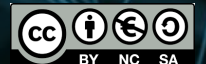# Managing Research Data
Principles, policies and practices

Laura Armstrong & Dr Sarah Hopkins
Centre for eResearch

June 2025

Karakia

Whakataka te hau ki te uru

Whakataka te hau ki te tonga

Kia mākinakina ki uta

Kia mātaratara ki tai

E hī ake ana te atakura

He tio, he huka, he hauhū

Tihei Mauri Ora!

Cease the winds from the west
Cease the winds from the south

Let the breeze blow over the land
Let the breeze blow over the ocean

Let the red-tipped dawn come with a sharpened air
A touch of frost, a promise of a glorious day.

Managing research data

Learning outcomes

At the end of this workshop, you will be able to:

- Describe the importance of **good research data management**

- Identify relevant **policies & processes** for research data management

- Apply **policies, data classification & best practices** to your research across the research data lifecycle

Activities

1. Introductions

- **Complete the Zoom poll to help us understand who's in the audience today**

- **Share in the Zoom chat what questions or issues you are hoping to explore or have answered in this workshop.**

What are research data?

The evidence that underpins the answer to a research question and can be used to validate findings regardless of its form (e.g., print, digital, or physical).

Data or artefacts, cultural taonga, research evidence, and digital representation of a physical item used in research.

> samples, interviews, images, surveys, observations, audio/visual recordings, medical records, maps, instrument data, spreadsheets, bibliographies, manuscript annotations...

Is software/code research data?

Maybe, it depends.

Software/code may be the focus or by-product of the research.

Does the software or code support:

- **Reproduction** - enabling others to find, access & run exactly same software, inputs & computational environment to verify/validate your results.

- **Replication** - using *similar* inputs, tools, environments, to arrive at mostly the same outputs and conclusions to justify the results.

The Turing Way is a handbook to reproducible, ethical and collaborative data science.

What is research data management?

Process of planning and undertaking the collection, organisation, management, storage, backup, preservation and sharing of data before, during and after the project.
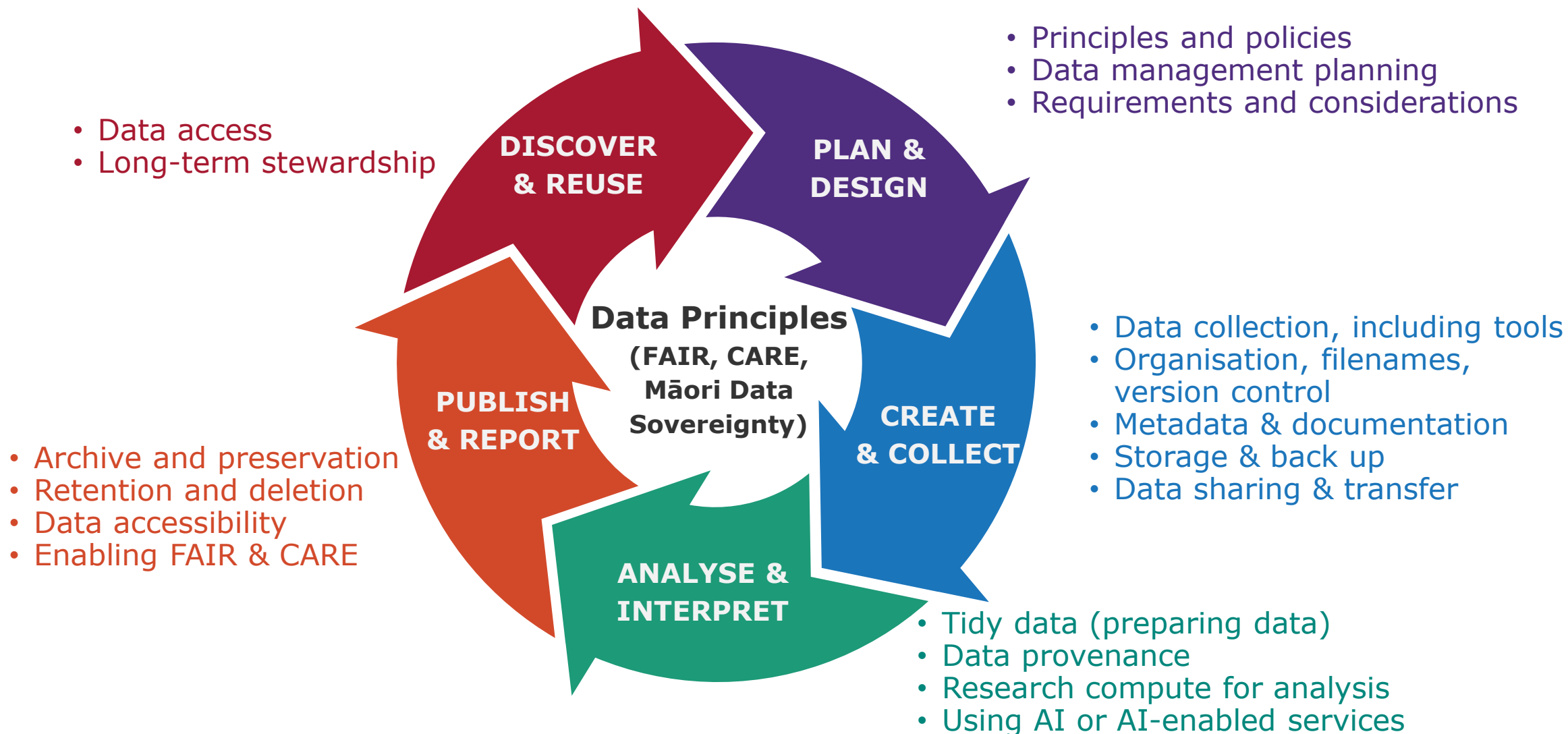
| Understanding |
| Integrity |
| Collaboration |
| Impact |
| Other environments |

A curated list of RDM resources for researchers and organisations (Mannheim University, Germany)
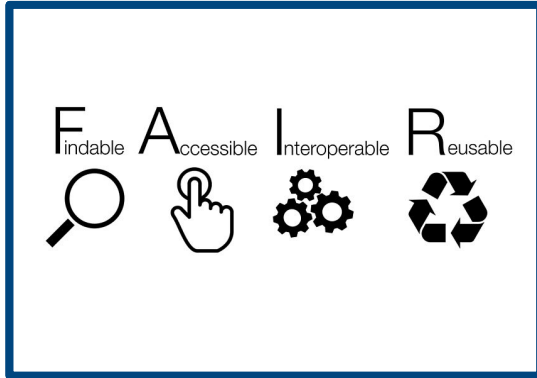
Research data lifecycle

- Data access
- Long-term stewardship

- Principles and policies
- Data management planning
- Requirements and considerations

DISCOVER & REUSE

PLAN & DESIGN

Data Principles (FAIR, CARE, Māori Data Sovereignty)

PUBLISH & REPORT

CREATE & COLLECT

- Archive and preservation
- Retention and deletion
- Data accessibility
- Enabling FAIR & CARE

- Data collection, including tools
- Organisation, filenames, version control
- Metadata & documentation
- Storage & back up
- Data sharing & transfer

ANALYSE & INTERPRET

- Tidy data (preparing data)
- Data provenance
- Research compute for analysis
- Using AI or AI-enabled services

Plan and Design

Plan for data management throughout the research data lifecycle.

A culture shift for RDM
What has driven changes over the last decade?

Expectations	Technology	Data sovereignty	Privacy & security
•"...maximise knowledge output from funding" • Reduce **duplication** and increase **reuse** of data • **FAIR principles** (Findable, Accessible, Interoperable, Reusable) • **Reproducibility**	•Increased capacity to generate, store & work with very large datasets, •Cloud computing •Artificial intelligence (AI) •Lower computing costs •**Digitisation**	•**Māori Data Sovereignty** principles •**CARE** data principles	•**Legal, ethical** and **protective security** •Managing risk

Researchers want to do the right thing but want clarity/direction on best practices, available services & support.

Policies related to research data

Plan and design | Data policies

Funder policies
Concerned with obtaining the best outcomes for the research they fund.
Examples: •MBIE Open Research policy

Publisher policies
Some publishers have policies or mandates related to data availability.
Examples: •Data Availability | PLOS One

Government requirements
Legislation or guidelines that impact the handling or management of research data.
Examples: •Trusted Research - Protective Security Requirements guidance

Professional codes of conduct
Concerned with the good conduct of research.
Examples: •Research Charter for Aotearoa New Zealand •Royal Society Professional Code of Conduct

Institutional policies related to research data management →

Data Management Planning

Data Management Planning is about preparing for data management across the research data lifecycle.

• Institutional, Ethics Committee or Funder requirement?
  • MBIE, NIH, Wellcome Trust
  • HDEC, University RDM Policy
• Risk-orientated and/or culture change approach
• Project specific
• Prompts conversations, captures decisions, clarifies roles and responsibilities and helps you to align with University policies and processes

Checklist for a Data Management Plan from the **Digital Curation Centre**

Requirements and Considerations

Plan and design | Considerations

Researchers should ensure that **legal, ethical, data sovereignty, protective security and commercial constraints** relating to research data are considered prior to data collection and adhered to throughout the research data lifecycle.

Summary: Plan and design

Plan and design | Summary

• Become familiar with the University's **Research Data Management policy** and other research-related data policies to understand the requirements for research data management.

• Use the new online **Data Management Plan (DMP)** to consider and document important decisions about managing research data throughout your project.

• Consider and document **ethical/legal** obligations, **sensitive data** considerations, **indigenous data sovereignty** and **contracts and copyright** responsibilities before collecting research data.

A culture shift for RDM

What has driven changes over the last decade?

Expectations

- "…maximise knowledge output from funding"
- Reduce **duplication** and increase **reuse** of data
- **FAIR principles** (Findable, Accessible, Interoperable, Reusable)
- **Reproducibility**

Technology

- Increased capacity to generate, store & work with very large datasets,
- Cloud computing
- Artificial intelligence (AI)
- Lower computing costs
- **Digitisation**

Data sovereignty

- **Māori Data Sovereignty** principles
- **CARE** data principles

Privacy & security

- **Legal**, **ethical** and **protective security**
- Managing risk

Researchers want to do the right thing but want clarity/direction on best practices, available services & support.

