

RPG seminar on DMP and data, theses and Turnitin submission

Jesse Xiao, Data and Scholarly Communication Librarian

Rebecca Tam, Science Librarian & Learning Services Coordinator

Tina Yang, Associate Librarian (Learning and Research Services)

April 2022



Topics covered

- RPG Data Management Plan (DMP)
- Dataset Submission via DataHub
- Thesis submission to Turnitin for similarity check
- Final Thesis Submission: MPhil & PhD Thesis

What do we mean by Research Data



Computer code



Spreadsheets



Audio tapes



Notebooks

Any information that has been collected, observed, generated or created to validate original research findings.

Why research data management (RDM)?

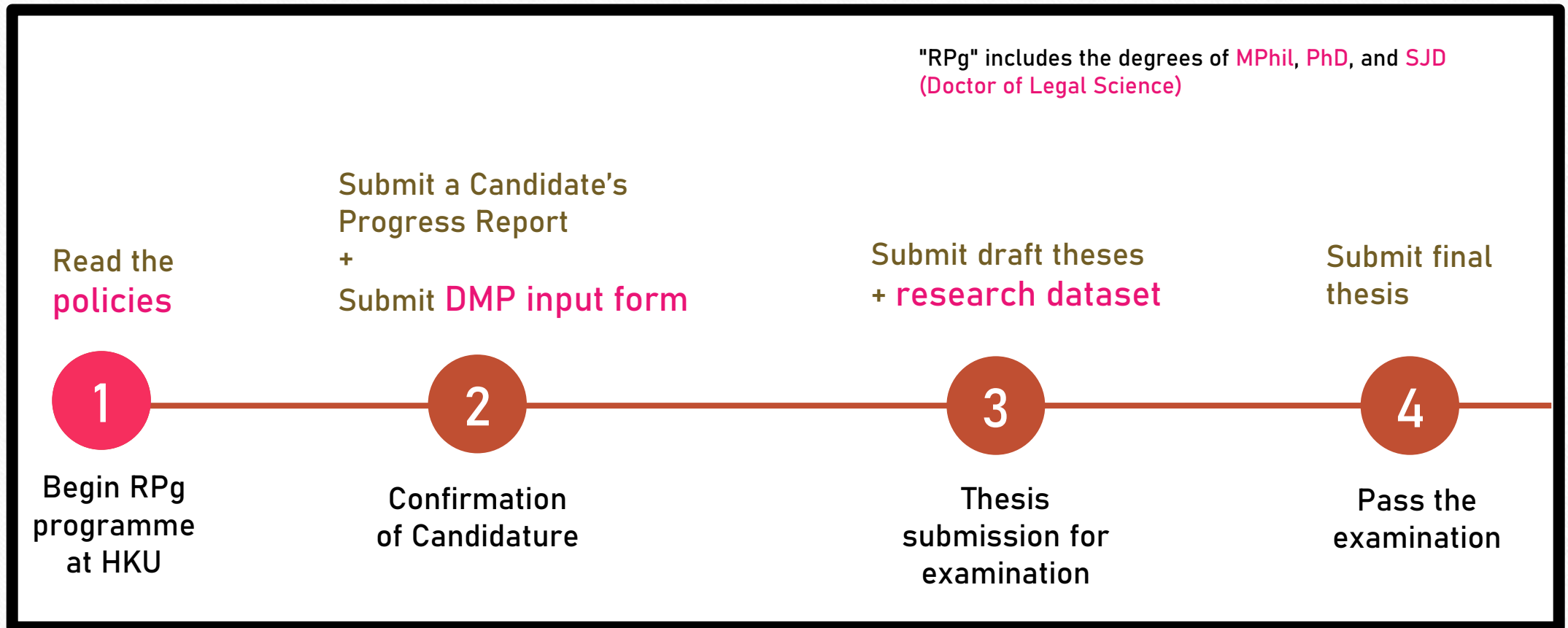
Compliance

- Compliance with University and School policies
- Compliance with funding organizations that require a data management plan and data sharing
- Compliance with journals that require to submit your data accompanying the article

Benefits

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

What actions to take?



"RPg" includes the degrees of MPhil, PhD, and SJD (Doctor of Legal Science)

HKU Policy on Research Data and Records Management

The University of Hong Kong


HKU Home | Research@HKU | Strategic Research Areas | Graduate School | Knowledge Exchange Office | Technology Transfer Office

 A+ A A-

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

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

About Us

Research Funding

IPR & Contracts

Research Integrity

- Responsible Conduct of Research
- [Research Data and Records Management](#)
- Research Ethics Compliance
- Safety Approval & Insurance

