NEW ZEALAND RESEARCH INFORMATION SYSTEM

DATA IN NZRIS

This information sheet provides a simplified summary of the full NZRIS data specifications, available at **www.mbie.govt.nz/nzris**. This information sheet is designed to be read in conjunction with our one-page diagram "Data in NZRIS at a glance".

MINISTRY OF BUSINESS, INNOVATION & EMPLOYMENT
HĪKINA WHAKATUTUKI

New Zealand Government

INTRODUCTION

THIS INFORMATION SHEET COVERS:

WHAT NZRIS DATA INCLUDES

DATA IN NZRIS RELATES TO THREE MAIN AREAS:

- Funding and non-monetary resources for research – referred to as “Asset Pool” data
- Research activity – referred to as “Research, Science and Innovation” or “RS&I” data
- Data about the person/s and organisation/s involved, and other contextual data – referred to as “Other” data.

Asset Pool data will generally include information about the asset pool such as the type of resource, how it’s being distributed and who the recipients are.

RS&I data will generally include information about the research activity, such as the project and its outputs, that are enabled by the asset pool.

Other data includes details of the person or people and organisation/s involved in the research, as well as submission and protection data.

MAKE-UP OF THE DATA

NZRIS data is made up of entities, sub-entities and elements, which can be mandatory or optional. All data must have an associated protection pattern(s).

A data entity is a commonly used term in data modelling, where data needs to be broken down into its smallest parts. An entity is usually a recognisable object, either concrete or abstract. An example of an entity in NZRIS would be “Asset Pool”.

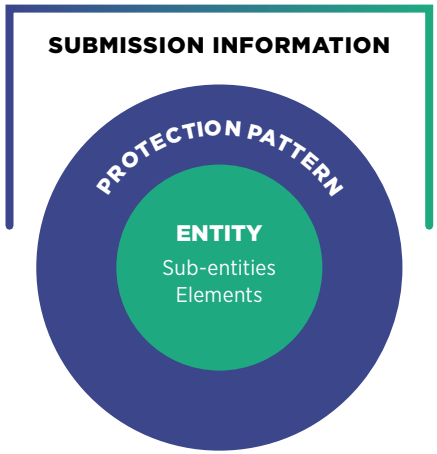
Entities are made up of sub-entities and elements. These provide context and meaning to the entity. “Asset Pool Name” and “Establishment Date” are examples of sub-entities and elements for the “Asset Pool” entity.

Entities, sub-entities and elements can be mandatory or optional. Mandatory means it must be included in the submission and optional means it does not have to be included. Some entities are mandatory only if they relate to the use of public funds.

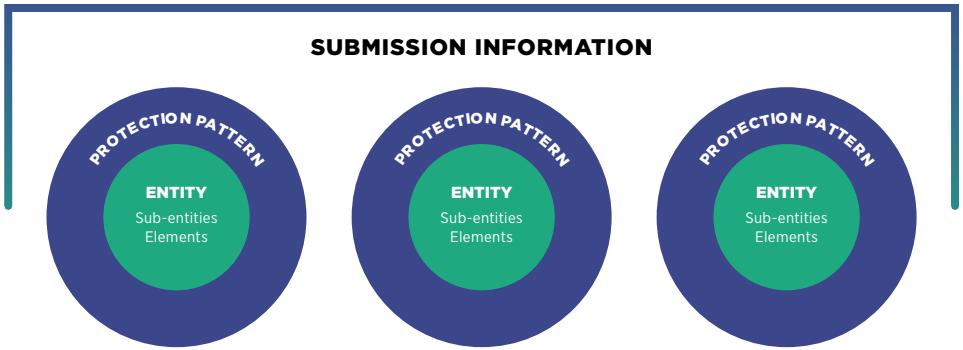
Protection patterns must be identified for each entity, to make it clear whether data needs to be kept private, or whether it can be published.

All data submitted to NZRIS must also be accompanied by submission information which provides context about the submission – such as the organisation submitting the data and what the data is about.

