Aotearoa New Zealand: Empowered by data

Long term insights about the role of data in wellbeing and economic advantage

Te Kāwanatanga o Aotearoa
New Zealand Government

Contents

1. What is a Long-term Insights Briefing?

Public services agencies have been asked to prepare Long-term Insights Briefings as part of their duty of stewardship under the Public Service Act 2020 (the Act). These briefings provide a platform for identifying and exploring the issues that matter for the future wellbeing of the people of New Zealand. The purpose of this LTIB is to make available into the public domain:

- information about medium and long-term trends, risks and opportunities that affect or may affect New Zealand and New Zealand society; and

- information, including possible policy options, for responding to these matters.

2. Introduction

Data has the power to change lives. It tells a story about us – how well we live, how well we take care of each other, how government services can be improved, and how our environment looks at any given time. Data also helps to grow our economy – supporting the design and development of new products, services, and business models, and the creation of new industries.

For our first Long-term Insights Briefing, Stats NZ Tatauranga Aotearoa has chosen the topic: *Aotearoa New Zealand: Empowered by data – Long term insights about the role of data in wellbeing and economic advantage.*

Stats NZ Tatauranga Aotearoa is New Zealand's national office of statistics. We are dedicated to improving the lives of New Zealanders today and for generations to come. We support and facilitate the release of social, economic, and environmental data for Aotearoa. We deliver statistics, data, advice, insights, and expertise to our customers, decision-makers, and the general public. Stats NZ is also the home of the Government Chief Data Steward – the system leader for data across government. The Government Chief Data Steward is responsible for:

- empowering agencies to use data more effectively while ensuring the trust and confidence of New Zealanders; and

- supporting the use of data as a resource across government to help deliver better services to New Zealanders.

As a leader in data and analytics, we are well placed to observe trends in the way that data is being managed and used, and to help inform public discussion on emerging issues, opportunities, and challenges. Our briefing explores how, as a country, we may approach data to ensure we can extract value for social and economic prosperity. Our briefing canvasses the strategic choices we may face on our way to ensuring that data creates positive outcomes for New Zealanders – making New Zealand a better place to live, work, visit, and do business.

How did we decide on our topic?

The world has changed significantly in recent years. The COVID-19 Pandemic has taught us a lot about the role data can play in how a country responds to, and rebuilds after, a crisis. While recent technological advancements have created a world where algorithms are increasingly present in everyday life, they can also be invisible at times. Algorithms can reinforce biases inherent in a great deal of our data at scale, exacerbating systematic inequalities if protections

aren't in place. In looking at New Zealand's future opportunities and challenges we wanted to understand more about how our nation can reap the significant benefits that are possible from data, while also ensuring that we protect New Zealanders from harm and preserve what New Zealanders value.

During December and January last year we consulted on the proposed subject and scope of our Long-Term Insights Briefing[1]. Many of the submissions received endorsed our selected topic and framing of the briefing. In particular, the concepts of openness and trust were widely supported. However, there were also points that had not been previously considered, or that our submitters felt were not given enough prominence in the consultation document, including different elements of the data value proposition. For example, engagement reinforced to us that New Zealanders care about more than just economic growth, and this briefing now seeks to also explore the role of data in social prosperity.
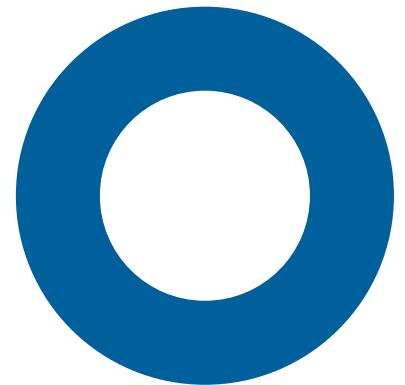
One of the driving factors behind the briefing is to understand New Zealand's many strengths and unique characteristics when it comes to realising the value of data. Traditionally, New Zealand's small and distant nature has limited our ability to connect to the rest of the world (with the exception of trade in natural factor endowments which has always been abundant, in-demand, and prosperous). However, over the last two decades, changes in technology and institutional structures have meant that global markets can be serviced more easily from a smaller number of locations[2]. This presents a data opportunity for New Zealand. Our small size and performance in key areas such as our institutional framework – for which we are globally recognised[3]; high levels of social capital, ingenuity, and trust; and potential for knowledge agglomerations due to our ability to cluster around innovative hubs, provides a strong environment to foster innovative new uses for data.

Another defining characteristic of the New Zealand data system is the intersection between data and Māori wellbeing. Success depends upon our ability to set up a system that treasures our cultural heritage and is informed by Te Tiriti o Waitangi. Data has an important role in preserving Māori identity and telling stories about Iwi-Māori and whakapapa. Often, the way that data is managed and used, is done according to Western frameworks, approaches, and values. Considering the increasing importance of data, we need to address some systematic issues where some communities in New Zealand cannot see themselves represented in the data we use, or where data might not reflect all aspects of the community – which could often be the case for Māori. A successful system is one where Māori interests are acknowledged and given effect to, where data can empower Māori communities. A successful system is also one where data informs and supports the growth of Māori Economies. This briefing will explore how we place Te Tiriti o Waitangi and Māori tikanga at the centre of the way data is managed and used in New Zealand.

What is the data system?

At its most basic, "data is any information in a form capable of being communicated, analysed or processed (whether by an individual or by computer or other automated means)."[4]

The data system is more than just the data used by government. Data systems enable the sharing of data between people and places within the right privacy settings. While the government data system is focused on the data government needs for decision making and service delivery, it is a subset of the wider data systems that New Zealanders interact with every day. In this briefing, we look at the wider data system across New Zealand, which encompasses government data, iwi-Māori data, research and scientific data, and data held or created by Non-Government Organisations (NGOs), academia, large institutions and private companies.

---

2 McCann, P (2003) Geography, Trade and Growth: Problems and Possibilities for the New Zealand Economy , PDF File – Working Paper 03-03 – Geography, Trade and Growth: Problems and Possibilities for the New Zealand Economy – The Treasury

3 NZ Treasury (2020) Institutional Framework and Economic Structure | New Zealand Debt Management – The Treasury

4 Australian Government; Department of the Prime Minister and Cabinet. (2022)

3. The Future of Data – the chances, choices, and challenges

Globally, there are several trends that are changing how societies live and creating challenges and opportunities that will be transformative. One of the most significant of these global trends is the rise in technology and the digitisation of how we live – and this is powered by data. Data is now at the heart of almost everything we do as we share and receive data every day to interact with the technologies that serve us.

As data continues to assume a fundamental and growing importance in all areas of human activity, a series of related trends are coming to the fore. There are emerging new issues that arise with more sophisticated use of data that many countries are grappling with, but an opportunity exists to be at the forefront – positioning New Zealand well to take economic and social advantage of data-driven innovation, delivering for and with Māori, meeting a range of data needs, and contributing multi-laterally as global conventions are developed.