Honours & Awards

RAE 2014 & Resources



- Read Handbook for Mphil & PhD Programmmes



Appendix XIX

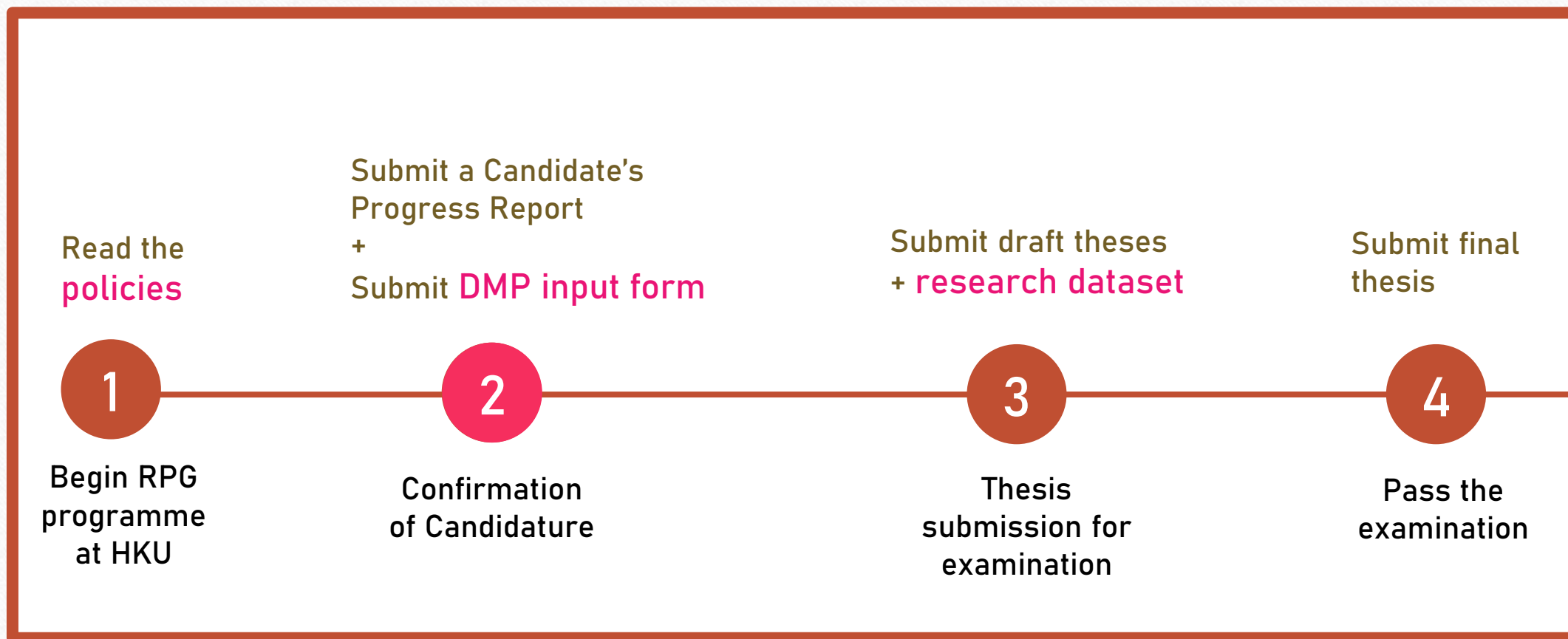
Procedures for the Degree of Master of Philosophy (MPhil)

Appendix XX

Procedures for the Degree of Doctor of Philosophy (PhD)

2

Submit DMP input form



"RPg" includes the degrees of MPhil, PhD, and SJD (Doctor of Legal Science)



RESEARCH DATA SERVICES

THE UNIVERSITY OF HONG KONG
LIBRARIES



RDM FOR RESEARCH
POSTGRADUATE (RPG) STUDENTS
- INSTRUCTIONS FOR DATA

LOGIN RPG INPUT FORM

RDM PLANNING

DATA MANAGEMENT PLAN (DMP)

RPG INPUT: DMP & DATASET

E-THESIS AND DATASET SUBMISSION



RDM FOR RESEARCH



RPG DATA MANAGEMENT PLAN (DMP) INPUT FORM

RPG STUDENTS



<https://hub.hku.hk/researchdata>

1. CHECK PROFILE

2. DATA OPTIONS

3. SUBMIT DMP

MY PROFILE

University No.:	3012345678
Name:	Mr CHAN Tai Man
Email:	u1234567@hku.hk
Degree:	Master of Philosophy
Department/Faculty:	School of Biomedical Sciences
Year of Admission:	2017
Hub RPg URL:	http://hub.hku.hk/cris/rpg/rpg123456

SUPERVISORS

Primary Supervisor:	Professor Chan Siu Man	(from 2017-09-01 to 2019-08-31)
Co-Supervisor:	Professor Wong Mei Yi	(from 2017-09-01 to 2019-08-31)

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.

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

After submitting this form, an email will go to your supervisors. If your primary supervisor rejects this choice, you will need to return to this form, and make a choice that is acceptable to your primary supervisor.

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

SUBMIT

LOGOUT

Option A

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.

The URL of your dataset retrieved online :

https://figshare.com/articles/_Genetic_Predisposition_to_Increased_Blood

Provide the URL & citation of the resource/data

Citation, including address of library or archive used (if not online) :

Proitsi, P, Lupton, MK, Velayudhan, L, Newhouse, S, Fogh, I, Tsolaki, M, Daniilidou, M, Pritchard, M, Kloszewska, I, Soininen, H, Mecocci, P, Vellas, B, Williams, J, Stewart, R, Sham, P, Lovestone, S, Powell, JF. (2014). Data from: Genetic Predisposition to Increased Blood Cholesterol and Triglyceride Lipid Levels and Risk of Alzheimer Disease: A Mendelian Randomization Analysis. [Data File]. The authors confirm that, for approved reasons, some access restrictions apply to the data underlying the findings. Data are from the Genetic and Environmental Risk for Alzheimer's Disease 1 (GERAD1) Consortium, the AddNeuroMed, the Dementia Case Register (DCR) and the Alzheimer's Disease Neuroimaging Initiative (ADNI) studies. Data access is available on request from the following: GERAD1, williamsj@cardiff.ac.uk. AddNeuroMed, simon.lovestone@psych.ox.ac.uk. DCR data, john.powell@kcl.ac.uk. ADNI data is publicly available (adni.loni.usc.edu).



Option B

YOU HAVE SELECTED OPTION B:

- 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.

The name of licensor, seller, or owner of your dataset :

Provide the name of licensor,
seller, or owner of the data

Proitsi, P, Lupton. Data are purchased from the Genetic and Environmental Risk for Alzheimer's Disease 1 (GERAD1) Consortium.



Option C

YOU HAVE SELECTED OPTION C:

