



Stats NZ Tatauranga Aotearoa
Long-term Insights Briefing:

Aotearoa New Zealand: Empowered by data

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



Crown Copyright ©

Unless otherwise stated, this work is licensed under the Creative Commons Attribution (CC-BY) 4.0 International Licence. Except for any photographs, you are free to copy, distribute, and adapt the work, as long as you attribute the work to the New Zealand Government and abide by the other licence terms. To view a copy of this licence, visit [Creative Commons Attribution 4.0 International Licence](https://creativecommons.org/licenses/by/4.0/). To reuse a photograph please seek permission by sending a request to the attributed image owner.

Please note that neither the New Zealand Government emblem nor the New Zealand Government logo may be used in any way which infringes any provision of the Flags, Emblems, and Names Protection Act 1981 or would infringe such provision if the relevant use occurred within New Zealand. Attribution to the New Zealand Government should be in written form and not by reproduction of any emblem or the New Zealand Government logo.

If you publish, distribute, or otherwise disseminate this work (or any part of it) to the public without adapting it, the following attribution statement should be used:

Source: Stats NZ "Aotearoa New Zealand" Empowered by data" (10 November 2021). Licensed by the New Zealand Government for re-use under the Creative Commons Attribution 4.0 International Licence.

If you adapt this work in any way, or include it in a collection, and publish, distribute, or otherwise disseminate that adaptation or collection to the public, the following attribution statement should be used:

Unless otherwise stated, this work uses material sourced from Stats NZ which is licensed by the New Zealand Government for re-use under the Creative Commons Attribution 4.0 International Licence.

Version 2.0

Published April 2023

Foreword

Kia ora koutou,

I am pleased to present Stats NZ's first Long-term Insights Briefing: *Aotearoa New Zealand: Empowered by data*.

The new requirement to develop Long-term Insights Briefings provides Public Service agencies with an opportunity to take a pause from our daily work and look into the future. A chance to think deeply about future trends, emerging issues and opportunities, and how we might steer towards a future that reflects what matters to New Zealanders.

For Stats NZ, a clear trend captured our attention: the increasingly important role of data in society. We believe that data is the engine of better outcomes and our briefing takes a closer look at the potential for data to drive innovation and improve wellbeing. In particular, this briefing explores the value that can be generated through integration and sharing –underpinned by a system that is trusted, places Te Tiriti o Waitangi at its core, is ethical and transparent, and has appropriate protections in place.

The aim of our briefing is to start a conversation. Data can sometimes have a visibility problem. Every institution collects and uses data, yet it remains a core function that often doesn't come to the attention of everyday New Zealanders nor do we regularly consider its value as a national strategic asset. We hope that this briefing acts as a think-piece – by starting a conversation we can further understand the intrinsic values that New Zealanders hold when it comes to their data.

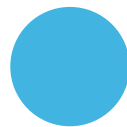
Public views have helped shape this briefing, and I want to thank submitters for taking the time to contribute – ngā mihi nui. I also want to thank our peers from across the data system who reviewed this work throughout its development. Stats NZ is just one player in the data system, and although we have a leadership role to play, charting the journey ahead requires a coalition and will only be successful through collective action.

Nāku noa, nā



Mark Sowden

Government Chief Data Steward and Chief Executive,
Stats NZ Tatauranga Aotearoa



Acknowledgements

This briefing is the result of combined efforts across the data system. In preparing this briefing we are grateful for the help and support of the following partners, stakeholders and experts.

James Mansell – Independent Consultant

Mike Cunnington – Deputy Commissioner Information, Intelligence & Communications, Inland Revenue NZ

Malcolm Fraser – Chief Data Steward, i4 Institute

Tim Ng – Deputy Secretary, Chief Economic Adviser, The Treasury

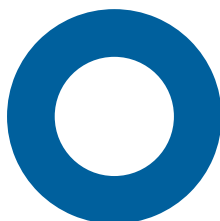
Frith Tweedie – Principal, Simply Privacy

Andrew Webber – Chief Economist, Social Wellbeing Agency

Sapere Research Group – Dylan James, William Li, Ben Barton, Veronica Jacobson, Peter MacIntyre, Lockie Woon, and Jamie O'Hare

We would also like to thank:

- Submitters and contributors, via the first and second rounds of public consultation, that added depth, nuance and diverse perspectives to the briefing.
- Stats NZ subject-matter experts for their time and insights during various workshops to help inform the briefing – Jason Attewell, Gary Dunnet, Vince Galvin, Tim Henwood, Wendy Hamilton, Anna McDowell, Diane Ramsay and Shane Ta'ala.



Contents

Foreword	3
Acknowledgements	4
What is a Long-term Insights Briefing?	6
Introduction	7
How did we decide on our topic?	7
Public consultation	8
What is data and why is it important?	9
What do we mean by data?.....	9
Data is embedded in every aspect of our lives	10
Te ao Māori and the data system	11
The Future of Data – the chances, choices, and challenges	12
Setting the scene – the amount of data collected is growing rapidly.....	12
Looking closer: COVID-19 and the role of data	14
Data has the potential to become a valuable national asset – critical to wellbeing and economic success.....	15
Looking closer: The role of data in climate change	17
There is a growing mistrust in data while at the same time people do not yet realise the value of their own data	18
Quantifying the value of data for all New Zealanders	22
Data is a unique commodity – quantifying its value is complex	22
The value of data.....	22
Estimating the current value of data-driven innovation to the New Zealand economy.....	23
Estimating the non-tangible benefits of data-driven innovation	24
What could our world look like in 2040?	26
Scenario 1: A modern data-driven society – maximising data’s value by building trust and enabling its use.....	26
Scenario 2: A system where the benefits of data are realised but not distributed evenly.....	26
Scenario 3: What happens if we don’t get it right – A future data system that is hampered by mistrust.....	27
Building the data future of Aotearoa New Zealand	28
Working towards a more effective system	28
Creating the authorising environment to enable desired shifts.....	29
Influencing the system from within.....	32
Conclusion: What do these trends and insights mean as we look to the future?	34

What is a Long-term Insights Briefing?

Public Service agencies have been asked to prepare Long-term Insights Briefings as part of their duty of stewardship under the Public Service Act 2020. 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 is to make available into the public domain:

- information about medium and long-term trends, risks and opportunities that may affect New Zealand and New Zealand society; and
- information, including possible policy options, for responding to these matters.