In deciding which trends are likely to have the greatest impacts on New Zealand's future data system, we looked across the data value chain to understand how the global trends impact each stage that data moves: through collection, publication, uptake, and impact[5].

Setting the scene – the amount of data collected is growing rapidly

The world's capacity to generate and store data is growing at a rate faster than ever before. By the end of this decade, it is predicted that there will be yottabytes of data[6] – a term new to many of our vocabularies, but an increasingly important one. It is estimated that by 2030, global data will grow by one yottabyte every year, computing power will see a tenfold increase, and AI computing power will increase by a factor of 500[7].

Every two years, we create more data than was previously created through all of history. Our hyper-connected world fuels this exponential increase in data – from wearable devices and smart appliances to electronic health records and autonomous vehicles.

The Davos Agenda 2021 (World Economic Forum)

Consequently, this means more real time information is available in our everyday life. As data becomes increasingly available in 'real-time', we are offered insights into the pulse of communities, the ability to distinguish between different groups, and the potential for an almost moment-by-moment picture of what is happening in the world at any given point in time. A recent example was the use of high-frequency job vacancy data, available through online job sites, by the International Monetary Fund to track the impact of the COVID-19 pandemic on the job market and calculate the differential effects of the pandemic on different segments of the population using data about population employment patterns[8].

---

5   World Economic Forum (2022), Advancing Digital Agency: The Power of Data Intermediaries Insight Report.
    WEF_Advancing_towards_Digital_Agency_2022.pdf (weforum.org)

6   Huawei Technologies (2021) Intelligent World 2030 – Building a Fully Connected, Intelligent World. Intelligent World 2030 (huawei.com)

7   Huawei Technologies (2021) Intelligent World 2030 – Building a Fully Connected, Intelligent World. Intelligent World 2030 (huawei.com)

8   Chen, W. (2022) Disparities in Real Time. International Monetary Fund.
    https://www.imf.org/en/Publications/fandd/issues/2020/12/value-of-real-time-data-in-covid-crisis-chen

LOOKING CLOSER:
COVID-19 and the role of data

The COVID-19 pandemic has led to a heightened focus on the power of data. We have been able to see, almost in real time, where the virus was spreading, how it mutated, and what effect it has had on economies across the world.

"If there is one thing more important than vaccines in this pandemic, it's data: data about transmission dynamics of the novel coronavirus, about symptoms, testing results and hospital admissions."

Dr Ali Okhowat, Chief Executive Officer, mHealth Global, Canada | Co-Lead of the World Health Organization Innovation Hub in Geneva

In some respects, the pandemic has accelerated the ways in which data can be used, and how people view the use of data. This is because the urgent need for evidence to understand the virus has at times made the pandemic a catalyst for data sharing. The Ada Lovelace Institute observes that the use of data through the pandemic has been a key mechanism by which 'the science' – the clinical, epidemiological, computational, behavioural evidence – has been followed and that COVID-19 has demonstrated the need for swift and accurate data-sharing[9].

The way that data is used as we recover from the pandemic may build upon this catalyst further, creating a new operating environment for data. Data will likely play a significant role in how countries experiment with ways to limit the spread of new variants, and in shaping and informing the economic response to cushion impacts and build resilience. It has become increasingly important to look for data that is held by trusted institutions as not all data is equal.

At the same time, in some countries across the globe, there have been questions about the collection, sharing, use and interpretation of data during the pandemic, particularly regarding privacy, ethics, misinformation and surveillance. The response to the pandemic has also surfaced the difficulties in sharing data – with some governments unable to access timely data, which has provided a better understanding of some of the barriers preventing effective access and sharing.

CASE STUDY: Using the Integrated Data Infrastructure to map the spread of COVID-19

The Integrated Data Infrastructure (IDI) links together Stats NZ surveys, data from other government agencies, and non-governmental organisations. It holds de-identified microdata about people and households, and includes information about education, income, benefits, migration, justice, and health. Data available in the IDI is used for research projects in the public good that provide insight into New Zealand's society and economy.

Soon after COVID-19 was detected in New Zealand, researchers from the University of Auckland used the IDI to create a model that used population insights, such as where people worked and attended school; the makeup of families and communities; and how people moved around inside and between regions to better understand how the virus might spread. Insights from this model provided the Ministry of Health and the
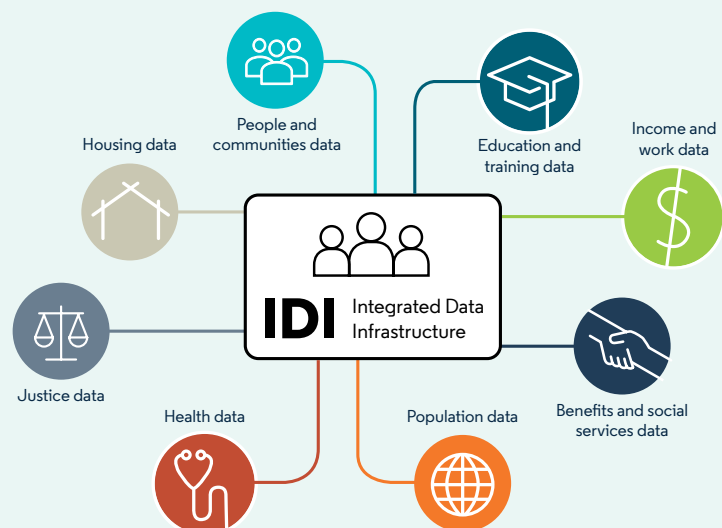
National Crisis Management Centre with information to make evidence-based decisions on how to control the spread of the virus.

9 Ada Lovelace Institute (2021) Learning data lessons: data access and sharing during COVID-19.
   The Royal Society Learning data lessons: data access and sharing during COVID-19 (royalsociety.org)

There are many drivers behind this growth trend. People have more tools to create and share information than ever before, a lot of which has been fuelled by the evolution of social media and the use of smartphones. Smartphones have steadily become integral to many of our lives and are now one of the most ubiquitous technology devices of all time with billions of users worldwide[10]. We carry smartphones everywhere we go and as such, there is a lot more data providing insights into everyday life and behaviour.

There has also been a shift away from people producing the most data, to machines and sensors contributing a wealth of data through the internet of things (IoT). From fitness trackers to smart heating systems, the IoT describes a growing network of internet-enabled devices. The World Economic Forum's State of the Connected World report predicted that in just a few years 41.6 billion devices will be capturing data on how we live, work, move through our cities, and operate and maintain the machines we depend on[11].

The acceleration in data storage capacity, lower cost computing power, and high-speed internet have also all led to the growth of data. The roll out of next generation broadband technology has begun in New Zealand and is set to become widespread in the coming years.

There is also an emerging role for synthetic data. Synthetic data is data that is generated artificially, rather than by real events. It retains the original data's statistical properties, but has no identifiable information, making it useful for research in areas where privacy is key, such as healthcare and finance[12]. It is created within a digital world from computer simulations or algorithms. Synthetic data technology enables practitioners to simply digitally generate the data that they need, on demand, in whatever volume they require, tailored to their precise specifications. Gartner suggests that, in just a few years, 60% of all data used in the development of AI will be synthetic rather than real[13].