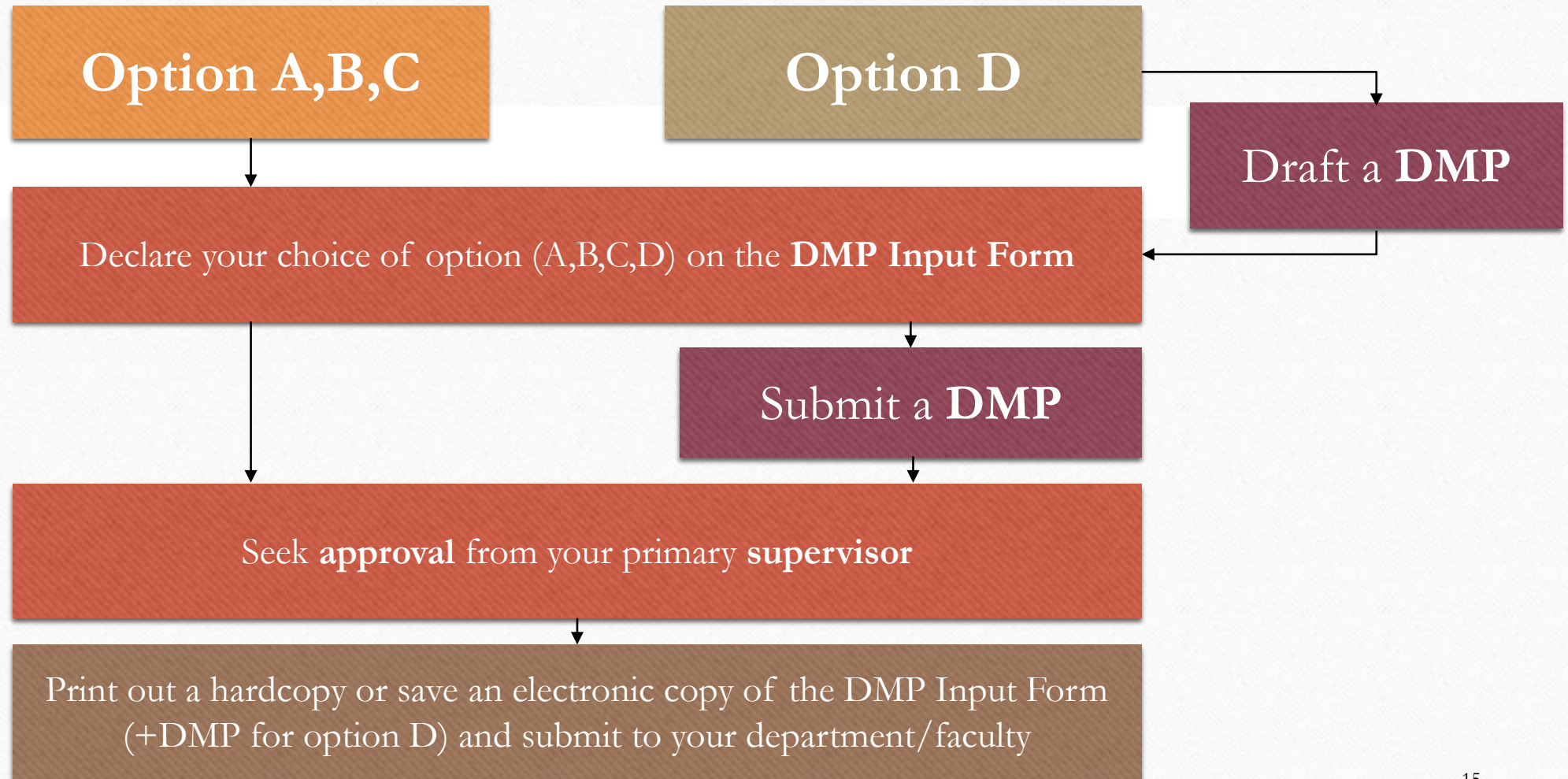
- C. No data was used in my research project for the creation of my thesis. DMP and Dataset submission is not needed. Primary supervisor approval will be sought.

No extra information required

Decide the Nature of Your Data

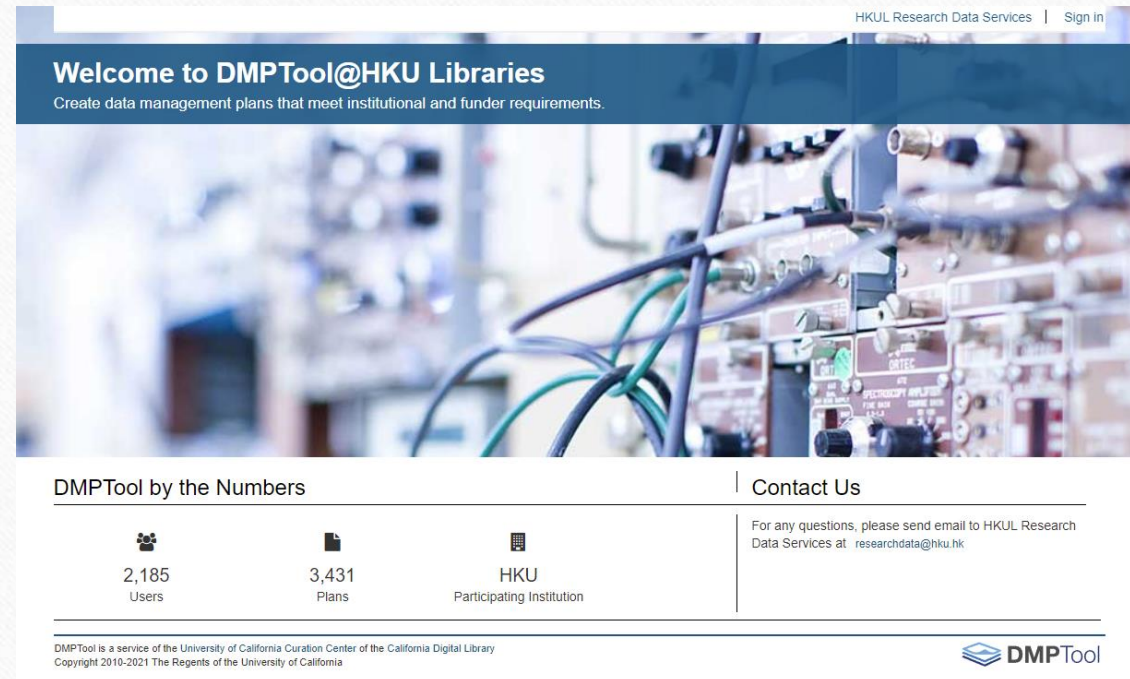


Submit the DMP Input Form



Writing a DMP

- Write a DMP using the template via DMPTool@HKUL
- <http://dmp.lib.hku.hk/>






HKUL Research Data Services | Sign in

Welcome to DMPTool@HKU Libraries

Create data management plans that meet institutional and funder requirements.


