard, Nicholas Mays "New Zealand's new health sector reforms: back to the future?" (2001) British Medical Journal 322 1171-1174.

Yao Dong *Privacy Act 2020* (2020) Auckland University Law Review 26.

Jamie Duncan "Data protection beyond data rights: governing data production through collective intermediaries" (2023) Internet Policy Review 12 3.

John Eckenrode, Elliott G. Smith, Margaret E. McCarthy, Michael Dineen "Income Inequality and Child Maltreatment in the United States" (2014) Pediatrics 133 3.

John Edwards, Privacy Commissioner "Privacy Commissioner's submission on the Privacy Bill to the Justice and Electoral Select Committee" (31 May 2018).

Lilian Edwards and Michael Veale “Slave to the Algorithm? Why a Right to an Explanation is Probably Not the Remedy You’re Looking for” (2017-2018) Duke Law and Technology Review 16.

Michael Eisenstein “Infection forecasts powered by big data” (2018) 555 Nature.

Susan L. Erikson “Cell Phones as an Anticipatory Technology: Behind the Hype of Big Data for Ebola Detection and Containment” (2018) Working Papers of the Priority Programme 1448 of the German Research Foundation Adaption and Creativity in Africa: technologies and significations in the making of order and disorder Nr.24 at 5.

Virginia Eubanks *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (St. Martin’s Press, New York, 2017).

European Court of Human Rights "Guide on Article 8 of the European Convention on Human Rights - Right to respect for private and family life, home and correspondence" (Council of Europe, 2020).

Jeffrey De Fauw et al. “Clinically applicable deep learning for diagnosis and referral in retinal disease” (2018) 24 Nature Medicine 1342-1350.

Joel Feinberg *Harm To Others: The Moral Limits of the Criminal Law* (Oxford University Press, Oxford, 1984).

Joel Feinberg *The Moral Limits of the Criminal Law: Harm to Self* (Oxford Scholarship Online, 1989).

Henry Fraser and José-Miguel y Villarino "Acceptable Risks in Europe's Proposed AI Act: Reasonableness and Other Principles for Deciding How Much Risk Management Is Enough" (2023) European Journal of Risk Regulation.

Katherine Freeman "Algorithmic Injustice: How the Wisconsin Supreme Court Failed To Protect Due Process Rights In State v. Loomis" (2016) North Carolina Journal of Law & Technology 75.

Hannah Fry *Hello World: How to Be Human in the Age of the Machine* (Transworld Publishers, Penguin, London, 2018).

Chlotia Garrison and Clovia Hamilton “A Comparative Analysis of the EU GDPR to the US’s Breach Notifications” (2019) Information and Communications Technology Law DOI:10.1080/13600834.2019.1571473.

Graham Garth, Marylynn Ostrowski and Alyse Sabina “Population health-based approaches to utilizing digital technology: a strategy for equity” (2016) Journal of Public Health Policy 37 (2).

Colin Gavaghan, Alistair Knott, James Maclaurin and Joy Liddicoat *Government Use of Artificial Intelligence in New Zealand: Final Report on Phase 1 of the New Zealand Law Foundation's Artificial Intelligence and the Law in New Zealand Project* (New Zealand Law Foundation, Wellington, 2019).

Elena Gil González and Paul de Hert "Understanding the legal provisions that allow processing and profiling of personal data - an analysis of GDPR provisions and principles" (2019) ERA Forum 19.

Ben Green "The flaws of policies requiring human oversight of government algorithms" (2022) Computer Law and Security Review 45.

Graham Greenleaf and Lee Bygrave "Not entirely adequate but far away: Lessons from how Europe sees New Zealand data protection" (2011) Privacy Laws and Business International Report 111.

Chufeng Gu, Yujie Wang, Yan Jiang, Feiping Xu, Shasha Wang, Rui Liu, Wen Yuan, Nurbiyimu Abudureyimu, Ying Wang, Yulan Lu, Xiaolong Li,Tao Wu, Li Dong, Yuzhong Chen, Bin Wang, Yuncheng Zhang, Wen Bin Wei, Qinghua Qiu, Zhi Zheng , Deng Liu, Jili Chen, "Application of artificial intelligence system for screening multiple fundus diseases in Chinese primary healthcare settings: a real-world, multicentre and cross-sectional study of 4795 cases" (2023) Br J Ophthalmol.

Audrey Guinchard "Our digital footprint under Covid-19: should we fear the UK digital contact tracing app?" (15 July 2021) International Review of Law, Computers & Technology 35 1.

Gehan Gunasekara and Erin Dillon “Data Protection Litigation in New Zealand: Processes and Outcomes” (2008) VUWLR 457-486.

Gehan Gunasekara and Jingyi Xiong "Lost in Translation? Privacy and Unfair or Deceptive Acts or Practices in Commerce in the United States" (2016) New Zealand Business Law Quarterly 22.

Matt Hancock Ministerial forward to the UK Department of Health and Social Care Policy Paper: *Data saves lives: reshaping health and social care with data (draft)* (updated 8 September 2021).

Peter Hanlon, Luke Daines, Christine Campbell, Brian McKinstry, David Weller and Hilary Pinnock “Telehealth Interventions to Support Self-Management of Long-Term Conditions: A Systematic Metareview of Diabetes, Heart Failure, Asthma, Chronic

Obstructive Pulmonary Disease, and Cancer" (2017) Journal of Medical Internet Research 19 (5).

Steve Harris et al. "Critical Care Health Informatics Collaborative (CCHIC): Data, tools and methods for reproducible research: A multi-centre UK intensive care database" (2018) International Journal of Medical Informatics 112 82-29.

Woodrow Hartzog "What is Privacy? That's the Wrong Question" (2021) 88 U. CHI. L. REV. 1677.

Nigel Hawkes "NHS data sharing deal with Google prompts concern" (2016) British Medical Journal 353 i2573.

Simon I. Hay and Peter W. Gething et al. "Mapping the global prevalence, incidence and mortality of *Plasmodium falciparum,* 2000-17: a special and temporal modelling study" (2019) 394 Lancet 322-331.

Richard Hillestad, James Bigelow, Anthony Bower, Frederico Girosi, Robin Meili, Richard Scoville, Roger Taylor "Can Electronic Medical Record Systems Transform Health Care? Potential Health Benefits, Savings, And Costs" (2005) 24 Health Affairs 5 1103-1117.

Hal Hodson "Google knows your ills" (2016) New Scientist 230 3072 22-23.

Doug Howe and Seung Yon Rhee et al "The Future of Biocuration" (2008) 455 Nature 47-50.

Philip Hunter "The big health data sale" (2016) Science and Society 17(8) 1103-1105.

Information Commissioner's Office "Letter to Sir David Sloman, Chief Executive, Royal Free NHS Foundation Trust" (3 July 2017).

Rebecca Janβen, Reinhold Kesler, Michail E. Kummer, Joel Waldfogel "GDPR and the Lost Generation of Innovative Apps" (2022) National Bureau of Economic Research, Cambridge, Massachusetts.

James S. Kahn, Veenu Aulakh, Adam Bosworth “What It Takes: Characteristics Of The Ideal Personal Health Record” (2009) 28 Health Affairs 2 369-376.

Frederike Kaltheune “Europe cannot afford to “shelter-in-place” on digital regulation” (2020) Data and Society.

Margot E. Kaminiski and Gianclaudio Malgieri "Algorithmic impact assessments under the GDPR: producing multi-layered explanations" (2021) International Data Privacy Law 11 2.

Cecilia Kang "Democratic Congress Prepares to Take On Big Tech" (26.1.21) New York Times.

Jaclyn C. Kearns, Emily R. Edwards, Erin P. Finley, Joseph C. Geraci, Sarah M. Gildea, Marianne Goodman, Irving Hwang, Chris J. Kennedy, Andrew J. King, Alex Luedtke, Brian P. Marx, Maria V. Petukhova, Nancy A. Sampson, Richard W. Seim, Ian H. Stanley, Murray B. Stein, Robert J. Ursano and Ronald C. Kessler "A practical risk calculator for suicidal behavior among transitioning U.S. Army soldiers: results from the Study to Assess Risk and Resilience in Servicemembers-Longitudinal Study (STARRS-LS)" (2023) Psychological Medicine, Cambridge University Press.

