

RPG PROCEDURES FOR RESEARCH DATA MANAGEMENT (RDM)

HKUL RESEARCH DATA SERVICES

<http://lib.hku.hk>

The screenshot shows the homepage of the HKUL Libraries website. At the top left is the University of Hong Kong logo and the text "The University of Hong Kong Libraries". A large banner image of a modern library interior with study carrels and tables is displayed, overlaid with a blue diagonal band containing the text "Introducing the Learning Hub". Below the banner are two search boxes: "Find@HKUL" and "Site Search". A "Search" bar with a magnifying glass icon is positioned below these. On the right side of the header, there are links for "High Contrast" and "A A A". Below the header, there are three main sections: "User Information" (with links for Student, Staff, Graduate, and Public), "Research Tools" (with links for FIND@HKUL, Electronic Resources, Dragon, HKUL Catalogue, HKU Scholars Hub, and HKUL Digital Initiatives, where "HKU Scholars Hub" is circled in red), and "Research Support" (with links for BrowZine, Subject Guides, Endnote, Training, Research Guides, and Turnitin). At the bottom right is a "ASK A LIBRARIAN" button with a question mark icon.

High Contrast A A A

Hours Contact Us e-Forms Services

User Information

Student Staff Graduate Public

MyAccount@HKUL

Research Tools

FIND@HKUL

Electronic Resources

Dragon, HKUL Catalogue

HKU Scholars Hub

HKUL Digital Initiatives

Research Support

BrowZine

Subject Guides

Endnote

Training

Research Guides

Turnitin

MAIN LIBRARY

Find@HKUL Site Search

Search

About HKUL

Hours Contact Us

Map Giving to HKUL

Policy Publications

ASK A LIBRARIAN

HKUL RESEARCH DATA SERVICES

<http://hub.hku.hk>

The University of Hong Kong
The HKU Scholars Hub 香港大學學術庫

Home Publications Researchers Organizations Grants Datasets Theses Patents Community Service

THE HKU SCHOLARS HUB AT THE CENTRE OF HKU

Deposit Data
HKUL Research Data Management

Information system of The University of Hong Kong. As a key vehicle of HKU's Knowledge Exchange Initiative, The Hub aims to enhance the visibility of HKU authors and their research, and to foster opportunities for collaboration.

Quick Search BETA Research Collaborations Thesis Supervisors Media Commentators

Search for Everything... **Search**

Featured Scholar

Professor Li, Yuguo

- Professor
- Associate Dean, Faculty of Engineering

Research Interests

- Infection control and airborne transmission
- Ventilation
- Building energy efficiency
- Computational fluid

Hub News

- Apr 2017: HKU Theses On Amazon and Other Online Retailers.
- Jul 2016: The Hub adds **Datasets**. New deposit page coming soon.
- Feb 2016: Updated to DSpace 5.2. Added **global search** capability under Quick Search tab.
- Jan 2016: The Hub is No. 1 in Asia and 41st worldwide, according to **Webometrics**.

More

Relevant Links

- HKU's Top 1%
- Open Access @HKU
- Usage Stats & Downloads
- HKU most cited articles in Scopus
- HKU research on Web of Science
- Research @HKU

HKUL RESEARCH DATA SERVICES

<https://lib.hku.hk/researchdata/rds.htm>

HKUL RESEARCH DATA SERVICES

RPG STUDENTS | RESEARCH STAFF | RDM | METADATA | DEPOSIT DATA | RESOURCES | HELP

WHAT IS RESEARCH DATA MANAGEMENT?

RDM FOR RESEARCH POSTGRADUATE (RPG) STUDENTS

RDM FOR RPG SUPERVISORS

CONTINUE READING

THE UNIVERSITY OF HONG KONG

LIBRARIES

香港大學 THE UNIVERSITY OF HONG KONG

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

[Quick Links](#)[Search this site](#)

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

Research Data and Records Management

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 integrity.

HKU recognises the importance of good practice in research data and records management and seeks to promote the highest standards. The University's *Policy on the Management of Research Data and Records* was approved by the Senate at its meeting on May 5, 2015, along with the establishment of a **Task Force on Management of Research Data and Records** to oversee the planning of the implementation of the Policy.

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

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

THE HONG KONG RESEARCH GRANTS COUNCIL (RGC)

General Research Fund (GRF) Explanatory Notes for 2016/17 applications

"PI should assess data archive potential and opportunities for data sharing. **Due additional weight** will be given to an application where the applicants are willing to **make research data available to others.**"

<http://www.ugc.edu.hk/doc/eng/rgc/form/GRF2.pdf>



authors & referees

Search

[authors & referees](#) > [Policies](#) > Availability of data, material and methods

Site content

[Homepage](#)

Policies

[Publication ethics](#)

[Bioethics & Biosecurity](#)

[Availability of data & materials](#)

[Peer-review policy](#)

[Embargo](#)

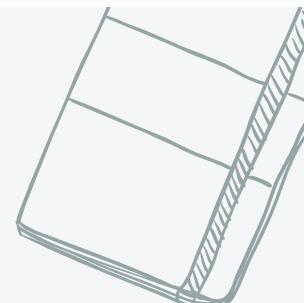
[Corrections](#)

Availability of data, material and methods

An inherent principle of publication is that others should be able to replicate and build upon the authors' published claims. A condition of publication in a Nature journal is that **authors are required to make materials, data, code, and associated protocols promptly available to readers without undue qualifications**. Any restrictions on the availability of materials or information must be disclosed to the editors at the time of submission. Any restrictions must also be disclosed in the submitted manuscript.

After publication, readers who encounter refusal by the authors to comply with these policies should contact the chief editor of the journal. In cases where editors are unable to resolve a complaint, the journal may refer the matter to the authors' funding institution and/or publish a formal statement of correction, attached online to the publication, stating that readers have been unable to obtain necessary materials to replicate the findings.

<http://www.nature.com/authors/policies/availability.html>





PUBLICATION POLICIES

Data and Materials Availability after Publication

After publication, all data and materials necessary to understand, assess, and extend the conclusions of the manuscript must be available to any reader of *Science*. All computer codes involved in the creation or analysis of data must also be available to any reader of *Science*. After publication, all reasonable requests for data or materials must be fulfilled. Any restrictions on the availability of data, codes, or materials, including fees and restrictions on original data obtained from other sources must be disclosed to the editors as must any Material Transfer Agreements (MTAs) pertaining to data or materials used or produced in this research, that place constraints on providing these data or materials. Patents (whether applications or awards to the authors or home institutions) related to the work should also be declared. Fossils or other rare specimens must be deposited in a public museum or repository and available for research. Unreasonable restrictions on data or material availability may preclude publication.



<http://www.sciencemag.org/authors/science-editorial-policies#unpublished-data-and-personal-communications>





advanced search

Data Availability

The following policy applies to all PLOS journals, unless otherwise noted.

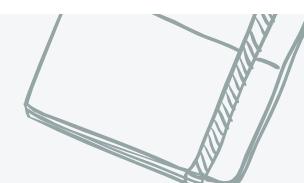
PLOS journals require authors to make all data underlying the findings described in their manuscript fully available without restriction, with rare exception.

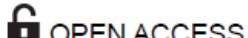
When submitting a manuscript online, authors must provide a *Data Availability Statement* describing compliance with PLOS's policy. If the article is accepted for publication, the data availability statement will be published as part of the final article.

Refusal to share data and related metadata and methods in accordance with this policy will be grounds for rejection. PLOS journal editors encourage researchers to contact them if they encounter difficulties in obtaining data from articles published in PLOS journals. If restrictions on access to data come to light after publication, we reserve the right to post a correction, to contact the authors' institutions and funders, or in extreme cases to retract the publication.

Methods acceptable to PLOS journals with respect to data sharing are listed below, accompanied by guidance for authors as to what must be indicated in their data availability statement and how to follow best practices in reporting. If authors did not collect data themselves but used another source, this source must be credited as appropriate. Authors who have questions or difficulties with the policy, or readers who have difficulty accessing data, are encouraged to contact the journal office (plosone@plos.org). If you have broader questions about the PLOS data availability policy, contact data@plos.org.