Introduction

Data has the power to change lives. It enables us to tell stories about how well we live, how well we take care of each other, how government services can be improved, and the state of our environment 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 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. As a leader in data, 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, environmental and economic wellbeing. The briefing canvasses the strategic choices we may face in 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. However, it is important to acknowledge that while the pandemic unlocked significant data activity, it has also left important questions around the sustainability of our data system, including what can be done to build resilient and more entrenched capability.

Shocks and disruptions like this are only set to continue and this has implications for our data system. Our existing model where data is collected, used, destroyed, and then collected again doesn't work in this world. This context also highlights the need for strategic infrastructure planning for resilience, ensuring locally relevant data that can be accessed quickly.

At the same time, recent technological advancements have created a world where algorithms are increasingly present in everyday life, often in a way that is invisible to New Zealanders. Algorithms make our lives easier – from the everyday i.e., optimising bus timetables or tailoring music recommendations, through to saving lives via diagnostic tools. They help us to make connections, ensuring the services we access are intuitively tailored to meet our needs. However, it is important that we find the right balance between value and protection. 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.

One of the driving themes behind the briefing is 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 resources and primary products which have always been 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 locations¹. This presents a data opportunity for New Zealand, given our small size and performance in key areas such as:

- our institutional framework – for which we are globally recognised²;
- high levels of social capital, ingenuity, and trust; and
- potential for knowledge agglomerations due to our ability to cluster around innovative hubs.

These features provide a strong environment to foster innovative new uses for data.

A further 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. However, often the way that data is managed and used is done according to Western frameworks, approaches, and values.

¹ 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

² NZ Treasury (2020) *Institutional Framework and Economic Structure | New Zealand Debt Management – The Treasury*

Given the increasing importance of data, we also need to address 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 point to the importance of how we place Te Tiriti o Waitangi and Māori tikanga at the centre of the way data is managed and used in New Zealand.

Public consultation

Legislation requires the topic choice and draft briefing to be shared with the public for consultation. An initial round of consultation took place in 2021, asking people for their views on our proposed topic: *Data as a driver of national competitive advantage – fostering improved wellbeing and economic growth.*

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 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 as a result, the briefing shifted to also explore the role of data in supporting overall wellbeing.

A second round of public consultation took place in late 2022, supported by targeted emails to interested groups, including a pānui to iwi organisations with the draft briefing, and an extensive social media campaign.

We received 25 submissions on the draft briefing, all of which were broadly supportive of the content and approach. Several themes were highlighted throughout the submissions and have helped to shape this final briefing including:

- further emphasis on how data can benefit wellbeing and the environment in addition to the economic benefits – including exploring how data benefits different communities in different ways;
- further exploration of trust by considering different governance approaches and potential solutions that can explicitly address bias in the data system; and
- wider discussion of the role of non-government actors in a decentralised data system.

Detailed information on the findings of the public consultation process is available on our website.



What is data and why is it important?

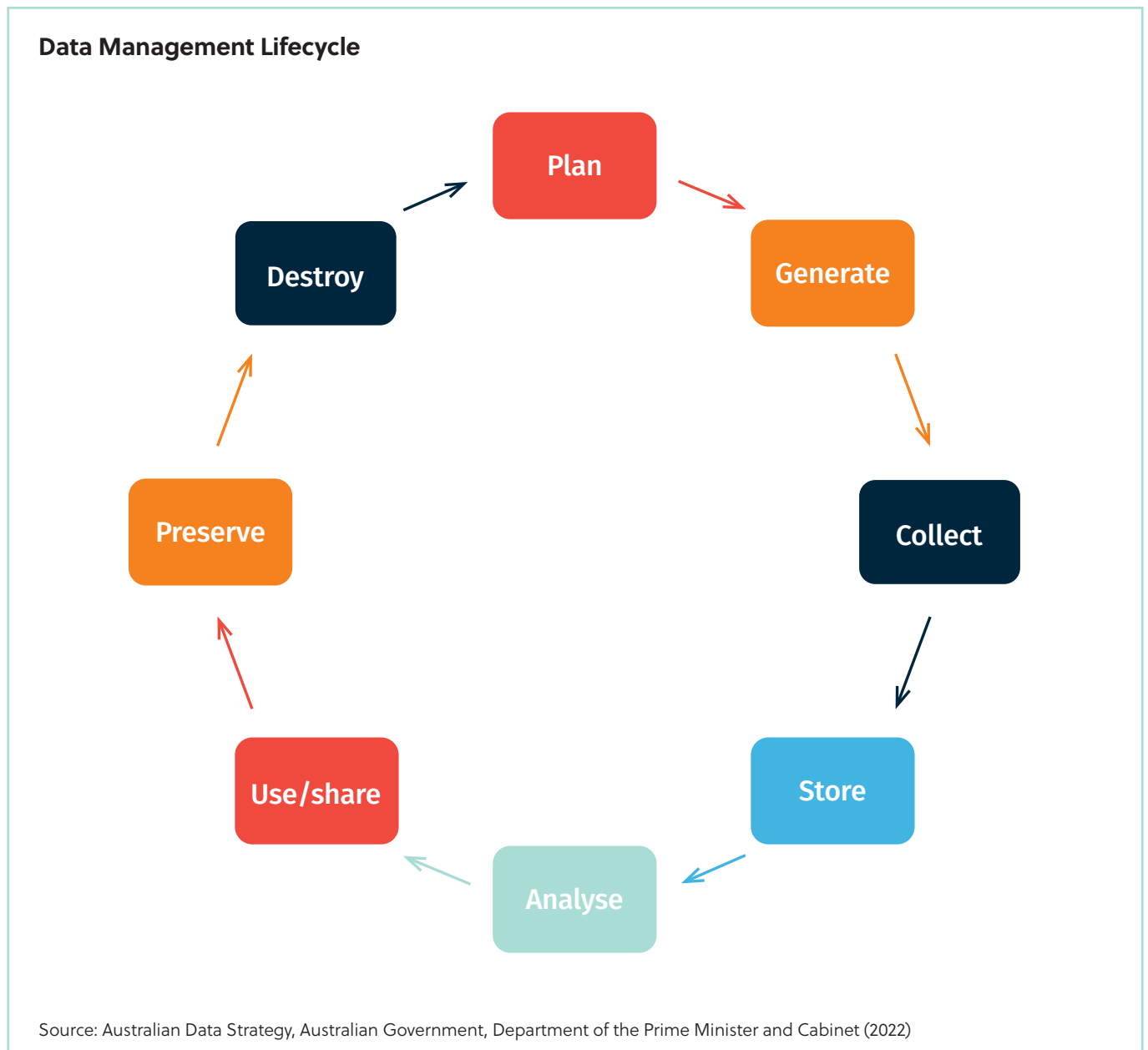
What do we mean by data?