The minimum data submission that could be made to NZRIS would be one entity (as long as it had all the mandatory elements). Here’s an example of how the different data parts would work together if only one entity was included in the submission:



It is more likely that a submission would contain more than one entity. Here’s an example of how a submission would look where more than one entity was part of the submission:



WHO PROVIDES THE DATA

NZRIS data is provided by organisations in the research, science and innovation sector known as “data providers”.

Data providers are defined according to the type of data they provide.

Organisations that provide asset pool data are referred to as “Asset Pool Managers”. They are organisations that carry out funding and resourcing of research. MBIE and the Health Research Council are examples of Asset Pool Managers.

Organisations that provide RS&I data are referred to as “Research, Science & Innovation Managers” or “RS&I Managers”. They are organisations that undertake research. Crown Research Institutes and universities are examples of RS&I Managers.

Most organisations in the research sector perform only one of these functions, but some may carry out both functions – for example, where an organisation receives funding to undertake research, and then provides a portion of that funding to another organisation to undertake additional research. In this case they would act as an Asset Pool Manager when providing Asset Pool data, and a RS&I Manager when providing RS&I data.

WHEN DATA NEEDS TO BE PROVIDED

NZRIS is being rolled out in three phases. Each phase builds on the previous one by incorporating more data providers and increasing the scope of the data.

Phase 1 runs until late 2019, Phase 2 from late 2019 until mid-2020 and Phase 3 from mid-2020 until 2021.

In Phases 1 and 2, NZRIS will hold only Asset Pool data, which means that only Asset Pool Managers need to provide data. For Phase 1, the Asset Pool Managers providing data include the Ministry of Business, Innovation and Employment, the Health Research Council, the Royal Society Te Apārangi and the Ministry for Primary Industries.

Phase 2 will focus on bringing two new Asset Pool Managers on board – Callaghan Innovation and the Tertiary Education Commission.

The focus of Phase 3 is including RS&I data as well as Asset Pool data, with RS&I Managers such as Crown Research Institutes, universities and other research organisations coming on board as data providers.

WHAT DATA IS IN EACH PHASE

This diagram describes what data NZRIS will hold in each phase.

DATA TYPE
- ASSET POOL DATA
- RS&I DATA
- OTHER DATA

REQUIREMENT
- MANDATORY
- MANDATORY IF USING PUBLIC FUNDS
- OPTIONAL

PHASE 1

In this phase, Asset Pool Managers are the only data providers. All entities within this phase are either mandatory, or mandatory if using public funds. Only mandatory sub-entities and elements need to be provided.

ASSET POOL
EG: ASSET POOL NAME, PURPOSE AND ESTABLISHMENT DATE.

AWARD GRANTED
EG: AWARD TITLE, DESCRIPTION, TYPE AND RECIPIENT.

RESOURCE DISTRIBUTED
EG: RESOURCE TYPE, RESOURCE QUANTITY, AND THE ASSET POOL AND AWARD IT IS LINKED TO.

PERSON
EG: NAME, SECTOR IDENTIFIER AND ORGANISATION AFFILIATION.

ORGANISATION
EG: ORGANISATION NAME, TYPE AND OTHER IDENTIFIERS.

SUBMISSION
EG: CONTACT EMAIL, COVERAGE, CAVEATS AND SOURCE.

PHASE 2

In this phase, Asset Pool Managers are still the only data providers. In addition, the data expands to include optional entities, sub-entities and elements. The diagram below shows the entities and selected examples of sub-entities and elements for each entity.

PHASE 1 DATA
ALL DATA FROM PHASE 1 IS REPEATED WITH THE ADDITION OF OPTIONAL SUB-ENTITIES AND ELEMENTS

APPLICATION
EG: APPLICATION ID, TYPE, PROPOSAL TITLE AND OUTCOME GOAL.

