

**Prepare a  
Data  
Management  
Plan (DMP)**

# Today's agenda:

- What is a Data Management Plan (DMP)
- Major components of a DMP
- Evaluate a DMP
- Prepare a DMP with DMPTool

1.

What is a Data  
Management  
Plan (DMP)



# Data management plan (DMP)

A document in which you describe

**what data** you will collect during your research project,

format, types, volume, ownership...

**how** you are going to store and manage the data during the project,

File organization, metadata, access...

And **what will happen** to the data after the project is finished.

Preservation and sharing...

Plan and Design

Discover,  
Reuse,  
Cite

## Research data life cycle

Collect and  
Capture

Share and  
Publish

Collaborate  
and  
Analyse

Store and  
Preserve

Source: [JISC](#)

# Example DMP

## DATA MANAGEMENT PLAN

The project will collect and analyze the following data:

- Conductivity and temperature from glider surveys.
- Horizontal currents from shipboard ADCP and the HDSS Doppler Sonars on the R.V. Revelle.
- LADCP/ CTD profiles from the R.V.Revelle.
- Moored ADCPs.
- CTD-u,v profiles from the McLane profilers.
- CTD profiles from the SIO Fast-CTD.
- Fine and microscale temperature from CHIPODs and moored thermistor chains.

## Quick-Response data management

The T-TIDE PIs have experience with this mix of data types from previous collaborative efforts, such as the ONR TWSE Experiment, 2010-11, in the S. China Sea. To guide both modeling and the Process experiment planning, quick-look Scout data will be centralized on a server at APL UW.

Scout Quick-look data responsibilities include:

J.Klymak	LADCp-CTD analysis.
S.Johnston	SIO glider analysis
L.Rainville	Co-operative CSIRO glider Tidal analysis
H.Simmons, J.Klymak	Ongoing model output predictions
R.Pinkel, J.Klymak	F-C TID site studies

The centralized data access will be maintained for the Process Experiment, with the McLane and thermistor chain data provided by the relevant PIs.

## Long Term data Archiving

Aside from the LADCP-shipboard CTD profiles, there are currently no established standards for archiving or data from many of the fine-scale sensors used in T-Tide. Archiving standards for glider data are evolving. This is a concern of the Climate Process Team on Ocean Mixing, of which many T-Tide PIs are members. We propose to work with the CPT to evolve formats for data and metadata suitable for archiving both sensor and (critically) model output from the experiment.

All field data collected under this program will be made available as per NSF guidelines within 2 years of collection via published manuscripts, publicly available final reports to NSF, and data archiving with NODC.

Data will be shared in matlab MAT file format and/or as netCDF files. Ultimate archival formats will be determined in consultation with NODC and with the CPT. Adequate archiving is anticipated to be an expensive, time-consuming task. All PIs have included funds for this effort in their budgets.

The primary T-TIDE models are all public domain. Published peer-reviewed manuscripts will document the simulations and forcing sufficiently. Recognizing that archiving high-resolution simulations at tidally resolving intervals can result in gigabytes-to-terabytes of data, every effort will be made by modeling PIs to archive model output and provide data and/or code to interested parties upon request. Model products and output will be available at the end of the grant period.

# To prepare a DMP

- Precise
- Realistic and workable
- Address major issues
- Revisit and revise

2.

Who has  
to submit a  
DMP

# Why research data management?

## Requirements

- Compliance with policies: HKU & funders
- Ensure data is accessible and shareable: journals requirement
- Demonstrate responsible practice

## Benefits

- Keep research safe and secure
- Increase research efficiency
- Improve research integrity
- Make research outputs more visible
- Enable collaboration

Source: [JISC](#)



# Research Services

Support and information for HKU researchers

[Sitemap](#) [Contact Us](#)
[Quick Links](#)
[Search this site](#)


[Home](#) > [Research Integrity](#) > [Research Data and Records Management](#)

## Research Data and Records Management

[About Us](#)

[Research Funding](#)

The management of research data and records refers to ways in which recorded information (in whatever form or medium) from research is organised, stored, maintained and accessed both during the lifespan of the research and in the long term. Effective research data and records management supports both high quality research and academic

5. Research data and records should be retained for as long as they are of continuing value to the researcher and the wider research community, and as long as specified by research funder, patent law, legislative and other regulatory requirements. The minimum retention period for research data and records is three years after publication or public release of the work of the research. In many instances, researchers will resolve to retain research data and records for a longer period than the minimum requirement.

[Policy on the Management of Research Data and Records](#)

[Honours & Awards](#)

[RAE 2014 & Resources](#)

1. The University of Hong Kong seeks to promote the highest standards in the management of research data and records (1) as fundamental to both high quality research and academic integrity, and acknowledges its obligations under research funders' data-related policy statements and codes of practice, where available (2), to ensure that sound systems are in place to promote best practice, including through clear policy, guidance, supervision, training and support.

2. The University recognises that accurate and retrievable research data are an essential component of any research project and necessary to verify and defend, when required, the process and outcomes of research. Research data are

# Timeline (RPG)



"RPG" includes the degrees of MPhil, PhD, and SJD.

Doctor of Legal  
Science

# Who has to submit a DMP

- A. data is freely available on the internet or in libraries, or
- B. license forbids deposit.
- C. **no data** was used in the research project for the creation of the thesis
- D. **data** was used in the research project**

# HKUL Research Data Services

<http://lib.hku.hk/researchdata/>

A screenshot of a Google search results page. The search query "hku research data" is entered in the search bar. Below the search bar, there are tabs for 全部 (All), 圖片 (Images), 新聞 (News), 地圖 (Maps), 影片 (Videos), and 更多 (More). On the right side, there are links for 設定 (Settings) and 工具 (Tools). The search results section shows a snippet for the HKUL Research Data Services page.

約 1,370,000 項搜尋結果 (0.49 秒)

## HKUL Research Data Services - HKU Libraries

[lib.hku.hk/researchdata/rds.htm](http://lib.hku.hk/researchdata/rds.htm) ▾ 翻譯這個網頁

HKU now requires **Research Data** Management (RDM). RDM is a general term covering how you organize, structure, store, and care for the information used or ...

## Research Data and Records Management - HKU Research Services

[www.rss.hku.hk](http://www.rss.hku.hk) > **Research Integrity** ▾ 翻譯這個網頁

The management of **research data** and records refers to ways in which recorded information (in whatever form or medium) from **research** is organised, stored, ...

# Who has to submit a DMP

<http://lib.hku.hk/researchdata/rpg4.htm>

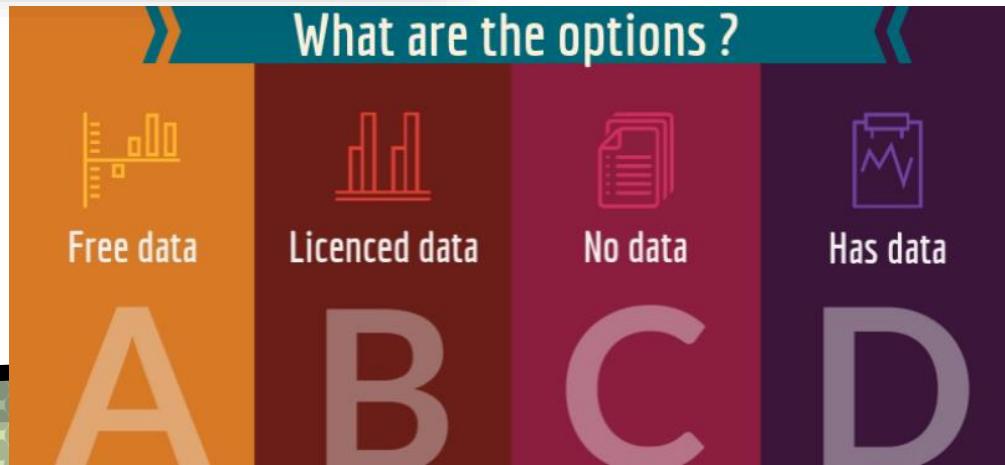
The screenshot shows the homepage of the HKUL Research Data Services. At the top, there is a navigation bar with the university's crest, the text "HKUL RESEARCH DATA SERVICES", and links for "RPG STUDENTS", "RPG SUPERVISOR", and "RDM". There is also a search icon and a menu icon. Below the navigation bar, there is a small image of a building and the text "RESEARCH DATA SERVICES". Underneath that, it says "THE UNIVERSITY OF HONG KONG LIBRARIES". A large white box in the center contains the text "RPG DATA MANAGEMENT PLAN (DMP) INPUT FORM". To the right of this box, there is a link "RPG STUDENTS" next to a green circular icon with a person silhouette. At the bottom of the page, there is a green button labeled "RPG LOGIN" inside an oval outline.

# Data Options

The screenshot shows the World Bank Data API interface. At the top, it says "THE WORLD BANK" and "Data". Below that is a search bar with "Poverty headcount ratio at \$..." and a dropdown menu set to "World". A placeholder text "Search data e.g. GDP, population, In" is visible. The main content area displays the results for "Poverty headcount ratio at \$1.90 a day (2011 PPP) (% of population)". It includes a chart and the text "World Bank, Development Research Group. Data are based on".

*"My dataset is downloaded from the World Bank at*

`https://data.worldbank.org/indicator/SI.POV.DDAY?end=2013&locations=1W&start=1992&view=chart`



# Data Options

## YOU HAVE SELECTED OPTION A:

- A. Data is freely available on the internet, in libraries or archives. DMP and Dataset submission are not needed. Primary supervisor approval will be sought.

Please give URL of your dataset retrieved online :

[http://dx.doi.org/10.5353/ds\\_dataset95133](http://dx.doi.org/10.5353/ds_dataset95133)

Please give citation, including address of library or archive used (if not online) :

Bacon-Shone, J, Bolton, K, Luke, KK. (2016). Data from LANGUAGE USE, PROFICIENCY AND ATTITUDES IN HONG KONG. (Dataset) The University of Hong Kong, Pokfulam, Hong Kong SAR. [http://dx.doi.org/10.5353/ds\\_dataset95133](http://dx.doi.org/10.5353/ds_dataset95133)

# Data Options

## Option A - Data citation

**Creator.** (Publication year). **Title.** Publisher. Identifier

Bacon-Shone, J, Bolton, K, Luke, KK. (2016). Data from LANGUAGE USE, PROFICIENCY AND ATTITUDES IN HONG KONG. (Dataset) The University of Hong Kong, Pokfulam, Hong Kong SAR.