DMPTool by the Numbers

 2,185 Users	 3,431 Plans	 HKU Participating Institution
---	---	---

Contact Us

For any questions, please send email to HKU Research Data Services at researchdata@hku.hk

DMPTool is a service of the University of California Curation Center of the California Digital Library
Copyright 2010-2021 The Regents of the University of California



RPG Data Management Plan (DMP)

- A DMP is:
- a digital document in which you describe **what data you will collect** during your research project, how you are going to **store and manage** the data during the project, and what will happen to the **data after the project is finished**.

Writing a DMP

- Follow the template and answer the questions under each section
- Export and download a .pdf or .docx document file when you completed writing your DMP.

The screenshot displays a web interface for writing a Data Management Plan (DMP). At the top, there are four tabs: 'Project Details', 'Plan overview', 'Write Plan' (which is the active tab), and 'Download'. Below the tabs, there is a progress indicator showing '0/14 answered'. The main content area lists seven sections, each with a plus icon and a progress count in parentheses:

- + Data Collection (0 / 2)
- + Documentation and Metadata (0 / 1)
- + Ethics and Legal Compliance (0 / 2)
- + Storage and Backup (0 / 3)
- + Selection and Preservation (0 / 2)
- + Data Sharing (0 / 2)
- + Responsibilities and Resources (0 / 2)

1. Data Collection

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

Project: Project Name Drosophila Genetics - BBSRC Example
https://www.gla.ac.uk/media/Media_418168_smxx.pdf

Data areas and data types

Outline the volume, type and content of data that will be generated e.g. experimental measurements, models, records and images

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.

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.

2. Documentation and Metadata

- What documentation and metadata will accompany the data?

Project: Atmospheric CO₂ Concentrations, Mauna Loa Observatory, Hawaii, 2011--2013.

https://www.dataone.org/sites/all/documents/DMP_MaunaLoa_Formatted.pdf

Metadata will be comprised of two formats—contextual information about the data in a text based document and ISO 19115 standard metadata in an xml file. These two formats for metadata were chosen to provide a full explanation of the data (text format) and to ensure compatibility with international standards (xml format). The standard XML file will be more complete; the document file will be a human-readable summary of the XML file.

Project: Work package (WP) 6 responsible innovation of the synthetic biology applications to water supply and remediation project

https://www.gla.ac.uk/media/media_441723_en.pdf

What documents will describe the data?

We will record **record metadata in a .txt file** the data collection. We will include information that will describe key aspects of the data, using DDI best practise, including:

- Principal investigator
- Funding sources
- Data collector/producer
- Sample and sampling procedures (if appropriate)
- Substantive, temporal, and geographic coverage of the data collection
- Data source(s)
- Variables
- Technical information on files
- Data collection instruments
- Citations to related publications
- Technical information on files, e.g., information on formats, file linking
- Data collection instruments
- Interviewer guide
- Coding instrument

3. Ethics and Legal Compliance

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

Project: Work package (WP) 6 responsible innovation of the synthetic biology applications to water supply and remediation project

https://www.gla.ac.uk/media/media_441723_en.pdf

How will you manage any ethical issues?

All interviewees and focus group participants will sign a Consent form agreed to by the School ethics committee.

We have guaranteed anonymity to our interviewees and focus group participants. Therefore we will not be depositing .wav files as this would compromise that guarantee. However, anonymised transcripts of the interviews and focus groups will be deposited.

We will make sure consent forms make provision for future sharing of data.

All identifying information will be kept in a locked filing cabinet and not stored with electronic files.

4. Storage and Backup

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

Project: Work package (WP) 6 responsible innovation of the synthetic biology applications to water supply and remediation project

https://www.gla.ac.uk/media/media_441723_en.pdf

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

Storage and back up will be in three places:

- On Laptop of [Researcher1]
- On a portable storage device (hard drive)
- On the University of Glasgow server.

[Researcher1] will be responsible for the storage and back up of data. This will be done weekly.

How will you manage access and security?

Both the laptop and external storage device will be password protected. The risks are that the computer will be hacked and the external drive stolen. The laptop will have anti-virus software installed which is updated daily. The external device is in a locked cabinet. Moreover, the data files will also be password protected.

The paper consent forms from interviewees/focus group participants will be kept in a locked cabinet in the university office of [Researcher1].

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?

Project: Atmospheric CO₂ Concentrations, Mauna Loa Observatory, Hawaii, 2011--2013.
https://www.dataone.org/sites/all/documents/DMP_MaunaLoa_Formatted.pdf

Our intent is that the long-term high quality final data product generated by this project will be available to policy communities in perpetuity. The raw supporting data will be available in perpetuity for use by researchers to confirm the quality of the Mauna Loa Record. The investigators have made arrangements for long-term stewardship and curation at the Carbon Dioxide Information and Analysis Center (CDIAC), Oak Ridge National Laboratory (ter of support). The standardized metadata record for the Mauna Loa CO₂ data will be added to the metadata record database at CDIAC, so that interested users can discover the Mauna Loa CO₂ record along with other related Earth science data. CDIAC has a standardized data product citation [5] including DOI, that indicates the version of the Mauna Loa Data Product and how to obtain a copy of that product.

6. Data Sharing

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

Project: Effects of temperature and salinity on population growth of the estuarine copepod, *Eurytemora affinis*.

https://www.dataone.org/sites/all/documents/DMP_Copepod_Formatted.pdf

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.

5. Roles and responsibilities

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

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?



Option D

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.

Attach the DMP document file here before submitting the form

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

YOU HAVE SELECTED OPTION D:

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

Option D

DATA MANAGEMENT PLAN SUBMISSION STATUS

Version 1	DMP Submitted on 20180113 15:04:34	Download DMP File
Approval status :	Approved by Professor Chan Ying Shing	
Reference No. :	2018011304	

Print this page :



If you would like to further update your DMP, you may revisit this page to upload and replace your previous DMP by clicking "**Update DMP**" button below.