Data has the potential to become a valuable national asset – critical to wellbeing and economic success

The vast volume of data described in the last section is already creating new forms of value that impact the lives of New Zealanders. The better the data, the better the services and products provided by the government, businesses, iwi-Māori organisations and community sectors.

Traditionally data has always been used to inform decision-making, whether it has been paper-based records held by small-businesses about the number and types of items sold each week, stream-lined ordering processes, or the use of customer surveys to determine ways to improve customer experiences. However, with the increasing data available, alongside our ability and capacity to analyse data (which is becoming increasingly sophisticated), data can now enable:

- new insights into how, as a society, we can perform better;
- new ways to measure, manage, and control; and
- smarter products and services that are tailored to individual needs.

Within each of these uses, data has begun to inform decision-making and choices in a way that stretches beyond a capture of the past and present, to provide a view of the future with increasing accuracy. At the centre of this future view is the knowledge that data gives us about how our world works and our place within it. This knowledge becomes richer when we combine data to drive value. Data, as pieces of information, are often not valuable in themselves; rather, the value lies in the ability to combine and integrate data.

10 Fleming, S. (2021) Your smartphone has become your home, anthropologists say. World Economic Forum.
    https://www.weforum.org/agenda/2021/05/how-we-interact-with-smartphones-report/

11 World Economic Forum (2022) Future of the Connected World: Global Action and Recent Progress.
    Publications > Future of the Connected World | World Economic Forum (weforum.org)

12 UK Government Office for Artificial Intelligence (2021) Algorithm: National AI Strategy ALGORITHM (publishing.service.gov.uk)

13 Castellanos, S. (2021, July 23) Fake It to Make It: Companies Beef Up AI Models with Synthetic Data Wall Street Journal.
    https://www.wsj.com/articles/fake-it-to-make-it-companies-beef-up-ai-models-with-synthetic-data-11627032601

Combining and integrating data means that decision-making can be informed by exponential, connected, and customised insights that can inform choices across government policy making including allocation decisions; individual consumer choices; collective decision making; iwi decision-making; and by industries to reduce risk, optimise processes and production, target effort and customers, and allocate resources.

Design is also becoming data-driven, with data informing production design, the design of new services and products, and the way institutions interact and deliver to their customers. The ability to capture more data, in real-time, that is integrated with other data, provides a comprehensive view of customer behaviour and feedback – which can enhance user experiences and impact on adoption rates.

Integrated data also has an influence on how individuals interact with the world around them. For example, integrated personal data can influence the online interactions that people have, with data collected through social media, online purchases and other ways people have an online presence (i.e., location data). This data is often used to determine interests – which influence targeted content.

The trend is only set to continue, as more experimentation takes place – with the use of new data sources, including big data, such as satellite imagery, scanner data, or mobile phone data. In this way, our data science will become more sophisticated, with the capability to extract meaningful insights enhanced by less resource intensive approaches such as machine learning and artificial intelligence.

It is also critically important that we acknowledge the powerful role data can play in fostering a more equal, socially cohesive, and inclusive society. High-quality, granular data provides communities and government with data to inform action. Data carries its weight when it comes to ensuring the visibility of more diverse populations – particularly when the use is supported by sound governance and the trust and confidence of communities it represents. Currently, too many people are invisible in data and too little data is routinely disaggregated. We know that within Aotearoa New Zealand, richer data is needed for Māori, Pacific Peoples, ethnic communities, women, gender diverse peoples, disabled people, seniors, children and geographic communities (e.g., rural and suburban communities).

LOOKING CLOSER:
The role of data in climate change

To translate our green commitments and climate goals into meaningful action, data is critically important. Data can help us to identify and locate harmful emissions and monitor progress by quantifying emissions across the board.

"In addition to measuring the extent of climate change, data can be used to paint a picture of what the world would look like if we didn't act, to tell very powerful stories that can encourage people to take action."

Dr Eileen Tipoe, senior lecturer in economics
at Queen Mary University of London

As with other areas, our use of data for this purpose is becoming more sophisticated. For example, if we combine data captured from satellite imagery, we can use artificial intelligence to monitor different land-use types and protected species. Better data capture and artificial intelligence also provides a better understanding of the causes of climate change; provides insights into innovative and sustainable new ways to mitigate the impact of climate change; and informs accurate predictions and modelling of extreme climate effects that may occur in the future[14].

In addition to playing a significant supporting role in environmental management and stewardship, data also has a role in contributing to more sustainable consumption and production processes via building awareness on the state of play, influencing consumer behaviour and informing markets.

However, it is also important to consider that data itself can also have a direct impact on the environment. The sheer scale of data being created requires storage resulting in increased electricity consumption and emissions.

While most climate change activists are focused on limiting emissions from the automotive, aviation and energy industries, the processing of digital data is already comparable to these sectors and is still growing. In 2020, digitisation was purported to generate 4% of global greenhouse gas emissions.

CASE STUDY:
Using online information commons to help farmers transition to regenerative agriculture

New Zealand company, Toha, has developed a platform to support marketplaces for impact data that reward and incentivise environmental and social outcomes. The first marketplace on the Toha platform is for climate and environmental data. Toha enables people to commit to environmental actions, provide data on their outcomes and receive financial rewards, all while retaining ownership and control of their data.

The system is designed to combine data from many different sources, including from individuals and big data sets and information from remote sensing. This 'data network' enables economies of scale around access to powerful new environmental and economic insights, as well as the ability to catalyse outcomes.

The system has been piloted in New Zealand with a cohort of farming families the first to receive revenue for their environmental data that shows positive outcomes they have achieved. This opens a whole new category of

economic activity, where farmers are adding new income to their P&L, generated solely from their environmental efforts and the accompanying creation of tradeable data assets. To find out more visit: www.toha.nz

14 Castellanos, S. (2021, July 23) Fake It to Make It: Companies Beef Up AI Models with Synthetic Data Wall Street Journal.
    https://www.wsj.com/articles/fake-it-to-make-it-companies-beef-up-ai-models-with-synthetic-data-11627032601

Data driven technologies are becoming increasingly mainstream

Artificial intelligence (AI) is about machines which act intelligently – typically making predictions or decisions about multiple aspects of the world in which we live[15]. The term AI covers a basket of technologies that use data and algorithms to allow machines to perform tasks or make decisions that would normally require human intelligence.

The acceleration in data growth has enabled the development of new technologies such as AI, robotics, and the internet of things (IoT). These technologies have the ability to change lives. The global market for these technologies alone already represents $US350 billion and has the potential to grow to $US3.2 trillion in just a few short years[16], and this figure doesn't capture the value generated within industries that deploy these technologies.