[http://dx.doi.org/10.5353/ds\\_dataset95133](http://dx.doi.org/10.5353/ds_dataset95133)

- Community standard only. If applicable, also include version and resource type.

# Data Options

*"I'm now an intern at the Transport Department, I'm going to make use of some datasets I'm working on as an intern at the Department for my research..."*

## What are the options ?



Free data



Licenced data



No data



Has data

# Data Options

*"I was granted the permission to use data from the Housing Department. And I'm going to conduct a survey or perhaps some interviews with my subjects..."*

## What are the options ?



Free data



Licenced data



No data



Has data

A

B

C

D

3.

Develop a  
DMP with  
DMPTool@  
HKUL

## Welcome to DMPTool@HKU Libraries

Create data management plans that meet institutional and funder requirements.



### DMPTool by the Numbers

**598**

Users

**511**

Plans

**HKU**

Participating Institution

### Contact Us

For any questions, please send email to HKUL  
Research Data Services at [researchdata@hku.hk](mailto:researchdata@hku.hk)

# 4.

## Major components of a DMP

# **7 Major Components of a DMP**

- 1. Data Collection**
- 2. Documentation and Metadata**
- 3. Ethics and Legal Compliance**
- 4. Storage and Backup**
- 5. Selection and Preservation**
- 6. Data Sharing**
- 7. Responsibilities and Resources**

# Administrative Data (RPG)

- Your Name
- University ID
- Email
- Degree
- Department/Faculty
- Field of Study
- Supervisor(s)
- Project Title and Description
- Date and Version

# Administrative Data (PI)

<b>Project Title:</b>	A cost-effectiveness threshold of body patients
<b>HKU Project Code:</b>	0201812
<b>Principal Investigator:</b>	Professor Chan Tai Man
<b>Co-Investigator(s):</b>	Dr Lam Mary Professor Smith Sandra Dr Cheung David
<b>Start Date:</b>	2018-07-05
<b>Completion Date:</b>	2020-07-04
<b>Grant Type:</b>	HMRF Research Fellowship Scheme
<b>Amount:</b>	2000000
<b>Funding Year:</b>	2018/2019

## Data Management Plan for Post-Graduate Research Projects

Name:	
Student ID:	
Email:	
Faculty/Department:	
Supervisor:	

### RESEARCH PROJECT TITLE

#### DATA COLLECTION

How will the data be collected or created?

What data will you collect or create?

#### DOCUMENTATION AND METADATA

What documentation and metadata will accompany the data?

#### ETHICS AND LEGAL COMPLIANCE

How will you manage copyright and Intellectual Property Rights (IPR) issues?

How will you manage any ethical issues?

#### STORAGE AND BACKUP

How will you manage access and security?

How will the data be stored and backed up during the research?

#### SELECTION AND PRESERVATION

What is the long-term preservation plan for the dataset?

Which data are of long-term value and should be retained, shared, and/or preserved?

#### DATA SHARING

Are any restrictions on data sharing required?

How will you share the data?

#### RESPONSIBILITIES AND RESOURCES

What resources will you require to deliver your plan?

Who will be responsible for data management?

Prepared by:	Endorsed by:
RPg student:	Supervisor:
Date:	Date:

RPGs only

# Research\_Project\_DMP.doc

# 1. Data Collection

- What data will you collect or create?

- Numbers
- Text – survey data, interview transcripts
- Multimedia – image, audio, video
- Software, programming scripts
- Models
- *in digital or physical forms*

# 1. Data Collection

- What data will you collect or create?
- How will the data be collected or created?

Type, format, volume of data

Expected rate of increase

How will you structure and name your folders and files?

What standards or methodologies will you use?

# 1. Data Collection

Data stage	Specification of type of research data	Data size/growth
Raw data		
Processed data		

# 1. Data Collection

## File Formats

It is important to plan for software obsolescence.

Formats more likely to be accessible in the future are:

- Non-proprietary
- Open, documented standard
- Standard representation (ASCII, Unicode)
- Unencrypted
- Uncompressed

Examples of preferred file format choices include:

TXT, or PDF/A,	not Word
CSV,	not Excel
MPEG-4,	not Quicktime
WAV,	not MP3
TIFF or JPEG2000,	not GIF or JPG

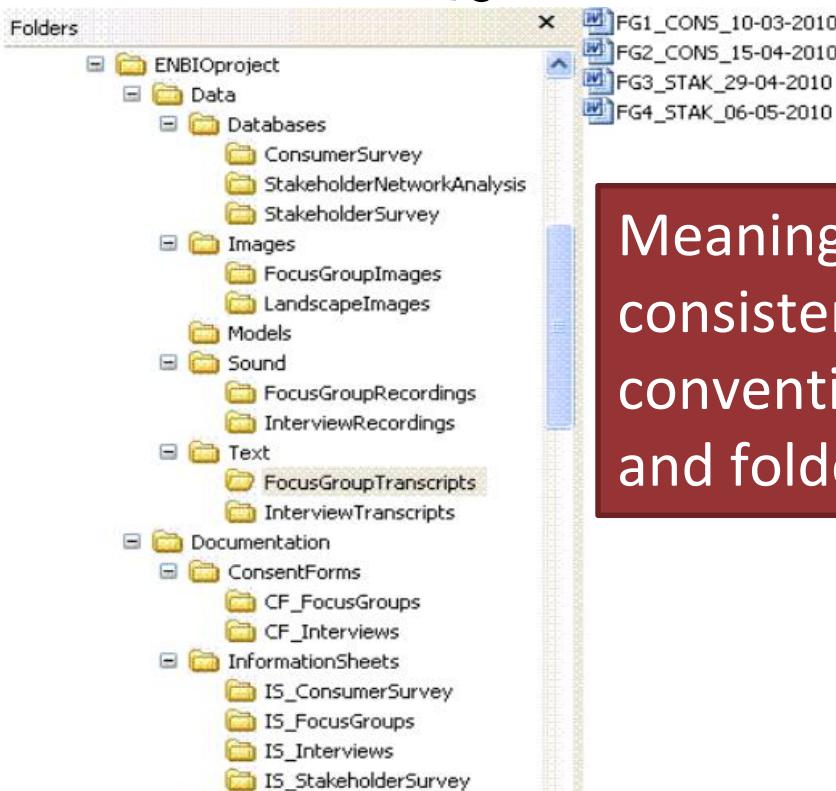
# 1. Data Collection

## Recommended formats

Type of data	Recommended formats	Acceptable formats
<b>Tabular data with minimal metadata</b> column headings, variable names	comma-separated values (.csv) tab-delimited file (.tab) delimited text with SQL data definition statements	delimited text (.txt) with characters not present in data used as delimiters widely-used formats: MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb), dBase (.dbf), OpenDocument Spreadsheet (.ods)
<b>Textual data</b>	Rich Text Format (.rtf) plain text, ASCII (.txt) eXtensible Mark-up Language (.xml) text according to an appropriate Document Type Definition (DTD) or schema	Hypertext Mark-up Language (.html) widely-used formats: MS Word (.doc/.docx) some software-specific formats: NUD*IST, NVivo and ATLAS.ti
<b>Image data</b>	TIFF 6.0 uncompressed (.tif)	JPEG (.jpeg, .jpg, .jp2) if original created in this format. GIF (.gif). TIFF other versions (.tif, .tiff) RAW image format (.raw). Photoshop files (.psd) BMP (.bmp). PNG (.png) Adobe Portable Document Format (PDF/A, PDF) (.pdf)
<b>Audio data</b>	Free Lossless Audio Codec (FLAC) (.flac)	MPEG-1 Audio Layer 3 (.mp3) if original created in this format Audio Interchange File Format (.aif) Waveform Audio Format (.wav)
<b>Video data</b>	MPEG-4 (.mp4) OGG video (.ogv, .ogg) motion JPEG 2000 (.mj2)	AVCHD video (.avchd)

# 1. Data Collection

## Example folder structure



Meaningful and consistent naming convention for files and folders

# 1. Data Collection

Best practice is to:

## File names

- create meaningful but brief names
- use file names to classify types of files
- **avoid using spaces, dots and special characters** (& or ? or !)
- use hyphens (-) or underscores (\_) to separate elements in a file name
- avoid very long file names
- reserve the 3-letter file extension for application-specific codes of file format (e.g. .doc, .xls, .mov, .tif)
- include versioning within file names where appropriate

## 2. Documentation and Metadata

- What documentation and metadata will accompany the data?

What information is needed for the data to be read and interpreted in the future?

How will you capture/create the documentation and metadata?

What metadata standards will you use?

# pollev.com/researchdata250

Poll Everywhere

Responding as abc

What is DATA1 about? Any other information you need in order to reuse this dataset?