<http://journals.plos.org/plosone/s/data-availability>





Citation: Hodgens C, Nimchuk ZL, Kieber JJ (2017) indCAPS: A tool for designing screening primers for CRISPR/Cas9 mutagenesis events. PLoS ONE 12(11): e0188406. <https://doi.org/10.1371/journal.pone.0188406>

Editor: Ruslan Kalendar, University of Helsinki, FINLAND

Received: October 3, 2017

Accepted: November 6, 2017

Published: November 15, 2017

Copyright: © 2017 Hodgens et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by a grant from the National Science Foundation (IOS-1238051 and IOS-1455607) <https://www.nsf.gov/div/index.jsp?div=IOS>. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

a cleaved amplified polymorphic sequence (CAPS) technique. If a restriction site is created or altered by the mutation such that only one allele contains the restriction site, a polymerase chain reaction (PCR) followed by a restriction digest can be used to distinguish the two alleles. However, in the case of most CRISPR-induced alleles, no such restriction sites are present in the target sequences. In this case, a derived CAPS (dCAPS) approach can be used in which mismatches are purposefully introduced in the oligonucleotide primers to create a restriction site in one, but not both, of the amplified templates. Web-based tools exist to aid dCAPS primer design, but when supplied sequences that include indels, the current tools often fail to suggest appropriate primers. Here, we report the development of a Python-based, species-agnostic web tool, called indCAPS, suitable for the design of PCR primers used in dCAPS assays that is compatible with indels. This tool should have wide utility for screening editing events following CRISPR/Cas9 mutagenesis as well as for identifying specific editing events in a pool of CRISPR-mediated mutagenesis events. This tool was

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Introduction

It is often necessary to genotype biological samples to select individuals from a large population with a desired genetic variant. Genetic variants generated by mutagenesis or natural variation can take the form of single nucleotide polymorphisms (SNPs) or insertions/deletions (indels). Suf-



<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0188406#>



Jordan, E., & G. R. Buckingham (2017) The effect of observing novice and expert performance on acquisition of surgical skills on a robotic platform. PLoS ONE 12(11): e0188233. <https://doi.org/10.1371/journal.pone.0188233>

Editor: Manabu Sakakibara, Tokai University, JAPAN

Received: September 5, 2017

Accepted: November 2, 2017

Published: November 15, 2017

Copyright: © 2017 Harris et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The full dataset is available for download at: https://osf.io/5z89v/?view_only=145ae36726f146b5bbb5ea6762a0d4f6

Funding: This research was supported by an Intuitive Surgical grant awarded to Dr G. Buckingham. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

enefit of observing not only expert performance but also error-strewn performance. The aim of this study was to determine which model (novice vs. expert) would lead to the greatest benefits when learning robotically assisted surgical skills.

Methods

120 medical students with no prior experience of robotically-assisted surgery completed a ring-carrying training task on three occasions; baseline, post-intervention and at one-week follow-up. The observation intervention consisted of a video model performing the ring-carrying task, with participants randomly assigned to view an expert model, a novice model, a mixed expert/novice model or no observation (control group). Participants were assessed for task performance and surgical instrument control.

Data Availability Statement: The full dataset is available for download at: https://osf.io/5z89v/?view_only=145ae36726f146b5bbb5ea6762a0d4f6.

Conclusions

Similar benefits were found when observing the traditional expert model or the error-strewn model, suggesting that viewing poor performance may be as beneficial as viewing expertise in the early acquisition of robotic surgical skills. Further work is required to understand, then

<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0114734>

Citation: Pinfield S, Cox AM, Smith J (2014) Research Data Management and Libraries: Relationships, Activities, Drivers and Influences. PLoS ONE 9(12): e114734. doi:10.1371/journal.pone.0114734

Editor: Pascal Launois, World Health Organization, Switzerland

Received: May 28, 2014

Accepted: November 13, 2014

Published: December 8, 2014

Copyright: © 2014 Pinfield et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that, for approved reasons, some access restrictions apply to the data underlying the findings. Data underlying this study cannot be made publicly available in order to safeguard participant anonymity and that of their organisations. Ethical approval for the project was granted on the basis that only extracts of interviews would be shared (with appropriate anonymisation) as part of publications and other research outputs. In order to share data with other researchers, the participants must be contacted and consent to this data release. Other researchers will be able to request the dataset by contacting the corresponding author or Chair of the University of Sheffield Information School Research Ethics Committee (ischool_ethics@sheffield.ac.uk).

Funding: The authors have no support or funding to report.

Competing Interests: The authors have declared that no competing interests exist.

Data Availability: The authors confirm that, for approved reasons, some access restrictions apply to the data underlying the findings. Data underlying this study cannot be made publicly available in order to safeguard participant anonymity and that of their organisations. Ethical approval for the project was granted on the basis that only extracts of interviews would be shared (with appropriate anonymisation) as part of publications and other research outputs. In order to share data with other researchers, the participants must be contacted and consent to this data release. Other researchers will be able to request the dataset by contacting the corresponding author or Chair of the University of Sheffield Information School Research Ethics Committee (ischool_ethics@sheffield.ac.uk).

GUIDELINES AND PROCEDURES FOR RPG

1.

GUIDELINES AND PROCEDURES

Beginning with the **September 2017 intake**, all **HKU research postgraduate (RPG) students** have responsibility for

1. using a **data management plan** (DMP), where applicable, to describe the use of data in preparation for, or in the generation of their theses, and
2. depositing, where applicable, a **dataset** in the HKU Scholars Hub.

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

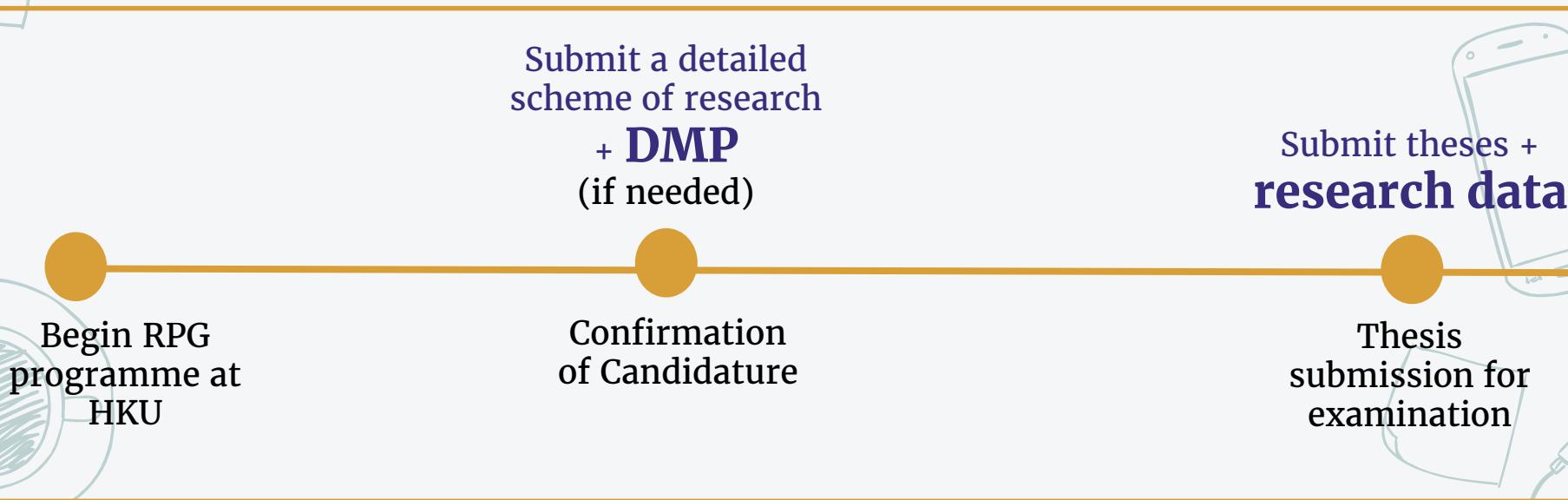
GUIDELINES AND PROCEDURES



- MPH5 & PHD5 **Probation and Confirmation of Candidature** – for description of a data management plan (DMP)
- MPH7 & PHD7 **Period of Study** – for describing when in the period of study, **a dataset, where applicable, is to be submitted**
- MPH14 & PHD14 **Submission of Thesis for Examination** – for description of **dataset submission**
- MPH15 & PHD15 **Thesis Examination** – for consideration of **DMP Entry results and dataset** if applicable, and if desired by the examiners