> "This is the age of artificial intelligence. Whether we know it or not, we all interact with AI every day – whether it's in our social media feeds and smart speakers, or on our online banking. AI, and the data that fuels our algorithms, help protect us from fraud and diagnose serious illness. And this technology is evolving every day."
>
> Secretary of State for Digital, Culture, Media, and Sport – HM Government, UK.

AI is considered one of the fastest growing technologies in the world, with huge potential to rewrite the rules of entire industries, drive substantial economic growth and transform all areas of life[17]. The upstream component of AI that often goes unacknowledged is data. Data is crucial to the effective use of AI. These technologies are powered by access to data, and they effectively utilise vast quantities of data to build and refine their algorithms to power 'learning' processes to achieve their desired outcomes. Each facet of AI, whether it is machine learning, machine reasoning, data analysis, natural language processing, or machine vision – has at its heart, good quality data.

There is a growing mistrust in data while at the same time people do not yet realise the value of their own data

Navigating the changing use of personal data

For the past two decades, the commercial use of personal data has grown. Personal data provides the foundation for some of the world's largest companies – harvested from personal devices, social media, and electronic transactions, it is turned into customer insights, market predictions, and personalised digital services[18]. As this activity has developed and grown, some people have begun to realise the value of their own data, its potential, and the possible harms that can come from its use if the rules aren't clear – but this is not widespread. Even companies themselves are starting to question what rules and structures should be put in place to protect individuals' data use.

---

15   The Alan Turing Institute (2018) What does AI mean for the Turing? What does AI mean for the Turing? | The Alan Turing Institute

16   United Nations Conference on Trade and Development (2021). Technology and Innovation Report 2021 (unctad.org)

17   UK Government Office for Artificial Intelligence (2021) National AI Strategy National AI Strategy (publishing.service.gov.uk)

18   Rahnama, H & Pentland, A (2022) The New Rules of Data Privacy. Harvard Business Review The New Rules of Data Privacy (hbr.org)

> *"Every time we swipe our bank card, upload a photo to Facebook, watch a video on YouTube, use a loyalty discount, hand over a prescription, read an article online, we're giving ourselves away. A detailed, digital record of our lives is a valuable commodity. Companies, governments, and even charities can use what we reveal about ourselves to make money. But they can also use it to make our lives easier, our cities smarter, and our societies fairer. This new marketplace of information exists without any broadly agreed master plan, or rules."*
>
> Katie Kenny, Stuff NZ.

The commercial value of personal data has prompted arguments that individuals should be allowed to commercialise their own data. It has led to important questions about what it might take to enable an environment where individuals can be a broker and beneficiary of their own personal data if they chose to, and conversely – how to ensure that the market does not create a model that exploits those who are more financially vulnerable and give away their ownership more readily.

At the same time, it is also important to acknowledge that some will not want control. Even with greater transparency and efforts to simplify a view of what data is collected and how it is used, its consideration may still be burdensome in today's busy lives and potentially lead to engagement fatigue. Many may just want to know that there is someone to look out for their interests on their behalf.

Perhaps the answer is instead in cultivating trust and building the right protections. On a daily basis, people, knowingly or unknowingly, trade their personal data for value – services that are specifically curated, bespoke, and convenient. However, this data is often extracted as part of an exchange where little is known about exactly how much data is collected, who gets to look at it, and what it's worth. How do we balance these innovative and exciting new markets and sophisticated insights that improve lives, while protecting individuals from the misuse of data? The emerging new role of data in today's society will likely be defined by our attitudes towards data ownership, ethics, and privacy.

Te Ao Māori and data

While data is widely recognised as a powerful tool and enabler, to some Māori, data is not a commodity to be exploited — to many it is a taonga. This understanding presents important questions to consider when looking towards a future state data system centred around Te Ao Māori.

If the right data is collected, and it is interpreted through a Te Ao view, data can lead to outcomes that are beneficial for Māori, particularly when it comes to data that can inform and improve Māori wellbeing. To enable this, we need to have a view of our future system. If our future system is designed well with, and for Māori, and based in mātauranga (knowledge), considerations, such as data sovereignty, ethics, and privacy, will become more easily resolved.

There are many examples of iwi and Māori organisations being empowered in their decision making through the effective use of data. We also see practices around Māori data sovereignty as a unique approach distinct from elsewhere in the world. To support improved levels of trust in how data is used in New Zealand it will be important for our settings to recognise and reflect te ao Māori, and for Māori to be enabled to participate as equal partners.

Enabling a new era of trust

There is a growing public awareness among some segments of society, of the volume and variety of data that is collected, stored, and used, particularly as the scale of algorithmic decision-making increases. While the Privacy Act and the incoming Consumer Data Right seek to provide New Zealanders with rights and protections, there is still the need to consider other associated issues, such as the quality and inclusivity of data used, the governance of data, and how it features in decision-making. Uses of data that may be legal, can still be considered irresponsible or unethical and undermine public trust more widely. For example, there is a growing public interest in the potential for adverse and discriminatory impacts of algorithms on individuals and society.

The growing use of AI and machine-learning will only fuel this concern about complex algorithms and their role in unfair or biased outcomes.

> "On the one hand, data has the potential to solve some of our most pressing societal challenges, from the pandemic to poverty to climate change. But that data can only be effectively leveraged if individuals and organizations trust those bodies that are responsible for handling it, rather than suspect that their privacy and wellbeing might be compromised without their knowledge or consent. This is where the challenges lie."
>
> World Bank (2021).
> World Development Report 2021: Data for better Lives
> https://www.worldbank.org/en/publication/wdr2021

Increasingly, trust is built on knowing that the technologies are safe, risks have been mitigated, and the data and information being used is robust. Decisions must be open and transparent, allowing for conflicting views to be aired and trade-offs to be made. As well as the decisions themselves, the process by which the decisions have been made needs to be transparent. In addition, Māori data experts have acknowledged that, given the constant evolution of the digital and data space, strong governance structures must be used to maintain trust in the system.

Shifts in the concentration of power

With the growing proliferation of data, there is a shift in the concentration of power when it comes to data. Knowledge is power, and any institution that controls a major, widely applicable source of raw personal data has the power to lock it down immediately and reap the proceeds by monetising it. When this happens, and data is disproportionately captured by just a few, it raises concerns for equality, the potential for innovation, sustainable development, use by social enterprises and 'reach' particularly in environments with significant power imbalances and inequalities of service provision.

4. Quantifying the value of data, including its value to people's wellbeing and the environment and climate, and why that matters for New Zealand

Data is a unique commodity – quantifying its value is complex

It is important to be able to quantify the value of current and future data-driven innovation. It gives a sense of the scale of opportunity. However, quantifying the true value of data and data-driven innovation is complex.

Every time we use data to improve systems or reduce costs it creates the potential for increased economic, societal, and/or environmental value. For example, an interoperable health system, where data is readily shared between parts of the system (such as between primary health care providers and hospitals), could result in a 10% reduction in hospital admissions for elderly New Zealanders due to earlier interventions, potentially result in cost savings of $97 million[19]. While the financial benefits of data are significant, the less financially quantifiable benefits to individuals are also significant. In this example, the service improvements from data sharing could mean these individuals receive better health services, possibly enabling individuals to remain in their homes and connected to their local communities.