You have not responded

Enter a response

Submit

# Dataset

File(s) held in the Hub

Title of Dataset	Data from: Benchmarking construction waste management performance using big data Data from: Analysis of the construction waste management performance in Hong Kong: the public and private sectors compared using big data Data from: The S-curve for forecasting waste generation in construction projects Data from: Identifying factors influencing demolition waste generation in Hong Kong
Author of Dataset	Lu, W  <sup>1</sup>
Contact	Lu, W  <sup>1</sup>
Date of Dataset Creation	2013-01-04
Description	Construction project and waste management performance
Citation	Lu, W. (2013). Data from: Benchmarking construction waste management performance using big data. (Dataset) The University of Hong Kong, Pokfulam, Hong Kong SAR. <a href="http://dx.doi.org/10.5353/ds_dataset103740">http://dx.doi.org/10.5353/ds_dataset103740</a>
Publisher	The University of Hong Kong, Pokfulam Road, Hong Kong SAR
Subject (RGC Codes)	E1 — Civil Engineering, Surveying, Building & Construction — 土木工程, 測量, 建造及施工 L 2101 — Building & Construction — 建造及施工
Subject (ANZSRC)	05 — ENVIRONMENTAL SCIENCES — 環境科學 L 0502 — ENVIRONMENTAL SCIENCE AND MANAGEMENT — 環境科學與管理 L 050205 — Environmental Management — 環境管理
Keyword	Construction waste management (CWM) Key performance indicator (KPI) Waste generation rate (WGR)

<http://hub.hku.hk/cris/dataset/dataset103740>

# Dataset

File(s) held in the Hub



Title of Dataset	Data from: Benchmarking construction waste management performance using big data Data from: Analysis of the construction waste management performance in Hong Kong: the public and private sectors compared using big data Data from: The S-curve for forecasting waste generation in construction projects
Subject (RGC Codes)	E1 — Civil Engineering, Surveying, Building & Construction — 土木工程, 測量, 建造及施工 L 2101 — Building & Construction — 建造及施工
Subject (ANZSRC)	05 — ENVIRONMENTAL SCIENCES — 環境科學 L 0502 — ENVIRONMENTAL SCIENCE AND MANAGEMENT — 環境科學與管理 L 050205 — Environmental Management — 環境管理
Keyword	Construction waste management (CWM) Key performance indicator (KPI) Waste generation rate (WGR) Benchmarking Big data Data mining Hong Kong
Subject (ANZSRC)	05 — ENVIRONMENTAL SCIENCES — 環境科學 L 0502 — ENVIRONMENTAL SCIENCE AND MANAGEMENT — 環境科學與管理 L 050205 — Environmental Management — 環境管理
Keyword	Construction waste management (CWM) Key performance indicator (KPI) Waste generation rate (WGR)

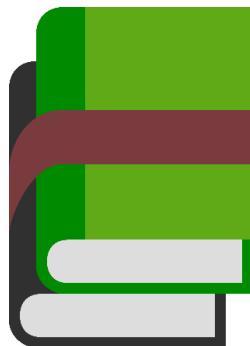
<http://hub.hku.hk/cris/dataset/dataset103740>

## 2. Documentation and Metadata

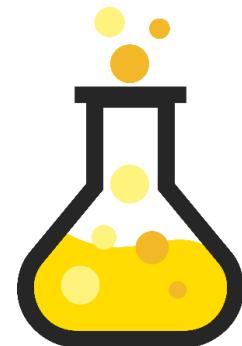
- What documentation and metadata will accompany the data?



Readme file



Codebook/  
Data dictionary



Electronic Lab  
Notebooks

## 2. Documentation and Metadata

Basic information needed to make the data reusable.

Title

Creator

Persistent identifier

Subject

Funders

Rights

Access information

Dates

File names

File format

Versions

**Software and version of the software required for its potential reuse.**

**Hardware and operation system requirements**

## 2. Documentation and Metadata

### Metadata Standards

Metadata (**data about data**) standards help to describe data in a consistent manner. Metadata can include descriptive information, provenance, quality and access/use of data.

# 2. Documentation and Metadata

**D|C|C** because good research needs good data

Contact us [search](#)

Home Digital curation About us News Events Resources Training Projects Community Tailored support

Home > Resources > Metadata Standards

Home > Resources > Metadata Standards > List

## List of Metadata Standards

**ABCD - Access to Biological Collection Data**  
The Access to Biological Collections Data (ABCD) Schema about specimens and observations (a.k.a. primary biodive structured, supporting data from a wide variety of databases that either (or both) atomised data and free-text can be at

Sponsored by Biodiversity Information Standards TDWG - modified in 2007.

**AgMES - Agricultural Metadata Element Set**  
A semantic standard developed by the Food and Agriculture resource discovery, interoperability and data exchange of nutrition and rural development.

Sponsored by the UN AIMS - Agricultural Information Management System

**AVM - Astronomy Visualization Metadata**  
The AVM<sup>®</sup> scheme supports the cross-searching of collected telescopic observations (also known as 'pretty pictures'). The scheme is compatible with the Adobe XMP<sup>®</sup> specification, so the metadata can be embedded within common image formats such as JPEG, TIFF and PNG.

In this section

- Briefing Papers
- How-to Guides & Checklists
- Developing RDM Services
- Curation Lifecycle Model
- Curation Reference Manual
- Policy and legal
- Data Management Plans
- Tools
- Case studies
- Repository audit and assessment
- Standards
- Disciplinary Metadata**
- DIFFUSE
- Publications and presentations
- Roles
- Curation journals
- Informatics research
- External resources
- Online Store

Disciplinary Metadata

While data curators, and increasingly researchers, know that good metadata is key for research data access and reuse, figuring out precisely what metadata to capture and how to capture it is a complex task. Fortunately, many academic disciplines have supported initiatives to formalise the metadata specifications the community deems to be required for data re-use. This page provides links to information about these disciplinary metadata standards, including profiles, tools to implement the standards, and use cases of data repositories currently implementing them.

For those disciplines that have not yet settled on a metadata standard, and for those repositories that work with data across disciplines, the General Research Data section links to information about broader metadata standards that have been adapted to suit the needs of research data.

Please note that a [community-maintained version of this directory](#)<sup>®</sup> has been set up under the auspices of the Research Data Alliance.

### Search by Discipline

 Biology

 Earth Science

 General Research Data

<http://www.dcc.ac.uk/resources/metadata-standards/list>

# 2. Documentation and Metadata

## The Data Documentation Initiative (DDI)



Standards ▾   Resources ▾   Training ▾   Community ▾   Publications ▾   About ▾

### Document, Discover and Interoperate

The Data Documentation Initiative (DDI) is an international standard for describing the data produced by surveys and other observational methods in the social, behavioral, economic, and health sciences. DDI is a free standard that can document and manage different stages in the research data lifecycle, such as conceptualization, collection, processing, distribution, discovery, and archiving. Documenting data with DDI facilitates understanding, interpretation, and use -- by people, software systems, and computer networks. Use DDI to Document, Discover, and Interoperate!



Specification



Tools



Learn



Collaborate

<https://www.ddialliance.org/>

# 2. Documentation and Metadata



## Dublin Core Metadata Initiative

[Home](#)   [News](#)   [DCMI Specifications](#)   [LRMI](#)   [Community and Events](#)   [Join / Support](#)   [About](#)

quick search....

### DCMI Specifications

- [Recommendations](#)
- [Proposed](#)
- [Working Drafts](#)
- [Superseded](#)
- [Community Specifications](#)
- [Recommended Resources](#)
- [Approval Processes](#)
- [Translations](#)

As part of its mission, the Dublin Core Metadata Initiative develops and maintains specifications in support of resource description. Specifications developed and reviewed in the context of DCMI's [formal approval process](#) are assigned a status (in ascending order of maturity and stability) of "DCMI Working Draft", "DCMI Proposed Recommendation", or "DCMI Recommendation". DCMI also provides pointers to guidelines and services developed outside of this formal review context ("Recommended Resources").

This selection highlights the specifications that currently attract the most attention in the Dublin Core community. Links to additional specifications (including superseded specifications) may be found at <http://dublincore.org/documents/>. Some of the specifications have been [translated](#) into one of twenty-five languages.

<http://dublincore.org/specifications/>

## 2. Documentation and Metadata

### Dublin Core Metadata Element Set (DCMES)

1. Title
2. Creator – Investigator, Photographer, Author, Composer
3. Subject
4. Description
5. Publisher
6. Contributor
7. Date – Collecting Date, Analysing Date, Create date
8. Type
9. Format
10. Identifier – HKID, Patient ID
11. Source
12. Language
13. Relation
14. Coverage
15. Rights Management

### 3. Ethics and Legal Compliance

- How will you manage any ethical issues?
- How will you manage copyright and Intellectual Property Rights (IPR) issues?

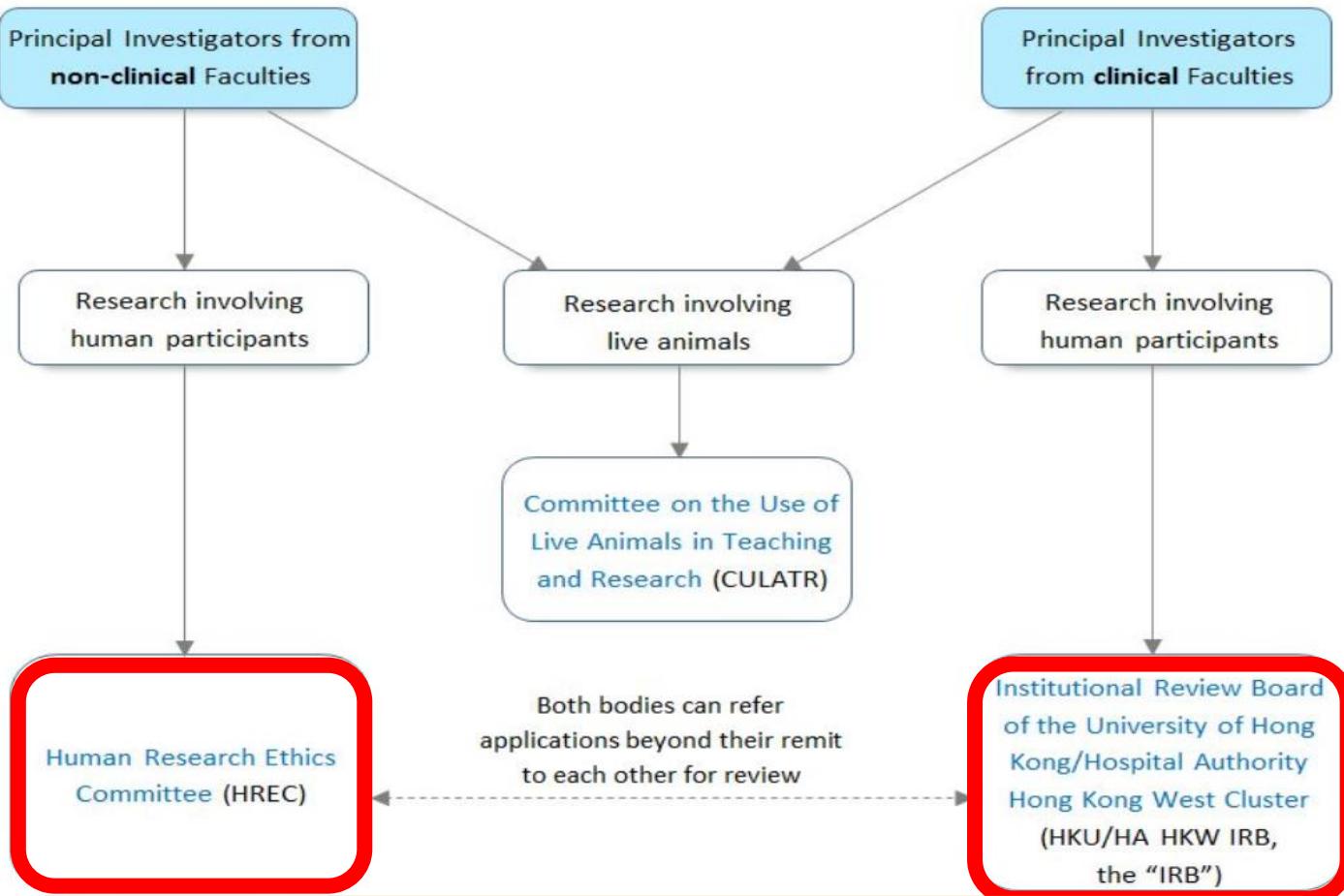
Have you gained consent for data preservation and sharing?