APPLICATION REVIEW
EG: REVIEW DATE, REVIEW METHOD AND REVIEW OUTCOME.

APPLICATION DECISION
EG: DECISION, DECISION DATE AND DECISION MAKERS.

PLANNED DISTRIBUTION
EG: RESOURCES TO DISTRIBUTE, ALLOCATION METHOD AND DISTRIBUTION PERIOD.

GROUP
EG: GROUP NAME, PURPOSE AND IDENTIFIER.

PHASE 3

In this phase, RS&I Managers can also become data providers, and the data in NZRIS expands to include RS&I data. The diagram below shows the entities and selected examples of sub-entities and elements for each entity.

PHASE 1 DATA
ALL DATA FROM PHASE 1 IS REPEATED WITH THE ADDITION OF OPTIONAL SUB-ENTITIES AND ELEMENTS

PHASE 2 DATA
ALL DATA FROM PHASE 2 IS REPEATED WITH THE ADDITION OF OPTIONAL SUB-ENTITIES AND ELEMENTS

PROJECT
EG: PROJECT TITLE, DESCRIPTION, START AND END DATES AND RELEVANT ANZSRC CODES.

AWARD RECEIVED
EG: AWARD TITLE, DESCRIPTION, TYPE AND RECIPIENT.

RESOURCE RECEIVED
EG: RESOURCE TYPE, RESOURCE QUANTITY, AND THE ASSET POOL AND AWARD IT IS LINKED TO.

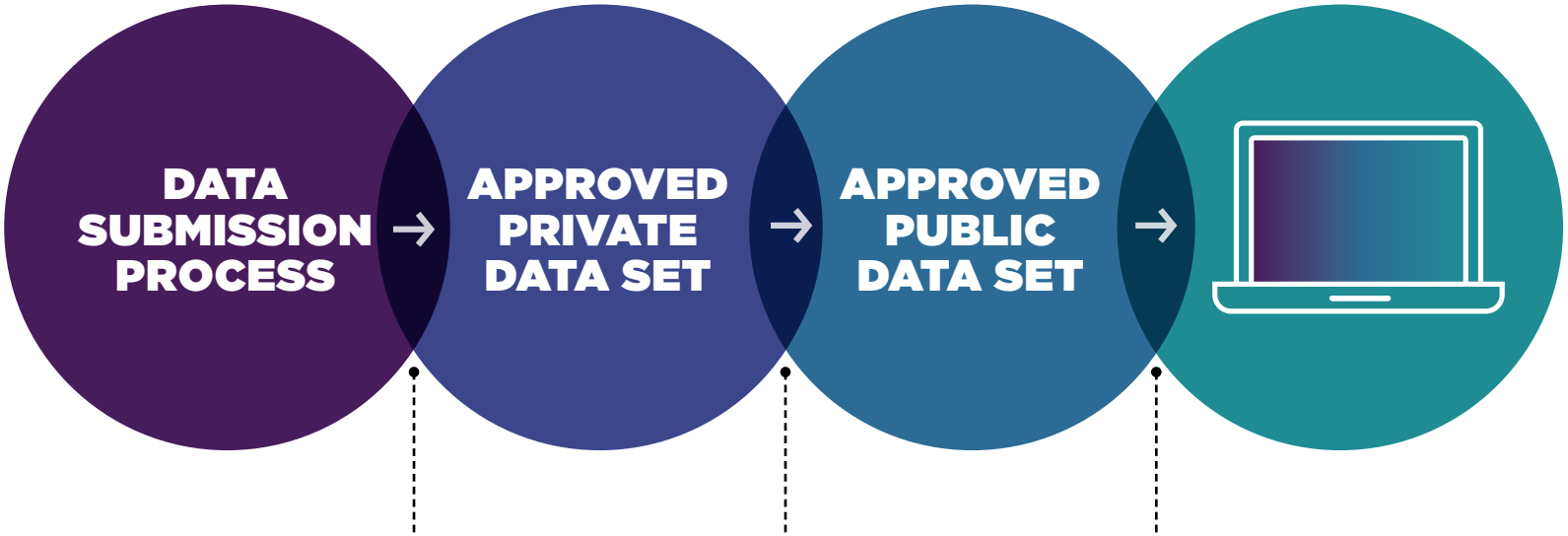
OUTPUT
EG: OUTPUT TITLE, OUTPUT TYPE AND OUTPUT DESCRIPTION.