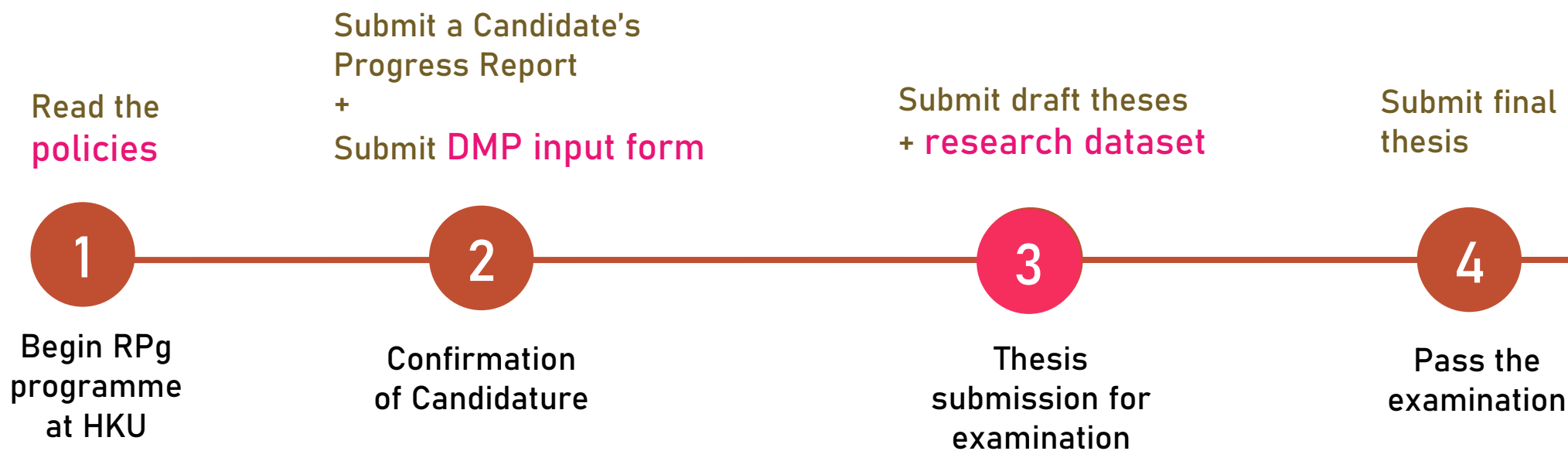
UPDATE DMP

You have not yet submitted your DATASET, click "**SUBMIT DATASET**" to submit!

SUBMIT DATASET

3

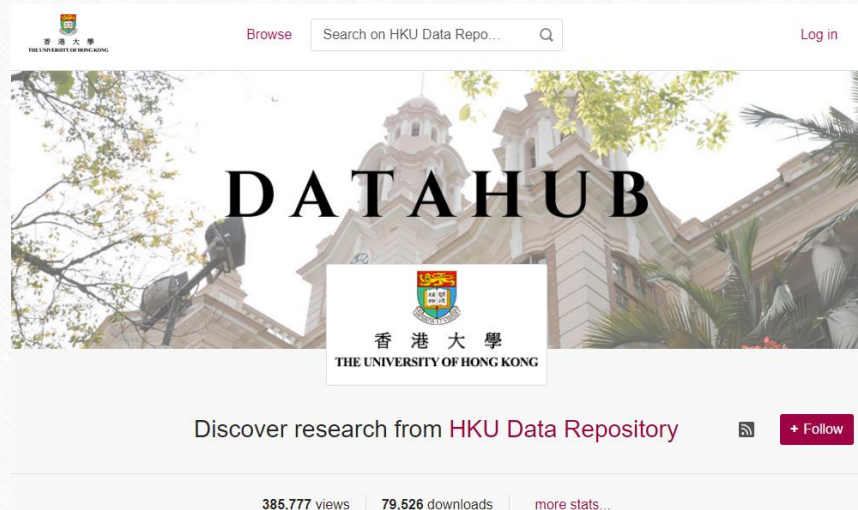
Submit your dataset



"RPg" includes the degrees of MPhil, PhD, and SJD (Doctor of Legal Science)

What is DataHub?

- DataHub powered by Figshare is the comprehensive **institutional data repository** for research data and scholarly outputs produced by researchers and students in the University of Hong Kong and their collaborators.



What to deposit?

1. All “**research data**” that are generated/collected/derived by you (under your ownership) and used in generating your thesis
2. A **README file** describing what your data are about and how to read them
3. If your data include personal data, please provide the **approval information** and documents as listed in the table

RPg students declared “Option A,B or C” in their DMP Input Form with their primary supervisors’ approval obtained are not required to submit their dataset

Personal data

Personal data from <u>clinical</u> research (i.e. Institutional Review Board (IRB) approved)	Personal data from <u>non-clinical</u> research (i.e. Human Research Ethics Committee (HREC) approved)
<ol style="list-style-type: none">1. Provide approval code, consent forms, ethical application form when available2. Please state the risk of re-identification from the different datafiles and how the risk has been minimised for any dataset intended for sharing.	<ol style="list-style-type: none">1. Provide approval code, consent forms, ethical application form2. Please state the risk of re-identification from the different datafiles and how the risk has been minimised for any dataset intended for sharing.

README file

- The README file aims to provide your supervisor, data curators, and potential readers (if published in open access) an overview of your data:
 - How your data are organized (e.g. folder structure and descriptions)
 - How the files are named (e.g. Experiment number_DD/MM/YYYY)
 - Relationships between the files
 - Details of your data (i.e. variables, number of entries, etc.)
 - Methodological information (i.e. details of the software required to open your files)

Preparing README file

- Prepare the README file following the [template](#)
- As descriptive as possible
- If you have too much detail, you may create individual README file for a specific folder (optional)

Procedures of RPg dataset submission

1. Read before you submit

- Please review the below:
 - Depositor's Agreement
 - What to Deposit
 - When to Redact and Anonymize
 - Open Access
 - What Can I Upload

2. Access restriction

- If your data contain **unpublished data** or **sensitive and confidential data** at time of submission, apply an embargo period for access restriction to your files
- Consult your supervisor on the suitability of your chosen access right option before submission
- Read the procedures for applying restricted access for RPg students

2. Access restriction

- Permanent embargo for dataset containing sensitive and confidential data
- Embargo period of your choice (years / specific date / permanent) for unpublished data

Supporting Data for "Implementing Victims' Rights A case study of India."

Reason: The files contain sensitive and confidential data such as interview scripts which require protection.

Supporting Data for "Implementation of Victims' Rights A case study of India."

Cite

Share

+ Collect

...