Policies related to research data

Funder policies

Concerned with obtaining the best outcomes for the research they fund.

Examples: •MBIE Open Research policy

Publisher policies

Some publishers have policies or mandates related to data availability.

Examples: •Data Availability | PLOS One

Government requirements

Legislation or guidelines that impact the handling or management of research data.

Examples: •Trusted Research - Protective Security Requirements guidance

Professional codes of conduct

Concerned with the good conduct of research.

Examples: •Research Charter for Aotearoa New Zealand •Royal Society Professional Code of Conduct

Institutional policies related to research data management →

Institutional data policies

RDM Policy ([link](#))

Outlines the responsibilities of the university community in the management of research data.

Supported by: •[Research data classification standard](#) • [Research data retention requirements](#) • [RDM policy guidance](#)

Research Integrity Policy ([link](#))

Outlines the responsibilities for conducting research in accordance with the highest standards of research integrity.

Supported by: •[Authorship and Publication Guidelines](#)

Privacy Policy ([link](#))

Sets out expectations for University members for the collection, access, use or disclosure of personal information.

Aligns with the Privacy Act 2020.

IP created by staff and students ([link](#))

Provides a framework for managing the IP created by staff and students within the University, including research data.

*Policies directing the use of IT resources (includes the conduct of research):
•[Generative AI usage standard](#) •[IT security policy](#) •[IT acceptable use policy](#)

What does an RDM policy look like?

A **research data management policy** sets out researcher and research student responsibilities for the management, preservation and sharing of research data.

- Common elements:
  - Data management planning
  - Support researchers to implement FAIR and CARE data principles "as open as possible, as closed as necessary"
  - Enable long-term stewardship
- Often paired with guidance on how to implement, e.g., apply governance and storage on a project-by-project basis

---

··· / **Research and innovation** / **Research data management** / Research Data Management Policy

Research Data Management Policy

Application

This policy applies from the commencement date (1 July 2023) to all **research staff**, **students**, **supervisors** and other members of the **University community** that are involved in the management of **research data**.

Purpose

To articulate the responsibilities of the University community for the management of research data. These responsibilities help to ensure that research data is managed in ways that are consistent with:

- international standards for **FAIR data** and open research that are increasingly required by funders, data providers and publishers
- the University's obligations under Te Tiriti o Waitangi and commitment to becoming a **Māori data sovereignty** organisation,
- the **CARE principles** for the governance of indigenous data, including **Pacific data,** and
- legal, ethical and protective security requirements for research data.

Policy

University responsibilities

The University is responsible for:

1. Defining University roles and responsibilities for the management of research data.

2. Communicating the requirements of this policy and facilitating its adoption through the provision of training and guidance to researchers.

3. Providing the necessary infrastructure and services to enable researchers to meet their responsibilities for research data throughout the research data lifecycle.

Researcher responsibilities

Researchers are responsible for:

Data security classification

Data Classification helps to understand constraints on research data and informs the appropriate tools and systems to be used.

- Aligns with national classification.

Data are commonly classified as:

- Public
- Internal / In-confidence
- **Sensitive**
- Restricted / Special
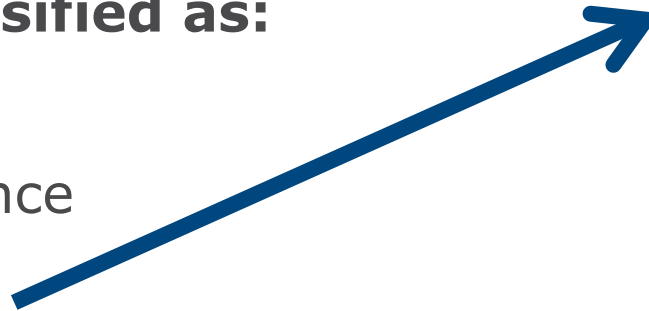
Sensitive research data could include data that:

- are **commercially sensitive**,
- are the subject of a **patent application,**
- contain **personally identifiable information (PII),**
- contain **human health data,**
- are subject to the New Zealand government **export control** regime,
- are subject to a **dual use / sensitive technology** risk rating by MBIE or other funder.

Data Management Planning

Data Management Planning is about preparing for data management across the research data lifecycle.

- Institutional, Ethics Committee or Funder requirement?
  - MBIE, NIH, Wellcome Trust
  - HDEC, University RDM Policy
- Risk-orientated and/or culture change approach
- Project specific
- Prompts conversations, captures decisions, clarifies roles and responsibilities and helps you to align with University policies and processes
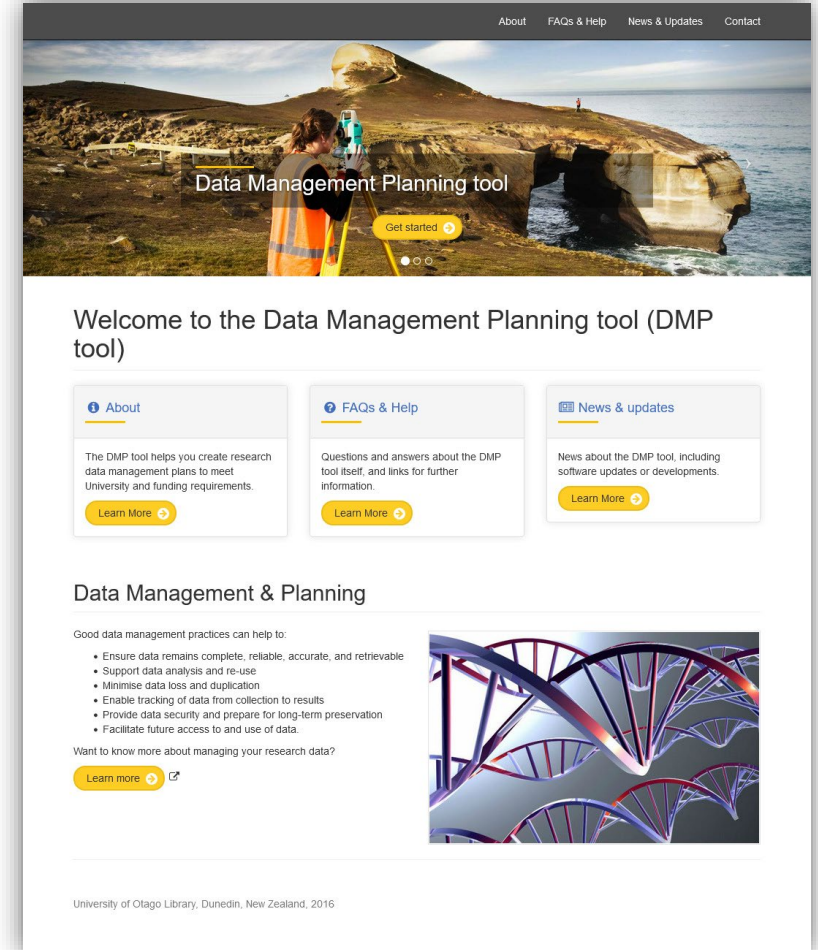
[Checklist for a Data Management Plan from the **Digital Curation Centre**](#)

Data Management Planning

Document decisions about:

- **Project information** - purpose, people, roles & responsibilities

- **Requirements and considerations** - ethical, legal, sovereignty, funder, etc.

- **Data** - collection, organisation & sharing, including access restrictions

- **Sharing and access** - storage locations, retention/ deletion, long term governance

- **Publication** - enabling FAIR

Data Management Planning
Monday 30 June, 1pm-2pm
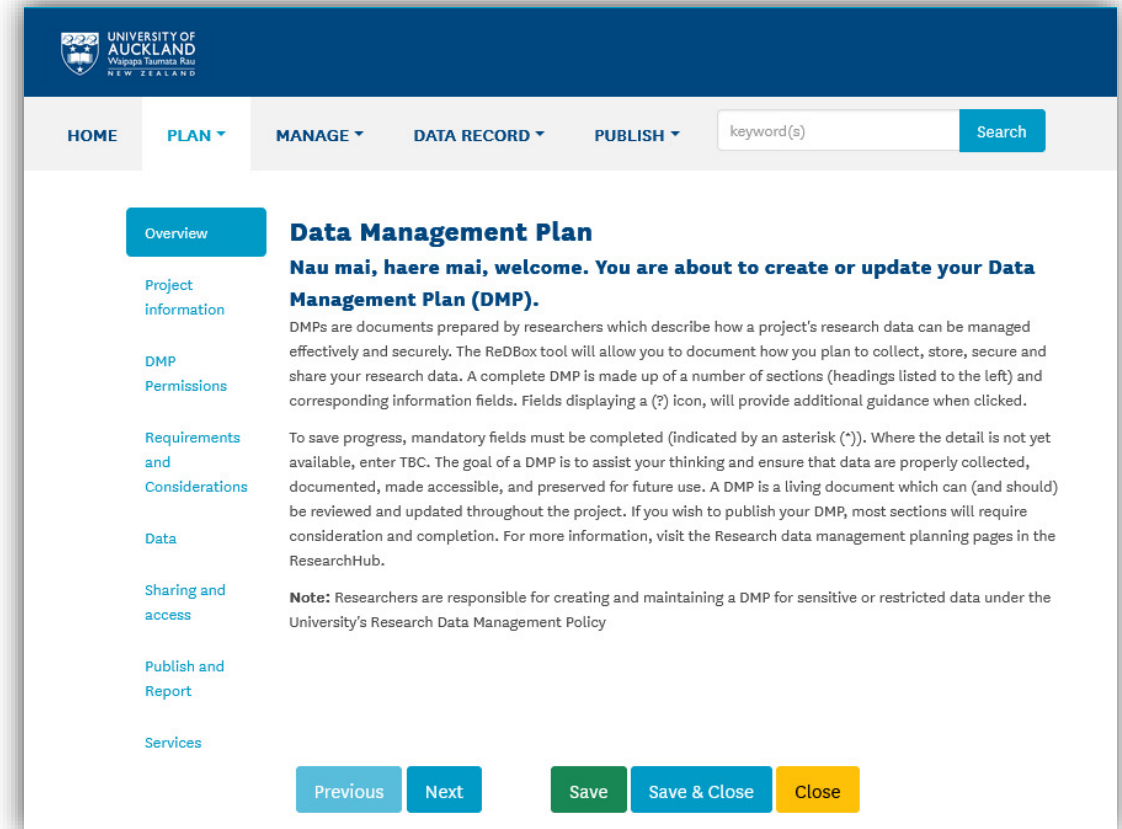
Review and update the DMP

PLAN & DESIGN	CREATE & COLLECT	ANALYSE & INTERPRET	PUBLISH & REPORT	DISCOVER & REUSE

1. DEVELOP research DMP	2. REVISE research DMP	3. UPDATE research DMP	4. FINALISE research DMP

❑ Understand legal & ethical requirements & data policies	❑ Update methods, processes & systems to collect or create data.	❑ Refine & extend metadata	❑ Archive & disseminate data and/or metadata
❑ Describe data to be created/collected	❑ Maintain data provenance & metadata	❑ Data access (secure transfer)	❑ Finalise datasets accessibility rules
❑ Plan metadata & documentation	❑ Implement storage ethics, privacy & security considerations.	❑ Document data analysis approaches	❑ Shared (published) data is licensed
❑ Establish security/access:			❑ Data custodianship in place
❑ Project team & roles			
❑ Ownership & custodianship			
❑ Identify storage solutions			

🤔 **2. Data management planning**

Requirements and Considerations

Researchers should ensure that **legal**, **ethical**, **data sovereignty**, **protective security and commercial constraints** relating to research data are considered prior to data collection and adhered to throughout the research data lifecycle.