HOW DATA IS PROTECTED

Not all data submitted to NZRIS will be visible on the public website. This is because some of the data needs to be protected from public view – such as some personal information, details about an award application, or details that are commercially sensitive.

Data providers are responsible for deciding what data needs to be kept private and what data can be published. This is done through the use of protection patterns, which the data provider needs to identify for each entity during the submission process. Data that is protected will remain in a specially-designated private data set that is visible only to the NZRIS custodian and the authorised data provider. Note that the data provider will only be able to view the data they have submitted – they cannot see protected data from other data providers. Data that is for public view is held in a different data set, before it progresses to the public website.

HERE'S HOW IT WORKS:



All data, once validated against the data specifications and business rules, and approved by the data provider and NZRIS custodian, progresses to the approved private data set.

All data that needs to stay protected will remain in the approved private data set. Data that can be published will progress to the approved public data set.

Data is periodically released from the approved public data set to the public website by the NZRIS custodian.

HOW PROTECTION PATTERNS WORK

THERE ARE 10 PROTECTION PATTERNS IN TOTAL FOR PHASE 1.

Each pattern has a unique and specific purpose. For example, the "Application Protection" pattern protects information about an application until such time as the application is confirmed as successful. This protection pattern, when applied to specific entities, will ensure that all application-related data within that entity is kept in the private data set and not released to the public data set. Another example is the "Commercial Financial" pattern which is designed to protect financial information that is considered to be commercially sensitive. This pattern can be applied to a range of entities, and ensures that any financial information for that entity is kept private.

As each entity potentially contains so many data elements, up to three protection patterns can be applied to some entities to create the desired effect. For example, a data provider that wants to keep a person's name and professional details private for the Person entity would apply the "Personal Identifiable" pattern as well as the "Personal Professional" pattern.

Protection patterns must be identified for all entities, even if the information doesn't need to be protected – in which case the "No protection needed" pattern would be used.

If the data provider is uncertain about which pattern to use or none of the patterns achieve the desired result, then there are two "catch-all" patterns that can be used – "Work in progress" protection pattern and "Unknown protection" pattern. These patterns allow the data to be protected until further notice, giving data providers time to work with the NZRIS custodian on the best way forward (which could be developing a new pattern).

The table below lists the different protection patterns and how they work.

PATTERN NUMBER	PATTERN NAME	PATTERN PURPOSE	HOW IT IS APPLIED	WHAT FIELDS ARE SUPPRESSED
01	Application protection	To protect information relating to an application	This protection pattern can be applied to the following entities: › Application › Application Decision › Person, Organisation and Group entities if they are linked to the Application and Application Decision entities	This protection pattern will suppress all data within the entity it is applied to.
02	Application review	To protect information about the review of an application, including scores	This protection pattern can be applied to the following entities: › Application Review › Person, Organisation and Group entities if they are linked to the Application Review entity › Organisation	This protection pattern will suppress all data within the entity it is applied to.

