Nau mai haere mai ki

Introduction to Data Management Part One

data.govt.nz

Am I in the right place?

The intended audience for this e-learning is:

- NZ government departments
- NZ not-for-profit organisations
- Iwi and Māori organisations

The content is for those who are new to data management.

It *is* assumed you understand basic data terms. If you are unsure, please look at the Introduction to Data module.

There is no assessment at the end of this e-learning.

What can I expect from this e-learning?

This e-learning will help you:

- Understand the lifecycle of a dataset.

- Learn what a data management plan (DMP) is and why it is important.

- Learn the parts of a good data management plan.

- Complete your own basic plan.

- Know where to go next for more information or help.

What is the lifecycle of data?

Data has a lifecycle.

Data management is a way of (positively) influencing how your data moves through this cycle.

A data management plan documents all the information and decisions made about the data.

Why manage data?

Demand

The increased demand for evidence-based decision-making has increased the demand for data. The uses for data are also rapidly changing, requiring more versatility.

Transparency and security

If people are going to continue to provide data, they need to know organisations are taking good care of it.

Use and re-use

Data has the potential to be used more than once, and not just for one purpose. This means the 'metadata' (which describes and gives the context for the data) is as important as the actual data.

Legislative requirements

The Public Records Act (2005) requires that *all* data collected as part of government business is managed, until it is archived or destroyed.

Data management is like running a library

Data management plan (DMP): the big picture view and record of how a dataset moves through its lifecycle.	The library's architectural plans and its strategy.
Datasets: the building blocks that will create statistics, maps, mobile apps, etc.	The books.
Metadata: the description of and information about the datasets.	The library's catalogue entries.
Statistics: the interpretation of the data. These results will be different if any of the parts are altered.	The creativity, knowledge and development that comes from reading the book.

A good DMP:

- Manages the dataset as well as describes (give information about) it.

- Is a gateway to everything to do with the dataset. It must be clearly linked to the dataset, the metadata and any other relevant documents/ records. If the dataset is small, the plan may even contain the metadata about the dataset.

Continued on the next slide...

What do I need to know about DMPs (cont.)?

- Should always be put in place, even if sometimes this means adding it in order to manage data retrieved from a long-existing system, like a taxation one.

- Can apply to more than one dataset if the datasets' governance and related documentation are the same.

- Shows that the data is being handled safely and securely.

- Identifies any legislative or contractual requirements for accessing or using the data.

Read on to learn about the parts of a good plan.

What are the parts of a good DMP?

1) Governance and access

    👥   Governance

    🔒   Access and security

2) Discovery, use and re-use

    📋   Data documentation

    📦   Data formats, volume and storage

3) Retention, preservation and disposal

Carry on to learn about each of the parts.

Parts: 1) governance and access

- *Governance* is about properly looking after the dataset, and applies to both individuals and organisations:

    - An *individual* is accountable for their datasets: knowing how to access them, how to keep them secure and how the datasets contribute value inside and outside their organisation.

    - An *organisation* is accountable for how it manages its data assets, so they are accessible, secure, usable and re-usable.

Go to the next slide for more on governance.

Parts: 1) governance and access

- Responsibilities for data management begin with the data creators/collectors. They need to be sure, for example:
  - Who will be responsible for the dataset.
  - How informed consent will be handled.
  - That legislative requirements will be met.
  - That relevant principles/frameworks are followed.
  - That documents regarding decisions are stored securely and are discoverable and accessible.

Go to the next slide to start creating your own DMP.

Parts: 1) governance and access

The easiest way to make sense of a data management plan is to create one on your own. Open this link, and then save the document to your own system.

Read the instructions on page 1 and then work through the section called Governance (page 3). Once you have finished that section, come back to this e-learning.

Parts: 1) governance and access

🔒 *Access and security* are important to those who provide data.

We need to be transparent about the openness

of data we collect, including access limits:

- How will we manage security?

- How will we make sure the data remains uncorrupted?

- What barriers might there be that would stop sharing?

Go back to your plan. Work through the section called Access and Security (page 5) and then come back to this e-learning.

🗎 *Data documentation*

- Describe how the dataset was extracted/created.

- Enable use and re-use by explaining what the data items mean.

- Use consistent names to identify the data through its lifecycle (i.e. raw, processed, final).

Go back to your plan. Work through the section called Data documentation (page 6) and then come back to this e-learning.

Parts: 2) discovery, use and re-use

The users of any particular dataset do not usually have the opportunity to talk with the creators. So describing the dataset is vital: it means the data can be discovered, used and potentially re-used.

- *Data formats, volume and storage:*

    - What form will the dataset be kept in?

    - What software is needed to use the dataset?

    - Where is the dataset stored?

    - What is the size of the dataset?

Go back to your plan. Work through the section called Data formats, volume and storage (page 7) and then come back to this e-learning.

Managing datasets well means both creators and users know that the data is being looked after. For example, they know:

- How long datasets will be kept and how long-term access will be managed.

- How decisions will be made about disposal.

- Reviews are scheduled in the data life cycle.

- The final archived data is read-only.

Go back to your plan. Work through the section called Retention, Preservation and Disposal (page 8) and then come back to this e-learning.
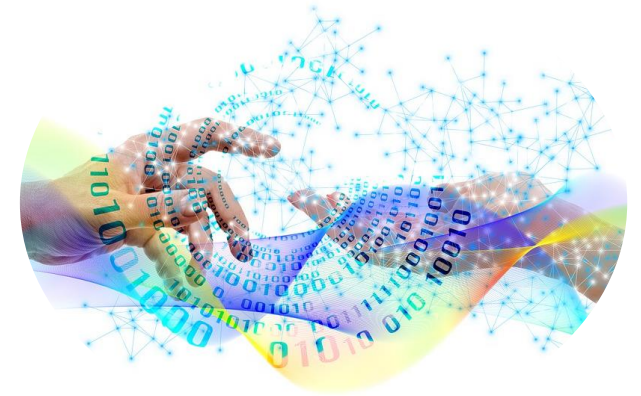
How do you practically implement a DMP?

Make sure that your data management plan (DMP) is:

- Easily accessible

- User-friendly

- Easy to maintain

If your organisation does not have a DMP template, you can use the one in this e-learning, or this Excel-based [example](#).

You can find other simple Word or Excel versions online and even software programmes that collect data, put it into a database and create interactive reports.

Where should a DMP live?

- All your DMPs should be stored together in a shared location.

- Each DMP should clearly link to its datasets, relevant documents and shared drives as appropriate.

How can I learn more?

This link is to a good guideline for data management (N.B. it is research-data focussed) if you would like to explore data management plans more. You'll need to create an account, but it is worth it.

The next introductory module on this site will be Introduction to Data Management Part Two. Keep an eye out!

Explore the data.govt.nz site further.

Learning....

Contact: datalead@stats.govt.nz