

Data Sharing Primer

January 2026

Contents

Document control	2
<i>Introduction</i>	<i>4</i>
<i>What is the IDI?.....</i>	<i>4</i>
<i>What data would need to be provided?</i>	<i>5</i>
Process	6
Data Extraction	8
Data Supply	9
<i>Next steps – working directly with Stats NZ</i>	<i>10</i>
<i>What do I get back?</i>	<i>10</i>

Introduction

A key part of Social Investment is ensuring government and the organisations that deliver social services - understand what works, and for whom. No matter how good your data collection is, you are not likely to have visibility over the wealth of data collected by government describing some of the experiences and outcomes of the individuals, whānau and families you are working with.

For example, you know who you are working with – but through data held in the IDI, we can understand who currently, is not being reached. We can support you to further understand:

- Insights about the individuals, whānau and families your initiative is reaching (aggregated into groups, not individuals) - such as what other supports they are receiving, their demographics and life circumstances visible in government data.
- Insights about other individuals, whānau and families (aggregated into groups, not individuals) - in the community your initiative is not yet reaching – helping your organisation understand who you might also reach.
- An estimate of the impact your initiative is having - how outcomes are changing for whānau, families and individuals participating in the initiative compared to similar groups who are not participants.

Using the Stats NZ Integrated Data Infrastructure (IDI) allows you to have visibility of this at a generalised (not individual) level, and where appropriate to learn from it and improve how you deliver support. This is provided on a regular basis through a dashboard created specifically for your initiative.

Our outcome-based agreements allow for organisations to define and focus on the goals and outcomes they include within their own theory of change. The IDI can support evaluation against agreed outcomes and provide valuable information that can be used to adapt overtime.

Partnering with SIA will also include building a shared impact and learning plan, we will support you along the way and ensure you have all you need to share minimal data to the IDI. This primer aims to answer some questions and help you on your journey.

What is the IDI?

The IDI is a large, trusted research database that holds de-identified information about people and households in Aotearoa.

The data is about life events like education, income, benefits, migration, justice, and health. It comes from government agencies, Stats NZ surveys and non-government organisations. The data is linked together, or integrated, to form the IDI.

Researchers (including within, and outside of, Government agencies) use the IDI to conduct cross-sector research that provides insight into our society and economy. The research can help answer questions about complex issues that affect New Zealanders.

Data in the IDI is de-identified. This means information like names, dates of birth, and addresses has been removed. Numbers that can be used to identify people, like IRD and NHI numbers, are encrypted (replaced with another number).

Stats NZ uses the '[Five Safes](#)' and '[Nga Tikanga Paihere](#)' frameworks to manage safe access to the rich source of information about New Zealand people, households, and Non-Government Organisations available within the IDI. Stats NZ ensures that it is only used safely in ways that benefit New Zealand.

After integrated data has had identifying information removed, only vetted and approved researchers can access selected, de-identified datasets for their specific project. This research must be for the public good. Users can only access the data in secure research facilities. Specifically, SIA will use this linked information for dashboard development to help provide insights back to you.

Importantly, even though data exists about de-identified individuals in the IDI environment, it can never leave that environment at the individual level. All data extracted from the IDI for analytical purposes is aggregated to remove any possibility of identifying individuals within the data.

What information is required to link data in the IDI?

You would need to supply the following information:

Personal Identifier(s)	Sufficient information for Statistics New Zealand to identify an individual, such as a unique identifier, for example NHI number, OR name (s), gender, and date of birth. This data is used only by Statistics New Zealand for linking in the IDI and is not shared with the Social Investment Fund.
Initiative Name	Name of the initiative (if you are delivering more than one initiative funded by SIA).
Start Date	The date the individual began working with your organisation.
Exit date	The date the individual ceased working with your organisation.

You can choose to supply other data at your own discretion as part of the IDI load – for example, you might have versions of your initiative where you want to be able to distinguish between different methodologies you are testing, in which case you might supply start and end dates for specific variations within your initiative.

Process

Documenting the information you collect and record is outlined below.

Step 1. Document how you record the details of those you work with

Example: *Every person who walks in the door is asked to sign up, using the following form. We also have an online version of the form. We save the information in our CMS tool / in a spreadsheet.*

Step 2. Document what you record about individuals you work with

Do you record key biographical information (e.g. Name, date of birth, gender) for everyone you work with or a unique identifier like National Health Index (NHI or Social Welfare Number (SWN) number?

Answer	Comment	Next steps
Yes - we collect at least one agency ID (like NHI or SWN) for all participants	Great! Linking is easy.	Briefly document how you get this Unique ID and whether you check or validate it.
Yes - we collect biographical details for all participants	Without IDs it is important your data is good quality, but it can be used to link.	Name, Date of Birth and Gender are used to link in the IDI. Briefly describe if you collect this information. On to the next question.
No	You will not be able to link the data into the IDI	If you don't current collect either of these, and wish to share data with the IDI, you will need to implement one of the above options.