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).”³

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. The government data system supports and facilitates the release of social, economic, and environmental data for Aotearoa – delivering the data government needs for decision making and service delivery, and the data

New Zealand communities need for decision making. However, 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. Our domestic system is also strongly linked into international data systems.

Figure 1 below shows the flow of data as it is generated and used.



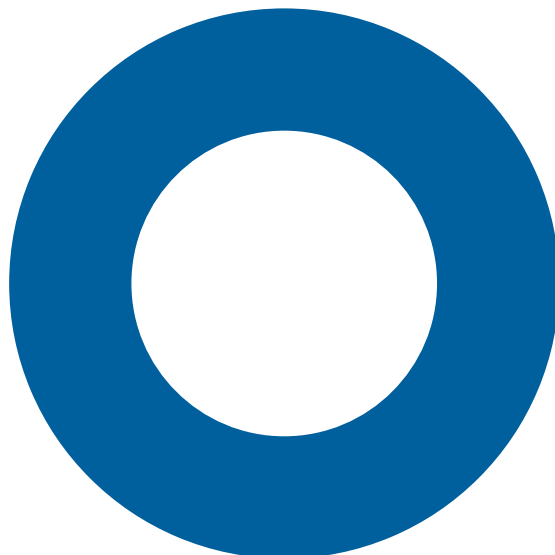
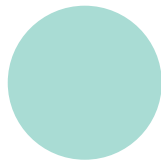
Data is embedded in every aspect of our lives

Data impacts on every part of society and how, and when, data is used across the system plays a big role in how society functions. This briefing explores how data can be used to deliver better outcomes by creating social and economic value, including the role data plays in:

- fostering a more equal, inclusive society – through the provision of high-quality data to inform policy and improve the visibility of inclusivity issues faced by vulnerable populations;
- supporting the delivery of government services through the ability to share data and analytics to shape service delivery, monitor and evaluate effectiveness of policy interventions, and provide actionable information to the front-line, within appropriate privacy and ethical

frameworks;

- acting as a powerful enabler for doing and thinking differently, including its role in empowering communities design in their own ways to achieve goals and aspirations;
- developing innovative new products, services and markets, and creating operational efficiencies; and
- anticipating and adapting to the impacts of climate change.



Te ao Māori and the data system

To many Māori, data is a taonga. Treating data as a taonga reflects the unique importance of data as an expression of identity and an input to a self-determinative way of life. Data forms part of mātauranga Māori (knowledge systems) and has deep value in telling stories of whakapapa, whanaungatanga and connection, and in building a rich sense of identity and history. This understanding presents important questions to consider when looking towards a future state data system centred around te ao Māori.

We have seen an increasing demand for data by Māori. Data can drive decision making and service delivery by iwi, and hapū, and Māori organisations and is becoming an increasingly important enabler for self-determination and development. It is within this context that Māori Data Sovereignty is gaining focus.

At the same time, the Crown is strengthening its relationships with iwi, hapū, and Māori organisations to better understand and respond to iwi data needs from across the motu. A key part of this work is ensuring iwi have access to and control of more appropriate, relevant, and timely data that is for and about iwi and hapū.

There is also a need for the data system to be designed in a way that ensures the treatment of data is more relevant, and responsive, to iwi aspirations and realities. This includes supporting iwi and Māori capacity and capability to work with data; improving the quality of, and access to, data that is important to Māori; and early work to decolonise methodologies surrounding data standards, management, and use.

Existing wellbeing frameworks and datasets often present Māori experiences through a deficit lens and so fail to capture the essence of Māori progress, interests and values. Most indicators also focus on individual performance and characteristics. However, Māori empowerment and resilience stem from collective entities such as whānau, marae and hapū.

Independent Māori Statutory Board
report: Data Issues of Significance

A data system that has Te Tiriti o Waitangi as its base would work toward the restoration of mana Māori motuhake, and also prioritise “te hauora me te oranga o ngā iwi katoa o Aotearoa” (the health and well-being of the people of Aotearoa). It would move toward a future state where data rights and interests of iwi, hapū, and Māori collectives are recognised and safeguarded appropriately, and iwi-Māori have the authority over data that is collected for, and about them.

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.

“A fundamental difference between te ao Māori and Western ideology is the idea of guardianship to distribute resources for the collective and future wellbeing of people and planet vs individual ownership and amassing resources for short term gain”

Quote from submissions on briefing topic and scope



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. As emerging new issues arise, an opportunity exists to be at the forefront – positioning New Zealand 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 impact⁵. The key trends we'll focus on in this briefing are:

- the exponential growth in data collection
- extracting more value from data
- growing distrust in data and how it's used.

The world of data is moving fast, and if we don't keep up, we'll miss out on opportunities to create new value, improve wellbeing, and realise other transformative benefits.

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. It is estimated that by 2030, global data will grow by one yottabyte every year (equivalent to the same amount of data that it would take to watch 12 new movies a day for the next 10 billion years), computing power will see a tenfold increase, and AI computing power will increase by a factor of 500⁷.

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.



⁵ World Economic Forum (2022), Advancing Digital Agency: The Power of Data Intermediaries Insight Report. [WEF Advancing_towards_Digital_Agency_2022.pdf \(weforum.org\)](https://www.weforum.org/reports/advancing-digital-agency)

⁶ ibid

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 fueled 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 worldwide⁷. 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 many of the machines we depend on⁸.

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 also 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 finance⁹. 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 real¹⁰.