Plan and design | Ethics

Ethical considerations

What are the **ethical considerations** surrounding the research data?

Informed consent
- Have you gained consent for data preservation and sharing?

Privacy considerations
- Are personally identifiable information being collected? What processes will you use to de-identify data to ensure confidentiality? What is the risk of re-identification?

Access controls, including storage and transfer of data
- Where and for how long will data be kept (including when collected)? With whom, how and for what purpose can it be shared? How will access be restricted, and on whose authority will this be controlled? What are the conditions of data sharing, including applying an appropriate license?

ResearchHub Home / Research ethics - Manage ethics and regulatory obligations

---

Plan and design | Sensitive data

Sensitive data

What data might need more **security**, **protection** or **access restriction**?

- Data from or about - human participants, health/clinical providers, environment, indigenous people, culture, politics, industry, defence/national security, animals … ?
- Data may become sensitive - unintended capture or context of use, or attitudes may change over time and place.
- Sensitive data is common.
- Data classification helps to understand constraints & meet requirements.

Consider impact on how data is governed, captured, stored, moved and shared, and future stewardship, etc.

ResearchHub Home / Managing research data and artefacts / Sensitive research data

---

Indigenous data sovereignty

Indigenous Peoples have inherent rights and responsibilities to **Indigenous data**.

- CARE principles for indigenous data sovereignty
  Collective Benefit, Authority to Control, Responsibility, and Ethics

- Māori Data Sovereignty principles
  Rangatiratanga (Authority), Whakapapa (Relationships), Whanaungatanga (Obligations), Kotahitanga (Collective benefit), Manaakitanga (Reciprocity), Kaitiakitanga (Guardianship)

- Pacific Data Sovereignty

Consider early as these impact the funding application, planning ethics application, consent, storage, metadata, sharing, and publishing of research findings and data throughout the research data lifecycle.

Global Indigenous Data Alliance
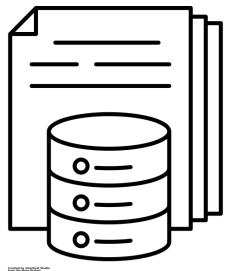
Tikanga, Māori Data Ethics and Māori Data Sovereignty
Monday 30 June, 3pm-4pm

Plan and design | Contracts & copyright
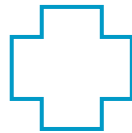
Contracts and copyright

- Contracts
- Intellectual Property / commercialisation
- Copyright of incoming data

My fancy data + External data source(s) → ? → New data set
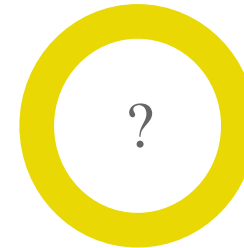
Ethical considerations

What are the **ethical considerations** surrounding the research data?

Informed consent

- Have you gained consent for data preservation and sharing?

Privacy considerations

- Are personally identifiable information being collected? What processes will you use to de-identify data to ensure confidentiality? What is the risk of re-identification?

Access controls, including storage and transfer of data

- Where and for how long will data be kept (including when collected)? With whom, how and for what purpose can it be shared? How will access be restricted, and on whose authority will this be controlled? What are the conditions of data sharing, including applying an appropriate license?

Privacy Act

Privacy principles covered by the Privacy Act 2020

Impacts use of AI tools

[National Ethics Advisory Committee](#)

[Principles for the safe and effective use of data and analytics](#), 2018
Stats NZ & the Privacy Commission

International: GDPR & HIPAA

Training Resources

- **Privacy 101**
- **Privacy 201**

Further details >> https://www.privacy.org.nz/assets/New-order/Privacy-Act-2020/Privacy-Act-2020/Privacy-Act-2020-information-sheets-full-final-set-A711970.pdf
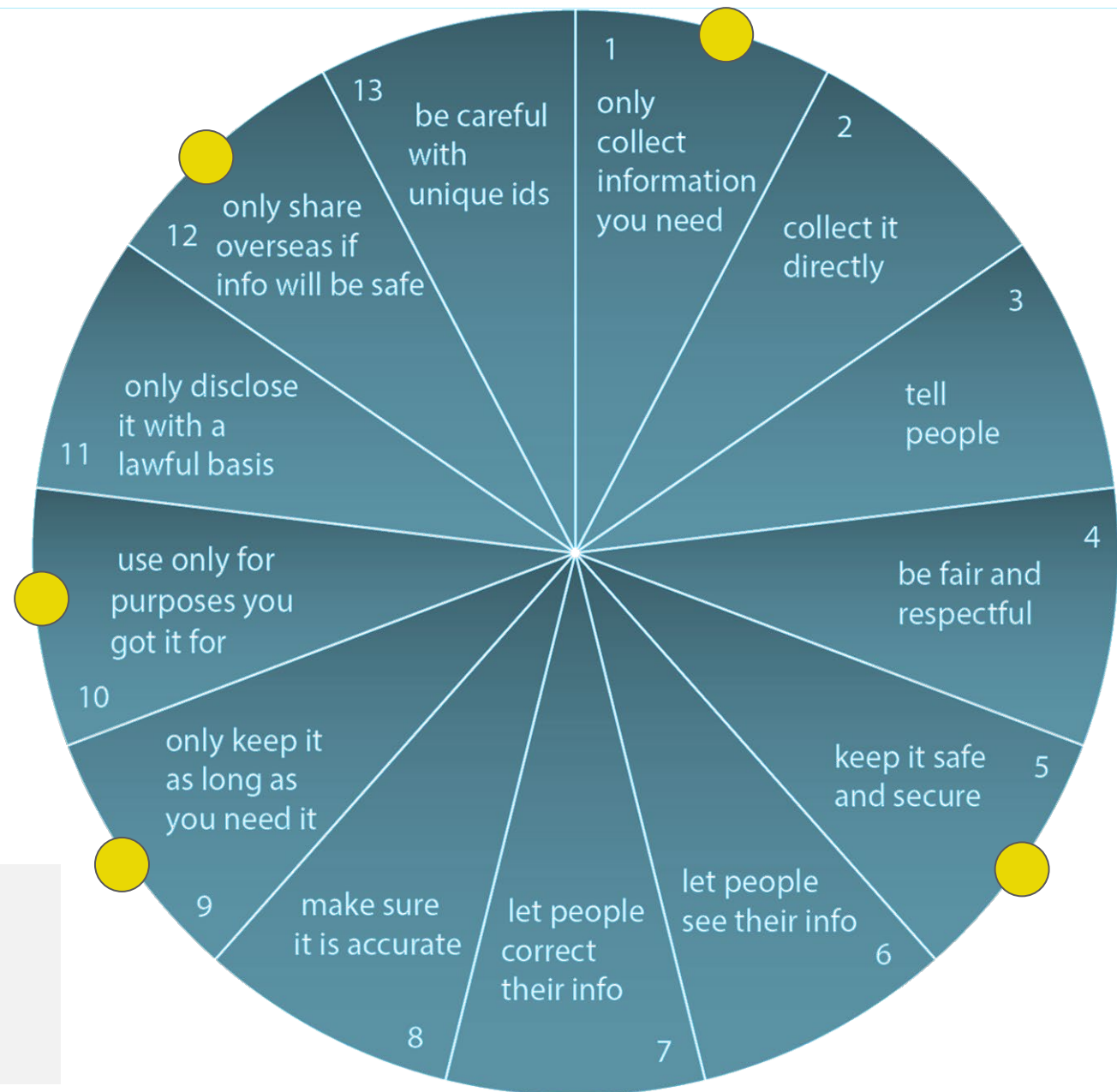
De-identifying data
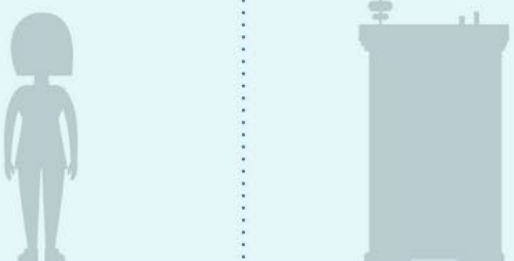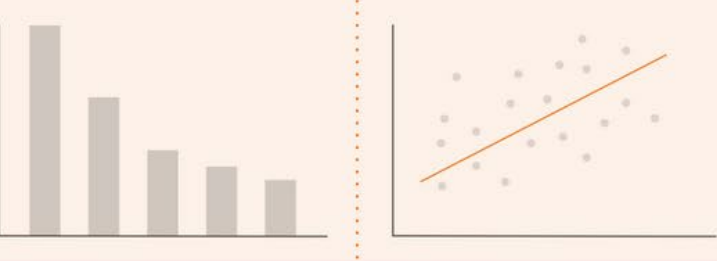
Identifiable
Data that directly or indirectly identifies an individual or business.