PATTERN NUMBER	PATTERN NAME	PATTERN PURPOSE	HOW IT IS APPLIED	WHAT FIELDS ARE SUPPRESSED
03	Commercial financial	To protect commercially-sensitive financial information	This protection pattern can be applied to the following entities: › Asset Pool | Other Resource › Application › Award Granted › Resource Distributed › Award Received › Resource Received	This protection pattern will suppress the elements of an entity that include commercially-sensitive financial information.
04	Topic identifiable	To protect information about topics of research	This protection pattern can be applied to the following entities: › Application › Award Granted › Project › Award Received › Output	This protection pattern will suppress the elements of an entity that identify the research topic.
05	Person identifiable	To protect personally identifiable data	Person	This protection pattern will suppress the elements of an entity that can be used to identify a person.
06	Person demographic	To protect information relating to a person's demographic	This protection pattern can be applied to the Person entity	This protection pattern will suppress the elements of an entity that relate to a person's demographic details.
07	Personal professional	To protect information about a person's profession and where they work	This protection pattern can be applied to the Person entity	This protection pattern will suppress the elements of an entity that identify the research topic.
08	No protection needed	To indicate that the information can be made public and doesn't need to be protected	This protection pattern can be applied to all entities	No fields
09	Work in progress protection	To protect information that isn't finalised or requires more work	This protection pattern can be applied to all entities	All fields
10	Unknown protection	For use when protection is required, but none of the other protection patterns apply	This protection pattern can be applied to all entities	All fields

Here's an example showing the impact of applying protection pattern 03, which is "Commercial Financial", to the "Award Granted" entity:

WITHOUT PROTECTION PATTERN

Award Granted Entity

ATTRIBUTE	VALUE
Data Provider ID	999
Data Owner ID	
Local Asset Pool ID	AP553-RCXS
Local Distribution ID	479
Local Application ID	PROP-2959-YMOYY
Local Award ID	CONT-20527-AZSYO
Local Contract ID	RN4481K70
Award Title	Funding for New Model of Sea Temperatures
Award Description	Funding for New Model of Sea Temperatures in the Pacific Ocean
Award Start Date	1/01/2019
Award End Date	1/01/2020
Award Status	Active
Award Type	Organisation
Organisation Identifier Type	NZBN
Organisation Identifier Value	9312345678920
Organisation Role	Lead Organisation
Person Identifier Type	ORCID
Person Identifier Value	0000123456789111
Role	Lead Contributor
Local Resource ID	123546
Resource Type	NZ Public Sector Financial Resource
Minimum Resource Quantity	100000
Maximum Resource Quantity	100000
Resource Measure	NZD
Resource Value	100000
Allocated Resource Start Date	1/01/2019
Allocated Resource End Date	1/01/2020
ANZSRC Type of Activity	Pure Basic Research
Share of ANZSRC Type of Activity	100%
ANZSRC Field of Research	040501
Share of ANZSRC Field of Research	100%
ANZSRC Socioeconomic Objective	960999
Share of Socioeconomic Objective	100%

WITH PROTECTION PATTERN

Award Granted Entity

ATTRIBUTE	VALUE
Data Provider ID	999
Data Owner ID	
Local Asset Pool ID	AP553-RCXS
Local Distribution ID	479
Local Application ID	PROP-2959-YMOYY
Local Award ID	CONT-20527-AZSYO
Local Contract ID	RN4481K70
Award Title	Funding for New Model of Sea Temperatures
Award Description	Funding for New Model of Sea Temperatures in the Pacific Ocean
Award Start Date	1/01/2019
Award End Date	1/01/2020
Award Status	Active
Award Type	Organisation
Organisation Identifier Type	NZBN
Organisation Identifier Value	9312345678920
Organisation Role	Lead Organisation
Person Identifier Type	ORCID
Person Identifier Value	0000123456789111
Role	Lead Contributor
Local Resource ID	123546
Resource Type	NZ Public Sector Financial Resource
Minimum Resource Quantity	[redacted]
Maximum Resource Quantity	[redacted]
Resource Measure	NZD
Resource Value	[redacted]
Allocated Resource Start Date	1/01/2019
Allocated Resource End Date	1/01/2020
ANZSRC Type of Activity	Pure Basic Research
Share of ANZSRC Type of Activity	100%
ANZSRC Field of Research	040501
Share of ANZSRC Field of Research	100%
ANZSRC Socioeconomic Objective	960999
Share of Socioeconomic Objective	100%