Who owns the data?

How will the data be licensed for reuse?

Any privacy or security issues? How are you dealing with them?

## Overview of Procedures



# 3. Ethics and Legal Compliance

The LSE Impact Blog header features the LSE logo and navigation links: Home, About, Latest, Our books, Series, Resources, LSE comment, Popular, and a search icon.

**The “long tail” of research impact is engendered by innovative dissemination tools and meaningful community engagement**

Share buttons: Facebook, Twitter, Email, LinkedIn, Print.

**Kip Jones and Lee-Ann Fenge** discuss what it takes to create meaningful community impact, highlighting a commitment to inclusive co-production and public engagement and the use of participatory research to create innovative dissemination tools.

**Research impact often tends not to happen in one emphatic, public moment but rather at more discrete points of the “long tail” of a research project. Achieving this depends largely on the tenacity of the research team but also on key allies such as the community members and service providers who have become energised by the work and inspired to continue it.**

**Creative Commons Attribution 3.0 Unported License**

This work is licensed under a **Creative Commons Attribution 3.0 Unported License** unless otherwise stated.

<https://blogs.lse.ac.uk/impactofsocialsciences/2018/02/20/the-long-tail-of-research-impact-is-engendered-by-innovative-dissemination-tools-and-meaningful-community-engagement/>

# 3. Ethics and Legal Compliance

## The Licenses



Attribution  
CC BY

This license lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered. Recommended for maximum dissemination and use of licensed materials.

[View License Deed](#) | [View Legal Code](#)



Attribution-ShareAlike  
CC BY-SA

This license lets others remix, tweak, and build upon your work even for commercial purposes, as long as they credit you and license their new creations under the identical terms. This license is often compared to "copyleft" free and open source software licenses. All new works based on yours will carry the same license, so any derivatives will also allow commercial use. This is the license used by Wikipedia, and is recommended for materials that would benefit from incorporating content from Wikipedia and similarly licensed projects.



Attribution-NoDerivs  
CC BY-ND

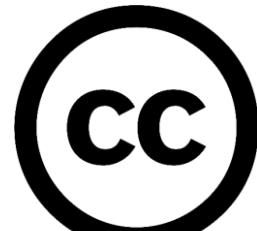
This license allows for redistribution, commercial and non-commercial, as long as it is passed along unchanged and in whole, with credit to you.

[View License Deed](#) | [View Legal Code](#)



Attribution-NonCommercial  
CC BY-NC

This license lets others remix, tweak, and build upon your work non-commercially, and although their new works must also acknowledge you and be non-commercial, they don't have to license their derivative works on the same terms.



<https://creativecommons.org/licenses/?lang=en>

### 3. Ethics and Legal Compliance

<https://creativecommons.org/choose/>

The screenshot shows the Creative Commons 'Choose a License' tool. At the top, there's a navigation bar with the Creative Commons logo, and links for 'Share your work', 'Use & remix', 'What we do', and 'Blog'. Below the navigation, a call-to-action banner says 'Help us build a vibrant, collaborative global commons' and features a 'Donate Now' button.

A red-bordered box contains a message about updated terms of service and privacy policy, dated November 7, 2017. It encourages users to review the changes before continuing.

Below this, there are links for new users: 'New to Creative Commons? [ Considerations before licensing ] [ How the licenses work ]', 'Explore the Creative Commons licenses.', '[ Want public domain instead? ]', and '[ Looking for earlier license versions, including ports? ]'.

The main panel is titled 'License Features' and states: 'Your choices on this panel will update the other panels on this page.' It asks: 'Allow adaptations of your work to be shared?' with three options: 'Yes' (selected), 'No', and 'Yes, as long as others share alike'. A question mark icon is also present.

## Data Confidentiality

Research records will be kept confidential, and access will be limited to the PI and primary research team members. For each testing session, the recorded data will have any identifying information removed and will be relabeled with study code numbers. A database which relates study code numbers to consent forms and identifying information will be stored separately on password-protected computers in a secured, locked office. These computers are housed in research facilities in the Psychology Building at Indiana University-Bloomington, and in the Psychology Department at UCSD. A list of the names of individuals who have participated in each study will be maintained in order to ensure that no individual is tested more than once on related studies. To maintain the privacy of the participants, any report of individual data will only consist of performance measures without any demographic or identifying information.

## 3. Ethics and Legal Compliance

- How will you manage any ethical issues?
  - How will you manage copyright and Intellectual Property Rights (IPR) issues?

# 4. Storage and Backup

- How will the data be stored and backed up during the research?
- How will you manage access and security?

How much storage?

How fast will the data grow?

Need to include costs for storage?

Who will be able to access or use your data?

How will you protect the confidentiality of your subjects?

# 4. Storage and Backup

Data stage	Storage location	Backup procedures (storage medium and location/ how often?)
Raw data		
Processed data		

# 4. Storage and Backup

**3-2-1**

**Rule**

3 copies

2 media

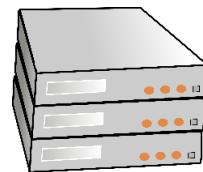
1 offsite



Personal  
computer



External  
hard  
drive



Networked  
drives



Cloud



CD/DVD



# 4. Storage and Backup

Storage solutions	Advantages	Disadvantages	Suitable for
<b>Personal computer &amp; Laptop</b>	Always available Portable	Drive may fail Laptop may be stolen	Temporary storage
<b>Networked drives File servers managed by your research group, university or facilities like a NAS-server</b>	Regularly backed up  Stored securely in a single place. Centralized storage makes it easier to maintain, backup.	Costs	Master copy of your data (if enough storage space is provided ...)
<b>External storage devices USB flash drive, DVD/CD, external hard drive</b>	Low cost  Portability	Easily damaged or lost	Temporary storage
<b>Cloud services Like Dropbox, SkyDrive, etc.</b>	Automatic synchronization between files online and folder on PC  Easy to use and access	Not sure whether data security is taken care of  You don't have direct influence on how often backups take place and by whom	Data sharing

## **2. Data Storage and Preservation**

Our short-term data storage plan, which will be used during the experiment, will be to save copies of 1) the .txt metadata file and 2) the Excel spreadsheet as .csv files to an external drive, and to take the external drive off site nightly. We will use the Subversion version control system to update our data and metadata files daily on the University of Alberta Mathematics Department server. We will also have the laboratory notebook as a hard copy backup that will be stored in a fire-proof cabinet.

The data set will be submitted to the Knowledge Network for Biocomplexity (KNB) data repository for long-term preservation and storage. The authors will submit metadata in EML format along with the data to facilitate its reuse. The data manager will be responsible for updating metadata and data author contact information in the KNB.

## **4. Storage and Backup**

- How will the data be stored and backed up during the research?
- How will you manage access and security?

# 5. Selection and Preservation

- Which data are of long-term value and should be retained, shared, and/or preserved?
- What is the long-term preservation plan for the dataset?

What data must be retained/destroyed for contractual, legal, or regulatory purposes?

Where or in which repository or data archive will the data be preserved (e.g. institution repository)?

# 5. Selection and Preservation

## Criteria for data appraisal

- Relevance to mission
  - Fit with goals and priorities of the institution
  - A legal requirement to keep the data
- Scientific or historical value
  - Data is scientifically, socially or culturally significant
- Uniqueness
  - The dataset is the only available copy
- Potential for redistribution
  - Reliable and usable for future use
- Nonreplicability
  - Costly to replicate
  - Derived from unrepeatable observations

Adapted from Whyte, Angus, and Andrew Wilson, 2010. "How to Appraise and Select Research Data for Curation"  
<http://www.dcc.ac.uk/resources/how-guides/appraise-select-data>

# 6. Data Sharing

- How will you share the data?
- Are any restrictions on data sharing required?

How will potential users find out about your data?

Access conditions, restrictions, embargoes etc.

Not a requirement that you share all of your data with anyone who asks

Indicate the criteria for deciding who can receive your data and whether or not you will place any conditions on their use

# 6. Data Sharing

## Not all data can be shared



Licensed  
data



Privacy,  
confidential,  
sensitive  
data



Data  
supporting  
a patent

# 6. Data Sharing



The University of Hong Kong

Research Services  
Support and information for HKU researchers

Sitemap Contact Us

Quick Links Search this site

Home > Research Integrity > Research Data and Records Management

Research Data and Records Management

## 4. Research data and records should be:

kept in a manner that is compliant with legal obligations and, where applicable, the requirements of funding bodies and project-specific protocols **approved by the Institutional Review Board (IRB) and Human Research Ethics Committee (HREC); and**

able to be made available to others in line **with appropriate ethical, data sharing and open access principles (3).**

# 6. Data Sharing

“Data should be made as widely and freely available as possible while safeguarding the privacy of participants, and protecting confidential and proprietary data.”

— *Final NIH Statement on Sharing Research Data February 26, 2003*

# 6. Data Sharing

## List of 18 HIPAA Identifiers

The Health Insurance Portability and Accountability Act (HIPAA) of 1996 specifies a number of elements in health data that are considered identifiers. If any are present, the health information cannot be released without patient authorization. Such data can be released for research purposes with approval of a waiver of patient authorization from an Institutional Review Board (IRB).