Emily Keddell "The ethics of predictive risk modelling in Aotearoa/New Zealand child welfare context: Child abuse prevention or neo-liberal tool?" (2004) Critical Social Policy.

Danny Keenan *Wars Without End: Ngā Pakanga Whenua O Mua New Zealand Land Wars - A Māori Perspective* (Penguin Random House, New Zealand, 2021).

Danielle Kehl, Priscilla Guo and Samuel Kessler *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing* (2017) at 13-14.

Wolfgang Kerber and Karsten K. Zolna "The German Facebook case: the law and economics of the relationship between competition and data protection law" (2022) European Journal of Law and Economics 54.

Ronald Kessler, Samantha L. Bernecker, Robert M. Bossarte, Alex R. Luedtke, John F. McCarthy, Matthew K. Nock, Wilfred R. Pigeon, Maria V. Petukhova, Ekaterina Sadikova, Tyler J. VanderWede, Kelly L. Zuromski, Alan M. Zaslavsky "The Role of Big Data Analytics in Predicting Suicide" in Passos I, Mwangi B, Kapczinski F (eds) *Personalized Psychiatry* (2019) Springer, Cham 77-98.

Diane Kirkby and Catherine Coleborne *Law, History, Colonialism: The reach of empire* (Manchester University Press, Manchester, 2001).

Tone Knapstad, "Breakups of Digital Gatekeepers under the Digital Markets Act: Three Strikes and You're out?" (2023) Journal of European Competition Law and Practice.

Tahu Kukutai and John Taylor (eds) *Indigenous Data Sovereignty: Toward An Agenda* (Australian National University Press, Canberra, 2016).

Julia Lane, Victoria Stodden, Stefan Bender and Helen Nissenbaum (eds) *Privacy, Big Data, and the Public Good: Frameworks for Engagement* (Cambridge University Press, New York, 2014).

Raeburn Lange *May the People Live: A History of Māori Health Development 1900-1920* (Auckland University Press, Auckland, 1999).

Eric B. Larson "Building Trust in the Power of 'Big Data' Research to Serve the Public Good" (2013) JAMA 309:23 2443-2444.

Johann Laux "Institutionalised distrust and human oversight of artificial intelligence: towards a democratic design of AI governance under the European Union AI Act" (2023) AI & Society at 2.

Johann Laux, Sandra Wachter and Brent Mittelstadt "Trustworthy artificial intelligence and the European Union AI Act: On conflation of trustworthiness and acceptability of risk" (2023) Regulation and Governance.

Law Commission *Review of the Privacy Act 1993: Review of the Law of Privacy Stage 4* (March 2010, Wellington).

David Lazer, Ryan Kennedy, Gary King, Alessandro Vespignani “The Parable of Google Flu: Traps in Big Data Analysis” (2014) 343 Science 1203-1205.

Choong Ho Lee and Hyung-Jin Yoon "Medical big data: promises and challenges" (2017) Kidney Research and Clinical Practice 36:3-11.

Newton Lee *Google It: Total Information Awareness* (Springer, Burbank USA, 2016).

Joy Liddicoat, Colin Gavaghan, Alistair Knott, James Maclaurin and John Zerilli “The use of algorithms in the New Zealand public sector” (2019) New Zealand Law Journal 26-30.

Mona Naomi Lintvedt "Putting a price on data protection infringement" (2022) International Data Privacy Law 12 1.

Wendy Lipworth, Paul H. Mason, Ian Kerridge and John P. A. Ioannidis “Ethics and Epistemology in Big Data Research” (2017) Bioethical Inquiry 14 489-500.

Clifford Lynch “How do your data grow?” (2008) 455 Nature 28-29.

Liz MacPherson, Deputy Privacy Commissioner, letter to Louise Wilsdon in response to Official Information Act request (4 July 2023).

Gianclaudio Malgieri "Automated decision-making in the EU Member States: The right to explanation and other "suitable safeguards" in the national legislations" (2019) Computer Law and Security Review 35.

Monique Mann and Tobias Matzner “Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination” (2019) Big Data and Society 1-11.

Luca Marelli, Elisa Lievevroux and Ine Van Hoyweghen “Fit for purpose? The GDPR and the governance of European digital health” (2020) Policy Studies 41:5 447-467.

Ken Mayhew and Samuel Willis "Inequality: an assessment" (2019) Oxford Review of Economic Policy 35 3.

Duncan McCann, *Digital Power Players: Power and Accountability Part 4: The Problem and Power of Tech Monopolies* (2018) New Economics Foundation.

Duncan McCann and Miranda Hall, *Blocking the Data Stalkers: Going Beyond the GDPR to Tackle Power in the Digital Economy* (2018, New Economics Foundation).

Aisling McMahon, Alena Buyx and Barbara Prainsack “Big Data Governance Needs More Collective Responsibility: The Role of Harm Mitigation in the Governance of Data Use in Medicine and Beyond” (2019) Medical Law Review.

John Stuart Mill *On Liberty* (Cambridge University Press, Cambridge, 2011).

Andrew D. Mitchell, Dominic Let, Lingxi Tang "AI Regulation and the Protection of Source Code" (2023) International Journal of Law and Information Technology 31 4.

Brent Daniel Mittelstadt and Luciano Floridi “The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts” (2016) 22 Sci Eng Ethics 303-341 at 320.

Kathryn Montgomery, Jeff Chester and Katharina Kapp “Health Wearables: Ensuring Fairness, Preventing Discrimination, and Promoting Equity in an Emerging Internet of Things Environment” (2018) Journal of Information Policy 8 37-77.

Evgeny Morozov *To Save Everything, Click Here: The Folly of Technological Solutionism* (Perseus Books Group, United States of America, 2013).

M.J. Mourby "'Leading by Science' through Covid-19: the NHS Data Store and Automated Decision-Making" (24 February 2021) International Journal of Population Data Science 5 2 03.

John Paul Mueller and Luca Massaron *Algorithms for dummies* (John Wiley & Sons, Inc., New Jersey, 2017).

New Zealand Government "Government Response to Law Commission Report on *Review of Privacy Act 1993* (2011)".

New Zealand Law Commission *A Conceptual Approach to Privacy* (NZLC, Wellington, 2007).

New Zealand Law Commission *Review of the Privacy Act 1993: Review of the Law of Privacy Stage 4* (NZLC, Wellington, 2011).

New Zealand Police "New Zealand Police Expert Panel on Emergent Technologies Terms of Reference" (May 2021).

W. Nicholson Price II and I. Glen Cohen "Privacy in the Age of Medical Big Data" (2020) Nature Medicine 25(1).

Helen Nissenbaum *Privacy In Context: Technology, Policy, and the Integrity of Social Life* (Stanford University Press, Stanford, California, 2010).

Safiya Umoja Noble "Algorithms of Oppression: How Search Engines Reinforce Racism" (New York University Press, New York, 2018).

OECD *The OECD Privacy Framework* (2013).

Paul Ohm "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization" (2010) UCLA Law Review

Cathy O'Neill *Weapons Of Math Destruction: How Big Data Increases Inequality And Threatens Democracy* (Penguin Random House, United Kingdom, 2016).

I. van Ooijen and Helena U. Vrabec "Does the GDPR Enhance Consumers' Control Over Personal Data? An Analysis from a Behavioural Perspective" (2019) Journal of Consumer Policy 42:1 91-107.

Organisation For Economic Co-operation And Development *The OECD Privacy Framework* (2013, OECD Publishing)

Kirsten Ostherr, Svetlana Bordina, Rachel Conrad Bracken, Charles Lotterman, Eliot Storer and Brandon Williams "Trust and Privacy in the Context of User-Generated Health Data" (2017) Big Data and Society.