<https://doi.org/10.25442/hku.16621561>

Dataset (Main folder).zip

Reason: The research work has not yet been published, as such the whole dataset should be under embargo for 2 years in order to allow for publication of the work.

— 1 — 10 — 23 —
year(s) month(s) day(s)

until file(s) become available

Supporting information for "Solution Combustion Synthesized Porous Ceria-Zirconia Supported Platinum-Ruthenium Catalyst for Simultaneous Oxidation of Propane and Carbon Monoxide"

Cite

Share

+ Collect

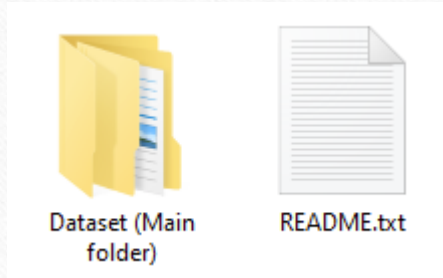
...

<https://doi.org/10.25442/hku.16680394>

3. Prepare your data

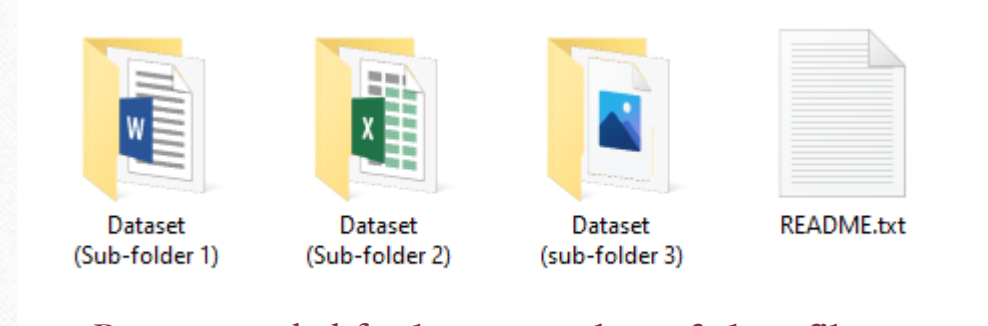
- Have your research data and README file prepared
- Organize your data files. If you have sub-folders for file classification, you may either:

Put all your files or sub-folders into **one** folder



OR

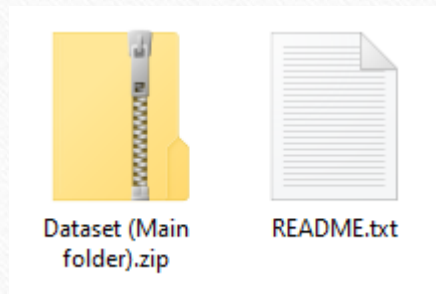
Separate your files into **multiple sub-folders**



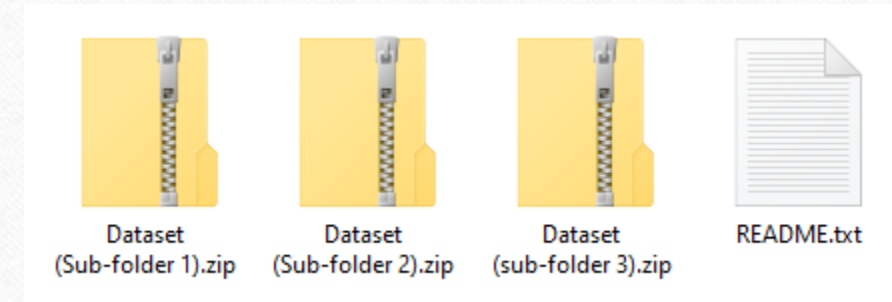
Recommended for large number of data files

3. Prepare your data

- Compress the folder(s) into **.zip*** folder(s)



OR



***Avoid** compressing into **.rar** as they may not be previewed on DataHub

3. Prepare your data

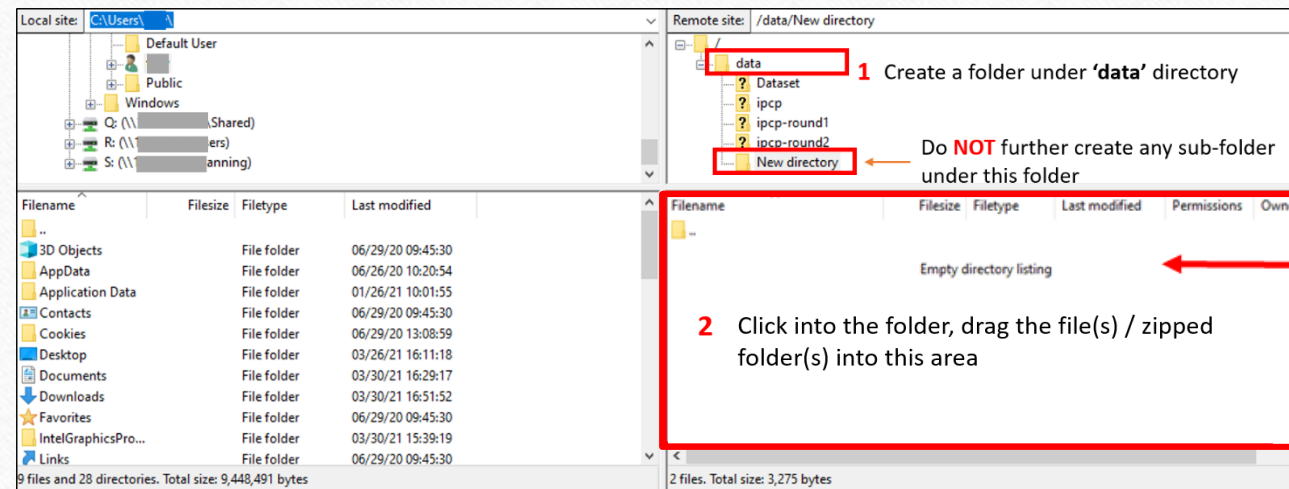
Each RPg student is assigned
100GB private storage on
DataHub by default.