Individual		Business	
Name	Hēni	**Name**	Puzzles
Gender	Female	**Type**	Paper Stationery Manufacturing
DOB	31/01/1985		
Address	28 My Road Postcode 6012 Wellington	**Employees**	34
		Expenditure	$398,000

De-identified
Data which has had information removed from it to reduce risk of spontaneous recognition.

Individual		Business	
Name	Unknown	**Name**	Unknown
Gender	Female	**Type**	Manufacturing
DOB	1985		
Address	Postcode 6012 Wellington	**Employees**	30 - 40
		Expenditure	$398,000

Confidentialised
Data which has had statistical methods applied to it to protect against disclosing unauthorised information.

Individual		Business	
Name	Unknown	**Name**	Unknown
Gender	Female	**Type**	Manufacturing
Age	30 - 40 years		
Address	Wellington	**Employees**	10 - 100
		Expenditure	Under $500,000

Degrees of identification in data – From data.govt.nz (Creative Commons Attribution 4.0 International)

Sensitive data

What data might need more **security**, **protection** or **access restriction**?

- Data from or about - human participants, health/clinical providers, environment, indigenous people, culture, politics, industry, defence/national security, animals … ?

- Data may become sensitive - unintended capture or context of use, or attitudes may change over time and place.

- Sensitive data is common.

- Data classification helps to understand constraints & meet requirements.

Consider impact on how data is governed, captured, stored, moved and shared, and future stewardship, etc.

Sensitive data

Available resources:

- [Sensitive data guide](#) (ARDC)

- [Data confidentiality principles & methods](#) (data.govt.nz)

- [NEAC National Ethical Standards](#)

- HDEC [template](#) for data/tissue management plan

- [Te Ira Kāwai, the Auckland Regional Tissue Bank](#) guidance for collection and storage of human tissue for research

Local resources and support?
- Ethics and Integrity
- Māori Responsiveness
- REDCap
- Genomics support
- Health Research / Clinical Trials

Activities

Checking in…

In the Zoom chat…

What types of sensitive data do you work with?

What does sensitive data look like to you?

and/or

What do you find most challenging
about managing sensitive data?

Indigenous data sovereignty

Indigenous Peoples have inherent rights and responsibilities to **Indigenous data**.

- CARE principles for indigenous data sovereignty
  Collective Benefit, Authority to Control, Responsibility, and Ethics

- Māori Data Sovereignty principles
  Rangatiratanga (Authority), Whakapapa (Relationships), Whanaungatanga (Obligations), Kotahitanga (Collective benefit), Manaakitanga (Reciprocity), Kaitiakitanga (Guardianship)

- Pacific Data Sovereignty

Consider early as these impact the funding application, planning ethics application, consent, storage, metadata, sharing, and publishing of research findings and data throughout the research data lifecycle.

Global Indigenous Data Alliance

ResBaz
RESEARCH BAZAAR

Tikanga, Māori Data Ethics and Māori Data Sovereignty
Monday 30 June, 3pm-4pm

Contracts and copyright

- Contracts

- Intellectual Property / commercialisation

- Copyright of incoming data

https://doi.org/10.5281/zenodo.11147887

My fancy data

External data source(s)

New data set

Data sharing agreements

Where research is shared with third parties, intellectual property rights, plans for data preservation and sharing, and legal responsibilities should be agreed in a formal **data sharing agreement**.

- Formal contract sets out:
  - sets out and agrees on the purpose of the data sharing and expected outcomes
  - describes what will to happen to the data at each stage (how the data will be transferred and stored)
  - sets standards and helps all the parties to be clear about their respective roles (access restrictions, stewardship)

- Parties demonstrate their accountability to legal, ethical, data sovereignty, etc. requirements.

More information on data sharing agreements from the UK Information Commissioner's Office

Summary: Plan and design

- Become familiar with the University's **Research Data Management policy** and other research-related data policies to understand the requirements for research data management.

- Use the new online **Data Management Plan (DMP)** to consider and document important decisions about managing research data throughout your project.

- Consider and document **ethical/legal** obligations, **sensitive data** considerations, **indigenous data sovereignty** and **contracts and copyright** responsibilities before collecting research data.

Create and Collect

Plan and document data collection.

Store data so that they are protected against corruption and loss.

Considerations for data creation

- What data* will be created or collected?
  (e.g., type, format, volume, whether pre-existing or new)?

  *Raw physical data (inputs) + raw digital data (outputs) + derived digital data + final datasets

- How will the data be collected/created?

- What standards or methodologies will be used for data collection?

- What quality assurance processes will be adopted?

- Do the chosen formats and software enable sharing and long-term access to the data?

Further information: Organising data (UK Data Service)

Organising research data

Project-based organisation

- CLEAR, CONCISE, CONSISTENT
- Folder hierarchy
  - short, descriptive folder names
  - avoid overlapping categories
  - limit size and depth of folders
- Consistent strategy prevents confusion
- Things are easy to find and to sort
- Document your strategy
- Set up and use databases if necessary

```
∨ 📁 project-name
    📁 code
    📁 data_clean
    📁 data_raw
∨ 📁 docs
    📁 admin
    📁 manuscripts
```

Further reading: Project structure by Danielle Navarro

Research metadata

Metadata is a list of information you expect will be needed for the data to be read and interpreted in the future.

- Ensures data can be shared, discovered and reused
- Facilitates reproducibility and scientific integrity

Metadata across the research data lifecycle:

- **Who** created the data?
- **What** does the data file contain?
- **When** were the data created?
- **Where** were the data created?
- **Why** were the data created?
- **How** were the data created?

ResearchHub Home / Managing research data and artefacts / Organise and describe research data / Research metadata

---

Digital research data storage

When selecting appropriate research data storage, you will need to take into account:

- Are you collecting personal (identifiable) data?
- Who needs access?
- How much storage space do you need?
- How will the data be protected against loss?
- How will you ensure the data is secure?
- What does your funder and/or organisation require?

Project specific
- People/access
- Legal, ethics etc. conditions
- Files, incl. README & DMP
- Security classification
- Governance
- Retention period

ResearchHub Home / Managing research data and artefacts / Research data storage / Choosing data storage

---

Data sharing and transfer

Secure option for getting data or files from or to others

FileSender

- Supports end-to-end encryption
- Approved for sensitive data
- Can resume interrupted uploads
- Logs maintained for 90d

ResBaz **Sharing Sensitive Research Data With FileSender**
Tuesday 1 July, 1pm-2pm

https://filesender.reannz.co.nz/

Summary: Create and collect

- **Plan and document** the data that will be generated and their associated data formats.

- Use **University-approved software** or tools whenever possible to create and/or collect data (includes AI or GenAI tools).

- Plan and document **folder hierarchies**, **file naming conventions** and **version control**.

- Ensure the data is accompanied by appropriate **research documentation** and **metadata** so that the data can be read and interpreted in the future.

- Ensure digital forms of research data are stored on a **University-managed research storage service**

- Use **FileSender** or another approved data-sharing service to receive or send research data.

Considerations for data creation

- What data* will be created or collected?
  (e.g., type, format, volume, whether pre-existing or new)?

  *Raw physical data (inputs) + raw digital data (outputs) + derived digital data + final datasets

- How will the data be collected/created?

- What standards or methodologies will be used for data collection?

- What quality assurance processes will be adopted?

- Do the chosen formats and software enable sharing and long-term access to the data?

Further information: Organising data (UK Data Service)

Considerations for data reuse

If you plan to use existing (secondary) data you will need to understand and comply with any terms of use under which the data may be used or shared.

Considerations:

1. Is it identifiable or re-identifiable data? If yes, ethics approval is required.

2. Confirm that participant consent included use for secondary analyses. If no, ethics approval and potential re-consent are required.

3. Check the quality of the data.

4. Check licenses and understand how the data can be used

5. If you reuse data, cite it.

Data collection software & tools

- **Use University-supported software or tools, whenever possible.** These applications are often security tested and approved ('Authority to Operate')

- Always consider where and with whom your data is being sent, stored, or shared during collection and processing.

- Is data backed up and is it secure?

Activity	Approved tools (examples)
Participant surveys	REDCap Qualtrics
Transcription of audio files	MS Word MS Teams Zoom

ResBaz RESEARCH BAZAAR

Digital tools and AI for Transcription
Thursday 3 July, 4pm-5pm

An overview of REDCap
Tuesday 1 July, 2pm-3pm

University of Auckland ResearchHub / Research software and computing

Generative AI usage standard

1. Select Data classification

3. Undertake Privacy Impact Assessment

6. Understand AI limitations and biases

Standards

1. The data classification of any Inputs submitted to GenAI tools must be established.

2. The choice of GenAI tools must be restricted to those suitable for the data classification level of inputs submitted:

   - Public Data: Any appropriate GenAI tool may be used.

   - Internal and Sensitive Data: Must only be used where a negotiated contract and service agreement exists between the University and the GenAI provider that establishes adequate protection for Inputs. For the avoidance of doubt, adequate protection will ensure that Inputs are not used for any other purpose by the provider, including further training of their public GenAI.

   - Restricted Data: Only services solely controlled by the University may be used.

3. A Privacy Impact Assessment must be completed before a GenAI tool is used with Personal Information.

4. The designated owner of a business function within the University is accountable and responsible for validating GenAI output prior to use of that output to inform business processes within their remit.

5. Users of GenAI tools should consult with the Office of the Pro-Vice Chancellor Māori where Māori data may be used in a GenAI tool, or use may impact Māori.

6. Users of GenAI tools should familiarise themselves with the limitations and/or the possibility of inherent bias within the tool prior to use.

7. Any content (including text, image, or video) intended for publishing or distribution where a substantial portion of the content has been created by a GenAI tool should be labelled as such.