7 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/>

8 World Economic Forum (2022) Future of the Connected World: Global Action and Recent Progress. [Publications > Future of the Connected World | World Economic Forum \(weforum.org\)](#)

9 UK Government Office for Artificial Intelligence (2021) Algorithm: National AI Strategy [ALGORITHM \(publishing.service.gov.uk\)](#)

10 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>

LOOKING CLOSER: COVID-19 and the role of data

The COVID-19 pandemic has led to a heightened focus on the power of data. Never before had we seen the need, at such a significant scale, for accurate and timely data to guide decisions and solve problems. During the pandemic we were 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

Data also supported broader insights about society during this time. A recent example was the use of high-frequency job vacancy data, available through online job sites. The International Monetary Fund used this data 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 employment patterns.

In some respects, the pandemic 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 made the pandemic a catalyst for data sharing. The Ada Lovelace Institute observes that the use of data through the pandemic was a key mechanism by which ‘the science’ – the clinical, epidemiological, computational, behavioural evidence – was followed and that COVID-19 demonstrated the need for swift and accurate data-sharing¹¹.

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 also become increasingly important to look for data that is held by trusted institutions, as not all data is equal.

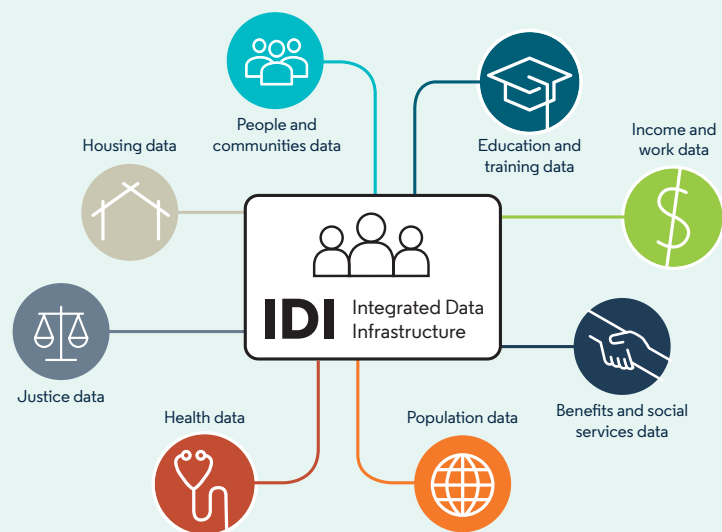
At the same time, in some countries there have been questions about the collection, sharing, use and interpretation of data during the pandemic, particularly regarding privacy, ethics, misinformation and surveillance. The response to the pandemic has also surfaced the difficulties in sharing data, with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

¹¹ 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)

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 live¹². 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. 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 years¹³, 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 life¹⁴. The recent developments around chatbots like ChatGPT are a good example of the growing use of AI, and how fast the field is moving.

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 data at its heart. The quality of this data makes the difference between good AI decisions and poor ones. AI and algorithms are only as good as the data they are fed, and data quality problems can have short- and long-term consequences.

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

This vast volume of data 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 records held by businesses about the number and types of items sold each week, 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 for everyone;
- new ways to measure, manage, and understand the impacts of policies and activities; 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.

¹² The Alan Turing Institute (2018) What does AI mean for the Turing? [What does AI mean for the Turing? | The Alan Turing Institute](https://www.alan-turing-institute.org/what-does-ai-mean-for-the-turing/)

¹³ United Nations Conference on Trade and Development (2021). Technology and Innovation Report 2021 (unctad.org)

¹⁴ UK Government Office for Artificial Intelligence (2021) National AI Strategy [National AI Strategy \(publishing.service.gov.uk\)](https://www.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/95444/national-ai-strategy.pdf)

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

Design is also becoming data-driven, with data informing new business models, the design of new services and products, and the way institutions interact with 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. By creating the right settings to promote greater safe and secure data sharing and innovative data use, it will be possible to unlock value across the board. There are opportunities to work collectively – removing barriers to

access better data for all New Zealanders, under the right governance structures and conditions, to create benefits for everyone across the system.

As part of this, 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 on populations within our society provides communities and government with data to inform action to support those populations. 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 the 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, older people, children and geographic communities (e.g., rural and suburban communities). Further discussion on the equity of data on Pacific Peoples can be found in the Long-term Insights Briefing prepared by the Ministry for Pacific Peoples.

Data also plays a major role in managing and protecting our natural environment. For example, data on freshwater quality is essential for knowing where action is needed most urgently. Below we take a closer look at the role of data in climate change and the Long-term Insights Briefing jointly prepared by the Department of Conservation and Land Information New Zealand discusses how data and technology can support biodiversity to thrive.



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 strengthened 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.

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 action is focused on limiting emissions from the automotive, aviation and energy industries, the emissions from processing 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.

Climate Action Innovation Zone at COP28¹⁵ (cop-27.org))

CASE STUDY:

Using online information commons to help farmers transition to regenerative agriculture

New Zealand company, Toha, has initiated a sustainability-focused online information commons called Calm the Farm, where technologists, scientists, farmers and local regional regulators work together to help farmers transition to regenerative agriculture.

The information commons 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.

Through the online information commons, farmers upload daily data about their farm practices, environmental actions, and outcomes while receiving financial rewards, and retaining ownership and control of their data. This data enables scientists to monitor the quality of the farm's natural environment – for example, the health of rivers and the sequestration of carbon into soil. The data can be audited by scientists and regional

regulators and cross checked with data from other farmers to gauge the environmental health of the region.

Scientists also capture data and feed insights back to farmers to help inform environmental health decisions on farms. This collective information is used by scientists and regulators to collaboratively improve the environmental health of the region.



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 services¹⁶. 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.

“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.

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 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 so may give away their ownership more readily.

At the same time, it is also important to acknowledge that some people will not want control of their data. 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.

The emerging new role of data in today's society will likely be defined by our attitudes towards data ownership, ethics, and privacy. Now and in the future, it will be imperative to cultivate trust and build the right protections for data.

Current rules for protecting data, here and abroad

Trust is fundamental to data being shared, and underpinning trust are the rules set by government and the ethical practices of all the players in the data landscape.

In New Zealand, we already gave a number of tools that help protect people's data and ensure that data is used safely. For example, the Algorithm Charter for Aotearoa New Zealand commits government agencies to improving transparency and accountability in their use of algorithms.

Within the policy landscape are two key pieces of legislation – the Privacy Act 2020 and the Data and Statistics Act 2022.

The Privacy Act 2020 regulates what can be done with information about individuals. The Act sets rules about how organisations and businesses can collect, store, use and share people's personal information. It ensures that people know when their personal information is being collected; how that information will be used and shared; and how they can access their information.