- Now look at the **total size of your data files**
- If a single folder is smaller than 5GB, proceed to DataHub web interface for data upload
- If a single folder is larger than 5GB or a very large total size of your data files, proceed to upload your data via FTP uploader

4. Upload your data

1. Uploading via FTP Uploader (Filezilla)

- Follow this step-by-step guide to set up the connection and upload your files:
<https://libguides.lib.hku.hk/researchdata/datahub/deposit#s-lib-ctab-21880673-1>



4. Upload your data

Please create only **ONE** item record for ALL of your folder(s) AND README file

2. Uploading via DataHub web interface

Login to [DataHub](#)

Drag **ALL** your zip folder(s) and README file here

The screenshot shows the HKU DataHub web interface. At the top, there is a navigation bar with 'Browse', 'Search on HKU Data Repository...', 'Upload', 'My data', and a user profile icon. Below this, there are tabs for 'My data', 'Projects', 'Collections', and 'Activity'. The 'My data' tab is selected and highlighted with a red box. Below the tabs, there is a '+ Create a new item' button, also highlighted with a red box. A table of existing data items is displayed below, with columns for 'Actions', 'STATUS', 'TYPE', 'CREATION ...', and 'SIZE'. The table contains three rows of data items, each with a checkbox in the 'Actions' column.

Actions	STATUS	TYPE	CREATION ...	SIZE
<input type="checkbox"/>		DATASET	29.10.2021 10:04	6.19 GB
<input type="checkbox"/>		PRESENTATION	13.9.2021 16:57	11.44 MB
<input type="checkbox"/>			2.9.2021 14:38	4.74 MB

The screenshot shows the DataHub upload form. At the top, there is a dashed box for file upload, highlighted with a red box, containing the text 'to upload, drag file(s) on the page or browse'. Below this, there is a checkbox for 'Metadata record only' and a 'Link file' button. The form contains several input fields: 'Title' (with 'Untitled Item' as the default), 'Authors' (with a search prompt), 'Categories' (with a dropdown menu), and 'Group' (with a dropdown menu). A 'Tips' section on the right provides instructions on how to use the form.

needed to publish ●

Title
Untitled Item

Authors
Search co-authors by name, full email or ORCID. Hit enter after each.

Categories
Select categories

Group
Research Postgraduates

Tips
Use this form to edit all information related to your data. Please be as descriptive as possible. The file upload is independent from the rest of the form, so you don't need to save an upload. This message will be replaced with helpful tips and suggestions as you begin interacting with the form.

4. Upload your data

3. If all your research data have been uploaded onto external repository, you are not required to upload them again on DataHub
- Login to DataHub > Create an item record without uploading the data files
 - Upload your README file only
 - Provide the URL(s) or DOI link(s) of your dataset under “Related Datasets”

Related Datasets

5. Complete the metadata

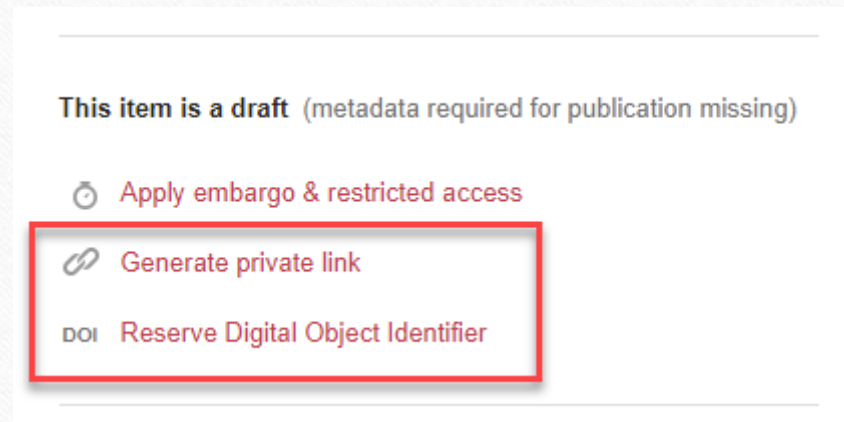
1. For FTP Uploader, return to DataHub and login, open the item record with your uploaded data under “My data” page, fill in the metadata
2. Continue to fill in the metadata

5. Complete the metadata

- Be reminded of the followings:
 - **Title** in standard format: Supporting data for “title of your thesis”
 - **Description** as descriptive as possible (e.g. introduction of your whole dataset), but without revealing any sensitive and confidential contents
 - **Keywords** are subject-related words related to your dataset/thesis
 - **Resource Title & Resource DOI** are optional, enter only if you have publications to be linked with this dataset
 - **License** is set as CC-BY-NC by default

5. Complete the metadata

- Important:
- **MUST** “Reserve a DOI” and “Generate private link” for your item record



5. Complete the metadata


- Important:

- Apply **embargo period** if applicable (refer to step 2)
- Follow the step-by-step guide:

https://libguides.lib.hku.hk/researchdata/datahub/rpg_data_submission#s-lib-ctab-21516215-3

This item is a draft (metadata required for publication missing)

 Apply embargo & restricted access

 Generate private link

DOI Reserve Digital Object Identifier

6. Submit your data record for review

- Submit your dataset record to the Libraries for review and curation
- Tick the “**publish**” box and press “**Publish item**” for submission*



*Your dataset is **not** yet submitted without clicking on the publish button

- The Libraries will contact you via email for confirmation or discuss with you directly if amendments are required

7. Dataset Submission Form

- **Once confirmed by the Libraries**, a link to the Dataset Submission Form will be sent to you
- Fill in the Dataset Submission Form with the DOI reserved for your dataset record
- After submission, you can print out / save an electronic copy of the form as **a proof of your dataset submission**
- Submit the document to your faculty/department

8. Review by your supervisor

- Your dataset submission will be sent to your supervisor for review
- If no comments were received from your supervisor, your dataset record will be released on DataHub

Example RPg Dataset Submissions

- Supporting data for “Genetic and mechanical interactions in the pathogenesis of adolescent idiopathic scoliosis”
<https://doi.org/10.25442/hku.13213547>
- Supporting data for “investigation on topological materials by first-principles calculations” <https://doi.org/10.25442/hku.16685161>
- Host Response to Influenza and Rhinovirus Infection: A Comparative Transcriptomics and Lipidomics Study
<https://doi.org/10.25442/hku.16533255>