Organising research data

Project-based organisation

- CLEAR, CONCISE, CONSISTENT

- Folder hierarchy

  - short, descriptive folder names

  - avoid overlapping categories

  - limit size and depth of folders

- Consistent strategy prevents confusion

- Things are easy to find and to sort

- Document your strategy

- Set up and use databases if necessary
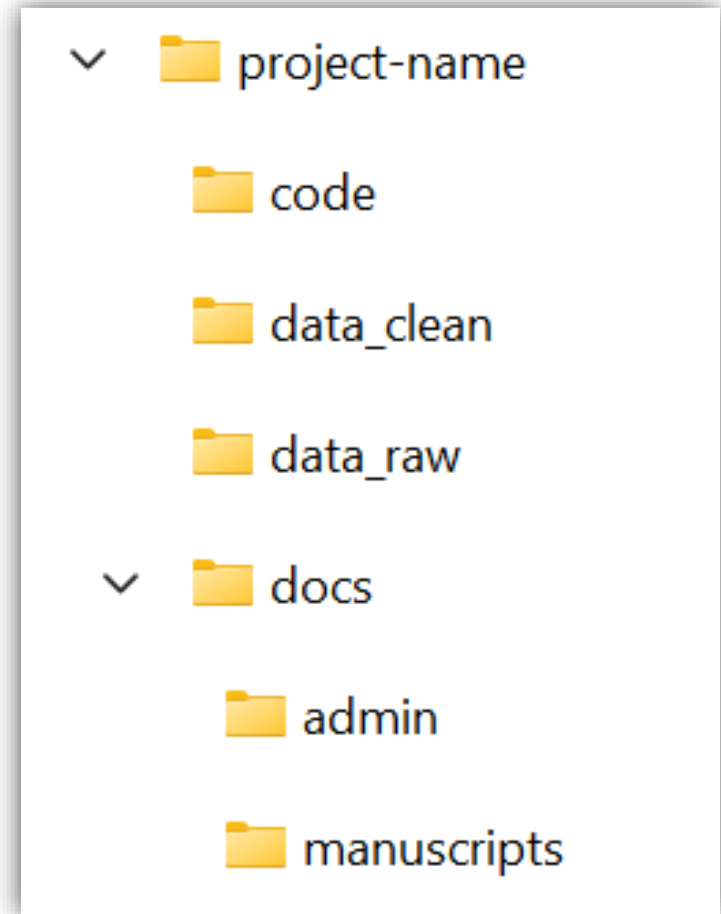
Further reading: Project structure by Danielle Navarro

File naming

- Create a template and document it

- Short, descriptive and use only important fields

- Avoid spaces or special characters and ambiguity
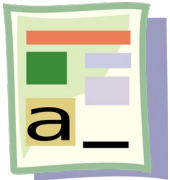
20170310-tmr-literature-review.docx
[date]-[creator]-[subject].[ext]

arthnz-rat-rbw-food-weights.xlsx
[project]-[animal model]-[creator]-[data type].[ext]
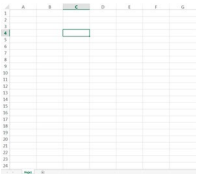
UCollege_AndersonM_Ped-Resp-Infection-Genomic-Determinants_Biosketch_20160125.pdf
[CTSA]_[InvestigatorLastNameFirstInitial]_[ProtocolShortTitle]_[Document]_[YYYYMMDD].[ext]

teko-van-kuyk_pineapple-41-white-purple-black_35x50_2017
[artist-name]_[artworks-name]_[length-X-height(depth in case of sculpture)]_[date or year]

fr3s-140623-129C-2653-w.jpg
[studysite,depth of water]-[yyymmdd]-[tile#,treatment]-[photo#]-[photo coverage].[ext]

Version control

1. **Semantic version control**

    Version numbers in the filename:
    coastline-erosion-manuscript-1.2.doc

2. **Collaborative writing tools**

    Office 365, Dropbox Paper, Google docs, etc.

3. **Version control systems**

    Originally designed for software development but widely used by researchers for collaborating on analyses and writing manuscripts.

    Git + GitHub
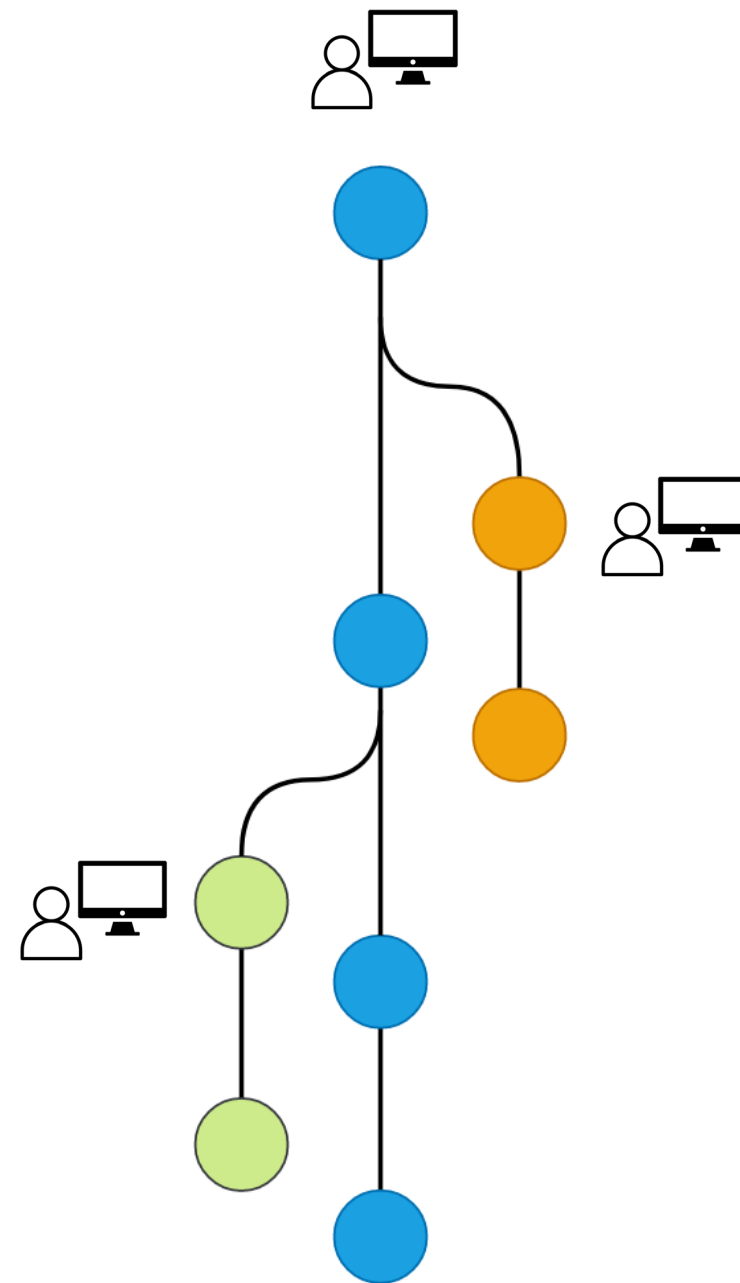
Further reading: Version control strategy and best practice (UK Data Service)

Research metadata

Metadata is a list of information you expect will be needed for the data to be read and interpreted in the future.

- Ensures data can be shared, discovered and reused
- Facilitates reproducibility and scientific integrity

Metadata across the research data lifecycle:

- **Who** created the data?
- **What** does the data file contain?
- **When** were the data created?
- **Where** were the data created?
- **Why** were the data created?
- **How** were the data created?
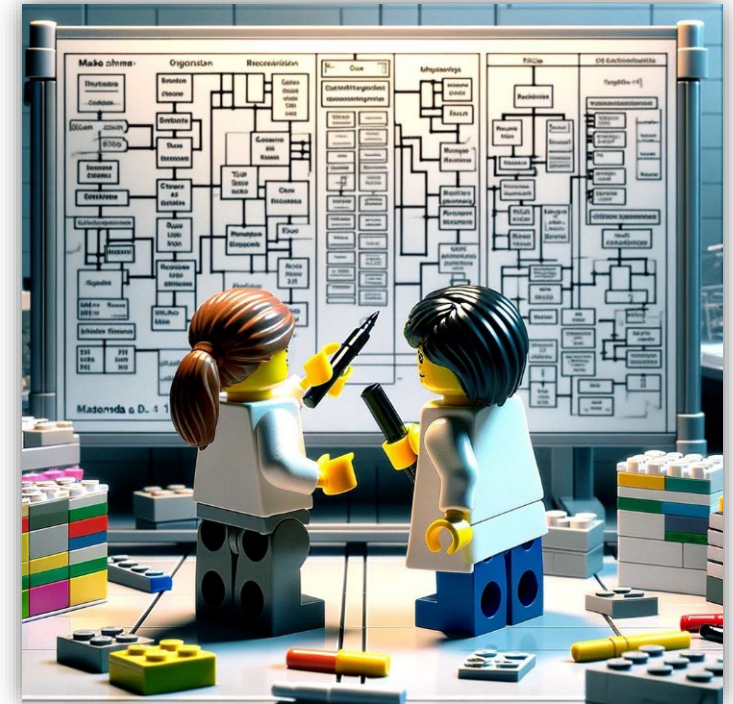
Metadata standards

Metadata standards specify data fields or elements for describing data that are machine-readable.

Some research fields have agreed metadata standards. **Example:** Clinical trials registry (*pictured*).

Find a disciplinary metadata schema:
- [Research Data Alliance (UK)](#)
- [FAIR sharing](#)
- [Digital Curation Centre (UK)](#)

Research data documentation

Data-level documentation

- Provides information on individual data objects, such as a variable in a data file or interview transcript.

- Should be brief and concise, include units of measurement, and use established classifications and standards, where appropriate.

Study-level documentation

- Provides high-level information on the research context and design, the data collection methods, data preparations and manipulations, etc.

