Data Management Planning for HDEC and HRC

Centre for eResearch

What is data?

Why manage your research data?

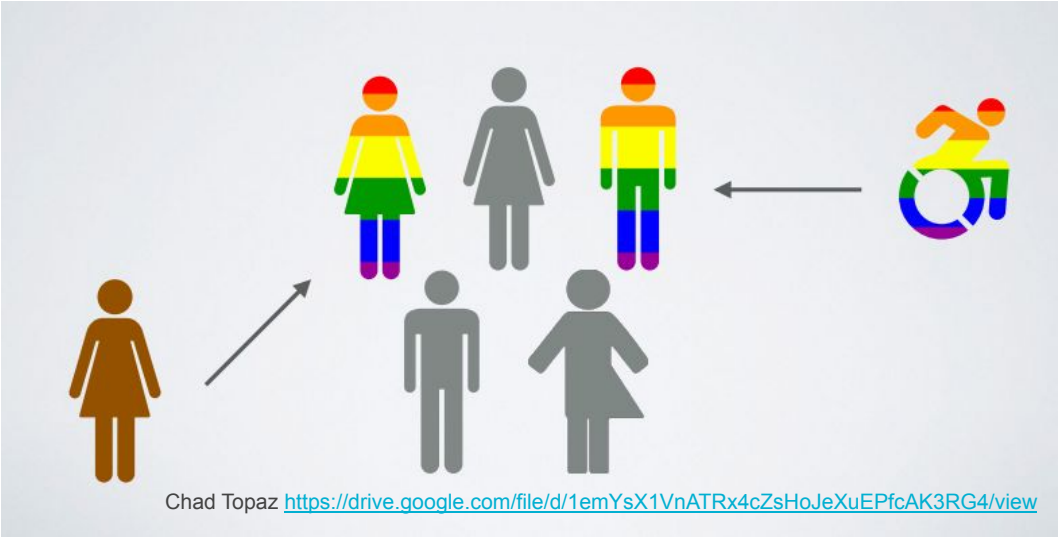
Best practices - planning, ethics, organising, and storing

Services

THE UNIVERSITY OF AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND

Valuing inclusion

Ensuring all individuals feel respected, accepted, and valued.



Chad Topaz https://drive.google.com/file/d/1emYsX1VnATRx4cZsHoJeXuEPfcAK3RG4/view

University Code of Conduct

- **We act with manaakitanga**: we show respect, care and support for others
- **We foster whanaungatanga**: making our **University community** a place in which all feel they belong.
- **We build kotahitanga**: teaching, learning and research is a partnership between our students and our staff.
- **We uphold kaitiakitanga**: we recognise our responsibilities as kaitiaki (guardians) to protect and respect our **environment**, traditions, knowledge, culture, languages and other taonga.

https://www.auckland.ac.nz/en/on-campus/life-on-campus/code-of-conduct.html

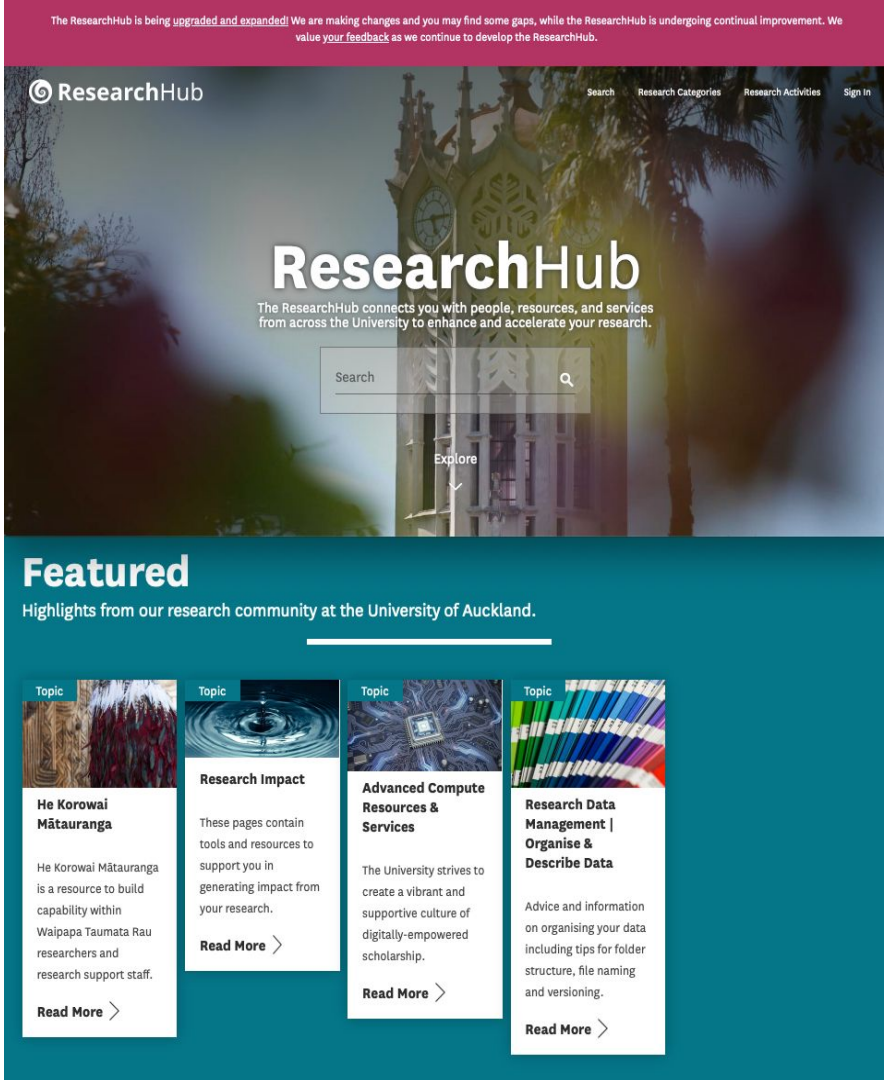
THE UNIVERSITY OF AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND

ResearchHub

research-hub.auckland.ac.nz

> New look and content
> He Korowai Mātauranga
> Managing research data resources
> Research storage - request

researchdata@auckland.ac.nz



WHY?

“So many boxes…”

Bears trying to share/reuse data.

Data Sharing and Management Snafu in 3 Short Acts

NYU Health Sciences Library
https://www.youtube.com/watch?v=N2zK3sAtr-4

NZ, professional & institutional RDM policy

Research Charter for Aotearoa New Zealand

- researcher and institutional responsibilities for data+ “safe and secure storage, management and access for future use”, 2020

Royal Society Professional Code of Conduct

8. To develop, and implement so far as they are reasonably able [2], a management plan to ensure the integrity, retention, secure storage, appropriate and transparent use of data and samples gathered or developed during their activities

Taumata Teitei – Vision 2030 and Strategic Plan 2025

p4, “We will forge enduring partnerships that inform and guide our progress towards becoming a Māori Data Sovereignty organisation. This will see transformations across our education, research and engagement practice and in how we work as an organisation.”

Research Code of Conduct*

4.5.Research Findings: Researchers should share data and findings openly and as promptly as possible, …

5.4 Researchers should keep records of all research in ways that will allow verification and replication of their work by others… Original research data should preferably be kept indefinitely. ...Data should be stored in a safe and secure location and manner. ...

Research data?

What is research data?
Anything created, collected or obtained in the course of research that underpins your research output (e.g. article, book, report, performance, exhibition, etc.)

Why manage it?