The true value of data is realised in how data is used. Data, along with other kinds of information goods, are deemed 'experience goods'[20], which means that their value can only be established after their use. In this sense, data needs to be used to be able to then establish its value. Accuracy and usability (when combined with other data) increases the value of data. Unlike other commodities, the more data is shared and used the more valuable it becomes. The value of data can also decline or depreciate over time, with the rate of depreciation dependent on the time-sensitivity of the data.

There are three properties that make data a unique type of asset:

- it can be linked together to improve its value

- you can use it and re-use it again and again and again with no loss of value

- a diverse range of people can re-use it to do different things

For these reasons, it needs to be joined up and pooled or reused by diverse interests to realise its latent value.

Technological advances will continue to improve the efficiencies of gathering, storing and sharing data, as well as enable higher-quality and more complex analysis. We will also continue to shift from resource intensive approaches (such as relying on high-quality information and manual data cleaning) to those incorporating machine learning and artificial intelligence, fuelled by mass datasets.

The potential reach of data-driven innovation will continue to expand, and it will be increasingly important to try and understand (and measure) the full ramifications of these impacts, beyond financial and economic benefits.

Estimating the current value of data-driven innovation to the New Zealand economy

As a way to begin to understand the value of data-driven innovation, Stats NZ engaged Sapere Consulting to undertake research to look at the impacts of data-driven innovation on seven broad sectors of the New Zealand economy. Combined, these sectors represent approximately 92% of GDP production in New Zealand.

Sapere's research estimated the value of current innovation arising from cost reductions and revenue increases now and into the future and included examining the pattern of adoption of innovation across the sectors.

The research estimates the current value of data-driven innovation to the New Zealand economy to be $5.1 billion per annum (approximately 1.6 per cent of NZ's GDP in 2021). By 2030, they estimate that the value of data driven innovation to the New Zealand economy could grow to between $13 billion to $36 billion (between 3.2 – 8.7 per cent of GDP)[21], and that, with higher adoption rates and increased levels of cost saving, this opportunity could be closer to $45 billion.

---

19 Sapere Research Group and Covec (2015) Data Driven Innovation in New Zealand Data_Innovation_Report_WEB.pdf (srgexpert.com)

20 Experience goods are those where price, quality or some other attribute remains unknown until purchase. When visiting a new restaurant or buying a new wine, for example, consumers show a willingness to take a risk on how satisfying the product will be.

21 Sapere Consulting (2022) Data-driven innovation in New Zealand. Unpublished research report

The chart below shows the estimated value of data driven innovation across different parts of the economy in 2021 and estimates its likely value in 2030.

Chart 1: Estimated value of data driven innovation by sector
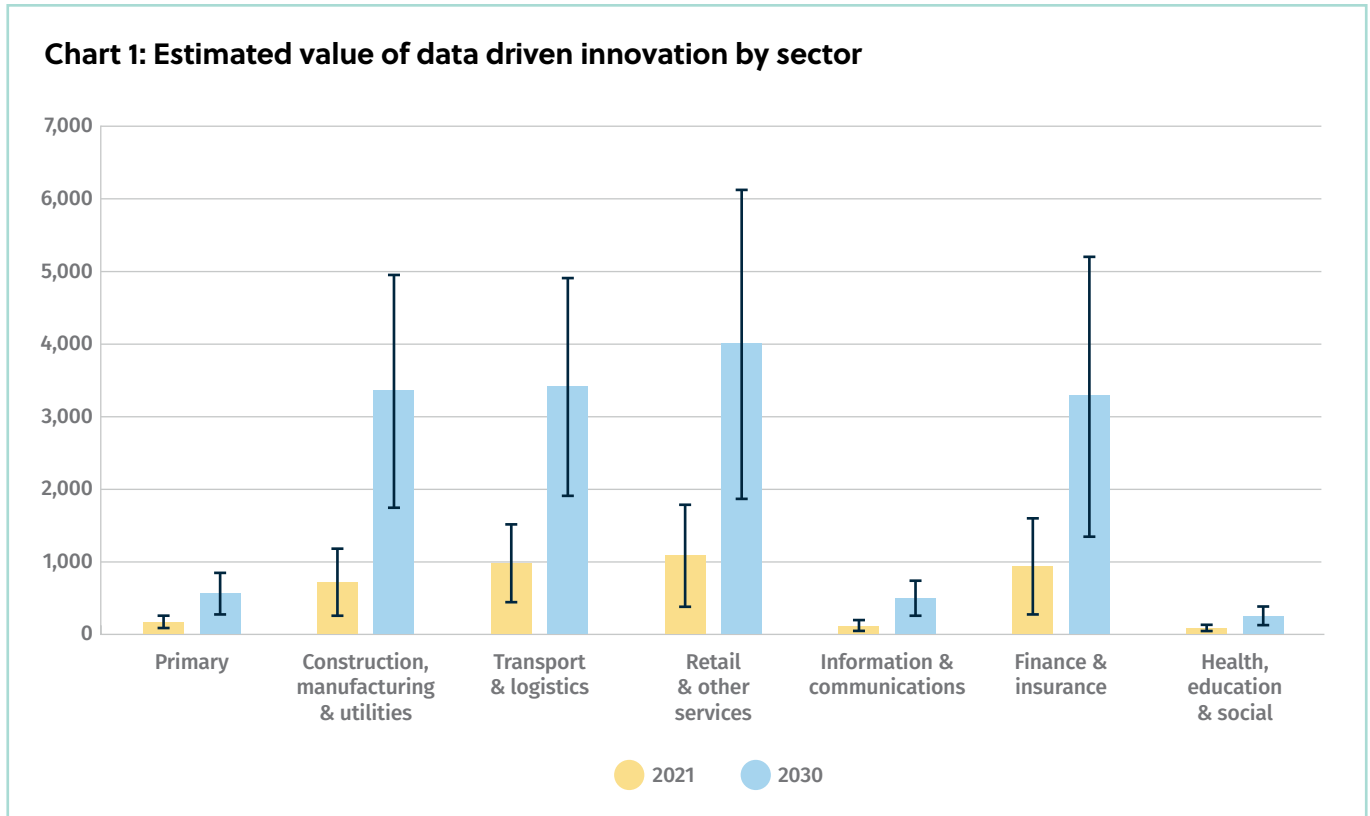
Data-driven innovation in the private sector delivers value by optimising a wide range of business operations such as improving value chain efficiency, building better customer relationships, development of new products, services and markets and optimising use of resources. This means that data-driven innovation can cover a range of applications, from improvements to business operations to whole-of-industry transformations.

In the Public sector data driven innovation can decrease the costs of service design, planning and delivery to organisations, and can also improve service delivery to the public and address social challenges leading to better health, education, and environmental outcomes.