Marion Oswald, Jamie Grace, Sheena Urwin and Geoffrey C. Barnes "Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality" (2018) Information & Communications Technology Law 27 2 223-250.

Michaela Padden and Andreas Öjehag-Pettersson "Protected how? Problem representations of risk in eth General Data Protection Regulation (GDPR)" (2021) Critical Policy Studies.

Hon Geoffrey Palmer, Minister of Justice *A Bill of Rights for New Zealand: A White Paper* (Wellington, 1985).

Abhishek Pandey and Katherine E. Atkins et al "Strategies for Containing Ebola in West Africa" (2014) 346 Science 6212 991-995.

Raja Parasuraman and Dietrich H. Manzey "Complacency and Bias in Human Use of Automation: An Attentional Integration" (2010) Human Factors 52 3.

Frank Pasquale “Redescribing Health Privacy: The Importance of Information Policy” (2014) Houston Journal of Health Law and Policy 14 (95).

Frank Pasquale *The Black Box Society* (Harvard University Press, London, 2015).

Ives Cavalcante Passos, Benson Mwangi, Flavio Kapczinski “Big data analytics and machine learning: 2015 and beyond” (2016) 3 Lancet 13-15.

David Peloquin, Michael DiMaio, Barbara Bierer and Mark Barnes “Disruptive and avoidable: GDPR challenges to secondary research uses of data” (2020) European Journal of Human Genetics 28: 697-705.

Steven Penk and Rosemary Tobin (eds) *Privacy Law In New Zealand* (Brookers Ltd, Wellington, 2010).

Scott Peppet “Regulating the Internet of Things: First Steps Toward Managing Discrimination, Privacy, Security and Consent” (2014) Texas Law Review 93 (1).

Christian Peukert, Stefan Bechtold, Michail Batikaas, Tobias Kretschmer "Regulatory Spillovers and Data Governance: Evidence from the GDPR" (2022) Marketing Science 41 4.

Marie V. Plaisime "Invited Commentary: Undiagnosed and Undertreated - the Suffocating Consequences of the Use of Racially Biased Medical Devices During the COVID-19 Pandemic" (2023) American Journal of Epidemiology 192 5 at 715.

Julia Powles and Hal Hodson “Google DeepMind and healthcare in an age of algorithms” (2017) Health Technology 7 351-367.

Jeffrey J. Rachlinski "Bottom-up versus Top-down Lawmaking" (2006) The University of Chicago Law Review 73 3.

Wullianallur Raghupathi and Viju Raghupathi “Big Data Analytics in Healthcare: Promise and Potential” (2014) Health Information Science and Systems 2:3.

John Rawls *A Theory of Justice* revised edition (Oxford, Oxford University Press, 1999).

Mark S. Rea and Andrew Bierman "Light source spectra are the likely cause of systemic bias in pulse oximeter readings for individuals with darker skin pigmentation" (2023) British Journal of Anaesthesia" 131 4.

Elizabeth M. Renieris *Beyond Data: Reclaiming Human Rights at the Dawn of the Metaverse* (The MIT Press, Cambridge, Massachusetts, 2023).

Annette Rid and Ezekiel J. Emmanuel “Why Should High-Income Countries Help Combat Ebola?” (2014) 312 JAMA 13 1297-1298.

Huw Roberts, Alexander Babuta, Jessica Morley, Christopher Thomas, Mariarosaria Taddeo, Luciano Floridi "Artificial intelligence regulation in the United Kingdom: a path to good governance and global leadership" (2023) Internet Policy Review 12 2.

Paul Roth and Blair Stuart *Roth's Companion to the Privacy Act 2020* (LexisNexis NZ Limited, Wellington, 2021).

Mark A. Rothstein “Ethical Issues in Big Data Health Research” (Summer 2015) Journal of Law, Medicine and Ethics 425-429.

Kathryn Rough and John T. Thompson "When Does Size Matter? Promises, Pitfalls, and Appropriate Interpretation of "Big" Medical Records Data" (2018) American Academy of Opthalmology at 1137.

Royal Free London NHS Foundation Trust and Google UK Limited “Information Sharing Agreement” (29 September 2015).

Omar M. Saad, Yungfeng Chen, Alexandros Savvaidis, Sergey Fommel, XiuXuan Jiang, Dino Huang, Yapo Abolé Innocent Oboulé, Shanshan Yong, Xinian Wang, Xing Zhang, Yangkang Chen "Earthquake Forecasting using Big Data and Artificial Intelligence: A 30-Week Real-Time Case Study in China" (2023) Bulletin of the Seismological Society of America 113 6.

Mattie Salim, Erilk Wahlin, Karen Dembrower, Edward Azavedo, Theodoros Foukakis, Yue Liu, Kevin Smith, Martin Eklund and Fredrik Strand "External Evalution of 3 Commercial Artificial Intelligence Algorithms for Independent Assessment of Screening Mammograms" (2020) JAMA Oncology 6(10).

Marijn Sax “Privacy from an Ethical Perspective” in B. Van der Sloot and A. De Groot (eds) *The Handbook of Privacy Studies: An Interdisciplinary Introduction* (Amsterdam University Press, 2018, Amsterdam).

Giulia Schneider “Disentangling health data networks: a critical analysis of articles 9(2) and 89 GDPR” (2019) International Data Privacy Law 9:4 253-271.

Daniel Schönberger "Artificial intelligence in healthcare: a critical analysis of the legal and ethical implications" (2019) International Journal of Law and Information Technology 27 171-203.

Ioannis Sechopoulos and Ritse M. Mann "Stand-alone artificial intelligence - The future of breast cancer screening?" (2021) Special issue: Artificial Intelligence in Breast Cancer Care Science Direct 56.

Matthias Schoror “Regulating to Support Privacy Disclosures: The First Step Towards Avoiding an Internet of Things Dystopia” (2018) LLB Honours dissertation, University of Otago.

Nayha Sethi and Graeme T Laurie “Delivering proportionate governance in the era of eHealth: making linkage and privacy work together” (2013) Medical Law International 13(2-3) 168-204.

Anna Sexton, Elizabeth Shepherd, Olive Duke-Williams and Alexandra Everleigh “A balance of trust in the use of government administrative data” (2017) Arch Sci 17 305-330.

Hetan Shah “The DeepMind debacle demands dialogue on data” (2017) Nature 547 259.

Abhinav Sharma et al. “Using Health Technology to Better Generate Evidence and Deliver Evidence-Based Care” (2018) Journal of the American College of Cardiology 71 (23).

Daniel Solove "Artificial Intelligence and Privacy" Florida Law Review (forthcoming Jan 2025) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4713111.

Daniel Solove "Conceptualizing Privacy" (2002) California Law Review 90 4.

Daniel Solove *Nothing to Hide: The False Tradeoff Between Privacy and Security* (Yale University Press, New Haven, 2011).

Daniel Solove *The Digital Person* (New York University Press, New York, 2004).

Sonja Starr "Evidence-Based Sentencing and the Scientific Rationalization of Discrimination". (2014) Stan. L. Rev. 66:4 at 863.

Stats NZ *Algorithm Assessment Report* (2018).

Stats NZ *Algorithm Charter for Aotearoa New Zealand* (July 2022).

Henri-Corto Stoeklé, Marie-France Mamzer-Bruneel, Guillaume Vogt, Christian Hervé “23andMe: a new two-sided data-banking market model” (2016) BMC Medical Ethics 17:19.

Alexander Szalay and Jim Gray “Science in an Exponential World” (2006) 440 Nature 413-414.

Graham Taylor and Paul Roth *Access to Information* (LexisNexis NZ Limited, Wellington, 2011).

Taylor Fry *Algorithm Charter for New Zealand: Year 1 Review* (20 December 2021).

Mark Taylor “Information Governance as a Force for Good? Lessons to be Learnt from Care.Data” (2014) 11 SCRIPTed 1.