Examples:

- Electronic lab notebook and protocol
- Data Dictionary
- Codebook

- Software syntax and output files
- Metadata record or database schema
- Links to reports and publications (DOIs)
- Provenance information (third-party data)

A README example

Title: Long-term variation of surface phytoplankton chlorophyll a in the Southern Ocean during 1965-2002

Authors:
  - TORU HIRAWAKE (INVESTIGATOR, TECHNICAL CONTACT)
  - DAVE CONNELL (DIF AUTHOR)

Owner: Commonwealth of Australia

Citation: Hirawake, T. (2005) Long-term variation of surface phytoplankton chlorophyll a in the Southern Ocean during 1965-2002, Ver. 1, Australian Antarctic Data Centre - doi:10.4225/15/5a384270f2b61, Accessed: 2020-07-02

Revision Date: 2017-12-18

Metadata URL: https://data.aad.gov.au/metadata/records/chlorophyll_65-02

Dataset API URL: https://data.aad.gov.au/s3/api/bucket/datasets/science/chlorophyll_65-02?export=json

Abstract: The variation in the phytoplankton biomass over a decadal time scale, and its relationship with the Antarctic Circumpolar Wave (ACW) and climate change, has been poorly interpreted because of the limited satellite chlorophylla (chl a) data compared with the physical parameters from satellite. We analysed a long-term chl a dataset along the Japanese Antarctic Research Expedition (JARE) cruise tracks since 1965 to investigate inter-annual variation of phytoplankton biomass. In the Southern Ocean, increasing trends of chl a and the spreading of higher chl a area to the north with 3-7 year cycles were found. Although relationships between the decadal change in chl a and climate change such as variation of sea ice extent and the El Nino are still obscure, large variation of primary production in proportion to the chl a is implied.

The chl a concentration of sea surface water has been measured routinely on board the icebreakers Fuji and Shirase during almost every cruise of the JARE.

The download file contains chlorophyll a data collected from ship tracks on JARE voyages between 1965 and 2002.

The field in this dataset are:

Date (local time)
Year
Latitude
Longitude
Corrected Chlorophyll a

See the attached paper for more details.

The publications on the data collected during the 1965-1976 and 1988-1993 cruises are listed in Fukuchi [1980] and Suzuki and Fukuchi [1997], respectively. For data on the 1977-1985 and 1994-1997 cruises, see [Kanda and Fukuchi, 1979; Fukuchi and Tamura, 1982; Tanimura, 1981; Watanabe and Nakajima, 1983; Ino and Fukuchi, 1984; Sasaki, 1984; Hamada et al., 1985; Fukuda et al., 1986; Hattori and Fukuchi, 1988; Midorikawa et al., 2000]. Data post 1998-2002 cruises is in Hirawake and Fukuchi [2004]. Data from the 1986-1987 will be published in the JARE data report of digital media, including all cruise data.

Auxiliary Material for paper 2004GL021394 Long-term variation of surface phytoplankton chlorophyll a in the Southern Ocean during 1965-2002. Toru Hirawake, Tsuneo Odate and Mitsuo Fukuchi (National Institute of Polar Research, Tokyo) Geophys. Res. Lett., Vol (Num), doi:10.1029/2004GL021394

All of the chl a data have been reported in the publications of the National Institute of Polar Research (NIPR).

Use Constraints: This data set conforms to the CCBY Attribution License (http://creativecommons.org/licenses/by/4.0/).

Please follow instructions listed in the citation reference provided at http://data.aad.gov.au/aadc/metadata/citation.cfm?entry_id=chlorophyll_65-02 when using these data.

ID: 1360
Metadata ID: chlorophyll_65-02
UUID: EBD86EDA-DE08-4C0F-AC94-AF171DBCE46A

README guide

Digital research data storage

When selecting appropriate research data storage, you will need to take into account:

- Are you collecting personal (identifiable) data?
- Who needs access?
- How much storage space do you need?
- How will the data be protected against loss?
- How will you ensure the data is secure?
- What does your funder and/or organisation require?

Project specific

People/access
Legal, ethics etc. conditions
Files, incl. README & DMP
Security classification
Governance
Retention period

Backing up your research data

University-managed storage

✓ Research Drive & Dropbox for Research are backed up automatically.

General principles

• What data/files need to be backed up? How often? Where? By whom?

• A least 2 people should have access to the data.

Be aware of the 3-2-1 rule

At least **3** copies | Using at least **2** different storage media | At least **1** copy offsite
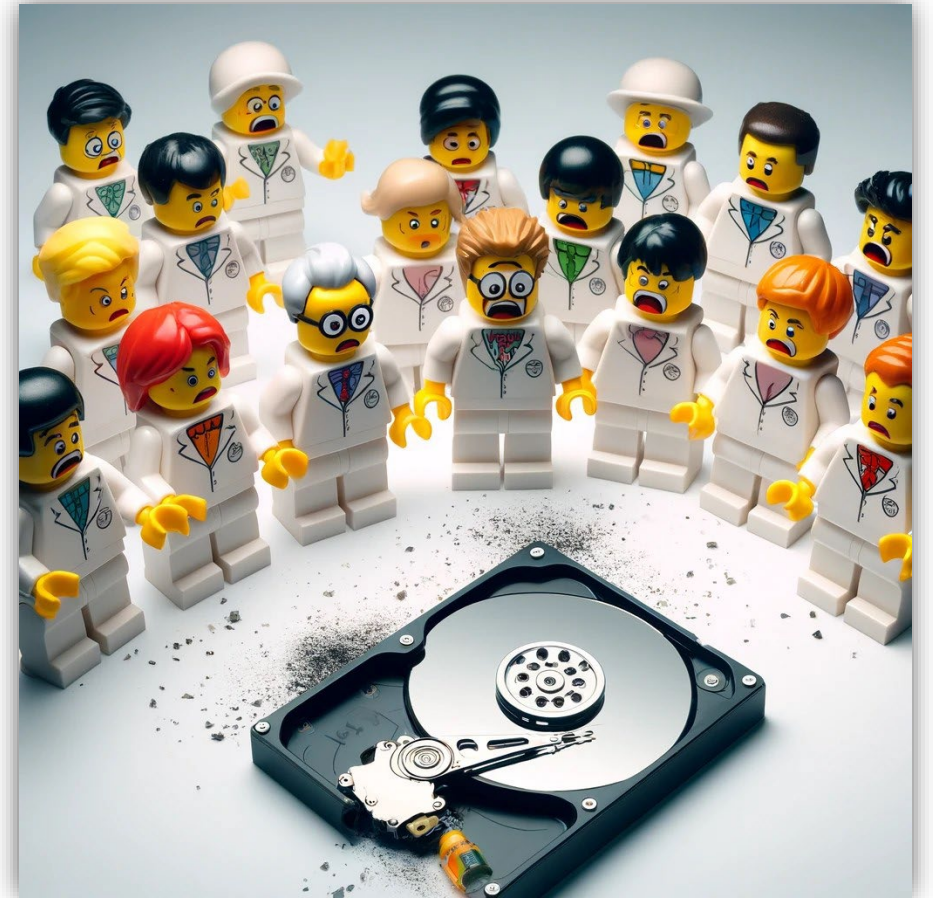
Research data storage options

Personal devices are usually not appropriate

Data capture

Self-service or managed

Analysis & visualisation with research compute - virtual machines and HPC

Storage

Network storage
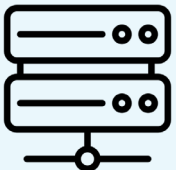
Dropbox (cloud storage)

MS OneDrive (personal), Teams, SharePoint (online)

Secure Research Environment
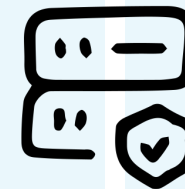
Sustainable use of storage

Data storage is **expensive** and has **environmental** costs to consider.

Tips to take control of your storage space:

1. Regularly review your files

2. Understand and identify **research data** vs. **research debris**

3. Enable future use e.g., open formats

4. Use different storage tiers/products – slower=cheaper

88% of organizations surveyed have no idea of the content in their stored data.

58% of organizations are keeping information indefinitely.

79% of organizations say too much time and effort is spent manually searching and disposing information that has met its retention requirements.

58% of organizations still rely on employees to decide how to apply corporate policies.

Source: The Information Explosion survey from the Council for Information Auto-Classification http:/infoautoclassification.org/

3. Data storage

Data sharing and transfer

Secure option for getting data or files from or to others

- Supports end-to-end encryption
- Approved for sensitive data
- Can resume interrupted uploads
- Logs maintained for 90d

https://filesender.reannz.co.nz/

Sharing Sensitive Research Data With FileSender
Tuesday 1 July, 1pm-2pm

Summary: Create and collect

- **Plan and document** the data that will be generated and their associated data formats.

- Use **University-approved software** or tools whenever possible to create and/or collect data (includes AI or GenAI tools).

- Plan and document **folder hierarchies**, **file naming conventions** and **version control**.

- Ensure the data is accompanied by appropriate **research documentation** and **metadata** so that the data can be read and interpreted in the future.

- Ensure digital forms of research data are stored on a **University-managed research storage service**

- Use **FileSender** or another approved data-sharing service to receive or send research data.

Analyse and Interpret

Plan and document data processing so that the end result could be interpreted, replicated and reused by others.

Tidy Data

1. Every column is a variable.

2. Every row is an observation.

3. Every cell is a single value.

country	year	cases	population
Afghanistan	1999	745	19987071
Afghanistan	2000	2666	20595360
Brazil	1999	37737	172006362
Brazil	2000	80488	174504898
China	1999	212258	1272915272
China	2000	213766	1280428583

- 1st row variable names (no spaces, simple).
- 1 sheet/table per file.
- Save in a 'plain text' format (.csv).
- Use a README.txt for background and context.

ResearchHub Home / Managing research data and artefacts / Organise and describe research data / Tidy data

---

Data Provenance

- Always keep a copy of the raw data untouched

- Have a separate copy which is your tidy dataset

- Keep a record of your 'recipe' (exact steps taken) to get from raw to tidy data

- Keep contextual information in a README

Identify resources for RDM

Researchers are responsible for **identifying resources** required for the management of research data, including storage, compute, unique or special infrastructure or governance requirements.

- Large or non-standard storage?
- Additional computer power? (access to GPUs, virtual machines, high performance computing [HPC], machine learning)
- Special infrastructure requirement? (software, hardware, technical expertise)
- Governance? (access committee, advisory groups)

ResearchHub Home / Research software and computing / Advanced compute resources and services

---

Summary: Analyse and interpret

- Use **Tidy Data standards** to prepare the data for analysis so that it is both human- and machine-readable and easy to interpret.

- Use **Tidy Data guiding principles** to ensure data provenance – keep an unaltered copy of the raw data, work off a separate tidy dataset, keep a record of your 'recipe' and maintain a README for contextual information.

- Know where to request **additional data compute** (virtual machines or high-performance computing) for more intensive computational analyses.