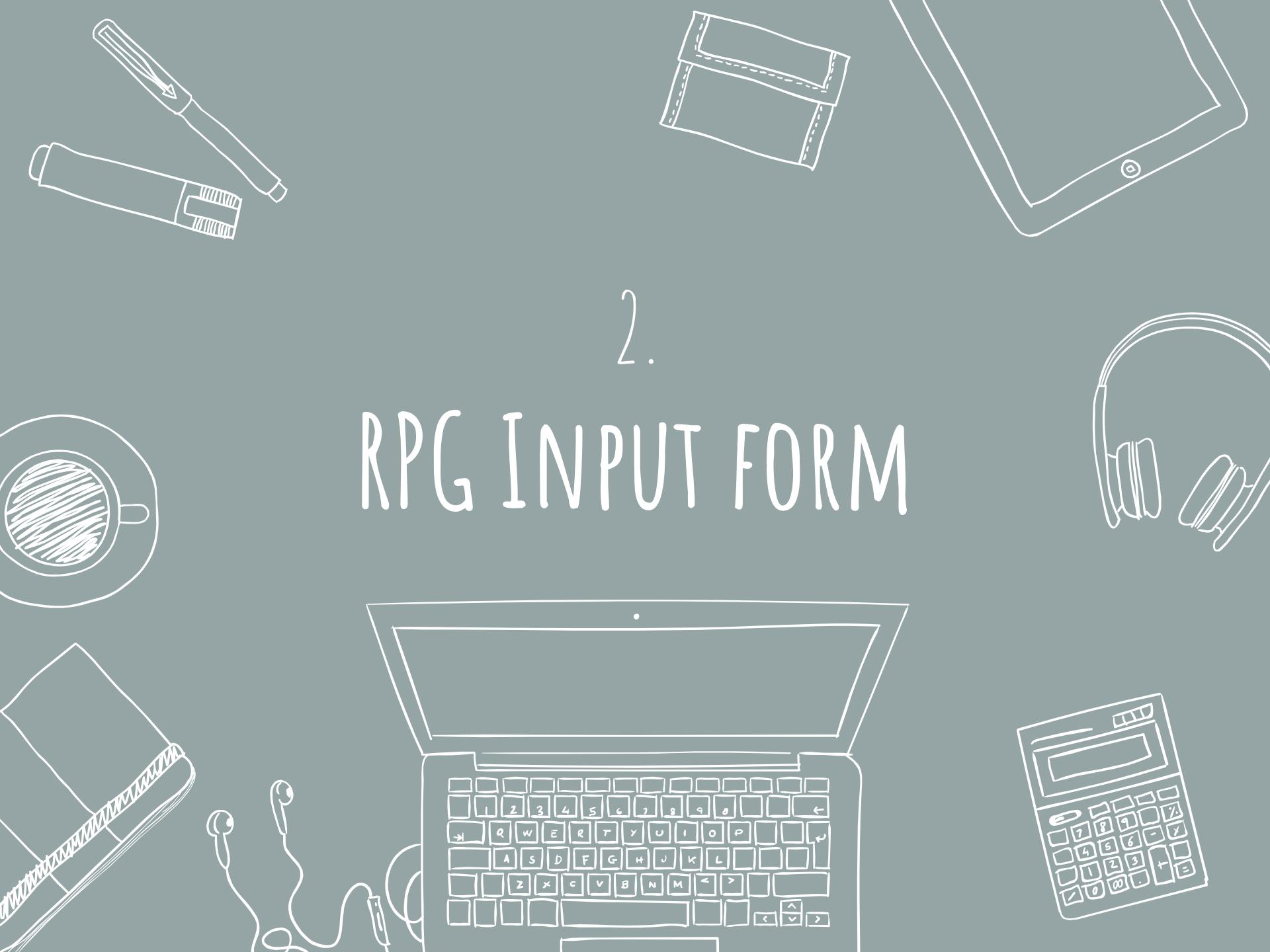
<https://www.gradsch.hku.hk/gradsch/current-students/handbooks>

GUIDELINES AND PROCEDURES



2.

RPG INPUT FORM





RPG Input: DMP & Dataset



DMP Entry

Dataset Submission

DMP Entry Result

DMP Entry

For instructions, please refer to <http://lib.hku.hk/researchdata/rpg.htm>

Student Number 3333333333

Thesis Title TBD



Author Chan Tai Man



Supervisor(s) Dr Chan Bo Hung



Professor Chow Kwok Ming

Degree Doctor of Philosophy



Field of Study Biology



Please choose ONE of the following: A, B, C, or D.

- 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.
- B. Data has been licensed, contracted for, or purchased with a license that explicitly forbids deposit in storage outside the student's or the primary supervisor's control. Primary supervisor approval will be sought.
- C. No data was used in my research project for the creation of my thesis. DMP and Dataset submission is not needed. Primary supervisor approval will be sought.
- D. Dataset Management Plan (DMP). Dataset will be uploaded later.

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

Save as Draft

Submit



Please choose ONE of the following: A, B, C, or D.

- 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.**
- B. Data has been licensed, contracted for, or purchased with a license that explicitly forbids deposit in storage outside the student's or the primary supervisor's control. Primary supervisor approval will be sought.**
- C. No data was used in my research project for the creation of my thesis. DMP and Dataset submission is not needed. Primary supervisor approval will be sought.**
- D. Dataset Management Plan (DMP). Dataset will be uploaded later.**

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

Save as Draft

Submit

Save as Draft

Submit



RPG Input: DMP

Please choose ONE of the following: A, B, C, or D.

- 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.**
- B. Data has been licensed, contracted for, or purchased with a license that explicitly forbids deposit in storage outside the student's or the primary supervisor's control. Primary supervisor approval will be sought.**
- C. No data was used in my research project for the creation of my thesis. DMP and Dataset submission is not needed. Primary supervisor approval will be sought.**
- D. Dataset Management Plan (DMP). Dataset will be uploaded later.**

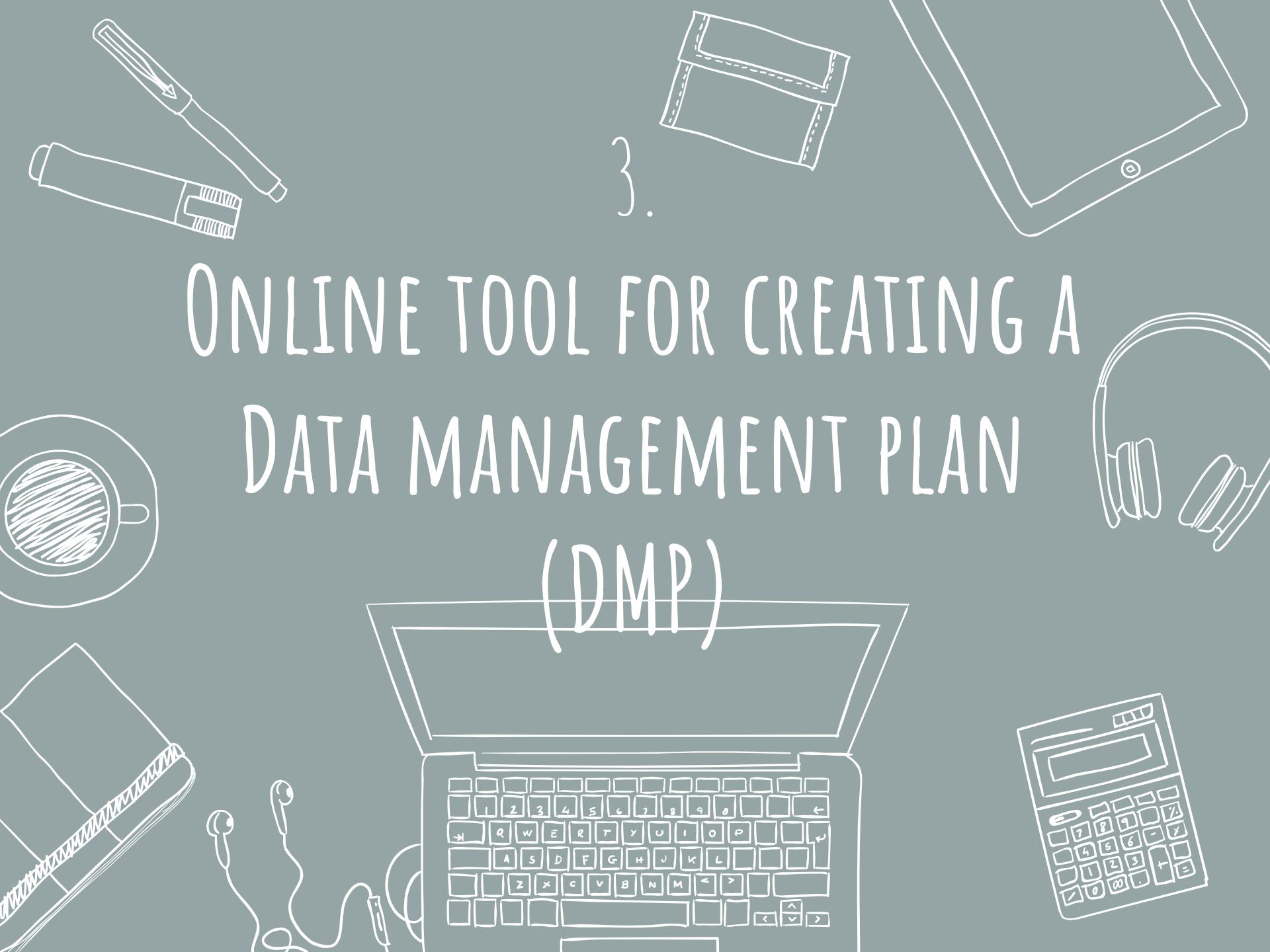
Drag and drop files here, or click in box to choose files.

Save as Draft

Submit

Save as Draft

Submit



ONLINE TOOL FOR CREATING A DATA MANAGEMENT PLAN (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 IWISE 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 TD 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.

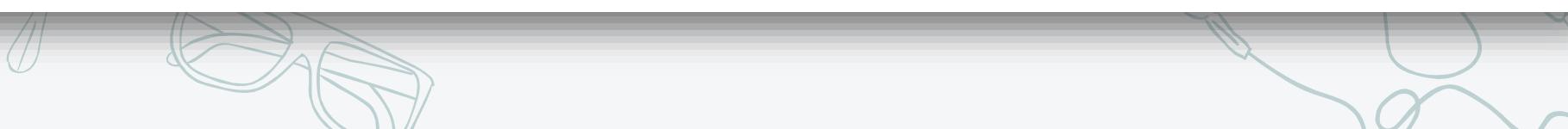
All field data collected under this program will be made available as per NSF guidelines within 2 years of collection via published manuscripts, publicly available final reports to NSF, and data archiving with NODC.