# 6. Data Sharing

## List of 18 HIPAA Identifiers

1. Names;
2. Address,
3. All elements of dates (except year) including birth date, admission date, discharge date, date of death;
4. Phone numbers;
5. Fax numbers;
6. Electronic mail addresses;
7. Social Security numbers;
8. Medical record numbers;
9. Health plan beneficiary numbers;
10. Account numbers;
11. Certificate/license numbers;
12. Vehicle identifiers and serial numbers, including license plate numbers;
13. Device identifiers and serial numbers;
14. Web Universal Resource Locators (URLs);
15. Internet Protocol (IP) address numbers;
16. Biometric identifiers, including finger and voice prints;
17. Full face photographic images; and
18. Any other unique identifying number, characteristic, or code (note this does not mean the unique code assigned by the investigator to code the data)

# 6. Data Sharing

## Indirect Identifiers

- Detailed geographic information (e.g. state, country, province)
- Organizations to which the respondent belongs
- Educational institutions (from which the respondent graduated and year of graduation)
- Detailed occupational titles
- Place where respondent grew up
- Exact dates of events (birth, death, marriage, divorce)
- Detailed income
- Offices or posts held by respondent

# 6. Data Sharing

## Options

- Institutional repository
- Interdisciplinary repository
- Domain/Subject specific repository
- Self-dissemination through website



MENDELEY DATA



figshare

zenodo

The  
**Dataverse**  
Project



R<sup>G</sup>

# Institutional repository

# 6. Data Sharing

The University of Hong Kong

**The HKU Scholars Hub 香港大學學術庫**

Home Publications Researchers Organizations Grants Datasets Theses Patents

 HELP HKU Login Guest Login

  
THE HKU SCHOLARS HUB  
AT THE CENTRE OF HKU

The HKU Scholars Hub is the current research information system of The University of Hong Kong. As a key vehicle of HKU's Knowledge Exchange, the Hub aims to increase the visibility of HKU authors and their research, and to foster opportunities for collaboration.

Quick Search META Research Collaborations Thesis Supervisors Media Commentators

Search for Everything...

**Featured Scholar**

  
**Dr Yiu, Siu Ming**  
Associate Professor  
Research Interests  
Cryptography  
Computer Security  
Bioinformatics  
[+ MORE](#)

**Hub News**

Dec 2018: HKU Theses On Amazon and Other Online Retailers (Update).  
Oct 2018: Report on the Sale of HKU Theses on Amazon and Other Websites.  
Oct 2018: Announcing the 2018 HKU's Top 1% Scientists.  
Sep 2018: HKU Theses On Amazon and Other Online Retailers (Update).

[More](#)

**Relevant Links**

- HKU's Top 1% Scientists
- Open Access @HKU
- Usage Stats & Downloads
- HKU research on Web of Science
- Research @HKU

connected to  
the researcher's  
institution

Source: <https://library.umassmed.edu/resources/necdmc/index>

# 6. Data Sharing

## Interdisciplinary repository

allows researchers from different disciplines to deposit and make their data available.



MENDELEY DATA



The  
**Dataverse**  
Project



figshare



# 6. Data Sharing

## Domain/Subject specific repository

The screenshot shows the ICPSR website homepage. At the top, there's a navigation bar with links: FIND DATA, START SHARING DATA, MEMBERSHIP, SUMMER PROGRAM, TEACHING & LEARNING, and DATA MANAGEMENT & CURATION. Below the navigation is a search bar with placeholder text "Search for studies, publications, variables, and webpages" and a "Search" button. A "Log In/Create Account" button is also visible. The main banner features the text "LOVE DATA WEEK" in large red letters, with "February 11-15" in smaller blue letters below it. To the left of the banner is the ICPSR logo with the tagline "Social science in the public interest". On the right side of the banner, there are two cartoon characters, one holding a heart and another holding a brain. Below the banner, there's a "News" section with several articles. One article is titled "Feb 11, 2019 New Releases through 2019-02-10" and includes a small icon of a document with a star. Another article is titled "Feb 11, 2019 New Data Release: Michigan Fish (ICPSR 37224)" and features an image of a fish. A third article is titled "Feb 6, 2019 Adopt a Dataset, deposit data, watch a webinar on education data, and more during Love Data Week (Feb. 11-15, 2019)" and includes an "Adopt a Dataset" logo with three cartoon characters.

discipline specific and often operated by a professional organization, a consortium of researchers, or some similar group.

[ICPSR](#) for Social Science

This is ICPSR



Source: <https://library.umassmed.edu/resources/necdmc/index>

# 6. Data Sharing

## Domain/Subject specific repository

The screenshot shows the NCBI homepage with a sidebar on the left containing links like 'NCBI Home', 'Resource List (A-Z)', and 'All Resources'. The main content area is titled 'How to: Submit data to NCBI' and includes sections for 'SEQUENCE DATA' and 'MICROARRAY DATA', each with a list of submission options.

**National Center for Biotechnology Information (NCBI)**  
for Health and Medical Sciences

Source: <https://library.umassmed.edu/resources/necdmc/index>

February 2019

# 6. Data Sharing

≡ **SPRINGER NATURE** Search Research Data at Springer  REGISTER SIGN IN

EDITOR RESEARCH DATA SUPPORT

SERVICES

## A new approach to supporting data sharing in breast cancer research

A partnership between Springer Nature and the Breast Cancer Research Foundation (BCRF) is enabling researchers to make their research data more accessible.



By Iain Hrynaszkiewicz on Feb 06, 2019

 1  0

<https://researchdata.springernature.com/users/11717-iain-hrynaszkiewicz/posts/43803-a-new-approach-to-supporting-data-sharing-in-breast-cancer-research>

# 6. Data Sharing

The screenshot shows the homepage of the *npj Breast Cancer* journal. At the top left is the navigation bar "nature > npj breast cancer". On the right, it says "a nature research journal". Below the header is a search bar with icons for "Search", "E-alert", "Submit", and "Login". The main content area features a large, detailed image of breast tissue under a microscope. Overlaid on this image is a white box containing the text "Supporting data sharing" and a quote from EiC Larry Norton about a research data pilot. Below this are three article cards:

- Article | 31 January 2019 | OPEN**  
pAKT pathway activation is associated with PIK3CA mutations and good
- Article | 18 January 2019 | OPEN**  
Assessment of HMGA2 and PLAG1 rearrangements in breast
- Article | 17 January 2019 | OPEN**  
MONARCH 3 final PFS: a randomized study of abemaciclib as initial

a research data pilot that the journal will provide to authors of accepted papers with enhanced editorial support to describe, share and link to the research data published.

<https://www.nature.com/npjbcancer/>

# 6. Data Sharing

## Before you share or publish your data

- Review the Depositor's Agreement, and Takedown Policy
- Perhaps you need to anonymize or redact your data before sharing?
- If you have created data which may have commercial value, please consult the **Technology Transfer Office**.

<http://www.tto.hku.hk/>

The screenshot shows the homepage of the Technology Transfer Office (TTO) at The University of Hong Kong. The header includes the university logo, navigation links for Home, About Us, Contact Us, and News & Events. Below the header, there are sections for Available Technologies (listing products, featured inventions, and intellectual property), Technology Transfer & Industry Partnerships (mentioning the HKU-CUPTI Partnership Program), and a featured invention for 'Automatic Acquisition of Lane-based Road Network Data for Driver Assistance'. The footer contains links for Researchers, Industry, and Featured Inventions.

#### **4. Data Dissemination and Policies for Data Sharing and Public Access**

We are required to share our data with the CAISN network after all data have been collected and metadata have been generated. This should be no more than 6 months after the experiments are completed. In order to gain access to CAISN data, interested parties must contact the CAISN data manager (data@caisn.ca) or the authors and explain their intended use. Data requests will be approved by the authors after review of the proposed use.

The authors will retain rights to the data until the resulting publication is produced, within two years of data production. After publication (or after two years, whichever is first), the authors will open data to public use. After publication, we will submit our data to the KNB enabling discovery and use by the wider scientific community. Interested parties will be able to download the data directly from KNB without contacting the authors, but will still be encouraged to give credit to the authors for the data used by citing a KNB accession number either in the publication's text or in the references list.

## **6. Data Sharing**

- How will you share the data?
- Are any restrictions on data sharing required?

# 7. Responsibilities and Resources

- Who will be responsible for data management?
- What resources will you require to deliver your plan?

Who is collecting the data? analysing the data?

Who is responsible for implementing the DMP, and ensuring it is reviewed and revised?

Who will be the contact person for questions regarding the research data?

Who can be contacted about the project after it has finished?

Project costs for data storage and costs for making the data accessible.

## **5. Roles and responsibilities**

The PI will be responsible for all data management during and after data collection.

[https://www.dataone.org/sites/all/documents/DMP\\_Copepod\\_Formatted.pdf](https://www.dataone.org/sites/all/documents/DMP_Copepod_Formatted.pdf)

## **5. Plans for Archiving and Preservation**

All original raw data files and data source processing programs will be versioned over time and maintained in a date-stamped file structure with text files documenting the provenance. The database will be preserved in perpetuity, housed initially at the New Mexico Interstate Stream Commission Central Office in addition to an off-site copy maintained at an NMISC field office and mirrored at the Consortium of Universities for the Advancement of Hydrologic Science (CUAHSI). We will also identify appropriate archiving institutions that might serve as a mirror repository. A data policy and stewardship plan will be established. In addition to archiving, each database table will be exported to a delimited text format to ensure accessibility of the data by other software programs. The data manager at the NMISC will be responsible for the management of long-term storage and archived data.

## **7. Responsibilities and Resources**

- Who will be responsible for data management?
- What resources will you require to deliver your plan?

[https://www.dataone.org/sites/all/documents/DMP\\_Hydrologic\\_Formatted.pdf](https://www.dataone.org/sites/all/documents/DMP_Hydrologic_Formatted.pdf)

#### **4. Policies for Re-use, Distribution**

Access to databases and associated software tools generated under the project will be available for educational, research and non-profit purposes. Such access will be provided using web-based applications, as appropriate.