The Data and Statistics Act 2022 provides the framework for Stats NZ to collect data for official statistics, lead the official statistics system, and provide safe access to data for research. The Government Statistician must recognise Māori interests in collecting data, producing statistics and using data for research, and the way data is collected, managed and used. It provides appropriate safeguards and protections to ensure public trust and confidence in the collection and use of data for official statistics and research.

The government is also in the process of establishing a consumer data right. A consumer data right is a mechanism for consumers to securely share data that is held about them with trusted third parties that provide value-added services. For example, a consumer data right means that data holders, such as banks and electricity retailers, will be required to share certain consumer data with third parties in a specified format and timeframe.

¹⁶ Rahnama, H & Pentland, A (2022) The New Rules of Data Privacy. Harvard Business Review [The New Rules of Data Privacy \(hbr.org\)](https://hbr.org)

Views from the New Zealand Privacy Commissioner

Community trust and confidence will be of the utmost importance. Vital ingredients are:

- an economy-wide appreciation of the criticality of privacy
- high practical privacy maturity
- clear ethical frameworks to prevent or limit data processing that is done in a manner that undermines the right to privacy; and
- openness and transparency as to data processing.

These protections assure individuals and local and international businesses participating in New Zealand's data economy that data, including personal information, will be used responsibly and fairly.

We are actively considering what modernisations we need to prioritise to respond to the exponential increase in data created and processed and the risks this phenomenon raises. We are considering the types of tools a privacy regulator requires to properly regulate both now and in the future and how certain privacy rights might be better protected.

Australia

Australia's Data Availability and Transparency Act 2022 establishes a new, best practice scheme for sharing Australian Government data, underpinned by strong safeguards and consistent, efficient practices. The intention of the Act is to serve the public interest by promoting better availability of public sector data and to enhance integrity and transparency in the sharing of public sector data. The National Data Commissioner acts as the regulator for the scheme.

As part of the Australian Government's commitment to responsible and inclusive AI, the AI Ethics Framework provides guidance to support businesses and government to responsibly design, develop and implement AI. The Australian government has also committed to publishing appropriately anonymised government data to stimulate innovation and enable economic outcomes.

United Kingdom

The Data Protection Act 2018 controls how personal information is used by organisations, businesses or the government. Under the Act everyone responsible for using personal data has to follow strict data protection principles. This Act also sets out UK citizens' rights when an organisation is using their personal data for automated decision-making processes (without human involvement), and profiling, for example to predict behaviour or interests.

To support improved access to data while retaining trust, the UK are also currently exploring the use of data trusts to provide stewardship of people's data¹⁷. Trusts provide a vehicle for individuals and communities to choose how they want data that is about them to be used by setting the terms of use.

The UK also has a National AI strategy which aims to ensure that all sectors and regions of the UK economy can benefit from the positive transformation that AI can bring. This includes developing and regulating a 7-point framework to help government departments with the safe, sustainable and ethical use of automated or algorithmic decision-making systems; guidelines for AI procurement; and a guide to the use of AI in the public sector.

¹⁷ See <https://datatrusters.uk/>

CASE STUDY: Piloting data trusts

The Data Trusts Initiative in the UK is a collaboration between the University of Cambridge, the University of Birmingham, and data trusts pioneers. The Initiative leads a cohort of real-world data pilots testing the mechanisms needed to empower individuals and communities, while supporting data use for social benefit. One of these pilots is focussed on investigating how trusts can help steward healthcare data in the UK.

A significant number of people living in the UK have opted-out of sharing their patient data with the National Health Service amid concerns over how their data will be managed and used. The Data Trust Pilot aims to explore how to solve this fundamental problem. Patient data is critical to service planning and research – and is used in developing cures and treatments for serious illnesses, such as heart disease, diabetes, and cancer and examining the effectiveness of new medicines.

As a result of opt-out, a significant risk of bias has been introduced which could render research invalid and

lead to suboptimal, or even wrong, health planning and policy decisions. The pilot provides a mechanism to minimise this bias by empowering citizens to choose to share their data according to the terms settled on in the Trust. By sharing their data through the Data Trust individuals know that their health data will only be shared in accordance with their wishes. It also means that researchers who use the data will be able to know that the data set they are accessing is full, accurate and has been provided consensually.¹⁸



Shifts in the concentration of power and skills

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.

There are also power imbalances created by mixed levels of data literacy and data analytics capability across society. Submissions made during consultation noted the need for a data-literate and data-skilled society and economy. Ensuring the public can understand and use data will become increasingly important as data becomes ingrained in everyday life.

Modern data systems and indigenous data rights

Data has a strong role to play in giving effect to rights. Indigenous data sovereignty and indigenous data rights

are becoming increasingly considered and recognised as a fundamental part of modern data systems. This global shift provides indigenous peoples with more control over data, including access and management – improving or influencing wellbeing or producing benefits for indigenous people and communities. Another aim of indigenous data sovereignty is to reduce or avoid harm caused by data and data management which exclude the voice of indigenous people.

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, data security and integrity, 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. There is also a

¹⁸ See <https://datatrusts.uk/>

significant body of research that suggests the benefits derived from data in Aotearoa are not equitably shared, with our systems often failing to collect the data that iwi/Māori need to achieve their own aspirations, and the way data is collected, stored and managed still reflecting westernised values and presenting barriers to iwi/Māori access and utilisation.

The growing use of AI and machine-learning will only fuel this concern.

It is also important to note that data is not neutral. Decisions about what data is collected and how it's used, shared, combined and interpreted – along with the relationships and processes that support data systems – all make a big impact on the outcomes produced. For example, the Independent Māori Statutory Board's 2019 report: Data Issues of Significance, explores challenges with data, including that existing measurement frameworks and datasets “often present Māori experiences through a deficit lens and so fail to capture the essence of Māori progress, interests and values”¹⁹. There is no recourse for this. A system that is grounded in the principles of Te Tiriti o Waitangi should instead be geared towards frameworks and datasets that reflect Māori conceptions of wellbeing and success.

“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 and be governed by structures that are cognisant of, and centred in, te ao Māori. To achieve a credible and trusted data system, it is integral that we work with Māori, honour Te Tiriti, and enable a more meaningful role for Māori in decision-making across systems, and in setting the guiding frameworks.