- Understand
- Integrity
- Credit
- Share
- Other environments

Data Management Plans (DMP)

Helps **you** manage, store & share your data during & after

DMP
- project details
- organisation & storage
- sharing
- retention & archive

Planning - active, dynamic, discuss with supervisor

Tissue and Software management plans

DMP template
https://doi.org/10.17608/k6.auckland.7268720
DMP guide
https://doi.org/10.17608/k6.auckland.7268729

Data Management Plan
Research Ideas and Outcomes 2: e10600 (27 Sep 2016)
https://doi.org/10.3897/rio.2.e10600
Reviewable v2

Data Management Plan for PhD Thesis "Climatic Limitation of Alien Weeds in New Zealand: Enhancing Species Distribution Models with Field Data"

Jennifer L. Pannell

Abstract

Background

This Data Management Plan (DMP) was created using the DMPTool. It describes all data collected and created as part of the recently submitted PhD thesis of the corresponding author, "Climatic Limitation of Alien Weeds in New Zealand: Enhancing Species Distribution Models with Field Data", undertaken at the Bio Protection Research Centre, Lincoln University, New Zealand. It is important to note that although the National Science Foundation (NSF) template was used for this DMP, the work was carried out under a grant from the Tertiary Education Commission (TEC). The NSF template was selected as it closely matched the requirements of the host institution.

Article metadata
Data and materials produced
- Data collected
- Materials produced

Standards, formats and metadata
Roles and responsibilities
Dissemination methods
Policies for data sharing and public access
Archiving, storage and preservation
References

https://doi.org/10.3897/rio.2.e10600

Software/code as research data

Software/code may be the focus of the inquiry/research

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

Reproduction - using *similar* inputs, tools, environments, arrive at (mostly) the same outputs and conclusions - more generalisable to justify your results.

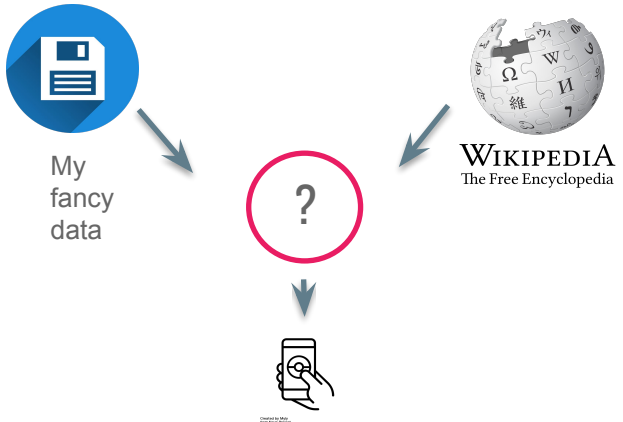
Software Sustainability Institute - licensing, publishing & packaging/sharing

Software Management Plan (SMP)
https://www.software.ac.uk/software-management-plans

Where to publish?
https://www.software.ac.uk/which-journals-should-i-publish-my-software

Researcher integrity: Legal issues

- Data Sharing Agreements? UniServices Contracts team
- Intellectual Property /commercialisation? UniServices IP team
- Copyright of incoming data? e.g.



Sensitive data and Ethics

- Different types of sensitive data
  - human participant, health, environmental, commercial/organisational, indigenous, animal
- UoA Ethics > Human Ethics - *future-proof scope, sharing & publishing*

  https://www.auckland.ac.nz/en/research/about-our-research/human-ethics/

- ARDC Sensitive data guide

  https://ardc.edu.au/resources/working-with-data/sensitive-data/
- REDCap - Liggins (supported, fee) and FMHS (self-service, free) instances

  https://wiki.auckland.ac.nz/display/ontrack/Accessing+REDCap

Privacy Principles

Privacy Act 2020, 1 December

Further details:
https://privacy.org.nz/assets/Privacy-Act-2020-content/2020-A-quick-tour-of-the-privacy-principles-Oct-2020.pdf

HDEC Data (and Tissue) Management Plan/

– – –

- Introduced late 2020
- Ethics questions structured as a Data Management Plan

Sections:

- 1. Policies? Research Code of Conduct. Privacy Policy. (IP Policy)
- (8. Who will have access to the identifiable/de-identified project data?)
- 9. Where are you storing the identifiable and/or de-identified data and how is it being kept secure/access restricted?
- T10. E.g. Te Ira Kāwai | Auckland Biobank https://www.biobank.ac.nz/
- 10/11. Māori Data Sovereignty principles https://research-hub.auckland.ac.nz/guide-to-managing-research-data/ethics-integrity-and-compliance/maori-data-sovereignty

Using identifiable data?

- UoA Privacy Centre - documents, service support & training modules.
- UoA's current policy documents Privacy Policy and Research Code of Conduct
- National Ethical Standards for Health and Disability Research and Quality Improvement (2019)
    a. Ch. 12 Health data, changing identifiable fields
    b. Further resources - Data.Govt
    c. Sensitive data resources from the ARDC - Working with Sensitive Data and De-identification guide
    d. Principles for the safe and effective use of data and analytics, 2018. Stats NZ & the Privacy Commission
- UoA Online Ethics module
- Health and Disability Ethics Committees (HDEC) new template for data/tissue management plans, October 2020

Specialist advice/support as required:

- Ethics and Integrity Team
- Te Kupenga Hauora Māori Responsiveness to Māori Team (R2M@auckland.ac.nz)
- ResearchHub
- Grafton Clinical Genomics
- Te Ira Kāwai, the Auckland Regional Tissue Bank
- Liggins Clinical Data Research Hub

ORGANISING: Data collection

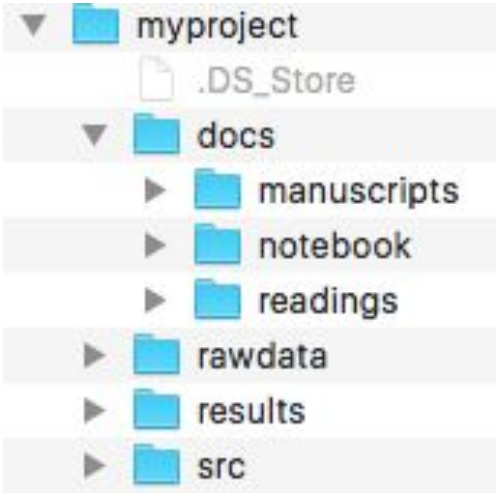
- What types of data ?
- How much ?

  *raw + (analysed * no. analyses) + (backup * redundancies)*
- Will it grow/ accumulate ?
- Will it change over time ?

- What file formats will you have?
- How will you organise it?
- Where will you store it?
- How will you document / describe it?
- How will you check it for errors?

Project organisation - slides.djnavarro.net/project-structure

- Clear. Concise. Consistent.
- Folder hierarchy
  - follow existing conventions
  - avoid overlapping categories
  - limit size and depth of folders
- Different data files are easily distinguishable
- Consistent strategy prevents confusion
- Things are easy to find and to sort
- Document your strategy
- Set up and use databases if necessary



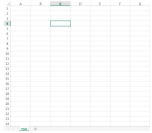
Files and folders naming

- Project/grant name/number
- Date of creation YYYYMMDD
- Initials of creator

- Description of content
- Collection method
- No spaces (no …%20..) or special characters
- Version number x.y

20170310-tmr-literature-review.docx
[date]-[creator]-[subject].[ext]