For example, the Ministry of Education used data from the Integrated Data Infrastructure (IDI) managed by Stats NZ to develop a student-focused statistical model to identify important socioeconomic factors that impact student achievement. This model led the Ministry to develop the Equity Index, providing a more accurate way of targeting a small proportion of school funding based on the actual level of disadvantage within a school. Transition to this

new funding system which replaces the decile system is underway, with the Equity Index being used to determine equity-weighted funding levels from January 2023[22].

Estimating the non-tangible benefits of data-driven innovation

The benefits of data driven innovation are usually articulated through the achievement of revenue increase and cost reductions. However, there are significant benefits from data driven innovation that are not able to be articulated in these terms. The ability for citizens to access and utilise open data has been shown to "increase the openness, transparency, and accountability of government activities and thus boost public trust in government". While these benefits are not readily measurable in financial terms, they are still valuable.

Data driven innovation can also directly and indirectly contribute to wellbeing, for example through innovation impacts on health, the environment and society. In the case of the public health system or social assistance, efficiencies realised through data driven innovation generally drive

22   See https://assets.education.govt.nz/public/Documents/our-work/changes-in-education/Introduction-to-the-new-Equity-Funding-system-for-schools-and-kura.pdf

improvements in the quality of care or service provided or allow organisations to reach more individuals with the same or less resourcing. These effects are often not as readily quantifiable as direct economic benefits.

Decision makers also often use data to tailor policy responses and target the available resources to the issues, people, and communities most in need. There is value in innovative, interoperable, open sharing of this data to improve outcomes for priority groups and to increase the transparency and trust in government use and sharing of data.

More generally, innovation and economic growth increase wellbeing because living standards rise, although the benefits may not be evenly distributed[23]. Key to the even distribution of benefits, is the ability to partner to collate and combine data sets, providing a more comprehensive view of our society.

The NZ Treasury Living Standards Framework incorporates a wider set of factors that contribute to wellbeing than have been typically applied in most economic analyses. The framework includes factors that affect individual and collective wellbeing, institutions and governance, and wealth. The framework reflects a broad view of wellbeing including te ao Māori and Pacific perspectives as well as children's wellbeing[24]. The wealth component of the Framework includes characteristics not captured in our system of national counts, for example, human capability or the natural environment. In addition, there are various frameworks that articulate and encapsulate a holistic view of wellbeing from a Māori perspective.

Data driven innovation influences all categories of the Living Standards Framework, either directly or indirectly and it is reasonable to assume that benefits will increase over time as data use and data driven innovation becomes more common across society.

23 Dolan, Paul and Metcalfe, Robert (2012) Measuring subjective wellbeing: recommendations on measures for use by national governments. Journal of social policy, 41 (2), pp. 409-427

24 Sapere Consulting (2022) Data-driven innovation in New Zealand. Unpublished research report

5. What could our world look like in 2040?

Four potential scenarios are set out below which map out the potential data future for Aotearoa New Zealand. These are not necessarily mutually exclusive, and it is possible that elements of each scenario may play out in the future.

Scenario 1: A modern data-driven society – maximising data's value by building trust and enabling its use

In this scenario, data is used, reused, and shared in effective, safe, and secure ways. This sharing provides for a system that is inclusive and integrated, and where innovation is encouraged and supported.

Institutional models are in place which allow organisations to come together with shared interests and for shared value. Organisations have access to platforms which enable them to safely share and combine data while meeting data security, reliability, transparency, and access control systems requirements. Participants in the data system are able to transact with each other knowing that commitments will be adhered to, and that processes are in place and are easy to access when things go wrong.

In this scenario, individuals have transparency of their personal data and know what is happening with their data and how it is being used, and what that data is worth. If they choose to, individuals can decide how their personal data is used, and by whom. There are rules and standards to ensure data is trusted and trustworthy, and data is used ethically across the system.

Individuals are also able to easily move their data from one place to another if they chose to, and there are rules in place to enable these processes.

Scenario 2: A modern Tiriti-based data system – delivering more equitable outcomes while maximising the value of data

A Tiriti-based data system would provide for a more meaningful balance between the exercise of rangatiratanga and kawanatanga in the data system and prioritise oritetanga. It would move toward a future state where the data interests of iwi, hapū, and Māori collectives are recognised and safeguarded appropriately, and where iwi-Māori have the opportunities and support to become brokers to, and beneficiaries of, the data system.

A more equitable data system would be designed to represent the nuances of te ao Māori; apply the Tiriti principles of active protection, participation, and partnership within the settings that govern the system; and ensure iwi, hapū and Māori collectives are resourced to:

- collect data – in a way that captures the depth of experiences and speaks to the reality of iwi and hapū across the motu;

- use data – for more improved governance, in reclaiming narratives for and about their people, to secure infrastructural, professional, and environmental development opportunities;

- store data – iwi, hapū, and/or Māori collectives have the physical infrastructure, technical skills, and cloud capacities (through partnerships, if necessary) to store, encrypt and keep secure data; and,

- manage their data – understanding how Māori view data, this would include decisions regarding the integrity of Māori data is maintained and protected.
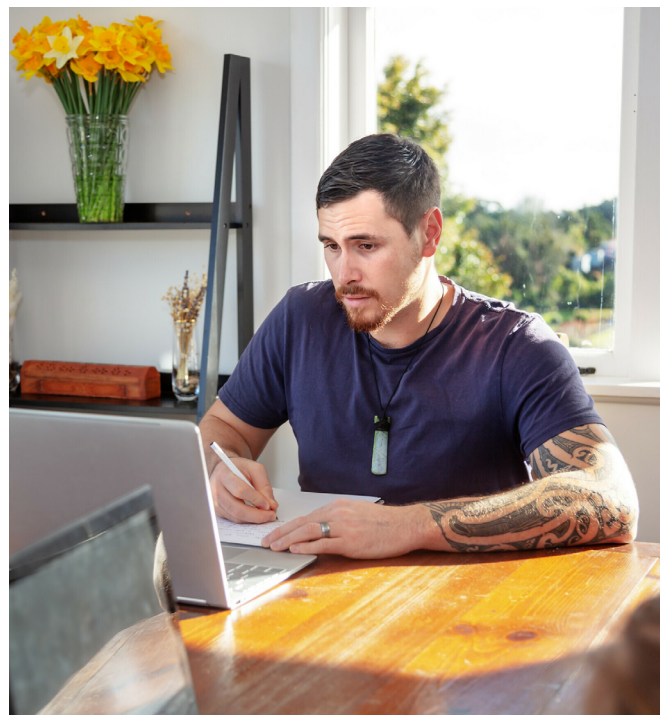
Scenario 3: A system where the benefits of data are realised but not distributed evenly

In this future scenario, data continues to improve our daily lives by providing us with better, more tailored products and services. However, it is one where data is being gathered by large, potentially multinational, organisations that create data monopolies – locking out small and medium sized players from accessing, and realising the value of data. Benefits that could be gained from greater personal autonomy, or sovereignty over data, are not being realised.