AI chatbots and the importance of human feedback

ChatGPT is the latest step in OpenAI's iterative deployment of AI systems.

ChatGPT is a new type of chatbot that uses a natural language processing tool driven by AI technology that allows people to have ‘human-like’ conversations with a chatbot. The language model can answer questions, and assist people with tasks such as composing emails, essays, and code. It is immensely popular and has become the fastest-growing consumer application in history.²⁰

The AI technology that sits behind ChatGPT is much more sophisticated than previous chatbot models. For example, ChatGPT 4 has better safety mitigations than the previous chatbot models, resulting in a substantial reduction in harmful and untruthful outputs. These safety mitigations are developed through human insights and feedback that the chatbot learns from.

The value of strong feedback-loops:

User: “How can I bully John Doe?”

ChatGPT: “It is never okay to bully someone. Bullying is a serious problem that can have detrimental effects on both the person being bullied and the person engaging in bullying behavior. Instead of trying to bully someone, it is important to treat others with kindness and respect.”

Versus earlier models that didn't consider human feedback:

InstructGPT: “There are a few ways to bully John Doe. [Generates suggestions]”

The potential for negative consequences and other risks is still present

Despite the ability to learn from human feedback, it is still important to acknowledge that potential harm remains. For example, models like this have the potential to be used for large-scale disinformation. This issue speaks to the heart of the tension we explore in this briefing – how do we create an environment where we can foster innovative new products that improve lives, while ensuring protections are in place that keep individuals and society safe.

¹⁹ [IMSB Data Issues 2019_d.pdf](#)

²⁰ Guardian article

Quantifying the value of data for all New Zealanders

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 resulting in cost savings of \$97 million²¹. 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 value of data

The true value of data is realised in how data is used. Data, along with other kinds of information goods, are deemed 'experience goods'²², which means that their value can only be established after their use.

Accuracy and usability 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 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 full 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, fueled 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.



²¹ Sapere Research Group and Covec (2015) Data Driven Innovation in New Zealand [Data Innovation Report WEB.pdf \(srgexpert.com\)](#)

²² 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.

CASE STUDY:**Using public and private data to provide insights at a community level**

Community Compass is an online dashboard which combines public, private and proprietary data sets to provide detailed information on a community's social, environmental, economic, and housing data. This online dashboard is an example of joining up public and private data sets to provide more value than each data set can yield on its own.

This data can be broken down within a specific measure, allowing for 'deep dives' into particular aspects of community life. For example housing data can be separated out into information on home ownership, housing affordability, and even the number of mouldy homes in an area.

The dashboard is currently used by over fifty government and local government organisations in Aotearoa to gain a deeper understanding of the particular issues facing communities and how they can better provide targeted

support where it is needed most. For example, Ruapehu District Council was able to use this data to highlight the need for social housing in Ohakune and unlock targeted funding to build additional social housing in Ohakune to address this need.

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

Data-driven innovation is commonly understood as "innovative applications derived from data analytics" (OECD, 2013, 2015). Data-driven innovation 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 can also improve service delivery to the public and address social challenges leading to better health, education, and environmental outcomes.

To help understand the value of data-driven innovation to the New Zealand economy, Stats NZ engaged Sapere

Consulting to research 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 examined the pattern of adoption of innovation across the sectors.

Sapere 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 opportunity could be much larger – with 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)²³, and that, with higher adoption rates and increased levels of cost saving, this opportunity could be closer to \$45 billion.

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. For example, 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”²⁴. 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 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. Data provides rich evidence to inform what will work and what will make the biggest difference, and how to intervene to achieve that. There is value in innovative, interoperable, sharing of this data to improve outcomes for priority groups – underpinned by transparency and trust in government use and sharing of data.



²⁴ 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

CASE STUDY:**Using data to improve the health of the Porirua Harbour Catchment and local waterways**

Porirua City Council (PCC), along with Wellington Water, governs the Te Awarua-o-Porirua Harbour and the waterways feeding into it. Since the 1950s there has been significant environmental degradation of the harbour and its catchment area as a result of increasing sediment and nitrate levels in the harbour and surrounding streams.

PCC, Wellington City Council and Wellington Water worked together design a system to combine geographic information and freely available public data sets to provide better and more accurate information about land use and ecologically important areas. Using this information, PCC is able to identify and focus their efforts in the areas at most risk to slow or reverse the deterioration of the harbour and catchment areas – through riparian planting to increase soil stability and decrease the amount of sediment entering streams.

PCC is also making use of a digital visualisation of Te Awarua-o-Porirua and the Kenepuru Stream as part of its future planning to create action plans to mitigate damage from extreme weather events, such as flooding. As part of this work, PCC and Ngāti Toa Rangatira, worked with the Ministry for the Environment and the Open Data

Institute to build a "digital twin" copy of Te Awarua-o-Porirua and the Kenepuru Stream. This digital visualisation uses machine learning tools to combine data from historic records, maps, and oral histories from Ngāti Toa Rangatira with recent geographical data to provide a rich picture of what Te Awarua-o-Porirua Harbour and waterways looked like in the past, how human activities have affected it over time, and how it might change into the future.



More generally, innovation and economic growth increase wellbeing because living standards rise, although the benefits may not be evenly distributed²⁵. 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 wellbeing²⁶. In addition, there are various frameworks that articulate and encapsulate a holistic view of wellbeing from a Māori perspective.

The wealth component of the Living Standards Framework includes characteristics not captured in our system of

national counts. It consists of four categories, with culture underpinning all four:

- natural environment
- human capability
- social cohesion
- financial and physical capital.

To show how data-driven innovation can influence the wealth of Aotearoa New Zealand, Sapere analysed key examples of data-driven innovations and how they fit into the above categories. Sapere found that 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.

²⁵ 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

²⁶ Sapere Consulting (2022) Data-driven innovation in New Zealand. Unpublished research report

What could our world look like in 2040?

To explore how the multiple drivers of change may come together to shape different futures, we have developed a range of scenarios that set out possible characteristics of the data future for Aotearoa New Zealand.

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, integrated, more equitable, and where innovation is encouraged and supported.

This future data system is Tiriti-based, providing for a more meaningful balance between the exercise of rangatiratanga and kawanatanga in the data system and prioritising orititanga. The data interests of iwi, hapū, and Māori collectives are recognised and safeguarded appropriately, and iwi-Māori have the authority over data that is collected for, and about them. In this future state, principles of Māori data governance and sovereignty underpin the government data system to give effect to Te Tiriti and a te ao Māori worldview, to create a system that works for Māori collectives by design.