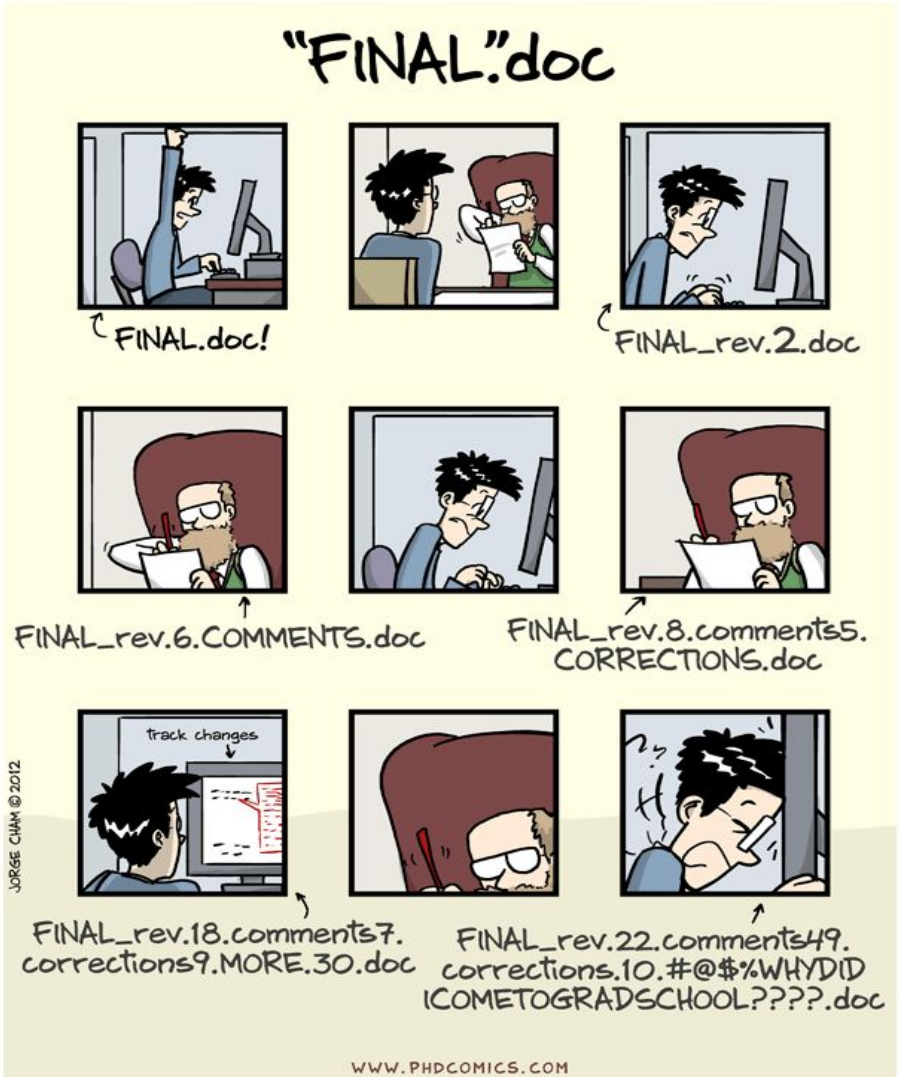
arthnz-rat-rbw-food-weights.xlsx
[project]-[animal model]-[creator]-[data type].[ext]



UCollege_AndersonM_Ped Resp Infection Genomic Determinants_Biosketch_20160125.pdf
[CTSA]_[InvestigatorLastNameFirstInitial]_[ProtocolShortTitle]_[Document]_[YYYYMMDD].[ext]
https://health.ucdavis.edu/ctsc/documents/Prior%20Approval_NCATS%20doc%20file%20naming%20conventions.pdf

fr3s-140623-129C-2653-w.jpg
[studysite,depth of water]-[yyymmdd]-[tile#,treatment]-[photo#]-[photo coverage].[ext]
https://library.stanford.edu/research/data-management-services/data-best-practices/best-practices-file-naming

THE UNIVERSITY OF AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND



Version control

- Semantic labelling system - *x.y*
  1.2, 1.3, 2.0, etc.

  20161201-tmr-literature-review-1.0.docx

  20161214-tmr-literature-review-1.1.docx

  headings formatted

  20170117-tmr-literature-review-2.0.docx

  new section, supervisor corrections

- Application – Word, Wikis, Google docs
- VCS - git

Documentation

(e.g. README.txt)

********** FRUBASE PACKAGE **********

VERSION 4.0. DEC 2007.

THE FRUBASE PACKAGE ACCOMPANIES:

Jordano, P. 1995. Angiosperm fleshy fruits and seed dispersers: a comparative analysis of adaptation and constraints in plant-animal interactions. American Naturalist 145: 163-191.

It contains a copy of the main data file exactly as used for this paper, as well as other accompanying files (see below).

Taxonomic arrangement follows:
Cronquist, A. (1981). An integrated system of classification of flowering plants. Columbia University Press.

Nomenclature follows Stevens, P. F. (2001 onwards). Angiosperm Phylogeny Website. Version 8, June 2007. http://www.mobot.org/MOBOT/research/APweb/. This scheme follows: A.P.G. [= Angiosperm Phylogeny Group] II. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. Bot. J. Linnean Soc. 141: 399-436.

Plant names and names of higher taxonomic categories have been checked with: Mabberley, D.J. 1987. The plant-book. A portable dictionary of the higher plants. Cambridge University Press, Cambridge, UK.

Please, contact me if you have suggestions, find errors, inconsistencies, or any other bug in the file. As well, please let me know about your uses of this data and send manuscripts and reprints when available. I'll be happy to help you in any case, as far as I can.

I am periodically updating this data base since I started writing my PhD thesis more than 20 years ago. Thus, I'd like to receive suggestions for new data sources and provide updated versions to those interested.

Please, use these data files for peaceful purposes, enjoy doing science with them as I have enjoyed writing the paper quoted above, and learn as much as you can with them. They are the result of splendid work by many people working with plant-frugivore interactions and are embedded in papers reporting very interesting results, descriptions and discussions on these interactions; please read them.

********** CONTENTS **********

All files are plain ASCII text files, with the exception of SUMMARY and FRUBASE.xls.
Those with data have TABs as their field delimiters so they can be readily imported in any statistical package or spreadsheet program. The FRUBASE.txt is readily imported by any spreadsheet application. Please, contact me if you need the files formatted in other ways (e.g., my original SAS datasets, or EXCEL worksheets).

1. README.txt - This file. Including a description of the variables and a listing of the literature sources with the numeric codes.
2. Summary.doc - A summary file (originally intended to appear as an Appendix in my 1995 paper) summarizing mean values for the main families and genera in the data base. This is a Microsoft WORD (version 6.0) file, which can be read directly either by the Mac or Windows versions of the program.
3. REFS.txt - A long list with the source reference used for each species in the data file. The file is TAB delimited and has a header line with variable names: FAMILY, GENUS, SPECIES, NEWREF, and REFERENCE (authors and year).
4. FRUBASE.txt - The data file itself. Missing data are indicated by dots (.). The file is TAB delimited and has a header line with variable names as in the list below. The file is sorted by FAMILY, GENUS, and SPECIES names, in ascending order.
5. FRUBASE.xls - The data file itself, now in Excel format for spraedsheets. See (4).

****

Variable names and descriptions in FRUBASE

CL	Class
SCL	SubClass
ORD	Order
FAM	Family
GEN	Genus
SP	Species
REF	Reference number - This is my maintenance code for updates.
NEWREF	New Reference number - These are the refs numbers in the files REFS and SUMMARY.
FAMLAB	Family Label - An 8-character label for family.
GENLAB	Genus Label - An 8-character label for genus.
SPLAB	Species Label - An 8-character label for species.
COD	Species code - A 5-character code for the species.
DISPCAT	Disperser type category - BIRDS, MIXED, MAMMALS.
DISP	Disperser type - Finer categorization. Not yet completed. Needs revision.
MEGAFAUNA	Whether the fruit species is associated with dispersal of seeds by megafauna.
AREA	Geographic area - Major geographic areas of the data sources. MEurope: Mediterranean Europe (also includes Israel and Morocco. NEurope: Temperate and Northern Europe. NAmerica: North America, excl. Southern Mexico.

https://dryad.figshare.com/articles/FRUBASE_dataset/4176363

README guide: https://data.research.cornell.edu/content/readme

File Formats

- Open. Standardised. In wide use.
  Easy to datamine, transform, or re-cast
- What software do you expect to use?
  Are you collaborating or sharing with others?
- Domain specific standards?
- Consider fidelity or quality issues if using compression