Materials generated under the project will be disseminated in accordance with University/Participating institutional and NSF policies. Depending on such policies, materials may be transferred to others under the terms of a material transfer agreement.

Those that use the data (as opposed to any resulting manuscripts) should cite it as follows:

Lind, E, E Borer and A Kay. yyyy. Grassland Arthropod abundance and stoichiometry associated with nutrient manipulation. [URL]; accessed on ddmm/yyyy.

This information will be described in the metadata.

Intended and foreseeable users of the data are NutNet collaborators and participants, as well as other scientists interested in arthropod-plant relationships. This data set could be used in combination with similar data sets from other NutNet sites or for meta-analysis.

#### **5. Plans for Archiving and Preservation**

We will preserve both arthropod datasets generated during this project (abundance and stoichiometry) for the long term in the Digital Conservancy at the U of M. We will include the .csv files, along with the associated metadata files. We will also submit an abstract with the datasets that describe their original context and any potentially relevant project information. Borer will be responsible for preparing data for long-term preservation and for updating contact information for investigators.

# pollev.com/researchdata250

Poll Everywhere

Responding as abc

What components and issues have been addressed in Example\_DMP\_1?

You have not responded

Enter a response

Submit

## Welcome to DMPTOOL@HKU Libraries

Create data management plans that meet institutional and funder requirements.



### DMPTOOL by the Numbers

**598**

Users

**511**

Plans

**HKU**

Participating Institution

### Contact Us

For any questions, please send email to HKUL  
Research Data Services at [researchdata@hku.hk](mailto:researchdata@hku.hk)

# Supervisor's Endorsement

## RESPONSIBILITIES AND RESOURCES

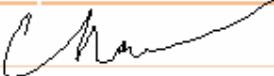
Who will be responsible for data management?

I will be responsible

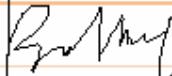
What resources will you require to deliver your plan?

Cost for acquiring external hard disks for off-site copy

Prepared by:

	Postgraduate student
Name:	
Date:	6-11-2017

Approved by:

	Supervisor
Name:	
Date:	6-11-2017

office lens

Cancel



Office Lens

Productivity

★★★★★ 208

OPEN

Easy

Neat

Powerful



DEVELOPER

Microsoft Corporation



Today



Games



Apps



Updates



Search

Office



Office Lens

**Data Management Plan**

Name	CHAN TAI MAN
UID	9999999992
Email	chantaiiman@hku.hk
Faculty	Social Science
Field of Study	
Degree	PhD
Project Title	Research Project example 1

Data will be stored in a CVS system and checked in and out for purposes of versioning. Variables will use a standardized naming convention consisting of a prefix, root, suffix system. Separate files will be managed for the two kinds of records produced: one file for respondents and another file for children with merging routines specified.

What data will you collect or create?

Quantitative survey data files generated will be processed and submitted to the [repository] as SPSS system files with DDI XML documentation. The data will be distributed in several widely used formats, including ASCII, tab-delimited (for use with Excel), SAS, SPSS, and Stata. Documentation will be provided as PDF. Data will be stored as ASCII along with setup files for the statistical software packages. Documentation will be preserved using XML and PDF/A.

Metadata will be tagged in XML using the Data Documentation Initiative (DDI) format. The codebook will contain information on study design, sampling methodology, fieldwork, variable-level detail, and all information necessary for a secondary analyst to use the data accurately and effectively.

For this project, informed consent statements will use language that will not prohibit the data from being shared with the research community.

The principal investigators on the project and their institutions will hold the intellectual property rights for the research data they generate but will grant redistribution rights to hku repository for purposes of data sharing.

1





**WHITEBOARD**   **DOCUMENT**   **BUSINESS CARD**

Will place a master copy of each digital file (i.e., research data files, documentation, and other related files) in Archival Storage, with several copies stored at designated locations and synchronized with the master through the Storage Resource Broker.

The data will be processed and managed in a secure networked environment. We have offline backup with portable hard drives that will be stored in secured, locked office. Those data can only accessed by my PI and the research team members.

By depositing data with HKU repository, our project will ensure that the research data are migrated to new formats, platforms, and storage media as required by good practice.

All raw and processed data.

No access restrictions.

The research data from this project will be deposited with hku repository to ensure that the research community has long-term access to the data.  
Question not answered.

During data analysis, the data will be accessible only by members of the project team. The research project will remove any direct identifiers in the data before deposit with hku repository.

All research data collected as part of this project is owned by the University. The Principal Investigator of this project will take responsibility for the collection, management, and sharing of the research data.

Prepared by:	Approved by:
Postgraduate student	Supervisor
	
Name: CHAN TAI MAN	Name: Dr WONG SUI MING
Date: 11-6-2018	Date: 11-6-2018

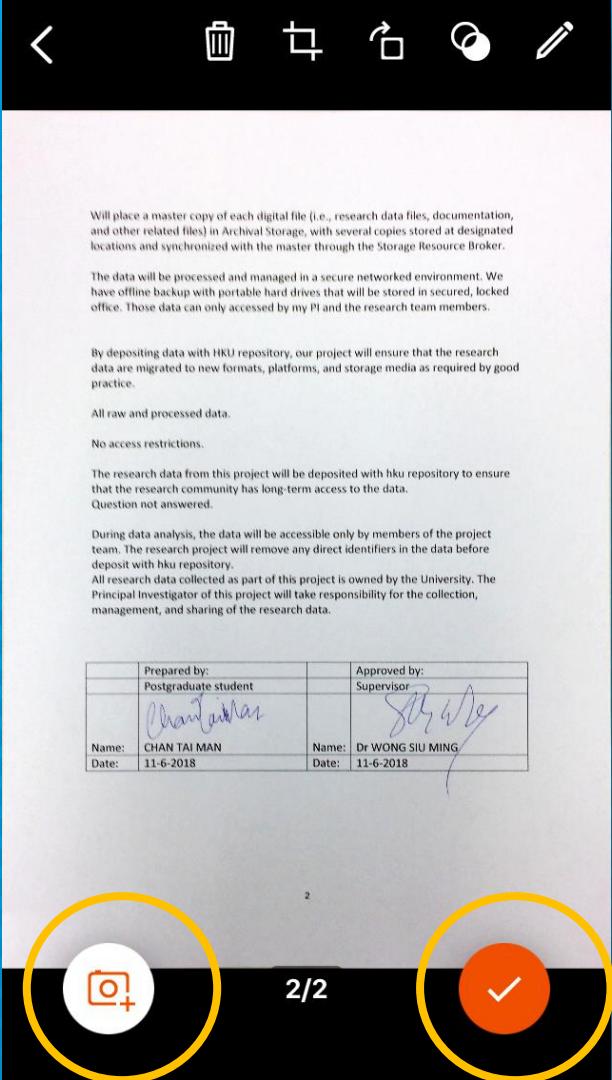
2







**WHITEBOARD**   **DOCUMENT**   **BUSINESS CARD**



SMC HK 4:28 PM 94%

Done 4-7-2018, 4:28 PM Off..

1 of 2

Data Management Plan

Name	CHAN TAI MAN
UID	9999999992
Email	chantaiman@hku.hk
Faculty	Social Science
Field of Study	
Degree	PhD
Project Title	Research Project example 1

Data will be stored in a CVS system and checked in and out for purposes of versioning. Variables will use a standardized naming convention consisting of a prefix, root, suffix system. Separate files will be managed for the two kinds of records produced: one file for respondents and another file for children for merging routines specified.  
What data will you collect or create?

Quantitative survey data files generated will be processed and submitted to the [repository] as SPSS system files with DDI XML documentation. The data will be distributed in several widely used formats, including ASCII, tab-delimited (for use with Excel), SAS, SPSS, and Stata. Documentation will be provided as PDF. Data will be stored as ASCII along with setup files for the statistical software packages. Documentation will be preserved using XML and PDF/A.

Metadata will be tagged in XML using the Data Documentation Initiative (DDI) format. The codebook will contain information on study design, sampling methodology, fieldwork, variable-level detail, and all information necessary for a secondary analyst to use the data accurately and effectively.

For this project, informed consent statements will use language that will not prohibit the data from being shared with the research community.

The principal investigators on the project and their institutions will hold the intellectual property rights for the research data they generate but will grant redistribution rights to hku repository for purposes of data sharing.

Will place a master copy of each digital file (i.e., research data files, documentation, and other related files) in Archival Storage, with several copies stored at designated locations and synchronized with the master through the Storage Resource Broker.

The data will be processed and managed in a secure networked environment. We have offline backup with portable hard drives that will be stored in secured, locked office. Those data can only accessed by my PI and the research team members.

Done 4-7-2018, 4:28 PM Off...



AirDrop. Tap to turn on Wi-Fi and Bluetooth to share with AirDrop.



Message



Mail



Add to Notes



WhatsApp



Copy



Print



Save to Files



More

Cancel

Cancel 4-7-2018, 4:28 PM... Send

To:

Cc/Bcc:

Subject: 4-7-2018, 4:28 PM Office  
Lens.pdf



4-7-2018,...fice Lens.pdf

Sent from my iPhone

**YOU HAVE SELECTED OPTION D:**

- D. Submit Data Management Plan (DMP). Dataset will be uploaded later.

You may proceed to submit your Data Management Plan (DMP) by drag and drop your file to the box below:

Drag and drop files here, or click in box to choose files.

After submitting your DMP, an email will go to your supervisors.

If you would like to update your DMP later, you may revisit this page to upload and replace your previous DMP with an updated version.

Please click "**SUBMIT DMP NOW**" button below to proceed, or click "**LOGOUT**" to exit form without submission.

**SUBMIT DMP NOW**

**LOGOUT**

# 5. Tools and other Resources

# Example DMP

THE LIBRARY

UC San Diego

Research & Collections

Borrow & Request

Computing & Technology

Visit

Ask Us

About

Hours

Q ▾

HOME / Research & Collections / Data Curation / UC San Diego Sample NSF Data Management Plans



## Sample NSF Data Management Plans

These examples from UC San Diego proposals are intended to provide a starting point for the development of other proposal-specific Data Management Plans.

We thank the UC San Diego investigators who gave permission to include their DMPs in this collection. If you have a DMP you'd be willing to have included here, please contact [Sharon Franks](#) or the library [Research Data Curation Program](#).