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.

<https://library.ucsd.edu/research-and-collections/data-curation/data-management/dmpssample/DMP-Example-Pinkel.pdf>

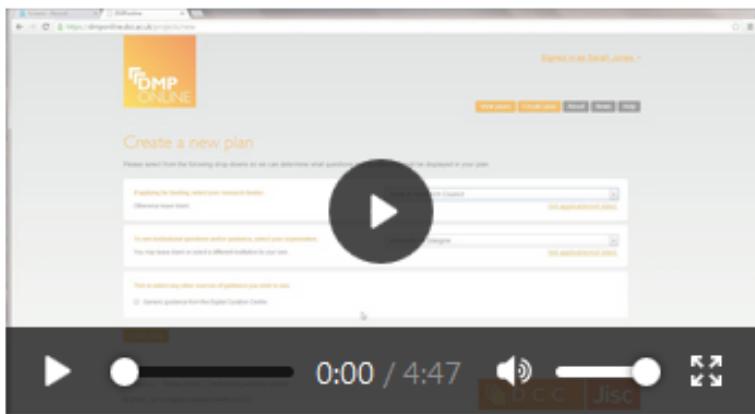
| Sections | Questions |
|--------------------------------|---|
| Data Collection | <p>What data will you collect or create? How will the data be collected or created?</p> |
| Documentation and Metadata | <p>What documentation and metadata will accompany the data?</p> |
| Ethics and Legal Compliance | <p>How will you manage any ethical issues? How will you manage copyright and Intellectual Property Rights (IPR) issues?</p> |
| Storage and Backup | <p>How will the data be stored and backed up during the research? How will you manage access and security?</p> |
| Selection and Preservation | <p>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?</p> |
| Data Sharing | <p>How will you share the data? Are any restrictions on data sharing required?</p> |
| Responsibilities and Resources | <p>Who will be responsible for data management? What resources will you require to deliver your plan?</p> |



Welcome.

DMPRoadmap has been jointly developed by the **Curation Center** to help you write data management plans.

Screencast on how to use DMPRoadmap



Sign in

Email address *

Password *

[Forgot your password?](#)

Remember me

Sign in

[Or, sign in with your institutional credentials](#) (UK users only)

Create account



New to DMPRoadmap? Create an account today.

Please fill in the basic project details below

Plan name Thesis Title 1

ID 3333333333

Your student Number

Grant number

Principal
Investigator/Researcher Chan Tai Man

Your name

Principal
Investigator/Researcher ID Dr Chan Bo Hung

Your supervisor's name

Plan data contact chanbohung@hku.hk

Your email address

Biology

Description

Save

Cancel



Plan details

Generic Data Management Planning Template

Share

Export

Data Collection (2 questions, 2 answered)

What data will you collect or create?

B *I*

This project will generate three main types of raw data.

1. Images from transmitted-light microscopy of giemsa-stained squashed larval brains.
2. Images from confocal microscopy of immunostained whole-mounted larval brains.
3. Western blot data.

Save

Share note

Share note with collaborators

B *I*

Save

Answered 5 minutes ago by eunice08@hku.hk

How will the data be collected or created?

B *I*

Measurements and quantification of the images will then be recorded in spreadsheets.

Micrograph data is expected to total between 100GB and 1TB over the course of the project.

Scanned images of western blots are expected to total around 1GB over the course of the project.

Other derived data (measurements and quantifications) are not expected to exceed 10MB.

Save

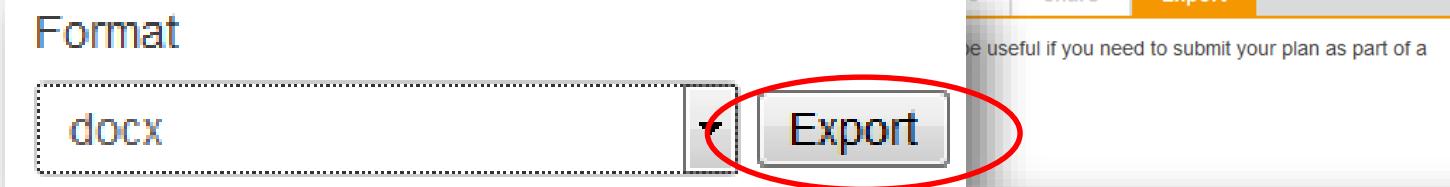
Share note

Share note with collaborators

B *I*

Save

Format



docx

Export Settings (Using default PDF formatting values)

File Name

File Name Thesis Title 1

Included Elements

Details

- Plan Name
- Plan ID
- Grant number
- Principal Investigator / Researcher
- Plan Data Contact
- Plan Description
- Funder
- Institution
- Your ORCID

Sections

Data Collection

What data will you collect or create?

How will the data be collected or created?

Documentation and Metadata

What documentation and metadata will acc...

Ethics and Legal Compliance

How will you manage any ethical issues?

How will you manage copyright and Intellect...

Storage and Backup

How will the data be stored and backed up ...

How will you manage access and security?

Selection and Preservation

Which data are of long-term value and sho...

What is the long-term preservation plan for ...

Data Sharing

How will you share the data?

Are any restrictions on data sharing require...

Responsibilities and Resources

Who will be responsible for data managem...

What resources will you require to deliver y...

Save Reset



THESIS TITLE 1

MY CURATION CENTER'S DEFAULT TEMPLATE

ADMIN DETAILS

Plan Name: My Curation Center's Default Template

Plan ID: 3333333333

Grant number: -

Principal Investigator / Researcher: Chan Tai Man

Plan Data Contact: chanbohung@hku.hk

Plan Description: Biology

Funder: -

Institution: HKU

Your ORCID: -

DATA COLLECTION

What data will you collect or create?

This project will generate three main types of raw data.

1. Images from transmitted-light microscopy of giemsa-stained squashed larval brains.
2. Images from confocal microscopy of immunostained whole-mounted larval brains.
3. Western blot data.

How will the data be collected or created?

Measurements and quantification of the images will then be recorded in spreadsheets.

Micrograph data is expected to total between 100GB and 1TB over the course of the project.

Scanned images of western blots are expected to total around 1GB over the course of the project.

Other derived data (measurements and quantifications) are not expected to exceed 10MB.

DOCUMENTATION AND METADATA

What documentation and metadata will accompany the data?

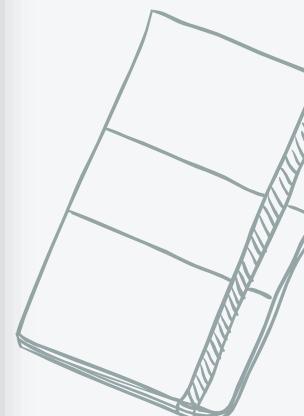
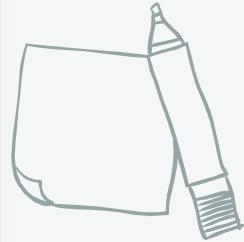
All samples on which data are collected will be prepared according to published standard protocols in the field. All microscopes used for sample examination are serviced and recalibrated regularly. All Drosophila lines used in experiments are checked periodically for phenotypic markers. Drosophila are maintained in live culture according to standard methods in the field.

Files will be named according to a pre-agreed convention. The dataset will be accompanied by a README file which will describe the directory hierarchy and filenames convention.

Each directory will contain an INFO.txt file describing the experimental protocol used in that experiment. It will also record any deviations from the protocol and other useful contextual information.

Microscope images capture and store a range of metadata (field size, magnification, lens phase, zoom, gain, pinhole diameter etc) with each image.

This should allow the data to be understood by other members of our research group and add contextual value to the dataset should it be reused in the future.





RESPONSIBILITIES AND RESOURCES

Who will be responsible for data management?

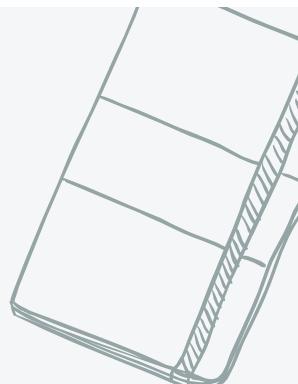
I will be responsible

What resources will you require to deliver your plan?

Cost for acquiring external hard disks for off-site copy

Prepared by:

| Postgraduate student | | Supervisor | |
|----------------------|-----------|------------|-----------|
| Name: | | Name: | |
| Date: | 6-11-2017 | Date: | 6-11-2017 |





RPG Input: DMP

Please choose ONE of the following: A, B, C, or D.