Combined data from Rangitoto trip 2015

16+ m island size

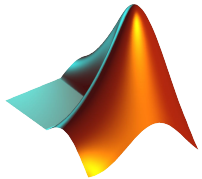
NORTH	Saturday															Sunday																		
Species	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	Mean	Std Dev	Sqrt n	Std Err
Asplenium flaccidum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0				
Asplenium flabellifolium	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0				
Asplenium oblongifolium	1	3	3	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	1	2	0	0	0	0				
Ctenopteris heterophylla	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0				
Hymenophyllum spp.	0	3	3	0	0	1	0	2	1	0	0	1	1	0	0	0	0	0	2	0	0	0	0	0	0	0	2	0	1	0				
Pellaea rotundifolia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0				
Microsorum pustulatum	1	2	0	3	3	3	5	3	2	3	2	2	2	2	0	3	3	3	2	1	1	0	2	2	2	2	1	1	2	1				
Pteridium esculentum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0				
Pyrrosia eleagnifoila	3	1	3	0	0	1	1	2	2	0	2	1	1	0	0	0	0	3	1	0	3	0	0	2	2	3	2	3	1	0				
Trichomanes reniforme	1	3	4	0	0	0	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
Acianthus sinclairii	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0				

CENTRE	Saturday															Sunday																		
Species	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	Mean	Std Dev	Sqrt n	Std Err
Asplenium flaccidum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0				
Asplenium flabellifolium	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	0	0	2	0	1	0				
Asplenium oblongifolium	0	0	1	0	0	2	3	1	2	0	2	1	1	0	0	0	0	0	0	0	1	1	0	1	1	0	3	3	1	0				
Ctenopteris heterophylla	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	2	0	0	0				
Hymenophyllum spp.	0	2	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2	0	0	0				
Pellaea rotundifolia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0				
Microsorum pustulatum	2	2	3	0	0	1	2	1	1	1	1	2	1	0	1	2	2	1	2	3	0	0	0	2	0	3	2	3	1	1				
Pteridium esculentum	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
Pyrrosia eleagnifoila	0	0	0	0	0	0	0	1	1	0	1	0	1	0	0	1	0	0	2	0	0	0	0	0	3	0	1	0	1	0				
Trichomanes reniforme	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	4	3	0	1	0	0	0	0	0	2	0	0	0				
Acianthus sinclairii	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	2	0	1	2					

SOUTH	Saturday															Sunday																		
Species	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	Mean	Std Dev	Sqrt n	Std Err
Asplenium flaccidum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0				
Asplenium flabellifolium	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	2	0	0	1	0	1	0	1	1	1	0	0	1	0				
Asplenium oblongifolium	0	0	0	0	0	0	0	1	1	0	1	2	0	0	0	0	0	0	0	1	0	0	0	1	1	1	0	0	1	0				
Ctenopteris heterophylla	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0				
Hymenophyllum spp.	2	2	3	4	4	3	6	0	2	1	2	1	0	0	0	0	0	0	3	0	3	1	4	3	3	3	2	0	1	2				
Pellaea rotundifolia	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0				
Microsorum pustulatum	1	3	2	0	0	1	2	1	2	1	2	2	0	2	1	0	1	2	1	2	3	1	3	2	2	2	2	0	1	2				

Tidy Data Principles

- Always keep a copy of the raw data
- Have a separate copy which is your tidy dataset
- Keep metadata record (codebook, readme.txt)
- Keep a record of your ‘recipe’ (exact steps taken) to get from raw to tidy data



Reading:: Hadley Wickham, *Tidy Data*, Vol. 59, Issue 10, Sep 2014, Journal of Statistical Software. http://www.jstatsoft.org/v59/i10

Tidy Data tips

- 1 piece of information per cell
- 1 variable (e.g. weight) 1 column
- 1 observation (e.g. patient) 1 row

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

table1

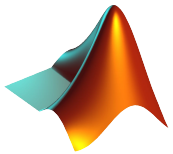
variables

observations

Tools and Training

- Tidy Data (online resources)
- Open Refine (https://openrefine.org/)
- Carpentries https://carpentries.org Lessons - Unix Shell, Git, Python, R & data management and analysis (Ecology, Genomics, Social Sciences, Geospatial, plus..)
- iNZight lite https://lite.docker.stat.auckland.ac.nz/
- nVIVO
- SPSS

OpenRefine
A free, open source, powerful tool for working with messy data

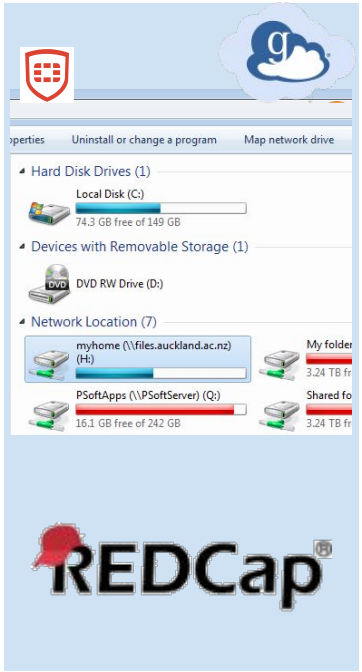



And more

Active data storage

Local

Network



University managed cloud

Dropbox

Google Drive
Keep everything. Share anything.

OneDrive

Virtual Machine (VM) & HPC

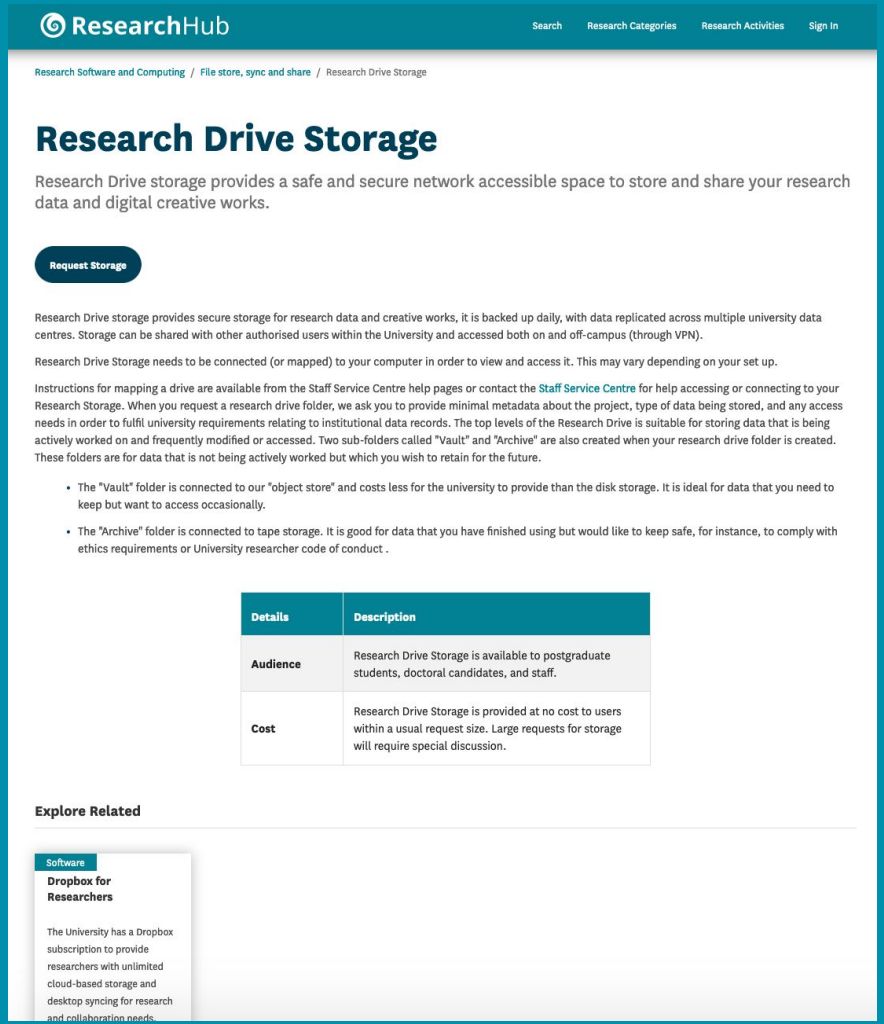
nectar

NeSI
New Zealand eScience Infrastructure

Third Party/ Personal cloud

amazon web services™

⚠ CAUTION



ResearchHub

Search Research Categories Research Activities Sign In

Research Software and Computing / File store, sync and share / Research Drive Storage

Research Drive Storage

Research Drive storage provides a safe and secure network accessible space to store and share your research data and digital creative works.