Nicolas P. Terry “Big Data Proxies and Health Privacy Exceptionalism” (2014) Health Matrix 24:65.

Rosemary Tobin, "The Common Law Tort of Invasion of Privacy in New Zealand" in Nikki Chamberlain and Stephen Penk (eds) *Privacy Law in New Zealand* (3rd ed.) (Thomson Reuters, Wellington, 2023).

Luca Tosoni "The right to object to automated individual decisions: resolving the ambiguity of Article 22(1) of the General Data Protection Regulation" (2021) International Data Privacy Law 11 2.

Rhema Vaithianathan, Tim Maloney, Nan Jiang, Irene De Haan, Clare Dale, Emily Putnam-Hornstein, Tim Dare "Vulnerable Children: Can Administrative Data be Used to Identify Children at Risk of Adverse Outcomes?" (Centre for Applied Research in Economics, University of Auckland, 2012).

Aysem Dike Vanberg “Informational Privacy Post GDPR – end of the road or the start of a long journey?” (2020) International Journal of Human Rights DOI:10.1080/13642987.2020.1789109.

Anton Vedder and Laurens Naudts "Accountability for the use of algorithms in a big data environment" (2017) International Review of Law, Computers and Technology 31:2, 206-224.

Carissa Véliz *Privacy Is Power: Why And How You Should Take Back Control Of Your Data* (Penguin Random House, London, 2020).

Venkat Venkatasubramanian *How Much Inequality Is Fair? Mathematical Principles of a Moral, Optimal, and Stable Capitalist Society* (Columbia University Press, New York, 2017)

Paraskevas Vezyridis and Stephen Timmons "Understanding the care.data conundrum: New information flows for economic growth" (2017) Big Data & Society.

Salomé Viljoen "A Relational Theory of Data Governance" (2021) Yale Law Journal 131 2.

Paul Voigt and Axel von dem Bussche *The EU General Data Protection Regulation (GDPR) A Practical Guide* (Springer International Publishing, Cham, Switzerland, 2017).

Sandra Wachter, Brent Mittlestadt and Luciano Floridi "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation" (2017) International Data Privacy Law 7 2.

Sara Wachter-Boettcher *Technically Wrong: Sexist Apps, Biased Algorithms, And Other Threats Of Toxic Tech* (W.W. Norton & Company, New York, 2017).

Ari Ezra Waldman *Privacy as Trust: Information Privacy for an Information Age* (Cambridge University Press, New York, 2018).

Ari Ezra Waldman *Industry Unbound* (Cambridge University Press, Cambridge, 2021).

Maggie Walter, Tahu Kukutai, Stephanie Russo Carroll and Desi Rodriguez-Lonebear (eds) *Indigenous Data Sovereignty And Policy* (Routledge, Abingdon Oxon, 2021).

Samuel D. Warren and Louis D. Brandeis “The Right To Privacy” (1890) Harvard Law Review 4 5.

Martin A. Weiss and Kristin Archick *U.S.-EU data Privacy: From Safe Harbor to Privacy Shield* (Congressional Research Service, United States Congress, 2016).

Amy Wesolowski, Nathan Eagle, Abdisalan Noor, Robert Snow, Caroline Buckee "Heterogenous Mobile Phone Ownership and Usage Patterns. in Kenya" (2012) PLoS ONE 7(4).

Amy Wesolowski, Nathan Eagle, Andrew J. Tatem, David L. Smith, Abdisalan M. Noor, Robert W. Snow and Caroline O. Buckee "Quantifying the impact of human mobility on malaria" (2012) 338 Science 267-270.

Kiri West, Maui Hudson and Tahu Kukutai "Data Ethics and Data Governance from a Māori World View" in Lily George, Juan Tauri and Lindsey Te Ata o Tu MacDonald (eds) *Indigenous Research Ethics: Claiming Research Sovereignty Beyond Deficit and the Colonial Legacy* (Emerald Publishing Limited, Bingley UK, 2020).

Richard Wilkinson and Kate Pickett *The Spirit Level: Why Equality is Better for Everyone* (Penguin Books, London, 2009).

Denise Wilson and Melinda Webber "The People's Report: The People's Inquiry into Addressing Child Abuse and Domestic Violence" (2014) The Glenn Inquiry.

Hannah Wise and David B. Solit "Precision Oncology: Three Small Steps Forward" (2019) 35 Cancer Cell 825-826.

C.Wu, Y.Zhang and S. Nie et al "Predicting in-hospital outcomes of patients with acute kidney injury" (2023) Nature Communications 14.

Karen Yeung and Lee A. Bygrave "Demystifying the modernized European data protection regime: Cross-disciplinary insights from legal and regulatory governance scholarship" (2021) Regulation and Governance.

Oscar A. Zarate, Julia Green Brody, Phil Brown, Mónica D. Ramírez-Andreotta, Laura Perovich and Jacob Matz “Balancing Benefits and Risks of Immortal Data: Participants’ Views of Open Consent in the Personal Genome Project” (2016) Hastings Centre Report.

John Zerilli, John Danaher, Colin Gavaghan, Alistair Knott, Joy Liddicoat and Merel Noorman *A Citizen's Guide To Artificial Intelligence* (Massachussetts Institute of Technology, U.S.A, 2021).

John Zerilli, Alistair Knott, John Maclaurin, Colin Gavaghan "Algorithmic Decision-Making and the Control Problem" (2019) Minds and Machines 29.

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

C. Internet Materials

ACC Media release "Investigation into sharing client information has been completed" (30 November 2021) https://www.acc.co.nz/newsroom/stories/investigation-into-sharing-client-information-has-been-completed

Advisory Board, Daily Briefing "10 years ago, IBM's Watson threatened to disrupt healthcare. What happened?" (July 2021) advisory.com https://www.advisory.com/daily-briefing/2021/07/21/ibm-watson

AI Forum New Zealand *Trustworthy AI in Aotearoa New Zealand* https://aiforum.org.nz/wp-content/uploads/2020/03/Trustworthy-AI-in-Aotearoa-March-2020.pdf

Airdoc https://world.airdoc.com/#/

Te Aka Māori www.maoridictionary.co.nz

Allegheny County "Allegheny Family Screening Tool: Predictive Risk Modeling in Child Welfare in Allegheny County" https://www.alleghenycounty.us/Human-Services/News-Events/Accomplishments/Allegheny-Family-Screening-Tool.aspx

Chris Anderson “The end of theory: the data deluge makes the scientific method obsolete” (2008) www.wired.com/2008/06/pb-theory

Julia Angwin, Jeff Larson, Surya Mattu, Lauren Kirchner "Machine Bias: There's software used across the country to predict future criminals and its's biased against blacks" (2016) Propublica https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

The Annie E. Casey Foundation "Juvenile Detention Risk Assessment: A Practice Guide to Juvenile Detention Reform" (2006) https://assets.aecf.org/m/resourcedoc/aecf-juveniledetentionriskassessment1-2006.pdf

Article 29 Data Protection Working Party "Opinion 11/2011 on the level of protection of personal data in New Zealand" (4 April 2011) 00665/11/EN WP 182. https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2011/wp182_en.pdf

Article 29 Data Protection Working Party "Transfers of personal data to third countries: Applying Articles 25 and 26 of the EU data protection directive" (WP12, DG XV D/5025/98, adopted on 24 July 1998) https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/1998/wp12_en.pdf

Associated Press “Healthmap software flagged Ebola 9 days before outbreak announced” (10 August 2014) https://www.cbc.ca/news/health/healthmap-software-flagged-ebola-9-days-before-outbreak-announced-1.2732464

Associated Press “Online Tool Nailed Ebola Epidemic” (9 August 2014) https://www.politico.com/story/2014/08/healthmap-ebola-outbreak-109881

Auror www.aura.co

Frederic D. Bellamy "U.S. data privacy laws to enter new era in 2023" (13 January 2023) Westlaw Thompson Reuters https://www.reuters.com/legal/legalindustry/us-data-privacy-laws-enter-new-era-2023-2023-01-12/

Joe Biden "Republicans and Democrats Unite Against Big Tech Abuses" (11.1.23) Wall Street Journal https://www.wsj.com/articles/unite-against-big-tech-abuses-social-media-privacy-competition-antitrust-children-algorithm-11673439411

José Manuel Blanco "Big Data and Energy: A Combination for Success" (2022) https://www.plainconcepts.com/big-data-energy/

Jordan M. Blanke "Protection for 'Inferences Drawn': A Comparison Between the General Data Protection Regulation and the California Consumer Privacy Act" (2020) Global Privacy Law Review I 2 81-92 at 88.