- 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.**
- B. Data has been licensed, contracted for, or purchased with a license that explicitly forbids deposit in storage outside the student's or the primary supervisor's control. Primary supervisor approval will be sought.**
- C. No data was used in my research project for the creation of my thesis. DMP and Dataset submission is not needed. Primary supervisor approval will be sought.**
- D. Dataset Management Plan (DMP). Dataset will be uploaded later.**

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

Save as Draft

Submit

Save as Draft Submit

4.

EXAMPLES



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 in collaborative efforts, such as the ONR IWIS Sea. To guide both modeling and the Process data will be centralized on a server at APL.

Scout Quick-look data responsibility

J.Klymak

LADCP-CTD an

S. Johnston

SIO glider ana

L.Rainville

Co-operative C

H. Simmons, J.Klymak Ongoing model output predictions

R.Pinkel, J. Klymak F-C TD site studies

1. Data Collection

What data will you collect or create?

How will the data be collected or created?

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

Metadata

Metadata will address attribution, discovery, and preservation. The following metadata standards may be used to describe and organize the project's data:

- The IEEE 1484.12.1 – 2002 Standard for Learning Object Metadata (LOM) – may be used to describe and make discoverable learning material data, so that it can be searched for and reused by others. Descriptors include format and pedagogical attributes.
- METS is the Metadata Encoding and Transmission Standard, which will be used to encode the various types of digital objects created by the project. The METS standard is expressed via XML, and is supported and maintained by the Library of Congress. It is used in a number of other dissemination frameworks, including the NJVID project (<http://www.njvid.net/index.php>).

Access and Sharing

The data generated by the project will be housed above. It will be placed in the repository in a timely manner. Seeking access to the various public-use data will be done directly with the LIU Library. Restricted users

Period of Retention

Selection and Retention – Long Island University will work with the project and ensure that it is accessible through various dissemination frameworks. LIU will archive all additional data in a timely manner.

Archiving and Preservation

The repository framework will serve to preserve the data and also make it discoverable regardless of changing technologies and formats.

Storage & Backup

2. Documentation and Metadata

What documentation and metadata will accompany the data?



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?

INTELLECTUAL PROPERTY RIGHTS (IPR)

- × Intellectual property rights (IPR) include copyright, patents, trademarks and design rights.
- × The HKU Intellectual Property Rights Policy sets out ownership of intellectual property rights of staff and students and explains the University position on both copyright and patents.

<http://www.rss.hku.hk/contracts/ipr>

INTELLECTUAL PROPERTY RIGHTS (IPR)

"If research data included in a thesis are obtained by a collaborative effort (including collaboration between the Student and a supervisor or other researcher at the University), such data may be the joint property of the Student and the collaborating party.

It is strongly advised that Students and supervisors/researchers make clear agreements in advance concerning the ownership and use of Intellectual Property Rights created in connection with a Student thesis."

<http://www.rss.hku.hk/contracts/ipr>



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. 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.

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?



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?

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 **Versitech**, or the **Technology Transfer Office**.

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

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

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.

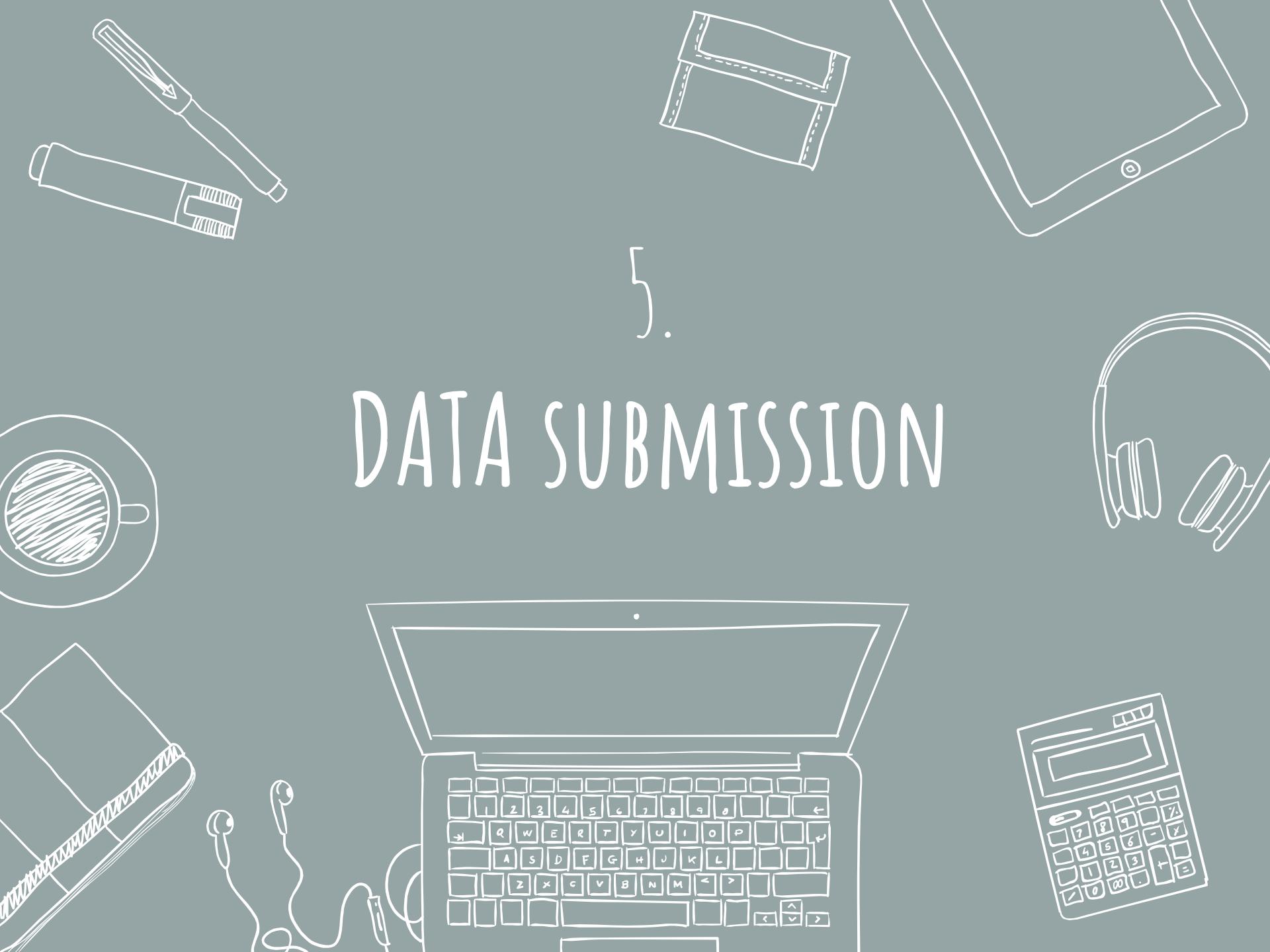
Intended and foreseeable users of the data are NutNet collaborators and participants, as well as other scientists interested in arthropod-plant relationships. This data set could be used in combination with similar data sets from other NutNet sites or for meta-analysis.