Tidy Data

1. Every column is a variable.

2. Every row is an observation.

3. Every cell is a single value.

- 1st row variable names (no spaces, simple).
- 1 sheet/table per file.
- Save in a 'plain text' format (.csv).
- Use a README.txt for background and context.

country	year	cases	population
Afghanistan	1999	745	19987071
Afghanistan	2000	2666	20595360
Brazil	1999	37737	172006362
Brazil	2000	80488	174504898
China	1999	212258	1272915272
China	2000	213766	1280428583

ResearchHub Home / Managing research data and artefacts / Organise and describe research data / Tidy data

Data Provenance

- Always keep a copy of the raw data untouched

- Have a separate copy which is your tidy dataset

- Keep a record of your 'recipe' (exact steps taken) to get from raw to tidy data

- Keep contextual information in a README

Identify resources for RDM

Researchers are responsible for **identifying resources** required for the management of research data, including storage, compute, unique or special infrastructure or governance requirements.

- Large or non-standard storage?

- Additional computer power? (access to GPUs, virtual machines, high performance computing [HPC], machine learning)

- Special infrastructure requirement? (software, hardware, technical expertise)

- Governance? (access committee, advisory groups)

Summary: Analyse and interpret

- Use **Tidy Data standards** to prepare the data for analysis so that it is both human- and machine-readable and easy to interpret.

- Use **Tidy Data guiding principles** to ensure data provenance – keep an unaltered copy of the raw data, work off a separate tidy dataset, keep a record of your 'recipe' and maintain a README for contextual information.

- Know where to request **additional data compute** (virtual machines or high-performance computing) for more intensive computational analyses.

Publish and Report

Prepare data for archive and preservation. Wherever possible, data should be shared (published) so that it can be checked and re-used by other researchers.

Prepare to archive

When you have finished working with your research data, you should...

- Evaluate your digital research files to identify data 🔍 from debris 🗑
- Ensure that data is stored safely, in a suitable file format, and accompanied by adequate and self-explanatory documentation (e.g., README)
- Digitise **non-digital research data** whenever possible (Guidance https://www.openaire.eu/non-digital-data-guide)

Data accessibility

Sharing research data means your work is more discoverable and citable and that research findings can be replicated and validated.

1. **What does the University/Organisation, funder and/or publisher require?**
2. **Am I enabling FAIR and CARE?**
3. **Is the data shareable?**
   - Is the data technically shareable? (format, metadata)
   - Is the data legally shareable? (deidentified)
   - Is it ethically sound? (Do I have participant consent to share the data?)
   - Is the data licensed?

ResearchHub Home / Managing research data and artefacts / Publish research data / Research data publishing and discovery

---

Summary: Publish and Report

- When you have finished working with the data, separate **'data'** from **'debris'**, and have a plan for **archive**.
- Be aware of the **minimum retention period** based on institutional policy or contractual obligations.
- Consider if research data should be **preserved** longer and evaluate associated resources & costs.
- Consider how you will **share your research data or metadata** while taking into any ethical, contractual, or indigenous data sovereignty sharing restrictions ("As open as possible, as closed as necessary").

Prepare to archive

When you have finished working with your research data, you should...

- Evaluate your digital research files to identify data 🔍 from debris 🗑

- Ensure that data is stored safely, in a suitable file format, and accompanied by adequate and self-explanatory documentation (e.g., README)

- Digitise **non-digital research data** whenever possible
  (Guidance https://www.openaire.eu/non-digital-data-guide)

Return, retention, deletion, & destruction

Researchers should ensure that research data are **returned**, **retained**, **deleted** and/or **destroyed** in accordance with legal, ethical, data sovereignty and commercial constraints.

- Is there a requirement to return research data?
- What is the minimum retention period? During this period, the data should be **archived** on University- or Organisation-managed storage and non-digital data held locally.
- Deletion of archived digital research data ✓
- Destruction of digital research data files stored on University-managed storage is not achievable in most cases. ✗

Data accessibility

Sharing research data means your work is more discoverable and citable and that research findings can be replicated and validated.

1. **What does the University/Organisation, funder and/or publisher require?**

2. **Am I enabling FAIR and CARE?**

3. **Is the data shareable?**

   - Is the data technically shareable? (format, metadata)
   - Is the data legally shareable? (deidentified)
   - Is it ethically sound? (Do I have participant consent to share the data?)
   - Is the data licensed?

F   A   I   R

Findable   **Accessible**   **Interoperable**   **Reusable**
(others know how to
access)

Metadata (descriptive information), DOI and process for access are external facing, human and machine readable E.g., internet search results, bibliographic databases.

Metadata is with data and disciplinarily specific. Combining and using data are enabled by format and file type(s).
E.g., Data Management Plan, Protocol, README.txt

Australian Antarctic Data Centre
Data management and spatial data services

Menu | Search | Search | Login | Support

Australian Antarctic Data Centre / Discover and Manage Data / Records / chlorophyll_65-02

Metadata details

⇄ Request DOI   Edit record

chlorophyll_65-02

View the full metadata record

Citation
Hirawake, T. (2005) Long-term variation of surface phytoplankton chlorophyll a in the Southern Ocean during 1965-2002, Ver. 1, *Australian Antarctic Data Centre* - doi:10.4225/15/5a384270f2b61, Accessed: 2025-04-07

Title
Long-term variation of surface phytoplankton chlorophyll a in the Southern Ocean during 1965-2002

Data Centre
Australian Antarctic Data Centre, Australia

DOI
doi:10.4225/15/5a384270f2b61

Created Date
2005-08-22

Revision Date
2017-12-18

Parent record
None

Unique persistent identifier (DOI) = **Findable**

Datasets and documents

chlorophyll_65-02
Long-term variation of surface phytoplankton chlorophyll a in the Southern Ocean during 1965-2002
⬇ Download dataset   📂 View dataset contents
Public   Submitted 22 Aug 2005

chlorophyll_65-02
Long-term variation of surface phytoplankton chlorophyll a in the Southern Ocean during 1965-2002
⬇ Download dataset
Released - AAD Only   Submitted 22 Aug 2005

Related links

⬇ Download point for the data - Excel spreadsheet
⬀ Download point for the data - papers - AAD Staff Only
⬀ Citation reference for this metadata record and dataset

Access

These data are publicly available for download from the provided URL. A copy of some of the referenced publications is available for download by AAD staff only.

Temporal Coverages

- **Start date:** 1965-11-23 - **Stop date:** 2002-12-08

Spatial Coverages

	Latitude		Longitude
Northernmost:		**Westernmost:**	
24.567		100.147	
Southernmost:		**Easternmost:**	
-54.985		137.95	

Describe access conditions = Accessible

Science Keywords

- EARTH SCIENCE > CLIMATE INDICATORS > ATMOSPHERIC/OCEAN INDICATORS > TELECONNECTIONS > ANTARCTIC OSCILLATION
- EARTH SCIENCE > CLIMATE INDICATORS > ATMOSPHERIC/OCEAN INDICATORS > TELECONNECTIONS > EL NINO SOUTHERN OSCILLATION (ENSO)
- EARTH SCIENCE > BIOSPHERE > ECOSYSTEMS > AQUATIC ECOSYSTEMS > PLANKTON
- EARTH SCIENCE > OCEANS > OCEAN CHEMISTRY > PIGMENTS > CHLOROPHYLL
- EARTH SCIENCE > BIOSPHERE > ECOLOGICAL DYNAMICS > ECOSYSTEM FUNCTIONS > BIOMASS DYNAMICS

Additional Keywords

- CHLOROPHYLL A
- JARE
- PHYTOPLANKTON
- SOUTHERN OCEAN

Locations

- OCEAN > INDIAN OCEAN
- OCEAN > SOUTHERN OCEAN
- OCEAN > PACIFIC OCEAN
- GEOGRAPHIC REGION > POLAR

Use Constraints

This data set conforms to the CCBY Attribution License (http://creativecommons.org/licenses/by/4.0/).

Please follow instructions listed in the citation reference provided at http://data.aad.gov.au/aadc/metadata/citation.cfm?entry_id=chlorophyll_65-02 when using these data.

Project	ISO Topic	Dataset Language
	- BIOTA - CLIMATOLOGY/METEOROLOGY /ATMOSPHERE - OCEANS	- ENGLISH

Orignating Centre	Dataset Progress	IDN Node
- JARE	- COMPLETE	- AMD/AU - CEOS - AMD

Publications

- Fukuchi, M. (1980) Phytoplankton chlorophyll stocks in the Antarctic Ocean, J. Oceanogr. Soc. Jpn., 36, 73-84
- Fukuchi, M., and S. Tamura (1982) Chlorophyll a distribution in the Indian sector of the Antarctic Ocean in 1978-1979, Antarct. Rec., 74, 143-162
- Fukuda, Y., M. Ohno, K. Iwanami, and H. Touju (1986) Chlorophyll a content in the surface and subsurface waters along the course of the Shirase to Antarctica in 1984-1985, Antarct. Rec., 30, 103-112
- Hamada, E., A. Taniguchi, M. Okazaki, and Y. Naito (1985) Report on the phytoplankton pigments measured during the JARE-25 Cruise to Syowa Station, Antarctica, November 1983 to April 1984, ARE Data Rep., 89, Natl. Inst. Polar Res., Tokyo, 103
- Hattori, H., and M. Fukuchi (1988) Report on the phytoplankton pigments concentrations, zooplankton and benthos sampling during the JARE-27 cruise, November 1985 - April 1986, JARE Data Rep., 28, Natl. Inst. Polar Res., Tokyo, 135
- Hirawake, T., and M. Fukuchi (2004) Chlorophyll a concentration of phytoplankton during the cruises of 40-44th Japanese Antarctic Research Expedition in 1998-2003, JARE Data Rep., 31, Natl. Inst. Polar Res., Tokyo, 279
- Ino, Y., and M. Fukuchi (1984) Report on chlorophyll a distribution along the course of the Fuji in 1981-1982, Antarct. Rec., 81, 38-44
- Kanda, H., and M. Fukuchi (1979) Surface chlorophyll a concentration along the course of the Fuji to and from Antarctica in 1977-1978, Antarct. Rec., 66, 37-49
- Midorikawa, T., K. Nomura, Y. Miyamoto, T. Odate, A. Ishikawa, N. Washiyama, T. Hirawake, M. Namiki (2000) Report on phytoplankton pigments measured during the JARE-36~-39 cruises to Syowa Station, Antarctica in 1994-1998, JARE Data Rep., 249, 36, Natl. Inst. Polar Res., Tokyo
- Sasaki, H. (1984) Distribution of nano- and microplankton in the Indian sector of the Southern Ocean, Mem. Natl. Inst. Polar Res. Spec. Issue, 32, 38-50
- Suzuki, T., and M. Fukuchi (1997) Chlorophyll a concentration measured with a continuous water monitoring system during the cruise to Syowa Station, Antarctica, JARE-27 (1985/86) to JARE-35 (1993/94), 60, Natl. Inst. Polar Res., Tokyo
- Tanimura, A. (1981) Distribution of the surface chlorophyll a along the course of the Fuji to and from Antarctica in 1979-1980, Antarct. Rec., 72, 35-48
- Watanabe, K., and Y. Nakajima (1983) Surface distribution of chlorophyll a along the course of the Fuji (1980/81) in the Southern Ocean, Antarct. Rec., 77, 33-43

chlorophyll_65-02

Metadata Entry ID: chlorophyll_65-02

Long-term variation of surface phytoplankton chlorophyll a in the Southern Ocean during 1965-2002