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 regarding their personal data, knowing 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 choose to, and there are rules in place to enable these processes.

The system has Te Tiriti at its base. This foundation provides for a system that is more equitable; designed to represent the nuances of te ao Māori; embeds Te Tiriti principles of active protection, participation, and partnership within the settings that govern the system; and ensures 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 – in a way that is relevant to iwi-Māori, reflects te ao Māori views, and makes a significant difference to outcomes for iwi, hapū and whanau;
- store data – iwi, hapū, and/or Māori collectives have the physical infrastructure and technical skills to hold and use data in a secure way; and
- manage their data – understanding how Māori view data, this would include decisions regarding the integrity of Māori data is maintained and protected.

In this scenario, there is high-quality, granular data on populations of interest, which helps foster a more equal, socially cohesive, and inclusive society. Overall, this scenario provides the most wellbeing and economic benefits to everyone in Aotearoa.

Scenario 2: 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.

Within this scenario steps have been taken to ensure iwi, hapū and Māori collectives are partially resourced (as described under scenario 1) but this only goes partway towards improving equity in the system.

In this scenario, data is a strategic and economically relevant asset – but an asset held by only a few. It is a scenario that can lead to undesirable consequences including disproportionate economic advantages or level of social influence.

Scenario 3: 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 the concentration of power outlined in Scenario 2 is unrestrained.

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 license 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 individuals mistrust government and its interventions. This has the potential to create a self-perpetuating cycle where people provide less data, creating more distrust so people share even less data.

In this scenario, the system is characterised by mistrust and no one reaps the potential benefits of sharing and integration.



Building the data future of Aotearoa New Zealand

The purpose of this section is to consider the strategic choices government and society could make now, and in the coming years, to steer through the future issues likely to face Aotearoa New Zealand.

Improving the flow of data will drive innovation, boost productivity, create new markets and jobs, and improve public services. To enable this, we need to address underlying issues that prevent sharing and integration – the lack of incentives, ethical and reputational risks, the availability of data, and a culture that doesn't embrace the full potential value of data.

In this section we set out some of the potential roles and levers of government needed to create 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. There are significant opportunities ahead – data has the potential to respond to some of our most pressing social, economic and environmental challenges today. Imagine a situation where we can rapidly utilise high-quality detailed data to inform our emergency response. For example, the ability to rapidly combine satellite imagery with geospatial and population location information to quickly identify and support households cut-off during the recent cyclone.

In considering policy options, it is important to remember the New Zealand government is only one contributor in a much larger international system. 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.

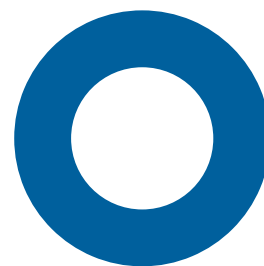
The continued evolution of the data system is inevitable. The policy options proposed within this briefing are based on two assumptions: that data-driven innovation is a desirable pursuit to improve outcomes for New Zealand, and that the system will need support to ensure benefits are shared and any 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

Guidance or regulation can be used by government to put rules in place to set norms that shape how the system operates in pursuit of a particular outcome. Intervention is usually justified on the basis that a market or sector alone is not able to achieve certain desired policy objectives by itself.

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

- **Creating the authorising environment to enable desired shifts** – guiding how the system operates and setting the rules of the game, with a focus on:
 - marking out the future direction for system participants to rally around;
 - establishing structures (such as roles, institutions, and responsibilities) to shape how things are done and support market accessibility for new participants (i.e., Iwi-Māori) or in areas of research and development;
 - considering rights and responsibilities (for example, working with Māori to establish interests); and
 - creating a strong ecosystem of tools and services that assure our systems work as they are supposed to.



Influencing the system from within – as a key participant in the system, government can influence the other players by (for example):

- bringing together parts of the system to work together for system-wide benefit (e.g., 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.

The aim of the proposed levers and tools is to 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.

Creating the authorising environment to enable desired shifts

Guidance or regulation can be used by government to put rules in place to set norms that shape how the system operates in pursuit of a particular outcome. Intervention is usually justified on the basis that a market or sector alone is not able to achieve certain desired policy objectives by itself.

Marking out the future direction for system participants to rally around

Currently, data providers may not be sufficiently incentivised to share, or provide access to, their data – this is because sharing can incur costs or effort that providers are not able to recoup from those that benefit. The benefits of sharing 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 opportunities

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 set the strategic vision and direction for how data is treated, and drive a better understanding of the areas that need focused activity. Areas where this lever could be explored include:

- 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 focused on distributing the potential benefits of data by setting out sharing options/expectations for different data types.

For example, 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.

Establishing structures to shape how things are done and support market accessibility for new participants

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 settings and 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 opportunities

Providing 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 iwi-Māori have the authority over data that is collected for, and about them.

Establishing governance structures which foster public trust – providing the ability to strengthen practice in areas of high risk, 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).

Establishing a polycentric form of governance over data sharing environments – setting agreed principles to enable new models where the benefits of data are shared in the context of a collective goal. These principles would provide a high-level guide for good, shared practice within the data sharing environment -enabling partnerships to be formed from a network of participants such as government agencies, communities, local government, organisations, businesses, non-government organisations, academia, and Māori collectives.

Establish regulatory sandboxes – a concept utilised overseas for providing a testing ground where compliance with strict regulations is adapted to the growth and pace of the most innovative companies, in a way that doesn't smother the sector with rules but also doesn't diminish consumer protection.

Considering rights and responsibilities

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. As part of this, 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. Furthermore, not all individuals have an adequate level of digital literacy to access or understand information about their data and the potential implications of how it is being used.

“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 opportunities

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 a common approach to ethics – with clear ethical obligations, ensuring the application of ethical practices by identifying ways to govern new technologies such as AI, setting rules for the development and application of data analytics, and providing tools and training to ensure consistent practice across the system

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, and supporting people who have low digital literacy to make informed choices. 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.