Request Storage

Research Drive storage provides secure storage for research data and creative works, it is backed up daily, with data replicated across multiple university data centres. Storage can be shared with other authorised users within the University and accessed both on and off-campus (through VPN).

Research Drive Storage needs to be connected (or mapped) to your computer in order to view and access it. This may vary depending on your set up.

Instructions for mapping a drive are available from the Staff Service Centre help pages or contact the Staff Service Centre for help accessing or connecting to your Research Storage. When you request a research drive folder, we ask you to provide minimal metadata about the project, type of data being stored, and any access needs in order to fulfil university requirements relating to institutional data records. The top levels of the Research Drive is suitable for storing data that is being actively worked on and frequently modified or accessed. Two sub-folders called "Vault" and "Archive" are also created when your research drive folder is created. These folders are for data that is not being actively worked but which you wish to retain for the future.

- The "Vault" folder is connected to our "object store" and costs less for the university to provide than the disk storage. It is ideal for data that you need to keep but want to access occasionally.
- The "Archive" folder is connected to tape storage. It is good for data that you have finished using but would like to keep safe, for instance, to comply with ethics requirements or University researcher code of conduct .

Details	Description
Audience	Research Drive Storage is available to postgraduate students, doctoral candidates, and staff.
Cost	Research Drive Storage is provided at no cost to users within a usual request size. Large requests for storage will require special discussion.

Explore Related

Software

Dropbox for Researchers

The University has a Dropbox subscription to provide researchers with unlimited cloud-based storage and desktop syncing for research and collaboration needs.

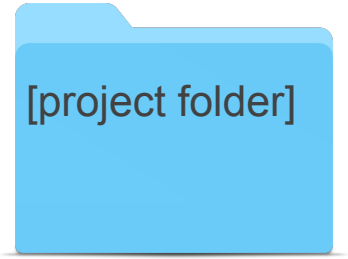
We recommend either or a combination of:

- *Research Drive* - ideal for collaboration within the university and systems, and sensitive data - remote access with VPN
- *University Dropbox* - unlimited, supports external collaborators, check suitability for sensitive data

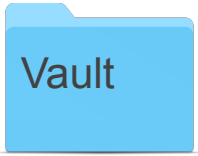
https://research-hub.auckland.ac.nz/research-software-and-computing/store-sync-share/research-storage

Sustainable storage

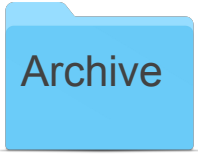
Research Drive - University managed storage, on-site, backed-up



Files, active data, *developing* README.txt
Fastest ($$$/GB)
Requires active ‘push’ to Vault

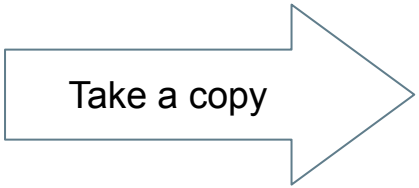


Files, semi-static, *refining* README.txt
Quick ($$/GB)
Requires active ‘push’ to Archive

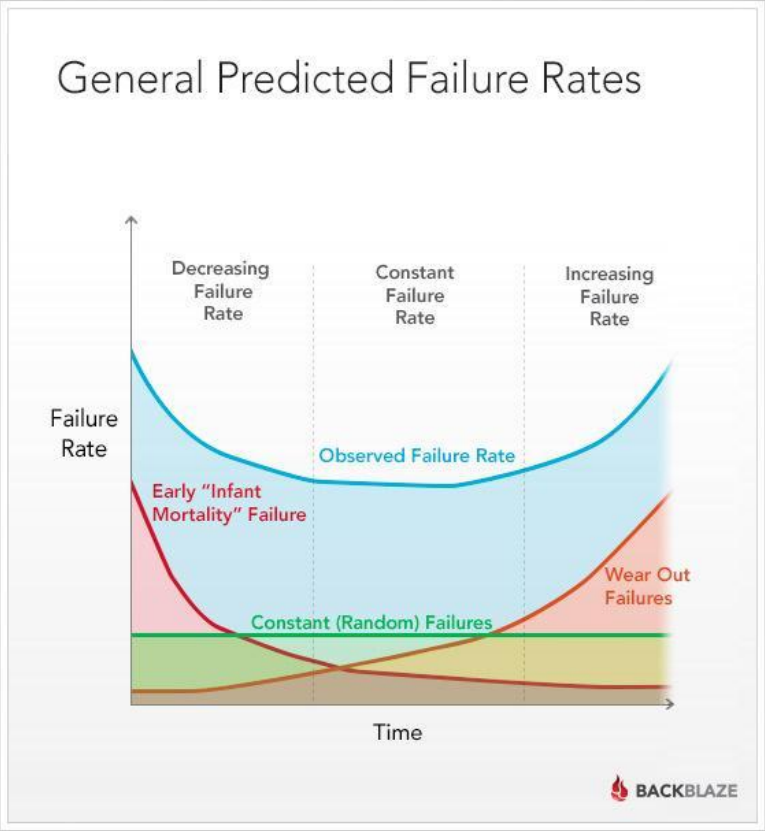


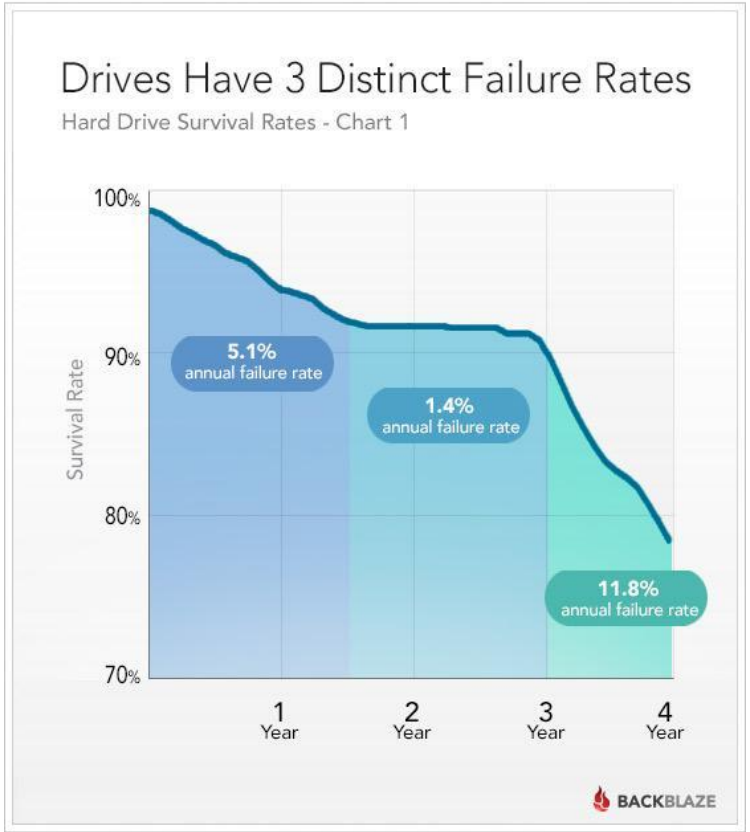
Files, static - not changing, README.txt
Slow ($/GB)
Long-term retention.