Anusha Bradley "ACC staffer breached client privacy after snooping in sensitive claim - review authority" (10 March 2022) RadioNZ https://www.rnz.co.nz/news/national/463044/acc-staffer-breached-client-privacy-after-snooping-in-sensitive-claim-review-authority

Britannica "Edward Snowden" https://www.britannica.com/biography/Edward-Snowden; National Whistleblower Centre "Edward Snowden" https://www.whistleblowers.org/whistleblowers/edward-snowden/

Dennis Campbell "NHS England faces lawsuit over patient privacy fears linked to new data platform" The Guardian (30 November 2023) https://www.theguardian.com/society/2023/nov/30/nhs-england-faces-lawsuit-patient-privacy-fears-new-data-fdp-platform?CMP=Share_iOSApp_Other

Foo Yun Chee “Google, Amazon among those targeted in EU unfair practices digital rules” (14 February 2019) Reuters https://www.reuters.com/article/us-eu-tech/google-amazon-among-those-targeted-in-eu-unfair-practices-digital-rules-idUSKCN1Q30ZY

Child Poverty Action Group www.cpag.org.nz

Lindsay Clark "Under threat of judicial review, UK.gov agrees to consultation before extending Palantir's NHS role beyond pandemic" (30 March 2021) The Register https://www.theregister.com/2021/03/30/ukgov_caves_over_nhs_palantir_lawsuit/
Theodore F. Claypoole "U.S Senate Takes Aim at Big Tech" (2021) National Law Review XI:40; Makena Kelly "All The Ways Congress Is Taking On The Tech Industry" (3.3.20) The Verge https://www.theverge.com/2020/3/3/21153117/congress-tech-regulation-privacy-bill-coppa-ads-laws-legislators

Clearview AI www.clearview.ai

Consumer Watchdog "FTC targets Google for antitrust investigation" https://www.consumerwatchdog.org/blog/ftc-targets-google-antitrust-investigation

Sam Corbett-Davies, Emma Pierson, Avi Feller, Sharad Goel "A computer program used for bail and sentencing decisions was labeled biased against Blacks. It's actually not that clear" (17 October 2016) The Washington Post https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/

Corporate Europe Observatory "Big tech lobbying is derailing the AI Act" (24 November 2023) https://corporateeurope.org/en/2023/11/big-tech-lobbying-derailing-ai-act

Cori Crider "Why Palantir's latest NHS land-grab is such bad news for patients" openDemocracy (17 March 2023) https://www.opendemocracy.net/en/palantir-foundry-faster-data-flows-nhs-cori-crider-foxglove/

Council of the European Union, *Proposal for a Regulation of the European Parliament and of the Council: laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Analysis of final compromise text with a view to agreement* (Brussels, 26.1.2024 2021/0106 (COD)) https://data.consilium.europa.eu/doc/document/ST-5662-2024-INIT/en/pdf

Daily Briefing, Advisory Board "10 years ago, IBM's Watson threatened to disrupt healthcare. What happened?" (July 2021) advisory.com https://www.advisory.com/daily-briefing/2021/07/21/ibm-watson

DataBreaches.net "WINZ privacy breach 'major stuff up'" (18 May 2013) https://www.databreaches.net/winz-privacy-breach-major-stuff-up/

data.govt.nz www.data.govt.nz

Data Protection Commission, Ireland "Data Protection Commission announces conclusion of inquiry into Meta Ireland" (22 May 2023) https://www.dataprotection.ie/en/news-media/press-releases/Data-Protection-Commission-announces-conclusion-of-inquiry-into-Meta-Ireland

DeepMind www.deepmind.com

Deloitte "The economic benefits of improving social inclusion" (August 2019) https://www2.deloitte.com/content/dam/Deloitte/my/Documents/risk/my-risk-sdg10-economic-benefits-of-improving-social-inclusion.pdf

Department of Corrections "Risk of Reconviction" https://www.corrections.govt.nz/resources/research/risk-of-reconviction

Department of Corrections "Risk of Reconviction: The ROC*ROI Measures - Explanatory Note" https://www.corrections.govt.nz/resources/research/risk-of-reconviction#:~:text=Hence%20the%20term%20ROC*ROI,the%20offender's%20Risk%20of%20Imprisonment.

Digitalhealth.net "BMA and RCGP issue joint letter to NHS Digital over GPDPR programme" (7 June 2021) https://www.digitalhealth.net/2021/06/bma-and-rcgp-issue-joint-letter-to-nhs-digital-over-gpdpr-programme/

Cory Doctorow "How to Destroy Surveillance Capitalism" OneZero (26 August 2020) https://onezero.medium.com/how-to-destroy-surveillance-capitalism-8135e6744d59.

Julia Dressel and Hany Farid "The accuracy, fairness, and limits of predicting recidivism" (2018) Science Advances https://www.science.org/doi/epdf/10.1126/sciadv.aao5580

Electronic Information Privacy Centre Press Release "BREAKING: Top Court in Europe Invalidates EU-U.S. Privacy Shield, Citing Lack of Privacy Safeguards and Overbroad U.S. Surveillance Laws" https://epic.org/privacy/intl/dpc-v-facebook/cjeu/RELEASE-EPIC-CJEU-July2020.pdf.

Equality and Human Rights Commission "What is the Charter of Fundamental Rights of the European Union?" https://www.equalityhumanrights.com/en/what-are-human-rights/how-are-your-rights-protected/what-charter-fundamental-rights-european-union

Equivant https://www.equivant.com

Euractiv "EU probe: Google may have abused advertising market dominance" (15.6.23) https://www.euractiv.com/section/digital/news/eu-probe-google-may-have-abused-advertising-market-dominance/

European Commission "Commission fines Google €2.42 billion for abusing dominance as search engine by giving illegal advantage to own comparison shopping service" (27.6.17) https://ec.europa.eu/newsroom/comp/items/104946/#:~:text=The%20European%20Commission%20has%20fined,Statement%20by%20Commissioner%20Vestager.

European Commission COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT AND THE COUNCIL “Data protection as a pillar of citizens’ empowerment and the EU’s approach to the digital transition – two years of application of the General Data Protection Regulation” COM(2020) 264 final (24.6.2020) (SWD(2020) 115 final). https://ec.europa.eu/info/law/law-topic/data-protection/communication-two-years-application-general-data-protection-regulation_en

European Commission High-Level Expert Group on AI, *Ethics Guidelines for Trustworthy AI* (2019) https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

European Commission press release "European Commission adopts new adequacy decision for safe and trusted EU-US data Flows" (Brussels, 10 July 2023) https://ec.europa.eu/commission/presscorner/detail/en/ip_23_3721

European Commission *Prohibition Decision (Art.102 Ex 82)* (27 June 2017) https://ec.europa.eu/competition/elojade/isef/case_details.cfm?proc_code=1_39740

European Commission "Report From The Commission To The European Parliament And The Council on the first review of the functioning of the adequacy decisions adopted pursuant to Article 25(6) of Directive 95/46/EU (15.1.2024). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52024DC0007

European Commission *Shaping Europe's Digital Future - AI Act* https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai#:~:text=The%20AI%20Act%20is%20the,play%20a%20leading%20role%20globally.&text=The%20AI%20Act%20aims%20to,regarding%20specific%20uses%20of%20AI

European Data Protection Supervisor www.edps.europa.eu

European Parliament "Briefing: Artificial intelligence act" (June 2023) https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf at 2.