In this scenario, organisations and government may be reluctant to share data because they are not sure what the boundaries are and whether they are able to share data safely. A conservative approach here may lead to data being locked away in silos, limiting our ability to connect data across institutions to better inform science, policy, research, and the creation of new services. This is because businesses with significant data holdings and analytical capabilities benefit from data integration enabling them to grow at scale. This may create barriers to market entry, with small businesses unable to use data to innovate.

Scenario 4: What happens if we don't get it right – A future data system that is hampered by mistrust

In this scenario, New Zealanders do not trust, or have confidence in, the way that data is managed and used. This may eventuate because we haven't protected New Zealanders from harm, held on to what New Zealander's value, or because of the concentration of power outlined in Scenario 3.

This scenario will likely exist when we haven't got the balance right between innovative and exciting new markets and sophisticated insights that improve lives; and protecting individuals from the misuse of data.

While there are strong protections set out in the Privacy Act 2020, these do not necessarily provide people with visibility of and autonomy over how, and when, their data is used and re-used. For this scenario to occur, it is likely that progress may have stalled in work to increase the level of custodianship and portability individuals have of their own data. There may also be a lack of transparency when it comes to how organisations use data. Additionally, if data is tied up in silos and held by just a few, there is a real possibility that New Zealanders may lose control of their data, or that data is used in a way that does not line up with the ethics and values of New Zealanders.

If the stewardship of the data system is inadequate, or if social licence is eroded, there is a real risk that the future data environment could be one characterised by mistrust, where data plays a role in undermining social cohesion, and individual mistrust in government and its interventions.

6. Building the data future of Aotearoa New Zealand

The purpose of a Long-term Insights Briefing is to focus attention and effort on future issues facing Aotearoa New Zealand and the strategic choices government and society could make now, and in the coming years, to steer through them.

This section sets out some of the potential roles and levers of government to support data-driven innovation. The policy options we've identified focus on creating an environment for innovation by seeking to improve the flow of data, increase our capability to use and value data, and ensure the right settings are in place to guide the responsible use of data and prevent harm.

In considering these options, it is important to note that data flows do not recognise geographical borders – New Zealand's data system is part of a much larger international system. This flow of data across jurisdictions has meant we are more interconnected than ever before and places data as a significant contributor to the global digital economy. This raises questions for the policy options we want to consider.

The New Zealand government is only one contributor, and it is important to acknowledge that in complex systems such as this, there are many institutions and players – both domestically and abroad. New Zealand, being a small nation, with high levels of trust, has the ability to build momentum and to achieve critical mass when it comes to domestic problems that need solving – but we also need to keep a strategic view of international shifts to ensure we can continue to dock into larger-scale systems and be seen as a valuable contributor on the international stage.

The continued evolution of the data system is inevitable. The policy options proposed within this briefing are based on two assumptions: that data-driven innovation is a desirable pursuit to improve outcomes for New Zealand, and that the system will need support to ensure benefits are shared and any potential harm is minimised. The scale of intervention will therefore depend on future choices that robustly consider both assumptions.

Working towards a more effective system

There are a range of tools and levers available to government to effect change across the data system are split into two categories:

- **Creating the authorising environment on the system** – Guiding how the system operates by (for example) setting the rules of the game; considering rights and responsibilities (for example, working with Māori to establish interests); establishing structures (such as roles, institutions, and responsibilities) to shape how things are done; investing to support market accessibility for new participants (i.e., Iwi-Māori) or in areas of research and development; and marking out the future direction for participants to rally around.

- **Influencing the system from within** – Bringing together parts of the system to work together for system-wide benefit (breaking down silos and connecting across organisations to build effective partnerships); leading and influencing by being an exemplar of best practice; and delivering goods and services that the market is not incentivised to deliver itself.

We think there are six potential areas for government to action, where a lever or tool of government can set in motion a shift that will help Aotearoa New Zealand progress towards a modern data-driven society. A society where data benefits all New Zealanders – helping people and businesses to make better decisions and creating opportunities for innovation. These potential areas are split between the two categories outlined above.
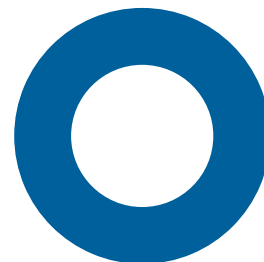
Creating the authorising environment on the system

Establishing data sharing environments in New Zealand

Data providers may not be sufficiently incentivised to share, or provide access to, their data – sharing can incur costs or effort that providers are not able to recoup from those that benefit. The benefits may have a wider public value, such as furthering knowledge or enabling innovation, but result in no commercial return to the provider. Organisations may be unaware of this public value or insufficiently motivated by it to incur the costs of sharing data.

There is a need to improve knowledge and understanding of data sharing with the intent to establish areas where there are shared interests, and to underpin this with an economic model that supports these interests. While doing so, there is a need to consider:

- how we ensure that the value of data and data-driven innovation is understood, and the value gained from data use is distributed fairly across the system;

- how, as a society, we account for the creation of, contributions to, and distribution of, various kinds of value arising from data sharing; and

- how we ensure the data reflects who we are, and is collected, stored, and used in ways that respect community and individual belief systems and values.

POTENTIAL LEVERS AVAILABLE FOR GOVERNMENT TO USE

Providing stewardship and leadership – Government is well-placed to signal the importance of an issue. For data to be treated as a valuable national asset, government could look to set the strategic vision and direction for how data is treated, and to drive a better understanding of the areas that need focussed activity. Areas where this lever could be explored include the creation of a national data strategy or vision, demonstrating incentives for, and benefits of, data sharing e.g., through testbeds and pilots; and identifying datasets or assets of national importance or public interest, to ensure appropriate stewardship activity is built around them.

Creating a model where the benefits of data are shared – One option, in response to the congregation of data power within just a few large technology firms, is for a shift towards collective ownership of data. This could be achieved through a model focussed on distributing the potential benefits of data by setting out sharing options/expectations for different data types (e.g., encouraging the pooling of data between specific partner stakeholders for collaboration, or an open access regime for some anonymised and aggregated data – within existing privacy settings).

Governance and institutional settings – Fostering confidence in the way we navigate evolving legal and policy considerations, ethics, and norms of a developing data economy

The perceived or actual risk of losing competitive advantage; breaching data protections; or suffering reputational damage from data uses that may breach trust, may deter providers from sharing data.

There will be a need to test whether our current governance umbrella can provide enough clarity to remove these perceived or actual risks. Of critical importance is governance systems and institutions that earn and maintain the trust of individuals. A lack of transparency undermines trust, which can also be further exacerbated by power

dynamics that exist within a data system (between an individual and a provider, but also when power across a system is concentrated within a single provider, which can occur when sharing is locked down and network effects are in place).