Leaving University?



Backup

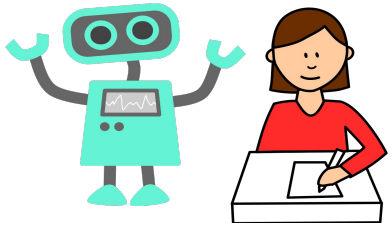




https://www.backblaze.com/blog/how-long-do-disk-drives-last/

What are *you* creating?

Metadata



Data/ information about your data

- Collaboration
- Discovery - human and machine readable
- Reuse - README.txt, ELN, data dictionary, code book
- Preservation

Project

- What is the study?
- Methodologies and instruments
- Bibliographic references

File/Database

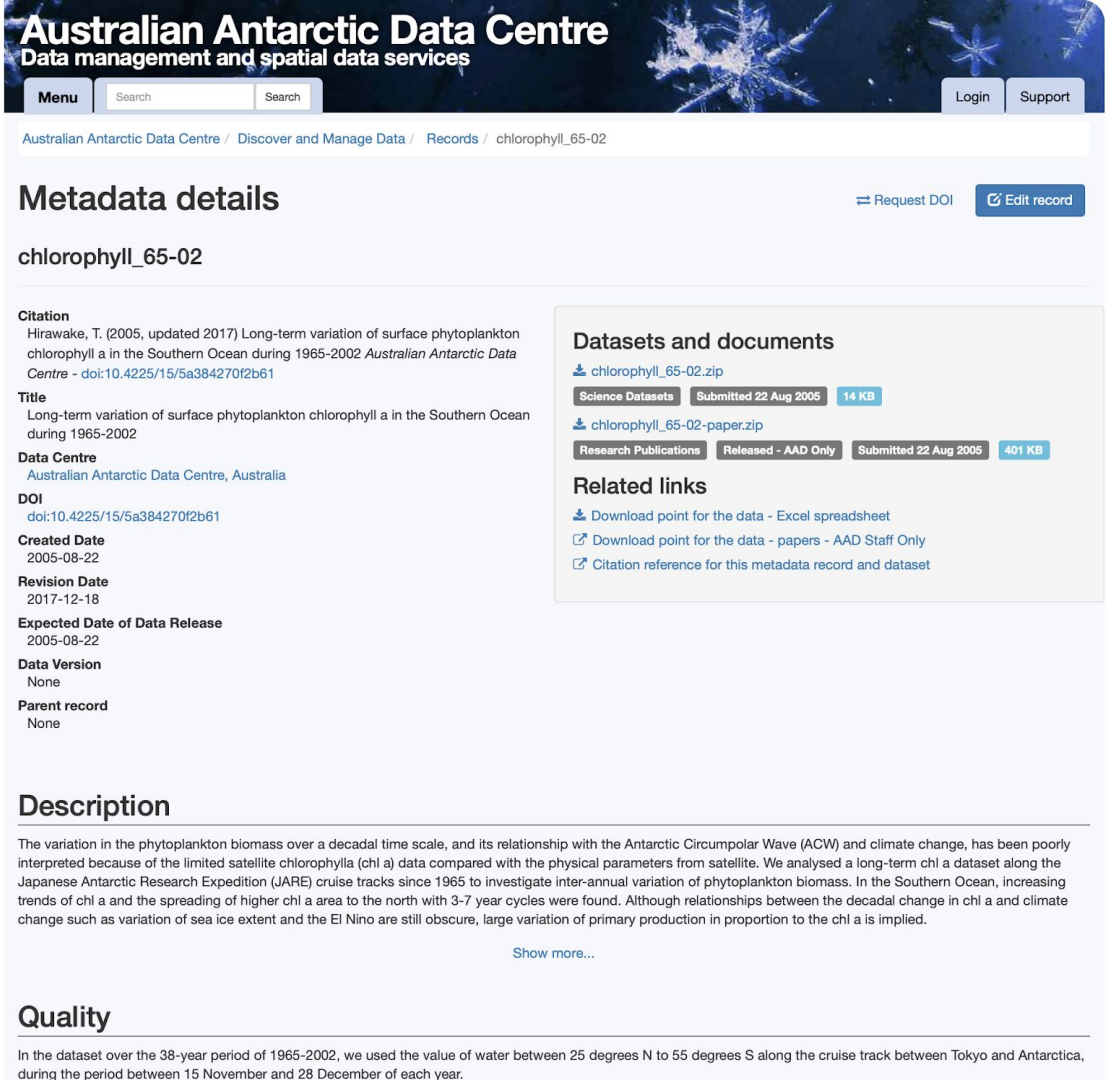
- How files or tables relate?
- What formats?
- README.txt

Item or variable

- Meaning or definition of variable terms

Metadata exemplar

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



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

Citation
Hirawake, T. (2005, updated 2017) Long-term variation of surface phytoplankton chlorophyll a in the Southern Ocean during 1965-2002 *Australian Antarctic Data Centre* - doi:10.4225/15/5a384270f2b61

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

Expected Date of Data Release
2005-08-22

Data Version
None

Parent record
None

Datasets and documents

chlorophyll_65-02.zip
Science Datasets | Submitted 22 Aug 2005 | 14 KB

chlorophyll_65-02-paper.zip
Research Publications | Released - AAD Only | Submitted 22 Aug 2005 | 401 KB

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

Description

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.

Show more...

Quality

In the dataset over the 38-year period of 1965-2002, we used the value of water between 25 degrees N to 55 degrees S along the cruise track between Tokyo and Antarctica, during the period between 15 November and 28 December of each year.

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:24.567	**Westernmost:**100.147
Southernmost:-54.985	**Easternmost:**137.95

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

Platforms

- SHIPS

Instruments

None

Researchers

- Hirawake, Toru (INVESTIGATOR,TECHNICAL CONTACT)
- Connell, Dave (DIF AUTHOR)

Use Constraints

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

- BIOTA
- CLIMATOLOGY/METEOROLOGY/ATMOSPHERE
- OCEANS

Dataset Language

- ENGLISH

Orignating Centre

- JARE

Dataset Progress

- COMPLETE

IDN Node

- 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

Metadata Revision History

2010-07-27 - record updated by Dave Connell to change URL Content Type. 2017-12-18 - record updated by Dave Connell - basic udpates.

Metadata standards

- Research Data Alliance
  http://rd-alliance.github.io/metadata-directory/standards/
- FAIR sharing
  https://fairsharing.org/
- Digital Curation Centre
  http://www.dcc.ac.uk/

Arts and Humanities

Encoded Archival Description (EAD) Edit

A standard for encoding archival finding aids using XML in archival and manuscript repositories, implementing the recommendations of the International Council on Archives ISAD(G): General International Standard Archival Description.

DDI (Data Documentation Initiative) Edit

A widely used, international standard for describing data from the social, behavioral, and economic sciences. Two versions of the standard are currently maintained in parallel:

- DDI Codebook (or DDI version 2) is the simpler of the two, and intended for documenting simple survey data for exchange or archiving. Version 2.5 was released in January 2014.
- DDI Lifecycle (or DDI version 3) is richer and may be used to document datasets at each stage of their lifecycle from conceptualization through to publication and reuse. It is modular and extensible. Version 3.2 was published in March 2014.

Both versions are XML-based and defined using XML Schemas. They were developed and are maintained by the DDI Alliance.

MIDAS-Heritage Edit

A British cultural heritage standard for recording information on buildings, archaeological sites, shipwrecks, parks and gardens, battlefields, areas of interest and artefacts.

Sponsored by the Forum on Information Standards in Heritage, MIDAS Version 1.1 was released in October 2012.