Please keep in mind that these examples are project-specific. PIs are encouraged to submit draft DMPs well in advance of the proposal deadline to OCGA to ensure compliance with University policy.

Data Curation

Data Management Best Practices

NIH Policy on Rigor and Reproducibility

Sample Data Management Plans

Obtain identifiers

Office of the Director (OD)

Office of Cyberinfrastructure (OD/OCI)

- [DMP Example Allan Snavely](#) From Allan Snavely's proposal to the Strategic Technologies for Cyberinfrastructure (STCI) program.

Example\_DMP\_1.doc

<https://libraries.ucsd.edu/research-and-collections/data-curation/dmp-samples.html>

# Example DMP

WAGENINGEN UNIVERSITY & RESEARCH  

About Wageningen Career Contact Login en|English ▾

Education & Programmes Research & Results Expertise & Services 

Home ▾ What is a Data Management Plan? ▾

## What is a Data Management Plan?

Research data exist in numerous forms, varying from logbooks with observations and protocols, to experimental and model data. A data management plan is a project document which tells the story of your research data. It outlines what research data were collected, how they were collected and what you will do with your data during and after your research.

To guide you in completing your DMP, Wageningen University provides you with a [template](#) with 10 questions about expected data formats, ownership, documentation, archiving and reuse. A [presentation](#) illustrates the template's questions in more detail. You can also have a look at some completed Plans.

› Example DMP 1  
› Example DMP 2

  
Contact  
Servicedesk Library  
[Contact form](#)



Example\_DMP\_2.pdf, Example\_DMP\_3.pdf

<https://www.wur.nl/en/show/What-is-a-Data-Management-Plan.htm>

# Example DMP



## Examples of DMP questions and answers

### Expert Tour Guide on Data Management

#### 1. Plan

Benefits of data management

Research data

Data in social sciences

FAIR data

European diversity

#### Adapt your DMP: Part 1

Sources and further reading

#### 2. Organise & Document

#### 3. Process

#### 4. Store

#### 5. Protect

#### 6. Archive & Publish

For inspiration of filled in DMPs look at some example DMPs we prepared. Both DMPs are based on a fictional research project with a basis in reality. For each topic of the DMP, there are example questions and answers where applicable. The examples are not country specific. Some of the information is generic.

#### Qualitative data



During this project, in-depth interviews with teachers in primary school will be held. The project has just started.

Click the link to view and download the DMP:

[DMPQuestionsQualitativeData.pdf](#) (165 KB)

#### Quantitative data



The project concerns a survey which is conducted in order to identify how the evolution of society affects attitudes and behaviour. The project is still running.

Click the link to view and download the DMP:

[DMPQuestionsQuantitativeData.pdf](#) (205 KB)

Example\_DMP\_4.pdf, Example\_DMP\_5.pdf

<https://www.cessda.eu/Research-Infrastructure/Training/Expert-Tour-Guide-on-Data-Management/1.-Plan/Adapt-your-DMP-Part-1>

# Example DMP



About Us Network Strategy Activities Join Login

Scholarly Communication

Digital Skills & Services

Research Infrastructure

Events

Menu

## Data Management Plan Catalogue

Home

About This DMP Catalogue

Current & Future Plans

The Review Process  
Explained

The Metadata Catalogue  
Explained

Project & Contact Details

Welcome to the Data Management Plan (DMP) Catalogue of the LIBER Research Data Management Working Group. This "DMP CAT" is an innovative approach to provide:

- A central hub for DMPs from different disciplines
- Quality reviews of the DMPs with assessments of the quality of the different parts

The purpose of this catalogue is to inspire researchers and others in the process of writing a Data Management Plan.

This DMP CAT provides the metadata and reviews in tabular form on this website and links to the citable and DOI-equipped PDF version of the DMPs on [Zenodo](#). This means that the DMP are officially published via Zenodo: authors and contributors are properly listed and linked to their ORCID-ID, the DMP received as persistent link and is citable.

All

Economics

Energy Management

Environmental Protection

Future Internet Informatics

Particle Image Velocimetry Psychology

Deploying Uncertainty Data Management Plan

[Read more](#)

IPER-MAN Data Management Plan

[Read more](#)

BEHAVE Data Management Plan

[Read more](#)

TWEETHER Data Management Plan

[Read more](#)

FUTURE TDM Data Management Plan

[Read more](#)

RAWFIE Data Management Plan

[Read more](#)

<https://libereurope.eu/dmpcatalogue/>

# Tools



because good research needs good data

[Home](#) [Digital curation](#) [About us](#) [News](#) [Events](#) [Resources](#) [Training](#) [Projects](#) [Community](#) [Tailored support](#)

[Home](#) > [Resources](#) > [External](#) > [Tools Services](#) > Managing Active Research Data

## Managing Active Research Data



Resources to help researchers manage data from its point of creation, facilitating its productive use in the present and future, as well as establishing the support structures necessary to ensure its future survival.

### [Active Data Storage \(4\)](#)

Flexible storage for evolving data.

### [Data Management Planning \(2\)](#)

Help in meeting funder requirements for data management plans.

### [Metadata Annotation \(1\)](#)

### [Persistent ID Assignment \(4\)](#)

Creating unique identifiers for digital objects.

### [Workflow and Lab Notebook Management \(8\)](#)

Digitising the traditional lab notebook and automating scientific workflows.

# Tools



Site DataONE Search

Search phrase

[About](#)   [News](#)   [Participate](#)   [Resources](#)   [Education](#)   [Data](#)[Home](#) » [Resources](#) » [Software Tools Catalog](#) » [Software Tools](#)

## Resources

---

### Tools

[Investigator Toolkit](#)  
[Data Management Planning](#)  
[Software Tools Catalog](#)

### Materials

[Publications](#)  
[Best Practices](#)  
[Data Life Cycle](#)  
[Librarian Outreach Kit](#)  
[Developer Resources](#)  
[Research Notebooks](#)

### Featured Tool

---

[Oracle](#)

## Software Tools

---

### CyberTracker



CyberTracker

Tags: [data entry](#), [metadata](#)

### dash

UNIVERSITY  
OF  
CALIFORNIATags: [collaboration](#), [data storage](#), [preservation](#)<https://www.dataone.org/all-software-tools?page=1>

September 2018

# Dataset Search

## Google Dataset Search Beta

Search for Datasets



Try [boston education data](#) or [weather site:noaa.gov](#)



ICPSR  
@ICPSR

ICYMI, Google just released its new Dataset Search feature, with a nod to ICPSR ([bit.ly/2oJVZm2](https://bit.ly/2oJVZm2))! See our news release for more details: [@umich @umisr @DDIAlliance @Google #GoTeam #Data #Research](https://bit.ly/2wONeeQ)

Following



NEWS • 05 SEPTEMBER 2018

## Google unveils search engine for open data

The tool, called Google Dataset Search, should help researchers to find the data they need more easily.

Google Dataset Search Beta

Supreme Court

- ACT Supreme Court - Judgments
- Spend over £25,000 in UK Supreme Court
- US Supreme Court Cases, 1794-2016
- Ideology Scores of Supreme Court Justices
- Supreme Court Case Book (also known as [Supreme Court Cause Book](#))
- Supreme Court Case Register
- ACT Supreme Court - Sentences
- United States Supreme Court Justices Biographical Data, 1789-1958
- Register of warrants on Supreme Court writs [Sheriff's Officers, Bega.]
- Supreme Court of the Australian Capital Territory - Court of Appeal 2002-

# Find a Data Repository

nature > scientific data > policies > recommended data repositories

a natureresearch journal

SCIENTIFIC DATA

Search E-alert Submit Login

Policies

Editorial & Publishing Policies

For Referees

Data Policies

Recommended Data Repositories

## Recommended Data Repositories

Scientific Data mandates the release of datasets accompanying our Data Descriptors, but we do not ourselves host data. Instead, we ask authors to submit datasets to an appropriate public data repository. Data should be submitted to discipline-specific, community-recognized repositories where possible, or to generalist repositories if no suitable community resource is available.

Repositories included on this page have been evaluated to ensure that they meet our requirements for data access, preservation and stability. Please be aware, however, that some repositories on this page may only accept data from those funded by specific sources, or may charge for hosting data. Please ensure you are aware of any deposition policies for your chosen repository. If your repository of choice is not listed please see our [guidelines for suggesting additional repositories](#).

<https://www.nature.com/sdata/policies/repositories>

# Find a Data Repository

Jisc Digital Resources > Open Access

<http://v2.sherpa.ac.uk/opendoar/>

## OpenDOAR

Browse Search Statistics Policy Tool Our APIs Suggest Admin

### Directory of Open Access Repositories

OpenDOAR is a global directory of Open Access repositories and their policies.

---

Search for a repository

---

OpenDOAR is the quality-assured global directory of academic open access repositories. It enables the identification, browsing and search for repositories, based on a range of features, such as location, software or type of material held. [Read more...](#)

---

**SERVICES** Open access services from Jisc

**SERVICE** SHERPA Services

**GUIDE** Managing open access costs

# Find a Data Repository

<https://repositoryfinder.datacite.org/>

## Repository Finder

Find a repository to upload your data.

Repository Finder, a pilot project of the [Enabling FAIR Data Project](#) led by the American Geophysical Union (AGU) in partnership with DataCite and the Earth, space and environment sciences community, can help you find an appropriate repository to deposit your research data. The tool is hosted by DataCite and queries the re3data registry of research data repositories.

Search [re3data](#) for a repository to upload your data

Type to search...

Search

or

See the repositories in re3data that meet the criteria of the Enabling FAIR Data Project.

# Resources

Karl W. Broman & Kara H. Woo (2018) Data Organization in Spreadsheets,  
The American Statistician, 72:1, 2-10, DOI: 10.1080/00031305.2017.1375989

The screenshot shows the Taylor & Francis Online journal page for *The American Statistician*. The top navigation bar includes links for Log in, Register, and Cart. The journal logo is on the left, and search fields for keywords, authors, DOI, etc., and This Journal are on the right. Below the header, the journal title and volume information are displayed, along with buttons for Submit an article and Journal homepage. The main content area shows an article titled "Data Organization in Spreadsheets" by Karl W. Broman & Kara H. Woo. It includes metrics like 47276 views, 1 CrossRef citation, and 912 Altmetric. The abstract is available, and there are links for Download citation, DOI, Check for updates, Full Article, Figures & data, References, Citations, Metrics, Licensing, and PDF. A "People also read" sidebar is visible on the right.