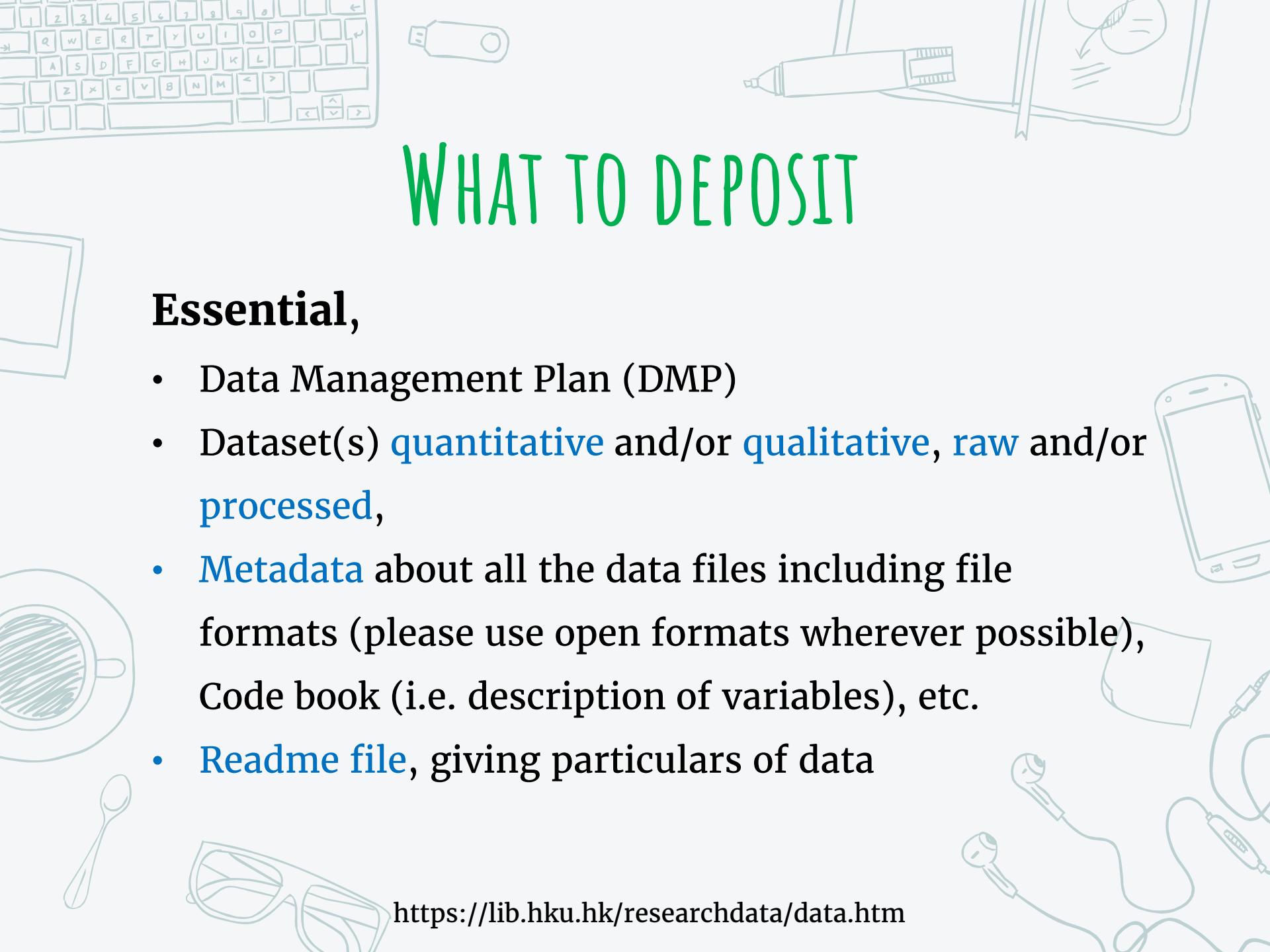
5. Plans for Archiving and Preservation

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.

5.

DATA SUBMISSION





WHAT TO DEPOSIT

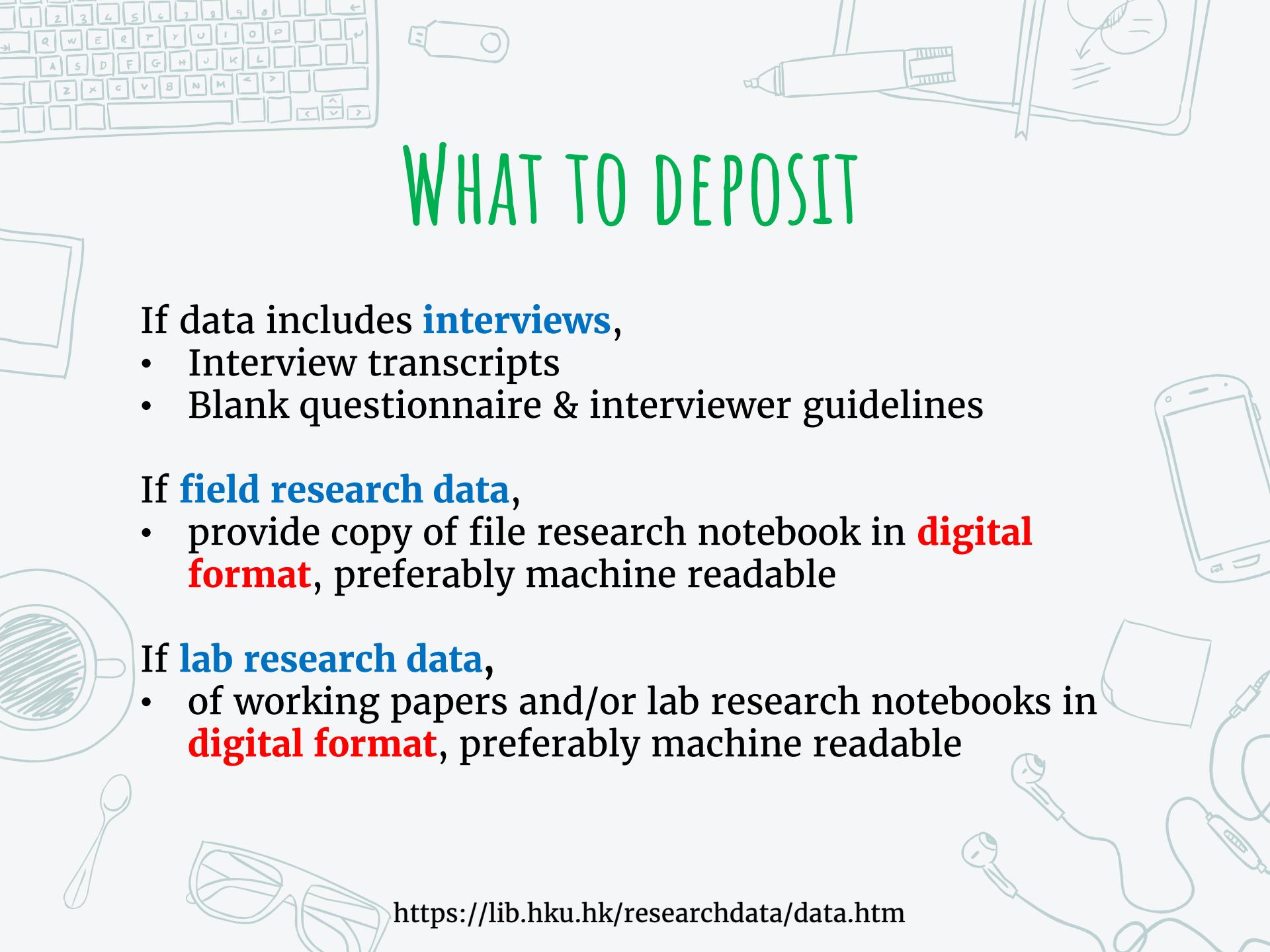
Essential,

- Data Management Plan (DMP)
- Dataset(s) **quantitative** and/or **qualitative**, raw and/or **processed**,
- **Metadata** about all the data files including file formats (please use open formats wherever possible), Code book (i.e. description of variables), etc.
- **Readme file**, giving particulars of data

WHAT TO DEPOSIT

If data includes **personal data**, the data should be put under **restricted access**,

- **Personal data from clinical research** (i.e. Institutional Review Board (IRB) approved) [<https://www.med.hku.hk/research/research-ethics/human-ethics>]
- provide approval code, consent forms, ethical application form when available, please state the risk of re-identification from the different datafiles and how the risk has been minimised for any dataset intended for sharing.
- **Personal data from non-clinical research** (i.e. Human Research Ethics Committee (HREC) approved) [<http://www.rss.hku.hk/integrity/ethics-compliance/hrec>]
- **provide approval code, consent forms, ethical application form**, please state the risk of re-identification from the different datafiles and how the risk has been minimised for any dataset intended for sharing.



WHAT TO DEPOSIT

If data includes **interviews**,

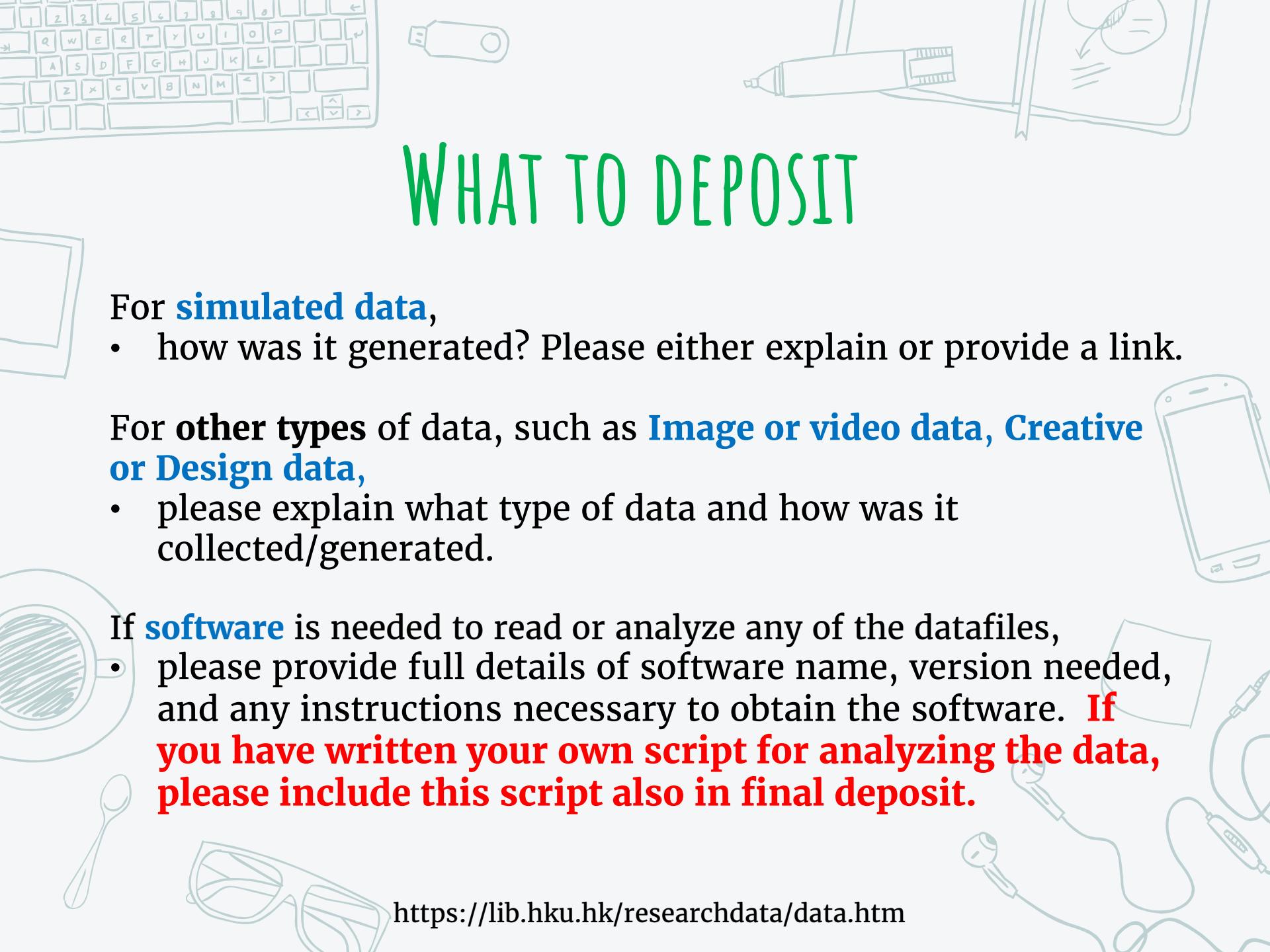
- Interview transcripts
- Blank questionnaire & interviewer guidelines