OAI-ORE (Open Archives Initiative Object Reuse and Exchange) Edit

The goal of these standards is to expose the rich content in aggregations of Web resources to applications that support authoring, deposit, exchange, visualization, reuse, and preservation. The standards support the changing nature of scholarship and scholarly communication, and the need for cyberinfrastructure to support that scholarship, with the intent to develop standards that generalize across all web-based information including the increasing popular social networks of "Web 2.0".

Engineering

CIF (Crystallographic Information Framework) Edit

A well-established standard file structure for the archiving and distribution of crystallographic information, CIF is in regular use for reporting crystal structure determinations to Acta Crystallographica and other journals.

Sponsored by the International Union of Crystallography, the current standard dates from 1997. As of July 2011, a new version of the CIF standard is under consideration.

CSMD (Core Scientific Metadata Model) Edit

A study-data oriented model, primarily in support of the ICAT data managment infrastructure software. The CSMD is designed to support data collected within a large-scale facility's scientific workflow; however the model is also designed to be generic across scientific disciplines.

Sponsored by the Science and Technologies Facilities Council, the latest full specification available is v 4.0, from 2013.

ISA-Tab Edit

Data Principles

FAIR
https://www.go-fair.org/fair-principles/

CARE - Global Indigenous Data Alliance
https://www.gida-global.org/care

Maori Data Sovereignty
https://www.temanararaunga.maori.nz

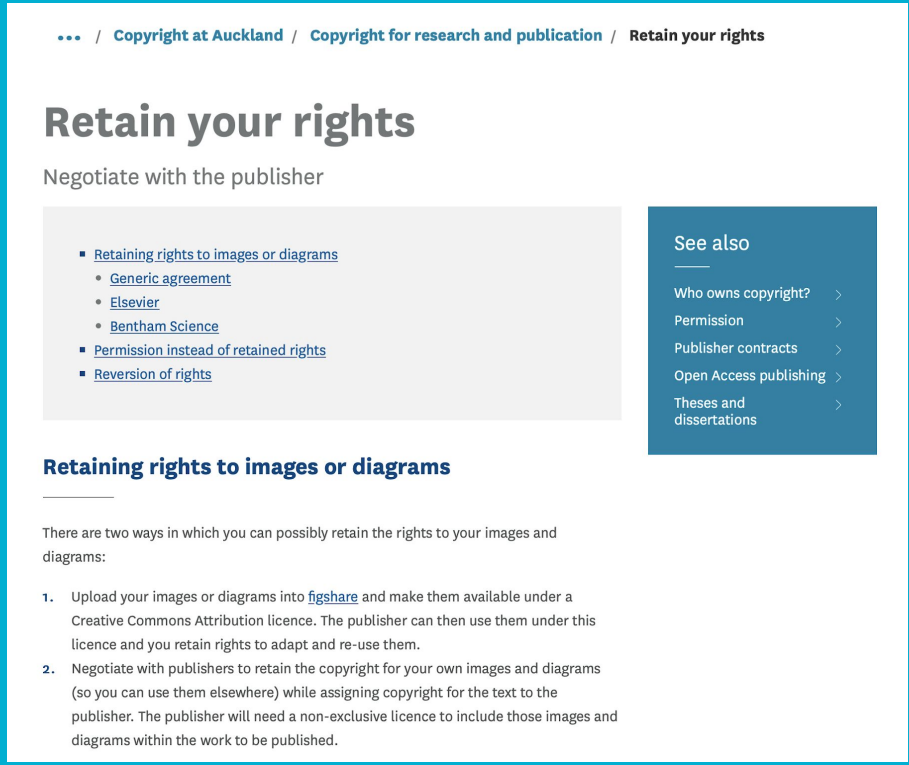
UoA' Taumata Teitei,
"becoming a Māori Data Sovereignty organisation." p4
https://www.auckland.ac.nz/en/about-us/about-the-university/the-university/official-publications/strategic-plan-development.html

University policy: research data retention

- Minimum 6 years
- Clinical trial – 10 years (or until children turn 26)
- Patent – 21 years from date of filing
- Ethics approved – check
- Community or heritage value – indefinitely

Who owns your data -holds the [copy]right?

- Copyright holder can license a work
- Choose open licences
- Check rights and permissions if re-using data
- Start ownership discussions early
- Explicit copyright display
- “As open as possible, as closed as necessary”



... / Copyright at Auckland / Copyright for research and publication / Retain your rights

Retain your rights

Negotiate with the publisher

- Retaining rights to images or diagrams
  - Generic agreement
  - Elsevier
  - Bentham Science
- Permission instead of retained rights
- Reversion of rights

See also

- Who owns copyright?
- Permission
- Publisher contracts
- Open Access publishing
- Theses and dissertations

Retaining rights to images or diagrams

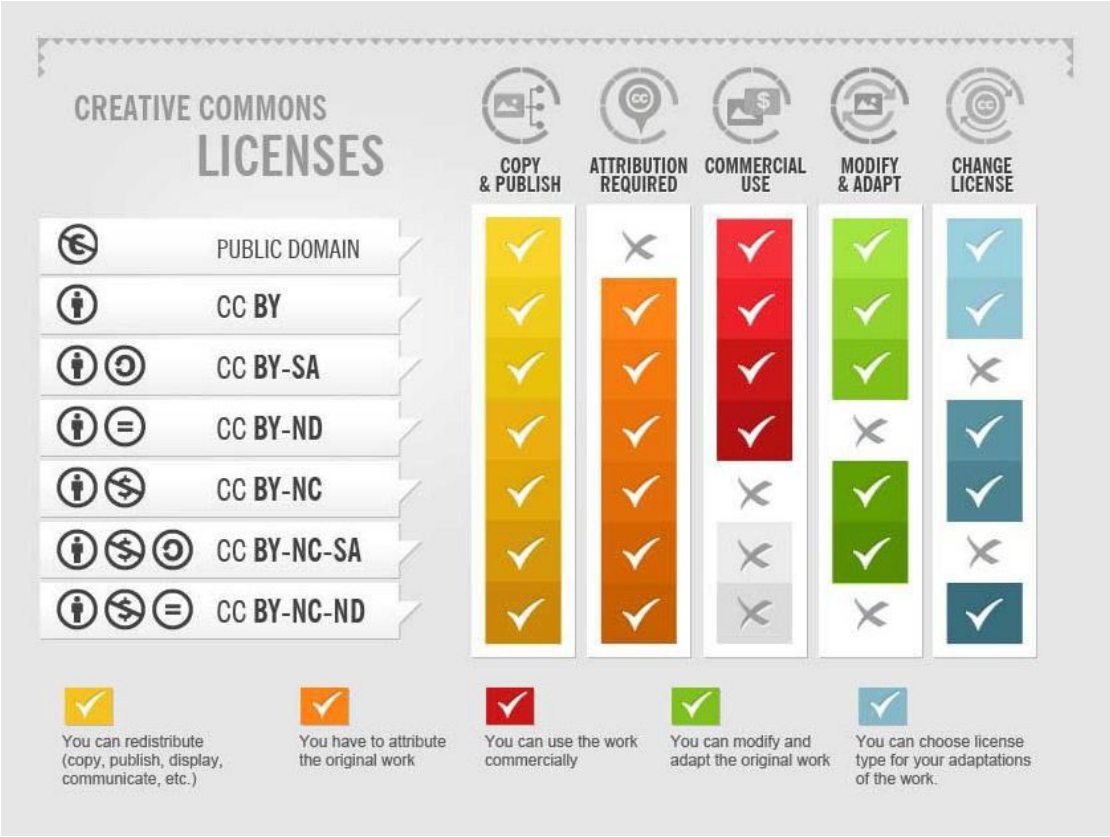
There are two ways in which you can possibly retain the rights to your images and diagrams:

1. Upload your images or diagrams into figshare and make them available under a Creative Commons Attribution licence. The publisher can then use them under this licence and you retain rights to adapt and re-use them.
2. Negotiate with publishers to retain the copyright for your own images and diagrams (so you can use them elsewhere) while assigning copyright for the text to the publisher. The publisher will need a non-exclusive licence to include those images and diagrams within the work to be published.

https://www.auckland.ac.nz/en/staff/learning-and-teaching/policies-guidelines-procedures/copyright-at-auckland/research-publication/retain-rights.html

Creative Commons Licenses

- Make it easy to allow reuse of your works by others
- Creative commons licences to choose from
- Most open licences allow adaptation, remix and sharing of materials



https://creativecommons.org/use-remix/cc-licenses/

https://medium.com/@adventureschool/understanding-copyright-laws-creative-commons-royalty-free-and-public-domain-what-do-they-all-871aaf2b1247

Data publishing - why?