Müge Fazlioglu "U.S. privacy legislation in 2023: Something old, something new?" (26 July 2023) IAPP https://iapp.org/news/a/u-s-federal-privacy-legislation-in-2023-something-old-something-new/

Foodstuffs "Foodstuffs North Island begins trialling facial recognition in select stores as part of its commitment to keep teams and customers safe by keeping previous offenders out" (8 February 2024). https://www.foodstuffs.co.nz/news-room/2024/Foodstuffs-North-Island-begins-trialling-facial-recognition-in-select-stores

Foxglove www.foxgove.org.uk

fyi.org.nz "RoC*RoI formula and briefings https://fyi.org.nz/request/501/response/4002/attach/html/4/Attachment%20C59531.PDF.pdf.html

Graham Greenleaf "Australia's APEC privacy initiative: the pros and cons of 'OECD lite'" [2003] Privacy Law and Policy Reporter 1. http://www8.austlii.edu.au/cgi-bin/viewdoc/au/journals/PLPR/2003/17.html

Glenn Greenwald "Edward Snowden: the whistleblower behind the NSA surveillance revelations" *The Guardian* (online ed, United Kingdom, 11 June 2013) https://www.theguardian.com/world/2013/jun/09/edward-snowden-nsa-whistleblower-surveillance

Google www.about.google

Google "Our Approach to Search" https://www.google.com/search/howsearchworks/our-approach/

Benjamin Harris "AI's future in healthcare is not entirely rosy" Healthcare IT News (23 August 2019) Healthcare IT News https://www.healthcareitnews.com/news/ai-s-future-healthcare-not-entirely-rosy

Hal Hodson "Revealed Google AI has access to huge haul of NHS patient data" (29 April 2016) New Scientist https://www.newscientist.com/article/2086454-revealed-google-ai-has-access-to-huge-haul-of-nhs-patient-data/#ixzz5xrJRhk3a

Nancy Huynh “How the ‘Big 4’ Tech Companies Are Leading Healthcare Innovation” (27 February 2019) Healthcare Weekly https://healthcareweekly.com/how-the-big-4-tech-companies-are-leading-healthcare-innovation/

The Indigenous World Indigenous Data Sovereignty 2021 https://www.iwgia.org/en/ip-i-iw/4268-iw-2021-indigenous-data-sovereignty.html

Insider Intelligence "BIG TECH IN HEALTHCARE: Here's who wins and loses as Alphabet, Amazon, Apple, and Microsoft target niche sectors of healthcare" (24.1.23) https://www.insiderintelligence.com/insights/big-tech-in-healthcare-report/

Cecilia Kang "Democratic Congress Prepares to Take On Big Tech" (26.1.21) New York Times https://www.nytimes.com/2021/01/26/technology/congress-antitrust-tech.html

Danielle Kehl, Priscilla Guo and Samuel Kessler *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing* (2017) https://dash.harvard.edu/bitstream/handle/1/33746041/2017-07_responsivecommunities_2.pdf?sequence=1&isAllowed=y

Makena Kelly "All The Ways Congress Is Taking On The Tech Industry" (3.3.20) The Verge https://www.theverge.com/2020/3/3/21153117/congress-tech-regulation-privacy-bill-coppa-ads-laws-legislators

Kidney Care UK www.kidneycareuk.org

Stacey Kirk "Children - 'not lab rats'. Anne Tolley intervenes in child abuse experiment" (30 July 2015) Stuff https://www.stuff.co.nz/national/health/70647353/children-not-lab-rats---anne-tolley-intervenes-in-child-abuse-experiment

Iga Kozlowska "Facebook and Data Privacy in the Age of Cambridge Analytica" (30 April 2018) University of Washington https://jsis.washington.edu/news/facebook-data-privacy-age-cambridge-analytica/

Akos Lada "Facebook: How Does News Feed Predict What You Want To See?" (2021) Meta https://about.fb.com/news/2021/01/how-does-news-feed-predict-what-you-want-to-see/

Joao Laranjeira "What is traffic prediction and how does it work?" (2020) TomTom https://www.tomtom.com/newsroom/behind-the-map/road-traffic-prediction/
Paul Lewis, David Comm and David Pegg "UK government using confidential patient data in coronavirus response" (12.4.20) The Guardian https://www.theguardian.com/world/2020/apr/12/uk-government-using-confidential-patient-data-in-coronavirus-response

The Legatum Prosperity Index 2023 "The Foundational Elements of Prosperity" https://www.prosperity.com/feed/foundational-elements-prosperity

Natasha Lomas "Google completes controversial takeover of DeepMind Health" (20 September 2019) Techcrunch https://techcrunch.com/2019/09/19/google-completes-controversial-takeover-of-deepmind-health/

Nessa Lynch and Andrew Chen "Facial Recognition Technology: Considerations for Use in Policing" (2021) commissioned by NZ Police https://www.police.govt.nz/sites/default/files/publications/facial-recognition-technology-considerations-for-use-policing.pdf

Vincent Manacourt "Oliver Dowden's 'hit squad' aims to replace UK civil service jobs with AI" (20.11.23) Politico https://www.politico.eu/article/dowdens-hit-squad-aims-to-replace-civil-service-jobs-with-ai/

Te Mana Raraunga: Māori Data Sovereignty Network https://www.temanararaunga.maori.nz

Kartikay Mehrotra, Laura Mahoney and Daniel Stoller "Google and other tech firms seeks to weaken landmark California data-privacy law" (4.9.19) Los Angeles Times https://www.latimes.com/business/story/2019-09-04/google-and-other-tech-companies-attempt-to-water-down-privacy-law

Meta "Our response to the decision on Facebook's EU-US data transfers" (22.5.23) https://about.fb.com/news/2023/05/our-response-to-the-decision-on-facebooks-eu-us-data-transfers/

Ministry of Business, Immigration and Employment "Consumer data right" https://www.mbie.govt.nz/business-and-employment/business/competition-regulation-and-policy/consumer-data-right

Ministry of Health "Tū Ora Compass Health cyber security incident - further information" (26 November 2019) https://www.health.govt.nz/about-ministry/information-releases/general-information-releases/tu-ora-compass-health-cyber-security-incident-further-information

Ministry of Justice "Constitutional issues and human rights: International Covenant on Civil and Political Rights" https://www.justice.govt.nz/justice-sector-policy/constitutional-issues-and-human-rights/human-rights/international-human-rights/international-covenant-on-civil-and-political-rights/

Ministry of Justice "Key initiatives: Broadening the Privacy Act's notification rules" https://www.justice.govt.nz/justice-sector-policy/key-initiatives/broadening-the-privacy-acts-notification-rules/

Ministry of Justice "Key initiatives: Political lobbying" https://www.justice.govt.nz/justice-sector-policy/key-initiatives/political-lobbying/

Ministry of Social Development "Identity checks and online verification" https://www.msd.govt.nz/about-msd-and-our-work/work-programmes/identity-check/identity-checks-and-online-verification.html

Ministry of Social Development "Privacy Human Rights and Ethics Framework" https://www.data.govt.nz/assets/data-ethics/algorithm/phrae-on-a-page.pdf

Ministry of Social Development "White Paper for Vulnerable Children" (2012) https://www.beehive.govt.nz/feature/white-paper-vulnerable-children

Michael Neilson "Te Wiki o Te Reo Māori: Beaten for speaking their native tongue, and the generations that suffered" NZ Herald (14.9.20) https://www.nzherald.co.nz/nz/te-wiki-o-te-reo-maori-beaten-for-speaking-their-native-tongue-and-the-generations-that-suffered/F7G6XCM62QAHTYVSRVOCRKAUYI

Newshub "Why cutting public service roles to save costs normally backfires, creates 'consultocracy' instead" (12 December 2023) https://www.newshub.co.nz/home/politics/2023/12/analysis-why-cutting-public-service-roles-to-save-costs-normally-backfires-creates-consultocracy-instead.html