Step 3. Document how you record the dates you provide support

Do you record the date that an individual or family started working with you and when they stop working with you?

Answer	Comment	Next steps
No	You may need to start capturing this information	Consider how you will collect and record this additional data collection.
Yes - we capture start and/or end dates	Great! This makes it easy to provide insights on how those you work with compare to others not working with you, and the impact of the support you provided to them while they were engaged.	Briefly document the process where this is recorded, and if and where you can see this in your system or spreadsheet. Example: When we 'create new person' in our system, it automatically captures a date, and we can see that in the 'Registration Date' field on the 'Enrolment' screen. When they age out of our service or we refer them to another social service organisation, we have an 'end service date' that we can manually enter.

Step 4. Document how you record the different services you provide and interactions with those you work with

Do you record each service or interaction with those you work with?

Answer	Comment	Next steps
No	If you offer more than one service, you may need to start capturing this information for Social Investment Outcomes Contracting, but even without it you can use the IDI to see great insights about your participants.	Consider whether you are comfortable with the additional data collection and whether it would be appropriate for your participants given the support you offer.
Yes - some initiatives	Excellent. You may need to capture initiative level information for initiatives you don't currently record if those initiatives are part of your Social Investment Outcomes Agreement.	Briefly document which initiatives are recorded, and how. Example: We capture engagement with young people we work with when they have a booking for a 1:1 session with one of our staff, but we don't capture it for group sessions. We have different activity types based on the type of staff member they see, e.g. GP visit, nurse visit. Activities have a

		<i>date and start time. We don't record duration and don't have a way to record those who do not show up, so there may be some activities that didn't actually happen but were scheduled.</i>
Yes - all initiatives	Great, that will likely make it easy to provide data for Social Investment Outcomes contracting	Briefly document how services are recorded. Example: We capture every engagement with those we work with and record the type of engagement (based on our internal work practices, e.g. Physical health, mental health, sexual health), the date and timestamp as well as the delivery channel (e.g. In person, phone, zoom). We also record when someone misses a scheduled engagement.

Data Extraction

If you are using a tool (e.g. Client Management System) of some kind (rather than a spreadsheet), do you have the ability to design your own queries and export data?

Answer	Comment	Next steps
No - we have a vendor	You will need to ask your vendor to either create an extract for you to share data with the IDI or ask them to provide you with a query tool and export function.	We can help work with your vendor to understand what will be needed
No - we built it ourselves but no longer have the capability in house	You will likely need a way to develop this capability.	You will need to look into how this might be possible – we are happy to provide some advice.
Yes - we can run predefined queries and export them	You may be able extract the data needed through a combination of one or more existing queries. Otherwise, you may need to ask your vendor to add a new query.	Check whether one or more (in combination) of the pre-set queries that would match the list of data covered above.

Yes, we can define our own queries and export them	Excellent, you are all set up to do the data sharing part of Social Investment Outcomes Agreement.	
---	--	--

Data Supply

Are you planning to send data directly to Stats NZ?

Answer	Comment	Next steps
Yes – directly	You will have a data sharing agreement with Stats NZ.	See the notes below about the agreement with Stats NZ.
Yes – via our vendor	You need to be comfortable that the data being sent is what you have agreed to. You will have a data sharing agreement with Stats NZ and your vendor will be sending on your behalf.	See the notes below about the agreement with Stats NZ
No – through an intermediary or commissioning organisation	If you are providing your data via a third party (e.g. your own head office for a collective, through a data brokerage, or through your commissioning organisation if that is where your contract is), you will have a data sharing agreement with Stats NZ.	You may need your own data sharing agreement with the intermediary – have a conversation with them about this.

Next steps – working directly with Stats NZ

If you are supplying your data directly to the IDI (through Stats NZ), they will expect you to sign a contract (data sharing agreement) and a confidentiality agreement and to complete a data specification.

The main **data sharing agreement** with Stats NZ describes the purpose of the data sharing and outlines the protections that Stats NZ has in place (along with any conditions agreed between your organisation and Stats NZ). The **data specification** has all of the technical information about the data being sent to Stats NZ, e.g. the information you have documented in the steps above. Stats NZ will help you document this in the specification, including key practical information like the frequency of data supply and what data specifically will be sent.

Stats NZ currently has a secure data transfer service that you will be set up with a login for. This service is fully accredited to ensure that it is safe and secure. They will test this secure data transfer with you to make sure you can upload data and that they receive it.

Before Stats NZ enter into a data sharing agreement with you, they will undertake an independent assessment. This assessment will be to check that the personal information collected is appropriate to be shared and integrated in the IDI. The SIA can work with you to provide the necessary information for this, but it is at the discretion of the Stats NZ to enter into this agreement. If they have any concerns, you may not be eligible to integrate your data into the IDI.