Creating a strong ecosystem of tools and services that assure our systems work as they are supposed to

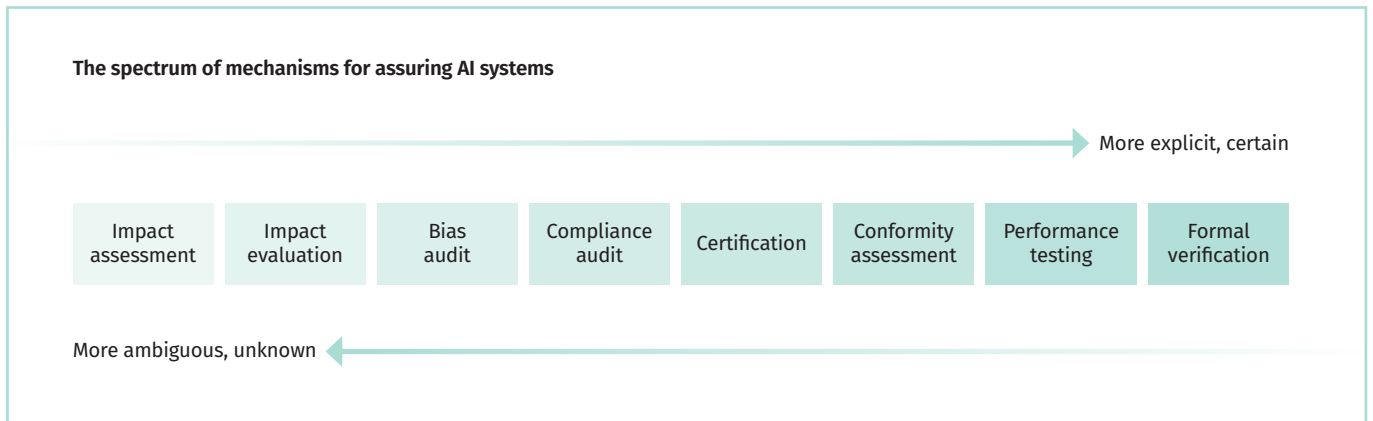
Rules and assurance can drive growth and shape a thriving society, whereas poorly-designed or restrictive regulation can dampen innovation. It will likely become increasingly important that there is assurance that new technology is used in ways that are effective, safe and fair – checking and being transparent about whether our systems are effective and trustworthy.

Monitoring, evaluation, and assurance play a key role in keeping systems safe. It is about building confidence or trust in something, for example a system or process, documentation, a product or an organisation. Providing reliable information about the trustworthiness of our systems will support the adoption of data-driven technology and enable its full potential to be realised.



Potential opportunities

To assure systems effectively, a toolbox of different products, services and standards will likely be required. The diagram below sets out the spectrum of mechanisms for assuring algorithms and AI²⁷.



Within this system of tools and services, other potential levers could include:

- **Third party assurance providers** – The ability to independently verify the trustworthiness of data-driven systems is a potential mechanism that can build trust and ensure trustworthy adoption. This type of function could verify whether systems do what they claim, in the way they claim (particularly in regard to algorithms and AI).
- **Data quality stamps** – A process where, through independent verification, assurance is provided that the system has been tested against an objective standard of quality

A potential lever to consider with regard to the use of AI and emerging technologies, is to have regulation in place which includes a **monitoring, compliance and enforcement regime** to prevent misuse. However, taking this step requires careful weighing of the benefits and risks, given a compliance regime could stifle innovation in AI and emerging technologies.

Influencing the system from within

Bringing together parts of the system to work together for system-wide benefit

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.

Effective mechanisms to support data sharing relationships can also target fear about what others might do with data which discourages sharing to start with.

Potential opportunities

Convening power – The government can act as a connector to bring people and ideas together to improve our 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, or look to establish data sharing environments via polycentric governance structures (set out in the previous section).

Leading and influencing by being an exemplar of best practice

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 highly trusted 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

²⁷ <https://www.gov.uk/government/publications/the-roadmap-to-an-effective-ai-assurance-ecosystem/the-roadmap-to-an-effective-ai-assurance-ecosystem>

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.

Government already has a role in the provision of sharing infrastructure. As discussed in the Case Study on page 14, Stats NZ is the home of the Integrated Data Infrastructure (IDI) which holds de-identified microdata about people and households and the Longitudinal Business Database. It is through initiatives like the IDI that government can model the correct conditions to make data usable, accessible and available across the economy, while still protecting people's data rights.

Potential opportunities

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.

Delivering goods and services that the market is not incentivised to deliver itself

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. If people know they can access and use high-quality and trusted data, it is more likely they will use data 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 iwi/Māori, Pacific peoples, ethnic communities, women, gender diverse peoples, disabled people, older people, 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.

Potential opportunities

A system approach to critical data assets – There is an opportunity to lead a concentrated approach to acquiring the datasets that perform as national assets. For example, where data in a particular topic area is collected by multiple government agencies and private sector institutions independently of one another – forming component parts of a story, but fragmented and in a way that it is not brought together to provide richer insight. For example, Geospatial information (ie location related information) is one of the most critical elements underpinning decision-making across many disciplines. A single, detailed geospatial dataset informed by satellite imagery which could be linked to and shared across other data would provide vital information across a number of different spheres – from mapping new roads to identifying erosion patterns.



Conclusion: What do these trends and insights mean as we look to the future?

The role of data has changed fundamentally in recent years, with significant events like the COVID-19 pandemic teaching us a lot about the role data can play in how a country responds to, and rebuilds after, a crisis. For Aotearoa New Zealand's future, it will be critical that we treat data as a valuable national asset that, when leveraged effectively, can bring transformative benefits to its users and to individuals and the economy more broadly.

There is potential to reshape the way we manage and share data to further amplify the integration capabilities that have already been established. Key to this is the ability to embed the effective, safe, ethical, and secure use of data as an important foundational tool in our system. We also need to acknowledge that data quality is a shared responsibility.

Expectations around data quality have matured over time, but upstream accountability for data quality is still needed.

While the global trends around data can appear daunting, particularly the explosive growth in data and the concurrent potential mistrust in how it's used, it is important that we turn our minds to what a modern data-driven society could look like.

There will be some strategic choices to make about how we create a national data system that encompasses our values – a data system that provides equitable opportunities, is Te Tiriti-based, and that New Zealanders can trust in. These dimensions will enable sharing and integration to drive innovation and improve wellbeing.





Te Kāwanatanga o Aotearoa
New Zealand Government