47276 Views  
1 CrossRef citations  
912 Altmetric

Article

**Data Organization in Spreadsheets**

Karl W. Broman & Kara H. Woo

Pages 2-10 | Received 01 Jun 2017, Accepted author version posted online: 29 Sep 2017, Published online: 29 Sep 2017

Download citation <https://doi.org/10.1080/00031305.2017.1375989> Check for updates

Full Article Figures & data References Citations Metrics Licensing PDF

**ABSTRACT**

Select Language ▾  
Translator disclaimer

Spreadsheets are widely used software tools for data entry, storage, analysis, and visualization. Focusing on the data entry and storage aspects, this article offers practical recommendations for organizing spreadsheet data to reduce errors and ease later analyses. The

People also read

Discussion  
50 Years of Data

<https://www.tandfonline.com/doi/abs/10.1080/00031305.2017.1375989>

# Resources

## Metadata

RDA | Metadata Directory

## [View the standards](#)

## View the extensions

## [View the tools](#)

[View the use cases](#)

## Browse by subject areas

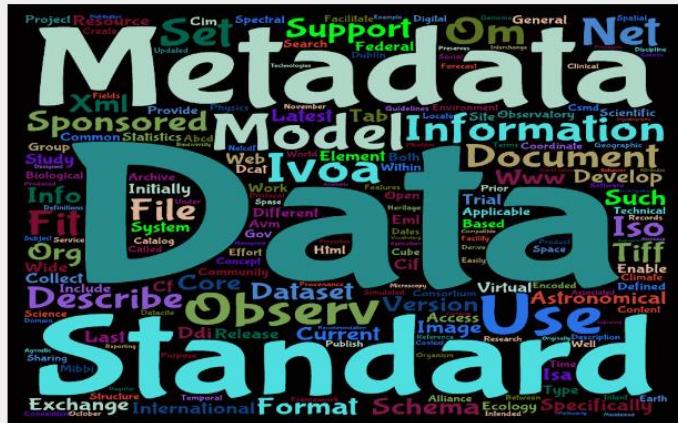
## Contribute

Add standards

## Add extensions

### Add tools

Add use cases



## Metadata Standards Directory Working Group

The RDA Metadata Standards Directory Working Group is supported by individuals and organizations involved in the development, implementation, and use of metadata for scientific data. The overriding goal is to develop a collaborative, open directory of metadata standards applicable to scientific data can help address infrastructure challenges.

The RDA Metadata Standards Directory is maintained by [Sean Chen](#), [Kate Anne Alderete](#), and [Alex Ball](#).

The theme is maintained by Dustin Allen.

This page was generated by GitHub Pages

# Resources



because good research needs good data

Search

Home | Digital curation | About us | News | Events | Resources | Training | Projects | Community | Tailored support

Home > Resources > How Guides > Develop Data Plan

## In this section

Briefing Papers

How-to Guides & Checklists

Appraise & Select Research Data for Curation

Cite Datasets and Link to Publications

Develop RDM Services

**Develop a DMP**

Discover Requirements

Five Steps to Decide What Data to Keep

Five Things You Need to Know About RDM and the Law

License Research Data

Track Data Impact with Metrics

Using RISE

Where to keep research data

## How to Develop a Data Management and Sharing Plan

This guide outlines the process of developing a data management and sharing plan. Planning for the effective creation, management and sharing of your data enables you to get the most out of your research. The guide outlines UK funder expectations for data management and sharing plans and provides practical guidance on how to meet these. It should be of relevance to researchers and those supporting them within higher education.

By Sarah Jones, Digital Curation Centre

Published: 8 September, 2011

Browse the guide below or [download the pdf](#).

**\*\* This publication is available in print and can be ordered from our online store  \*\***

Please cite as: Jones, S. (2011). 'How to Develop a Data Management and Sharing Plan'. DCC How-to Guides. Edinburgh: Digital Curation Centre. Available online: <http://www.dcc.ac.uk/resources/how-guides>



The Collaborative Assessment of Research Data Infrastructure and Objectives (CARDIO) toolkit helps HEIs assess and plan to improve their data management activity, infrastructure and support.

[Read more](#)

<http://www.dcc.ac.uk/resources/how-guides/develop-data-plan>

# Resources

FIND DATA ▾ START SHARING DATA ▾ MEMBERSHIP ▾ SUMMER PROGRAM ▾ TEACHING & LEARNING ▾ DATA MANAGEMENT & CURATION ▾ 

## ICPSR

Log In/Create Account

### Data Management & Curation

 QUALITY PRESERVATION ACCESS CONFIDENTIALITY CITATION TOOLS & SERVICES

## Guidelines for Effective Data Management Plans

Many federal funding agencies, including NIH and most recently NSF, are requiring that grant applications contain data management plans for projects involving data collection. To support researchers in meeting this requirement, ICPSR is providing guidance on creating such plans.

These guidelines can also be [downloaded as a single PDF](#).

### Creating a Data Management Plan

- [Elements of a plan](#)
- [Framework for creating a plan](#)

### Depositing Data with ICPSR

- [Sample plan](#)
- [How to deposit](#)

# Resources

cessda eric

About ▾ Consortium ▾ Projects ▾ Research Infrastructure ▾ Contact



[Home](#) / [Research Infrastructure](#) / [Training](#) / Expert Tour Guide on Data Management

## Expert Tour Guide on Data Management



### About this expert tour guide

This tour guide by CESSDA ERIC (the Consortium of European Social Science Data Archives European Infrastructure Consortium) aims to put social scientists like yourself at the heart of making their research data findable, understandable, sustainably accessible and reusable.

You will be guided by European experts who are - on a daily basis - busy ensuring long-term access to valuable social science datasets, available for discovery and reuse at one of the [17 CESSDA social science data archives](#). With this guide and the training events being held across Europe, we want to accompany and inspire you in your journey through the research data life cycle.

<https://www.cessda.eu/Research-Infrastructure/Training/Expert-Tour-Guide-on-Data-Management>

# Resources

**F1000Research**  
Open for Science

Search 

**SUBMIT YOUR RESEARCH** 

BROWSE GATEWAYS HOW TO PUBLISH ABOUT BLOG MY RESEARCH SIGN IN

## How to Publish

- Submit your Research
- Submissions
- Article Guidelines
- Posters and Slides Guidelines
- Data Guidelines**
- Article Processing Charges
- Finding Article Referees

**Data Guidelines**

1. Background
  - 1.1 Open Data Policy
  - 1.2 Fair Data Principles
2. Share Your Data in 3 Steps
  - 2.1 Prepare Your Data for Sharing
  - 2.2 Select a Repository
  - 2.3 Add a Data Availability Statement to Your Manuscript

**1. Background**

This page provides information about data you need to include when publishing an article in *F1000Research*, where your data can be stored, and how your data should be presented. In accordance with our [data policies](#), authors will be required to submit their data or provide details of where their data is hosted upon submission (excepting ethical, data protection or confidentiality considerations). Please note that adherence to our data policies is only compulsory for articles, not posters or slides.

A large number of [journals and publishers](#) have confirmed that they welcome research articles reporting analysis and conclusions that are based on previously published datasets: They do not consider the publication of a dataset with a DOI and associated protocol information as a 'prior publication' that would preclude subsequent publication of new results obtained from such a dataset.

---

**About F1000Research**

---

How it Works

---

For Referees

---

Our Advisors

---

Policies

---

Glossary

---

FAQs

---

Contact

<https://f1000research.com/for-authors/data-guidelines>

# Resources

## FAIR Principles

[Home](#) > FAIR Principles

### FAIR Principles

- › F1: (Meta) data are assigned globally unique and persistent identifiers
- › F2: Data are described with rich metadata
- › F3: Metadata clearly and explicitly include the identifier of the data they describe
- › F4: (Meta) data are registered or indexed in a searchable resource
- › A1: (Meta) data are retrievable by their identifier using a standardised communication protocol

In 2016, the '**FAIR Guiding Principles for scientific data management and stewardship**' were published in *Scientific Data*. The authors intended to provide guidelines to improve the findability, accessibility, interoperability, and reuse of digital assets. The principles emphasise machine-actionability (i.e., the capacity of computational systems to find, access, interoperate, and reuse data with none or minimal human intervention) because humans increasingly rely on computational support to deal with data as a result of the increase in volume, complexity, and creation speed of data.

#### Findable

The first step in (re)using data is to find them. Metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services, so this is an essential component of the **FAIRification process**.

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

**F2. Data are described with rich metadata**

**F3. Metadata clearly and explicitly include the identifier of the data they describe**

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

<https://www.go-fair.org/fair-principles/>

# We Welcome your feedback



HKU Libraries Application Manager

- HKU Libraries Applications
  - Internet Explorer
  - Windows Explorer
  - Notepad
  - Calculator
  - Set PIN
  - Library Course Survey
  - FIT Survey
  - EndNote Support workshop
- Bibliographic
  - EndNote X
- Microsoft Office XP
  - Microsoft Word XP - E-Lab
- Network Printer Connection
  - Connect Network Printer
  - Disconnect Network Printer
  - Check value card balance
- Scanning / Utilities
- Staff Training
  - Millennium
  - Millennium CityU
  - Millennium Lingnan
  - Telnet
  - Remote Desktop Connection Client
- Staff Only

Please double click to launch applications.

Internet Explorer  
Windows Explorer  
Notepad  
Calculator  
Set PIN  
Library Course Survey  
FIT Survey  
EndNote Support workshop  
Bibliographic  
EndNote X  
Microsoft Office XP  
Microsoft Word XP - E-Lab  
Network Printer Connection  
Connect Network Printer  
Disconnect Network Printer  
Check value card balance  
Scanning / Utilities  
Staff Training  
Millennium  
Millennium CityU  
Millennium Lingnan  
Telnet  
Remote Desktop Connection Client  
Staff Only

Library Course Survey

# THANKS!

Any questions?

You can find me at

**researchdata@hku.hk**