There is also a need to test how we specifically consider, and give rise to, Māori interests in how data is collected, managed, and used across New Zealand (and as it moves beyond our border).

If we think there is a need for shifts within our existing governance umbrella, this will require consideration of the types of roles and functions needed, how we separate powers and create checks and balances, what monitoring systems may be appropriate, and what the data-ownership models are that we would like to incentivise.

POTENTIAL LEVERS AVAILABLE FOR GOVERNMENT TO USE

Setting the rules of the game – Regulatory powers can be used by government to put rules in place and to set norms that shape how the system operates in pursuit of a particular outcome. Regulation is usually justified on the basis that a market or sector alone is not able to achieve certain desired policy objectives without a form of intervention. We may like to consider options which:

- **provide for the data interests of iwi-Māori** – enabling a more meaningful balance between the exercise of rangatiratanga and kawanatanga in the data system, recognising and safeguarding the interests of iwi-Māori and providing opportunities for Māori to become brokers to, and beneficiaries of, the data system.

- **establish governance structures which foster public trust** – providing the ability to strengthen practice across the data lifecycle in a way that is consistent across a system (e.g., the establishment of new oversight or governance bodies to guide component parts of the system);

- **establish a common approach to the protection of data** – with clear ethical obligations, ensuring the application of ethical practices by identifying ways to govern new technologies, setting rules for the development and application of data analytics, and providing tools to ensure consistent practice across.

Choice, control, and privacy settings – Strengthening transparency and autonomy

It will become increasingly important to provide individuals and communities with reassurance that their data is being treated with respect and is protected from misuse – and we need to examine what the right mechanisms are that both encourage data innovation and protect personal data rights.

In every region of the world, strengthening meaningful transparency in the ways that data is collected, stored, and used has been widely recognised as a priority. However, it is important to note that strengthened transparency, without controls and awareness, can also overwhelm individuals with too much information.

"It is important to understand that the huge untapped potential in data can only be successfully utilised if we are able to create a system where individuals can feel in charge of their data. This is the central question in making sure we can use data in the future to solve society's biggest problems or create new economic value."

Jan Vapaavuori, Urban Activist, former Mayor of Helsinki (the City of Helsinki has a Strategy to become the most functional city in the world that makes the best use of digitalisation).

POTENTIAL LEVERS AVAILABLE FOR GOVERNMENT TO USE

Setting requirements for the use of personal data – To strengthen transparency and autonomy, new requirements could be set that aim to embed a transparent approach to the management of data, and provide individuals with the ability to access, correct, erase or restrict the use of their data. This will ensure that participants in the data system have the choice to genuinely control their data relationships and decide what to share with whom, including reversing their decisions if trust breaks down.

Establishing an intermediary role that can act as digital agents for people – Intermediaries can be utilised to support data sharing and increase access to data. They can also play a role in driving responsible and ethical data use – supporting improved coordination and trust across the system and providing assurance through the role they play in making choices with individual needs in mind. There may be a need to consider the need for a centralised intermediary role or professional codes of conduct for controllers of data.

Influencing the system from within

Commitments and resolution – Supporting effective data collaborations

If, as a country, we are committed to data sharing as a foundation for innovation, then it will be important to have mechanisms in place that support sharing arrangements. The focus here is on the tensions that can result from power imbalances and when competing values arise, e.g., autonomy versus interdependence, inclusivity versus efficiency[25]. Effective mechanisms to support data sharing relationships can also target fear about what others might do with data which discourages sharing to start with.

25 Ruijer, E (2021) Designing and implementing data collaboratives: A governance perspective.
    Government Information Quarterly v38:4 Designing and implementing data collaboratives: A governance perspective – ScienceDirect

POTENTIAL LEVERS AVAILABLE FOR GOVERNMENT TO USE

Convening power – The government can act as a connector to bring people and ideas together to improve New Zealander's capability and capacity to use, and benefit from, data and emerging data technologies. Through this ability, government can also improve knowledge and understanding of data sharing – helping to establish networks and build momentum. For example, government could set up communities of practice to provide a forum to share knowledge about best practice in data sharing and ethical and innovative data use.

Data interoperability settings – Coordinating and supporting interoperable systems

Data integration brings together multiple datasets about people, places, or events to support research, policy analysis and service delivery. If this is done well it can lead to high trust free-flowing information systems that benefit both the individuals within the system and the system as a whole through the ability to generate greater insights and target service delivery.

Interoperability, findability, and the accessibility of data is also key to reducing the cost of data sharing. The goal here is to create systems that enable high velocity sharing and joining of data from diverse sources and contexts in a way that can scale and keep costs low, and where we can trust the interpretation or application of the data.

POTENTIAL LEVERS AVAILABLE FOR GOVERNMENT TO USE

Government already has a role in the provision of sharing infrastructure. Stats NZ is the home of the Integrated Data Infrastructure which holds de-identified microdata about people and households and the Longitudinal Business Database. Other potential roles may include:

- **supporting the establishment of integrated data systems outside of government** – by coordinating shared goals, setting operating principles, supporting decisions around the apportionment of benefits and commercial value, and ensuring various stakeholders, including vulnerable populations are visible across outcomes.
- **influencing the market** – as a significant purchaser within the New Zealand economy, government can use its purchasing power to encourage the behaviours we want to see fostered across the data system – for example, procuring technology services which support interoperability.

Consideration also needs to be given to how:

- we can accommodate multiple diverse platforms;
- we ensure that platforms meet security, reliability, transparency and access control requirements;
- platform technologies can be scaled at low-cost; and
- we ensure that platforms don't centralise power and control, or reduce trust in other ways.

Provision of public goods

The government's role as a producer of data services – be it social, economic or environmental – is only likely to become more important in the future. One of the key roles is ensuring access to trusted data products and services in a world where misinformation can be prevalent. In a context where people know that they can access, and use, high-quality and trusted data, it is more likely that data will be used to inform their decision-making. Government can play a role in creating and ensuring access to reliable, trustworthy data both within government and across the wider data system.

Government also has a role in ensuring that data represents the communities it serves through the provision of granular and up-to-date data on Pacific peoples, ethnic communities, women, gender diverse peoples, disabled people, seniors, children, and geographic communities (e.g., rural and suburban communities). A data system that reflects this diversity and intersectionality, and enables people to see themselves in the data, is a public good.

7. Continuing the discussion

We welcome your views on this important topic. Following this consultation period, we will take your feedback and work with our stakeholders and partners to develop a final briefing.

We will publish a summary of the submissions early in 2023. This summary may include quotes or extracts from submissions received.

If your submission contains any information that is confidential or other information you do not wish for us to publish, please:

- indicate this on the front of the submission, with any confidential information clearly marked within the text; and

- provide a separate version excluding the relevant information for publication.

Submissions remain subject to requests under the Official Information Act 1982. Please note, in your submission if there are any objections to the release of any information, including which parts you believe should be withheld, along with the reasons for withholding the information. We will take these objections into account and will consult with submitters when responding to requests under the Official Information Act.