If **field research data**,

- provide copy of file research notebook in **digital format**, preferably machine readable

If **lab research data**,

- of working papers and/or lab research notebooks in **digital format**, preferably machine readable



WHAT TO DEPOSIT

For **simulated data**,

- how was it generated? Please either explain or provide a link.

For **other types** of data, such as **Image or video data, Creative or Design data**,

- please explain what type of data and how was it collected/generated.

If **software** is needed to read or analyze any of the datafiles,

- please provide full details of software name, version needed, and any instructions necessary to obtain the software. **If you have written your own script for analyzing the data, please include this script also in final deposit.**

6.

REFERENCES



In this section

[Briefing Papers](#)[How-to Guides & Checklists](#)[Developing RDM Services](#)[Curation Lifecycle Model](#)[Curation Reference Manual](#)[Policy and legal](#)[Data Management Plans](#)[Checklist](#)[DMPonline](#)[FAQ on DMPonline](#)[FAQ on Data Management Plans](#)[Funders' requirements](#)[Guidance and examples](#)[Tools](#)[Case studies](#)

Example DMPs and guidance

[UCSD Example Data Management Plans](#)

Over 20 example plans submitted to the National Science Foundation (NSF) in the United States by academics at UC San Diego

[Colorado School of Mines examples](#)

A variety of US example DMPs from Mines and elsewhere

[NSF data management plans](#)

5 DMPs submitted to the NSF, shared by the DataOne initiative

[Biology and chemistry DMPs](#)

Three example DMPs from the USA shared by NECDMC, an instructional tool for teaching RDM to undergraduates, graduate students, and researchers in the health sciences, sciences and engineering.

[Data Management](#)[Follow Best Practices](#)[Write an Effective Data Management Plan](#)[Sample NSF Data Management Plans](#)[NIH Policy on Rigor and Reproducibility](#)

UC San Diego 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.



Contact the [Research Data Curation Program](#) with questions about our services or to provide feedback on our new website.

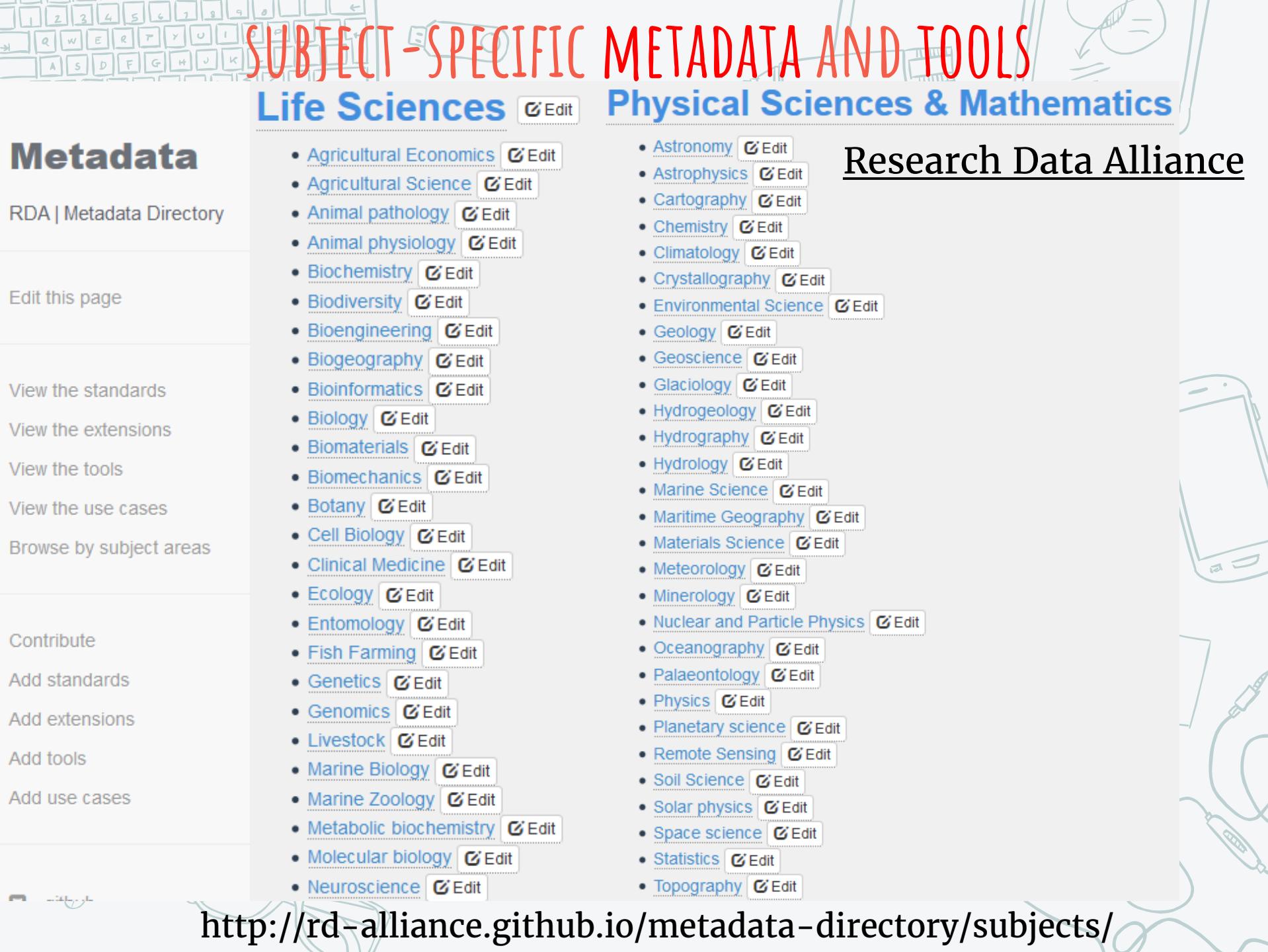
Office of the Director (OD)

Office of Cyberinfrastructure (OD/OCI)

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

Office of Integrative Activities (OD/OIA)

[DMP Example Todd Martz SIO.pdf](#) From Professor Todd Martz's proposal entitled "MRI: Development of an instrument for testing and calibration of autonomous sensors for



SUBJECT-SPECIFIC METADATA AND TOOLS

Metadata

RDA | Metadata Directory

Edit this page

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

Life Sciences

[Edit](#)

- Agricultural Economics [Edit](#)
- Agricultural Science [Edit](#)
- Animal pathology [Edit](#)
- Animal physiology [Edit](#)
- Biochemistry [Edit](#)
- Biodiversity [Edit](#)
- Bioengineering [Edit](#)
- Biogeography [Edit](#)
- Bioinformatics [Edit](#)
- Biology [Edit](#)
- Biomaterials [Edit](#)
- Biomechanics [Edit](#)
- Botany [Edit](#)
- Cell Biology [Edit](#)
- Clinical Medicine [Edit](#)
- Ecology [Edit](#)
- Entomology [Edit](#)
- Fish Farming [Edit](#)
- Genetics [Edit](#)
- Genomics [Edit](#)
- Livestock [Edit](#)
- Marine Biology [Edit](#)
- Marine Zoology [Edit](#)
- Metabolic biochemistry [Edit](#)
- Molecular biology [Edit](#)
- Neuroscience [Edit](#)