View Metadata Record | **Download Dataset**

Contents

Resource

- LICENSE
- README
- chlorophyll_65-02.csv
- chlorophyll_65-02.xml

4 records

Apply a license = Reusable

Metadata is with data = Interoperable

Use open file formats, = Interoperable

Citation

Hirawake, T. (2005) Long-term variation of surface phytoplankton chlorophyll a in the Southern Ocean during 1965-2002, Ver. 1, *Australian Antarctic Data Centre* - doi:10.4225/15/5a384270f2b61, Accessed: 2025-04-08

Use Constraints

This data set conforms to the CCBY Attribution License (http://creativecommons.org/licenses/by/4.0/). Please follow instructions listed in the citation reference provided at http://data.aad.gov.au/aadc/metadata/citation.cfm?entry_id=chlorophyll_65-02 when using these data.

Data repositories

- Designed to store, preserve, and provide access to research data (enables FAIR) or a metadata record (balance of FAIR/CARE)

- Multidisciplinary or discipline specific options

- Items are assigned a permanent, resolvable and citable Digital Object Identifier (DOI)

- Track views, downloads, citations for impact.
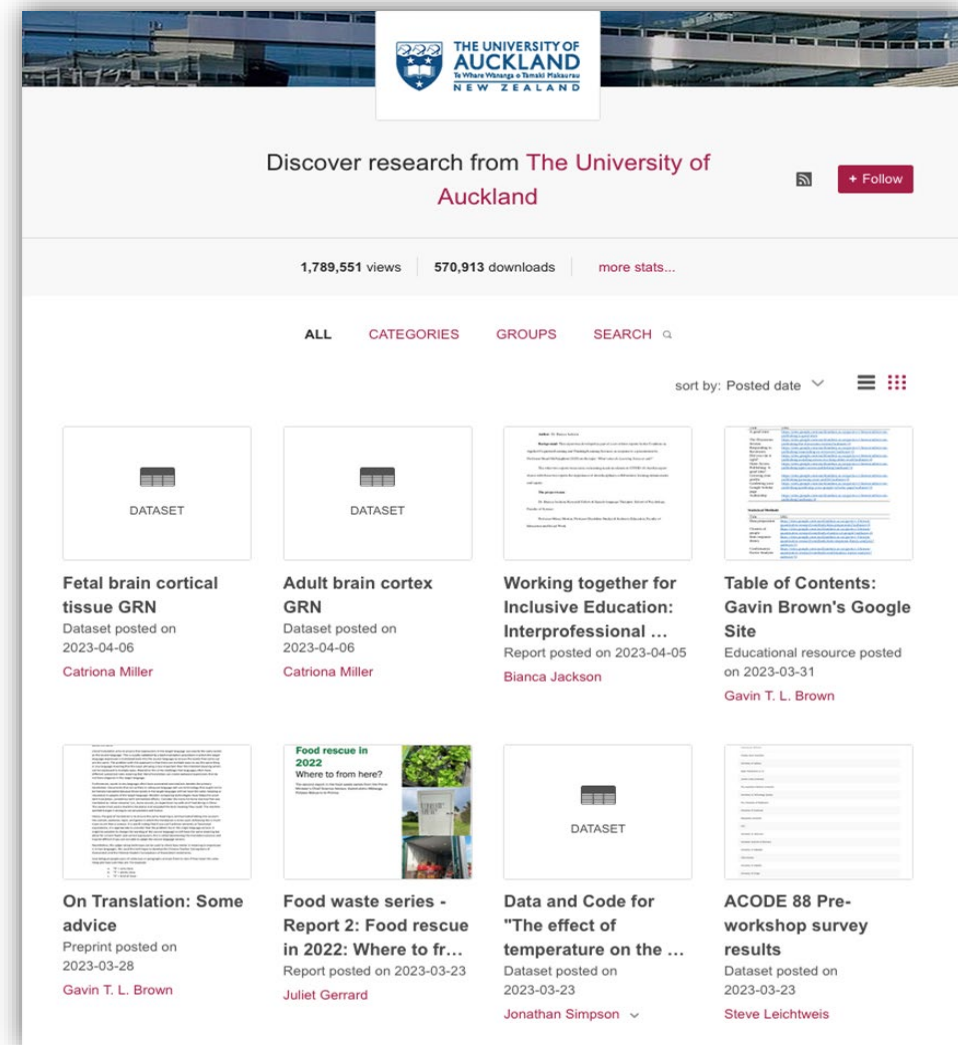
C
Collective benefit

A
Authority to control their data

R
Responsibility to engage
respectfully with those communities

E
Indigenous Peoples' **ethics** should
inform the use of data across time

F
Findable

A
Accessible

I
Interoperable

R
Reusable

Everyone

Specific people
and purpose

Balancing FAIR and CARE
(or other constraint to publishing data)

Be FAIR
Findable Accessible Interoperable Reusable

and

CARE
Collective Benefit Authority to Control Responsibility Ethics

Practical steps:

1. Publish a **descriptive or metadata-only record**

2. Create a **mediated access process**

3. Use a **data sharing agreement**

4. Produce a **data availability statement** linking data DOI to research outputs

Carroll, S.R., Herczog, E., Hudson, M. *et al.* Operationalizing the CARE and FAIR Principles for Indigenous data futures. *Sci Data* **8**, 108 (2021). https://doi.org/10.1038/s41597-021-00892-0

Data availability statements

Data Availability Statement: All data generated or analyzed during this study are included in this paper and its Supporting Information files, except the sound .wav files, which are available through Figshare (https://auckland.figshare.com/articles/media/Sound_wav_files_use/20103734; DOI: 10.17608/k6.auckland.20103734).

Summary: Publish and Report

- When you have finished working with the data, separate **'data' from 'debris'**, and have a plan for **archive**.

- Be aware of the **minimum retention period** based on institutional policy or contractual obligations.

- Consider if research data should be **preserved** longer and evaluate associated resources & costs.

- Consider how you will **share your research data or metadata** while taking into any ethical, contractual, or indigenous data sovereignty sharing restrictions ("As open as possible, as closed as necessary").

Discover and Reuse

Data that are available for discovery and access may be reused, either to substantiate findings or to generate new insights.

Governance for the re-use of data

Data that are made available for discovery and access may be reused by other researchers, either to substantiate or reproduce original findings or to generate new insights.

Governance is required to:
- Ensure compliance with original ethics obligations
- Prevent damage to the original researcher's IP
- Prevent harm to study participants (e.g., reidentification)
- Enable **Rangatiratanga** (authority to control) and **Kaitiakitanga** (guardianship)
- Provide processes to manage access and sharing, **Data Access Groups/Committees** developed to review proposed use of data prior to data being released.

When researchers leave...

Ensure appropriate arrangements are made if researchers require AND have rights to continued access to research data after leaving the project or moving to another research organisation/institution.

Handover of research data management, including:
- Updating the Data Management Plan (DMP)
- Off-boarding meetings to discuss arrangements for ensuring ongoing access
- Update all agreements and ethics approvals

May also include:
- Data Transfer Agreement to transfer the research data to another institution (where permitted by ethics approvals and agreements),
- Establishing external collaborator status for the departing staff member or student where ongoing access to research data held at the University is required and permitted.

What happens when a PhD candidate hands in their thesis?
- Students generally retain 'ownership' of data created for postgraduate thesis submission.
- Ethics restrictions regarding where, and for how long, sensitive data is retained still apply.

Governance for the re-use of data

Data that are made available for discovery and access may be reused by other researchers, either to substantiate or reproduce original findings or to generate new insights.

Governance is required to:

- Ensure compliance with original ethics obligations
- Prevent damage to the original researcher's IP
- Prevent harm to study participants (e.g., reidentification)
- Enable **Rangatiratanga** (authority to control) and **Kaitiakitanga** (guardianship)
- Provide processes to manage access and sharing, **Data Access Groups/Committees** developed to review proposed use of data prior to data being released.

When researchers leave...

Ensure appropriate arrangements are made if researchers require AND have rights to continued access to research data after leaving the project or moving to another research organisation/institution.

Handover of research data management, including:
- Updating the Data Management Plan (DMP)
- Off-boarding meetings to discuss arrangements for ensuring ongoing access
- Update all agreements and ethics approvals

May also include:
- Data Transfer Agreement to transfer the research data to another institution (where permitted by ethics approvals and agreements),
- Establishing external collaborator status for the departing staff member or student where ongoing access to research data held at the University is required and permitted.

What happens when a PhD candidate hands in their thesis?
- Students generally retain 'ownership' of data created for postgraduate thesis submission.
- Ethics restrictions regarding where, and for how long, sensitive data is retained still apply.

What haven't we discussed?
(Add your questions into the chat
or raise your hand)

Questions? Get in touch…

✉ researchdata@auckland.ac.nz

Research data are a treasure. Managing data is about caring for data to reflect this.

Thank you