- Research(er) integrity
- Funder or **Publisher requirements**
- Institutional obligation
- Collaboration
- Innovation and reuse
- Impact
- Preservation
- Teaching
- Public record

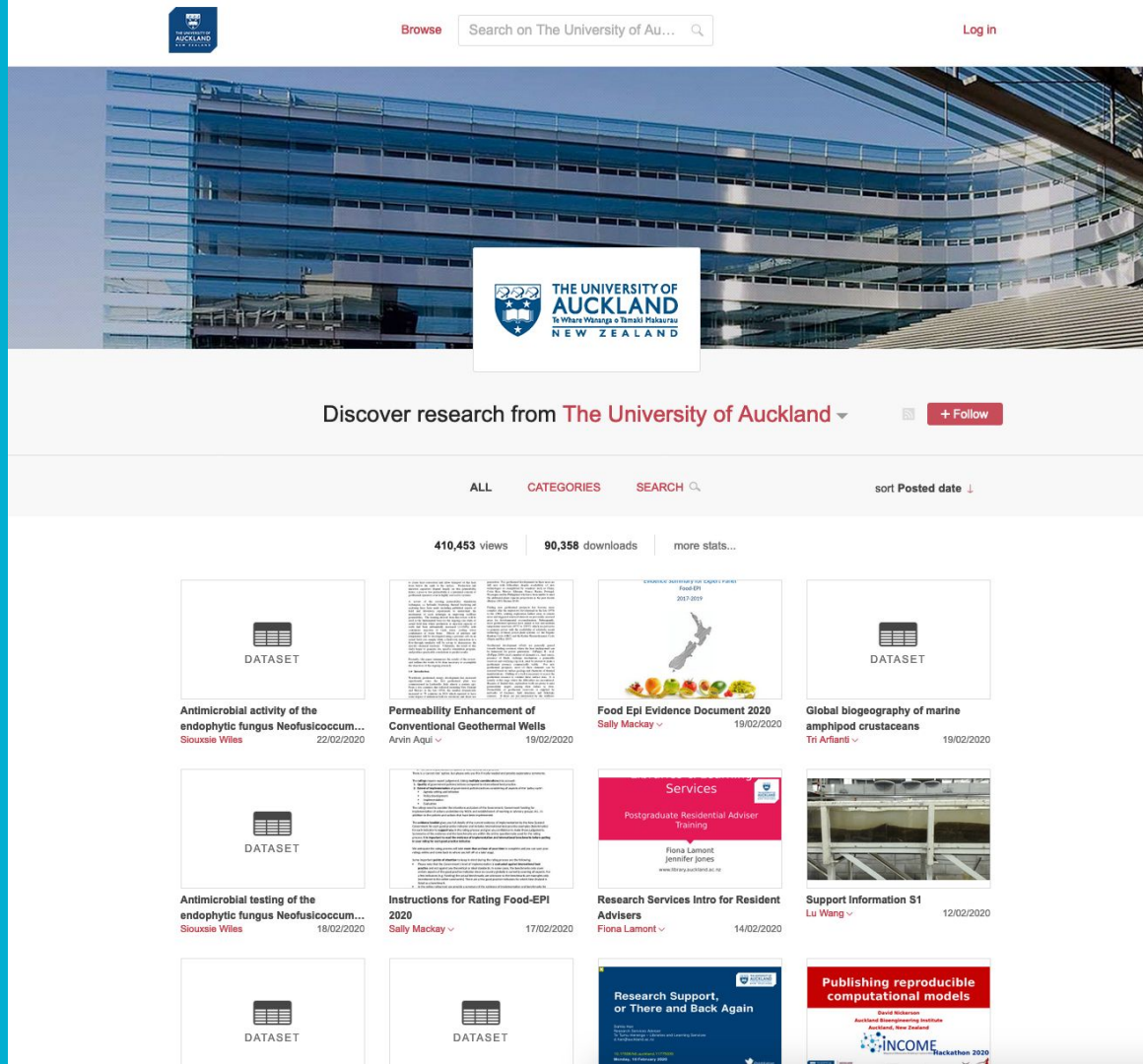
1. Deposit (DOI)
2. Citation
3. Linking or Data Availability Statement
4. Peer review of data

Elsevier
SpringerNature
Wiley

Data Publishing and Discovery Service

https://auckland.figshare.com

Publish data
OR
metadata-only record



F

Findable

F1. (meta)data are assigned a globally unique and persistent identifier

F2. data are described with rich metadata (defined by R1 below)

F3. metadata clearly and explicitly include the identifier of the data it describes

F4. (meta)data are registered or indexed in a searchable resource

A

Accessible

A1. (meta)data are retrievable by their identifier using a standardized communications protocol

A1.1 the protocol is open, free, and universally implementable

A1.2 the protocol allows for an authentication and authorization procedure, where necessary

A2. metadata are accessible, even when the data are no longer available

I

Interoperable

II. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

I2. (meta)data use vocabularies that follow FAIR principles

I3. (meta)data include qualified references to other (meta)data

R

Reusable

R1. meta(data) are richly described with a plurality of accurate and relevant attributes

R1.1. (meta)data are released with a clear and accessible data usage license

R1.2. (meta)data are associated with detailed provenance

R1.3. (meta)data meet domain-relevant community standards

https://www.force11.org/group/fairgroup/fairprinciples

Archive and preservation

- Passive vs active.
- Publishing as archive.
- What are you trying to achieve or enable?
- Where have you left your data?
- Who is responsible, who is the data steward?
- Will the future be able to open and make sense of it?

https://library.si.edu/research/best-practices-storing-archiving-and-preserving-data
https://www.ands.org.au/working-with-data/data-management/data-preservation

Allow time for data management AND document in DMP

Tuesdays 3-4pm, https://auckland.zoom.us/my/hackyhour

Website: https://uoa-eresearch.github.io/HackyHour/

Acknowledgement & references

In the creation of the workshop we have taken inspiration and adapted some ideas and materials from a number of existing resources.

Research Data Management: File Organization
Katherine McNeill & Helen Bailey
http://libraries.mit.edu/data-management/files/2014/05/file-organization-july2014.pdf
(CC-BY-NC-SA)

Melbourne_MANTRA
University of Melbourne and University of Edinbrugh
http://library.unimelb.edu.au/digitalscholarship/training_and_outreach/mantra2
(CC-BY)

Research Data Management: 101 The Lifecycle of a Dataset
Katherine McNeill
http://libraries.mit.edu/data-management/files/2014/05/research-data-management-iap2014.pdf
(CC-BY-NC-SA)

Escaping Datageddon - Dorothea Salo and Ryan Schryver - University of Wisconsin
http://researchdata.wisc.edu/wp-content/uploads/EscapingDatageddon1.pdf
(CC-BY)

Managing and Sharing Data: Best Practices for Researchers. Veerle Van den Eynden, Louise Corti, Matthew Woollard and Libby Bishop
http://www.data-archive.ac.uk/media/2894/managingsharing.pdf
(CC-BY-NC-SA)

Australian National Data Service website at http://ands.org.au/guides/data-citation-awareness.html
Accessed 8 December 2015
(CC-BY)

Tidy Data http://vita.had.co.nz/papers/tidy-data.pdf