Physical Sciences & Mathematics

Research Data Alliance

- Astronomy [Edit](#)
- Astrophysics [Edit](#)
- Cartography [Edit](#)
- Chemistry [Edit](#)
- Climatology [Edit](#)
- Crystallography [Edit](#)
- Environmental Science [Edit](#)
- Geology [Edit](#)
- Geoscience [Edit](#)
- Glaciology [Edit](#)
- Hydrogeology [Edit](#)
- Hydrography [Edit](#)
- Hydrology [Edit](#)
- Marine Science [Edit](#)
- Maritime Geography [Edit](#)
- Materials Science [Edit](#)
- Meteorology [Edit](#)
- Mineralogy [Edit](#)
- Nuclear and Particle Physics [Edit](#)
- Oceanography [Edit](#)
- Palaeontology [Edit](#)
- Physics [Edit](#)
- Planetary science [Edit](#)
- Remote Sensing [Edit](#)
- Soil Science [Edit](#)
- Solar physics [Edit](#)
- Space science [Edit](#)
- Statistics [Edit](#)
- Topography [Edit](#)

DATA REPOSITORIES



The University of Hong Kong

The HKU Scholars Hub 香港大學學術庫



HELP



HKU Login
Guest Login



Home

Publications

Researchers

Organizations

Grants

Datasets

Theses

Patents

Community Service

Dataset

Browse

- Faculties
- Title
- Keyword
- RGC Code
- ANZSRC

Title

Enter Search Terms Here...



Search

- [HKU Scholars Hub](#)
- [Re3data](#)
- [Open Access Directory](#)

Discovery - Top 10

Department

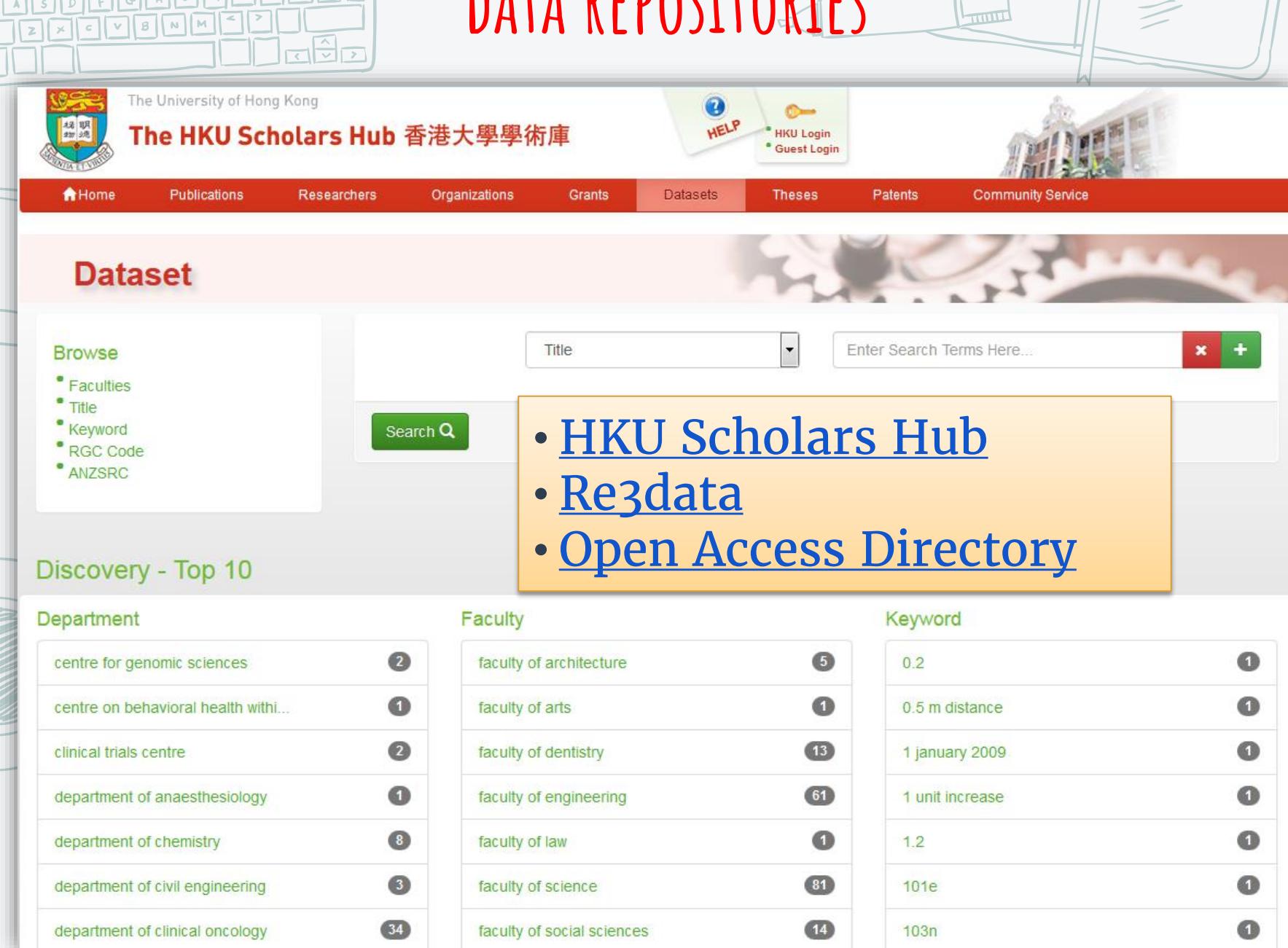
| | |
|-------------------------------------|----|
| centre for genomic sciences | 2 |
| centre on behavioral health with... | 1 |
| clinical trials centre | 2 |
| department of anaesthesiology | 1 |
| department of chemistry | 8 |
| department of civil engineering | 3 |
| department of clinical oncology | 34 |

Faculty

| | |
|----------------------------|----|
| faculty of architecture | 5 |
| faculty of arts | 1 |
| faculty of dentistry | 13 |
| faculty of engineering | 61 |
| faculty of law | 1 |
| faculty of science | 81 |
| faculty of social sciences | 14 |

Keyword

| | |
|-----------------|---|
| 0.2 | 1 |
| 0.5 m distance | 1 |
| 1 january 2009 | 1 |
| 1 unit increase | 1 |
| 1.2 | 1 |
| 101e | 1 |
| 103n | 1 |



Recommended Repositories

PLOS requires that authors comply with field-specific standards for preparation and recording of data and select repositories appropriate to their field, for example deposition of microarray data in ArrayExpress or GEO; deposition of gene sequences in GenBank, EMBL or DDBJ; and deposition of ecological data in Dryad. Authors are encouraged to select repositories that meet accepted criteria as trustworthy digital repositories.

PLOS has identified a set of established repositories below, which are recognized and trusted within their respective communities. For further information on environmental and biomedical science repositories and field standards, we suggest utilizing FAIRsharing; we have also created a FAIRsharing page of PLOS recommended data repositories. Additionally, the Registry of Research Data Repositories (Re3Data) is a full scale resource of registered repositories across subject areas. Both FAIRsharing and Re3Data provide information on an array of criteria to help researchers identify the repositories most suitable for their needs (licensing, certificates and standards, policy, etc.).

Authors are encouraged to select the repository most appropriate for their research. PLOS does not dictate repository selection for the data access policy. If authors use repositories with stated licensing policies, the policies should not be more restrictive than the Creative Commons Attribution (CC BY) license. More information about the content license can be found in our licenses and copyright policy.

If no specialized community-endorsed open repository exists, institutional repositories that use open licenses permitting free and unrestricted use or public domain, and that adhere to best practices pertaining to responsible data sharing, sustainable digital preservation, proper citation, and openness are also suitable for data deposition.

Cross-disciplinary repositories

- [Dryad Digital Repository](#)
- [figshare](#)
- [Harvard Dataverse Network](#)
- [Open Science Framework](#)
- [Zenodo](#)

<http://journals.plos.org/plosone/s/data-availability>

Repositories by type

| | | |
|---------------------|-------------------|--------------------------------|
| Biochemistry | Neuroscience | Social Sciences |
| Biomedical Sciences | Omics | Structural Databases |
| Marine Sciences | Physical Sciences | Taxonomic & Species Diversity |
| Model Organisms | Sequencing | Unstructured and/or Large Data |

Biochemistry

- [caNanoLab](#)
- [Kinetic Models of Biological Systems \(KiMoSys\)](#)
- [Mass spectrometry Interactive Virtual Environment \(MassIVE\)](#)
- [PubChem](#)
- [Standards for Reporting Enzymology Data \(STRENDA DB\)](#)

Biomedical Sciences

- [The Cancer Imaging Archive \(TCIA\)](#)
- [Influenza Research Database](#)
- [National Addiction & HIV Data Archive Program \(NAHDAP\)](#)
- [National Database for Autism Research \(NDAR\)](#)
- [PhysioNet](#)
- [SICAS Medical Image Repository](#)

Marine Sciences

- [SEA scieNtific Open data Edition \(SEANOE\)](#)

<http://journals.plos.org/plosone/s/data-availability>