New Zealand Government "Coalition Government 100-day plan" https://www.beehive.govt.nz/sites/default/files/2023-12/100%20Day%20Plan%20%281%29.pdf

New Zealand Government, *Data Ethics Advisory Group Terms of Reference* v2 (June 2023) at 4. https://data.govt.nz/assets/Uploads/Data-Ethics-Advisory-Group/Terms-of-Reference-DEAG.pdf

New Zealand Herald "IRD says sorry for privacy breach" (28 September 2021) https://www.nzherald.co.nz/business/ird-says-sorry-for-privacy-breach/54MDLADNVTX2HHIEKPI3URJBEQ/

New Zealand Herald "Predicting Trouble: Child abuse database raises eyebrows" (20 October 2012) https://www.nzherald.co.nz/nz/predicting-trouble-child-abuse-database-raises-eyebrows/UXLTJHIEKWAF5JDSZA7MPAHKQM/?c_id=1&objectid=10841709

New Zealand Parliament, Economic Development, Science and Innovation Committee *Cabinet Economic Development Committee Minute of Decision* (26 July 2023) DEV-23-MIN-0157

New Zealand Police Assurance Group "Audit of New Zealand Police's use of Automatic Number Plate Recognition Technology (ANPR) Platforms" (December 2022) at 1. https://www.police.govt.nz/sites/default/files/publications/police-use-anpr-platforms-audit-report.pdf

New Zealand Police "Mid-term health check of the expert panel" https://www.police.govt.nz/about-us/programmes-and-initiatives/police-use-emergent-technologies/expert-panel-emergent

New Zealand Police "New Technology Framework" https://www.police.govt.nz/sites/default/files/publications/new-technology-framework.pdf

New Zealand Police "Police release ANPR audit findings" (26 April 2023) https://www.police.govt.nz/news/release/police-release-automatic-number-plate-recognition-audit-findings#:~:text="It%20involved%20examining%20and%20cross,to%20use%20ANPR%20data%20responsibly.

New Zealand Police *We Asked: Expert Panel on Emergent Technology: Proposal to Trial Zavy* https://www.police.govt.nz/sites/default/files/publications/zavy-proposal-we-asked-they-said-we-did.pdf;

New Zealand Parole Board "Decisions" https://www.paroleboard.govt.nz/decisions

NHS Digital www.digital.nhs.uk

NHS England www.england.nhs.uk

NHSX "The COVID-19 Data Store: putting data at the centre of decision making" https://www.nhsx.nhs.uk/key-tools-and-info/data-saves-lives/improving-health-and-care-services-for-everyone/the-nhs-covid-19-data-store-putting-data-at-the-centre-of-decision-making/

NZ Trade and Enterprise "New Zealand's best tech companies are taking on the world" (25 June 2023) https://www.nzte.govt.nz/blog/new-zealands-best-tech-companies-are-taking-on-the-world

OECD Legal Instruments "Recommendation of the Council on Artificial Intelligence" (amended on 3 May 2024) OECD/LEGAL/0449 https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

Office of the Privacy Commissioner *A potential biometrics code of practice: discussion document' - summary of submissions* https://www.privacy.org.nz/assets/New-order/Resources-/Publications/Guidance-resources/Biometrics/Biometrics-November-2023/Summary-of-submissions-on-OPC-discussion-document.pdf

Office of the Privacy Commissioner *Briefing to the Incoming Minister of Justice* (4 December 2023) at 1 https://www.privacy.org.nz/assets/New-order/Resources-/Publications/Reports-to-Parliament-and-Government-/BIMs/BIM.pdf

Office of the Privacy Commissioner *Exposure draft of a biometric processing code of practice: consultation paper* (April 2024) https://www.privacy.org.nz/news/consultations/biometrics/

Office of the Privacy Commissioner *Position Paper on the Regulation of Biometrics* https://www.privacy.org.nz/assets/New-order/Resources-/Publications/Guidance-resources/2021-10-07-OPC-position-on-biometrics.pdf

Office of the Privacy Commissioner *Principles for the Safe and Effective Use of Data and Analytics* (218) https://www.privacy.org.nz/assets/New-order/Resources-/Publications/Guidance-resources/Principles-for-the-safe-and-effective-use-of-data-and-analytics-guidance3.pdf

Office of the Privacy Commissioner "Privacy Commissioner to keep a close eye on Foodstuffs' North Island FRT trial" (8 February 2024) https://www.privacy.org.nz/publications/statements-media-releases/privacy-commissioner-to-keep-a-close-eye-on-foodstuffs-north-island-frt-trial/

Office of the Privacy Commissioner www.privacy.org.nz

Oxford Big Data Institute “Big Data: challenges and opportunities for human health” (30 November 2012) https://www.bdi.ox.ac.uk/news/newsitem-3

Phil Pennington "Audit reveals new facial recognition tech tools in police's digital armoury" (5 November 2020) NZ Herald https://www.nzherald.co.nz/nz/audit-reveals-new-facial-recognition-tech-tools-in-polices-digital-armoury/FR7VXHHGE4QUBFQKJ5IRXYJDJU/

Phil Pennington "'Data is the new gold' - Warning New Zealand at risk with reliance on foreign firms" (4 July 2023) RadioNZ https://www.rnz.co.nz/news/business/493161/data-is-the-new-gold-warning-nz-at-risk-with-reliance-on-foreign-firms

Phil Pennington, "Facial recognition: Government rolls out new tech despite racial bias concerns" (15 November 2023) RNZ https://www.rnz.co.nz/news/national/502445/facial-recognition-government-rolls-out-new-tech-despite-racial-bias-concerns#:~:text=The%20government%20begins%20the%20roll,of%20Internal%20Affairs%20(DIA).

Phil Pennington "Head in the clouds? Call for NZ to take control of data storage" (1.8.22) RadioNZ https://www.rnz.co.nz/news/national/471967/head-in-the-clouds-call-for-nz-to-take-control-of-data-storage

Phil Pennington "Legal challenges to Police use of automated number plate recognition cameras" (26 October 2023) RadioNZ https://www.rnz.co.nz/news/national/501012/legal-challenges-to-police-use-of-automated-number-plate-recognition-cameras

Phil Pennington "Police drop technology designed to predict motorists? complete" RadioNZ (15 July 2021) https://www.rnz.co.nz/news/national/446957/police-drop-technology-designed-to-predict-motorists

Phil Pennington "Police made false report to use ANPR cameras to track women who triggered Northland lockdown" (28 September 2022) RadioNZ https://www.rnz.co.nz/news/national/475662/police-made-false-report-to-use-anpr-cameras-to-track-women-who-triggered-northland-lockdown#:~:text=28%20Sep%202022-,Police%20made%20false%20report%20to%20use%20ANPR%20cameras,women%20who%20triggered%20Northland%20lockdown&text=Just%20one%20month%20after%20the,so%20they%20could%20track%20it.

Phil Pennington, "Police step up surveillance activity, tap into CCTV footage from other businesses" (23 September 2022) Radio NZ https://www.rnz.co.nz/news/national/475342/police-step-up-surveillance-activity-tap-into-cctv-footage-from-other-businesses

Phil Pennington, "New police visual search tech not assessed for privacy impacts" (4 October 2023) RadioNZ https://www.rnz.co.nz/news/national/499415/new-police-visual-search-tech-not-assessed-for-privacy-impacts

Phil Pennington "Revealed: Amazon's efforts to work its way into the NZ healthcare system" (6 September 2023) RadioNZ https://www.rnz.co.nz/news/national/497352/revealed-amazon-s-efforts-to-work-its-way-into-the-nz-healthcare-system

Predpol www.predpol.com

Pulitzer www.pulitzer.org

Ben Quinn, “Google given access to healthcare data of up to 1.6 million patients” *The Guardian* (international edition) 4 May 2016

https://www.theguardian.com/technology/2016/may/04/google-deepmind-access-healthcare-data-patients

RadioNZ "Independent panel to advise on police use of new technology" (11 March 2021) https://www.rnz.co.nz/news/national/438147/independent-panel-to-advise-police-on-use-of-emerging-technology; https://www.police.govt.nz/about-us/programmes-and-initiatives/police-use-emergent-technologies/advisory-panel-emergent

RadioNZ "IRD privacy breach raises data handling concerns" (24 October 2018) https://www.rnz.co.nz/national/programmes/ninetonoon/audio/2018668093/ird-privacy-breach-raises-data-handling-concerns

RadioNZ "Police investigating staff misuse of number plate software" (27 April 2023) https://www.rnz.co.nz/news/national/488750/police-investigating-staff-misuse-of-number-plate-software

Reckon "Ancient Egyptians - the original bookkeepers" (2018) https://www.reckon.com/reckon-blog/ancient-egyptians-the-original-bookkeepers1/

Dillon Reisman, Jason Schultz, Kate Crawford, Meredith Whittaker "Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability" (2018) AI Now https://ainowinstitute.org/aiareport2018.pdf

Matt Reynolds "DeepMind's new AI predicts kidney injury two days before it happens" (2019) Wired https://www.wired.com/story/deepmind-streams-ai-algorithm-kidney-injury

Matt Reynolds “If you can’t build it, buy it: Google’s biggest acquisitions mapped” (25 November 2017) Wired https://www.wired.co.uk/article/google-acquisitions-data-visualisation-infoporn-waze-youtube-android

Casey Ross and Ike Swetlitz "IBM pitched its Watson supercomputer as a revolution in cancer care. It's nowhere close" (2017) STAT news https://www.statnews.com/2017/09/05/watson-ibm-cancer/

Royal Free NHS Trust www.royalfree.nhs.uk

Jathan Sadowski “How ‘Smart Tech’ Masks an Emerging Era of Corporate Control” (9 March 2020) OneZero https://onezero.medium.com/how-smart-tech-masks-an-emerging-era-of-corporate-control-779c96b05f85

SaferCities www.safercities.com

Dan Satherley "Work and Income accidentally reveals more than 100 clients' personal info to each other" (7 August 2021) https://www.newshub.co.nz/home/new-zealand/2021/08/work-and-income-accidentally-reveals-more-than-100-clients-personal-info-to-each-other.html

Sheppard Mullin Eye on Privacy "Maryland, the Old Line State, Creates New Lines with Consumer Privacy Law" (20 May 2024) https://www.eyeonprivacy.com/2024/05/maryland-the-old-line-state-creates-new-lines-with-consumer-privacy-law/

Sheppard Mullin "U.S. State Comprehensive Privacy Laws" https://discover.sheppardmullin.com/us-state-comprehensive-privacy-laws/p/1

Mackenzie Smith "Police 'stocktake' surveillance tech after Clearview AI facial recognition trial" (18 May 2020) RadioNZ https://www.rnz.co.nz/news/national/416913/police-stocktake-surveillance-tech-after-clearview-ai-facial-recognition-trial

Geoff Spencer “AI and preventative healthcare: Diagnosis in the blink of an eye” (17 September 2018) Microsoft Asia News Centre https://news.microsoft.com/apac/features/ai-and-preventative-healthcare-diagnosis-in-the-blink-of-an-eye/

StatCounter Global Stats "Search engine market share" (August 2019) https://gs.statcounter.com/search-engine-market-share

State of California Department of Justice, Office of Attorney-General "California Consumer Privacy Act (CCPA)" https://oag.ca.gov/privacy/ccpa

John Stephens "California Consumer Privacy Act" (American Bar Association, 2019) https://www.americanbar.org/groups/business_law/publications/committee_newsletters/bcl/2019/201902/fa_9/

Eliza Strickland "How IBM Watson Overpromised and Underdelivered on AI Healthcare" (2019) IEEE Spectrum https://spectrum.ieee.org/how-ibm-watson-overpromised-and-underdelivered-on-ai-health-care#toggle-gdpr

stuff.co.nz "A child should never be punished for speaking the language of their people" (29.9.21) https://www.stuff.co.nz/national/politics/opinion/300416107/a-child-should-never-be-punished-for-speaking-the-language-of-their-people

Mustafa Suleyman and Dominic King “Using AI to give doctors a 48-hour head start on life-threatening illness” (31 July 2019) DeepMind https://deepmind.com/blog/article/predicting-patient-deterioration

Ming Tang, NHS blog "Data integration - driving improvements in patient care" (18 December 2020) https://www.england.nhs.uk/blog/data-integration-driving-improvements-in-patient-care/

Te Ara: encyclopaedia of New Zealand www.teara.govt.nz

Techstack "Blog: Implementing Big Data Solutions for Transforming the Renewable Energy Sector: Techstack Case Included" (2023) https://tech-stack.com/blog/implementing-big-data-solutions-for-transforming-the-renewable-energy-sector-techstack-case-included/

TechTarget "3V's (volume, velocity, and variety)" https://www.techtarget.com/whatis/definition/3Vs#:~:text=The%203%20V's%20(volume%2C%20velocity,number%20of%20types%20of%20data.

UK Department for Science, Innovation and Technology "A pro-innovation response to AI regulation" (March 2023) https://assets.publishing.service.gov.uk/media/64cb71a547915a00142a91c4/a-pro-innovation-approach-to-ai-regulation-amended-web-ready.pdf

UK Department for Science, Innovation and Technology "Consultation Outcome: A pro-innovation approach to AI regulation: government response" (6 February 2024) https://www.gov.uk/government/consultations/ai-regulation-a-pro-innovation-approach-policy-proposals/outcome/a-pro-innovation-approach-to-ai-regulation-government-response

UK Government "A pro-innovation approach to AI regulation" (2023) https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper

United Nations Conference on Trade and Development https://unctad.org/page/data-protection-and-privacy-legislation-worldwide

University of Bristol *Avon Longitudinal Study of Parents and Children* https://www.bristol.ac.uk/alspac/

University of Michigan Institute for Healthcare Policy and Innovation "An AI model predicting acute kidney injury works, but not without some tweaking" (20 January 2023) https://ihpi.umich.edu/news/ai-model-predicting-acute-kidney-injury-works-not-without-some-tweaking

University of Oxford Big Data Institute "Big Data: challenges and opportunities for human health" Oxford-Stanford Conference on Big Data, November 2012 https://www.bdi.ox.ac.uk/news/newsitem-3

University of Oxford https://catalogofbias.org/biases/confounding-by-indication/

The Verge "Everything you need to know about PRISM" https://www.theverge.com/2013/7/17/4517480/nsa-spying-prism-surveillance-cheat-sheet

Zack Whittaker "Silicon Valley is terrified of California's Privacy law. Good" (20.9.19) TechCrunch https://techcrunch.com/2019/09/19/silicon-valley-terrified-california-privacy-law/

Mike White, "Huge increase in police drone use: Sensible solution or spies in the sky?" (6 September 2023) Dominion Post https://www.thepost.co.nz/nz-news/350066875/huge-increase-police-drone-use-sensible-solution-or-spies-sky

Anna Whyte "More job cuts at Housing Ministry, StatsNZ, DIA" The Post (2 May 2024) https://www.thepost.co.nz/politics/350265201/more-job-cuts-housing-ministry-stats-nz-dia

Laura Wiltshire "Qualifications authority running trials into automated NCEA essay marking" (24 February 2021) Stuff https://www.stuff.co.nz/national/education/124328937/qualifications-authority-running-trials-into-automated-ncea-essay-marking

WIRED Magazine "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete" (2008) https://www.wired.com/2008/06/pb-theory/

World Health Organisation www.who.int

World Health Organisation "Health topics: Ebola" https://www.who.int/health-topics/ebola/#tab=tab_1

World Health Organisation “Statement on the 1st Meeting of the IHR Emergency Committee on the 2014 Ebola Outbreak in West Africa” (8 August 2014) https://www.who.int/mediacentre/news/statements/2014/ebola-20140808/en/

YouGov YouGov.co.uk