DataHub LibGuide & Videos

- DataHub LibGuide (RPg dataset submission)
- https://libguides.lib.hku.hk/researchdata/datahub/rpg_data_submission
- Watch the 2 videos (Part 1 & 2) for the submission procedures

EXTRA

Apart from dataset submission for your thesis, you may upload **other scholarly outputs** (e.g. preprints, data for your publications, conference presentation materials, etc.) onto DataHub

Example

- **Preprint:**

Effect of Height Perception on State Self-Esteem and Cognitive Performance in Virtual Reality
(<https://doi.org/10.25442/hku.16551963.v1>)

- **Data for journal article publication:**

Data for the manuscript "Revealing the Importance of Capillary and Collisional Stresses on Soil Bed Erosion Induced by Debris Flows" (<https://doi.org/10.25442/hku.14254403.v1>)

- **Data for conference paper:**

REMOT: A Hardware-Software Architecture for Attention-Guided Multi-Object Tracking with Dynamic Vision Sensors on FPGAs (<https://doi.org/10.25442/hku.17284643.v2>)

Thesis submission to  turnitin[™]
for similarity check

Understanding Plagiarism

- Definition: The action or practice of **taking someone else's work, idea,** etc., and **passing it off as one's own**; literary theft.



Source: <https://www.oed.com/view/Entry/144939?redirectedFrom=plagiarism&>

Plagiarism – Regulations at HKU

- **Regulations Governing Students' Academic Conduct Concerning Assessment**

<http://www.rss.hku.hk/integrity/rcr/student-regulations>

- **Policy on Research Integrity**

<http://www.rss.hku.hk/integrity/rcr/policy>

- **What is Plagiarism**

<https://tl.hku.hk/plagiarism/>

Other useful resources

- **Types of Plagiarism:**

http://www.turnitin.com/assets/en_us/media/plagiarism_spectrum.php

- **ReadingList@HKUL: Plagiarism:**

https://julac-hku.alma.exlibrisgroup.com/leganto/public/852JULAC_HKU/lists/9907341570003414?auth=SAML

- **Information Literacy Training: Academic Honesty**

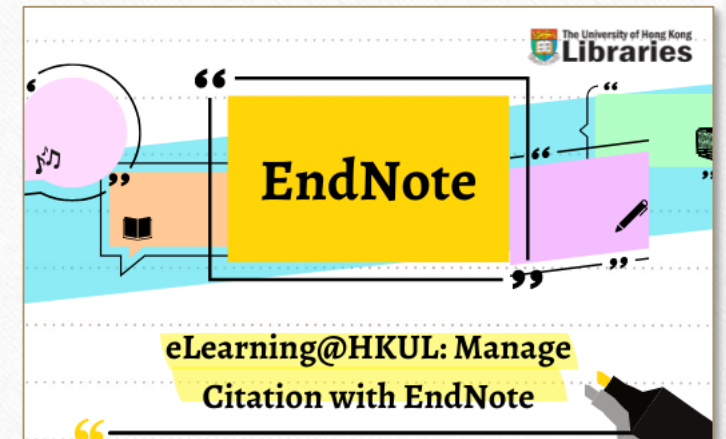
<https://learning.hku.hk/catalog/course/ilt01/>

Assist you in proper citation

- Thesis writing classes from various departments
- Readinglist@HKUL:
Citation Styles:
https://julac-hku.alma.exlibrisgroup.com/leganto/public/852JULAC_HKU/lists/9928326850003414?auth=SAML

Citation Management Tools

- Endnote
<http://lib.hku.hk/endnote/>
- Others:
https://libguides.lib.hku.hk/endnote/other_citation_management_tools



Access the course:

<https://learning.hku.hk/catalog/course/ilt02/>

Turnitin @ HKU



- **University policy**

Compulsory check of Research Postgraduate theses upon formal submission since 2011

https://www.gradsch.hku.hk/gradsch/f/page/394/Workflow_PlagiarismCheck_Nov2016.pdf

- **Departmental / Faculty policy**

Compulsory or random check of theses or student assignments

Global student papers database

USE TO:

Discourage student collusion, the most common form of plagiarism, when comparing against 15 years' worth of student papers.

1.8 student papers
BILLION

1 papers are submitted daily during peak period
MILLION

Submissions from around the world

Internet archive

USE TO:

Compare matches against individual internet sources to easily see potential copy and paste plagiarism.

22 new pages added daily
MILLION

99.3 current and archived web pages
BILLION

102 pages crawled and updated daily
MILLION

Premium scholarly publications

USE TO:

Researchers and students can easily compare original work against published works from around the world.

89.4 subscription articles
MILLION

56 subscription journals
THOUSAND

>1.3 publishers from around the world
THOUSAND

13 open access repositories
THOUSAND

95% journals are included in our database
OF THE TOP 10,000

Turnitin database content

Source: <https://www.turnitin.com/about/content>

Turnitin vs Other Free Plagiarism Check Websites

Turnitin

- **HKU subscription** since 2004
- Turnitin has a well-established privacy policy:
https://help.turnitin.com/Privacy_and_Security/Privacy_and_Security.htm#Privacy_Policy

Free Plagiarism Check Websites

- Free
- No guarantee to what to do with your paper. **It may be resold to other students for profit making.**

Get started – Create a Student Account

Access via www.turnitin.com

1. Instructor (Faculty Office/ Supervisor) added you in a class > receive a welcome email from Turnitin
2. You are given a set of Class ID and Class Enrollment Key

Access via Moodle

Once a Turnitin assignment is created by instructor, you can submit papers to Turnitin via Moodle. No need to create a student account separately.

Get started – Create a Student Account

- Go to <https://www.turnitin.com/>
- Click "Create Account"
- Select "Student"
- Complete the form with Class ID and Class Enrollment Key
- **Use your HKU email address to register**

For students who have submitted any Turnitin assignments via Moodle before, go to **Login > Forgot your password? Click here** to reset your password.

Already have a Student Account...

- Login at <https://www.turnitin.com/>
- You can **enroll to more than one class** in Turnitin with different sets of Class ID and Enrollment Key.

The screenshot shows the Turnitin user interface. At the top, there is a navigation bar with four tabs: 'All Classes', 'Enroll in a Class', 'What is Plagiarism?', and 'Citation Help'. The 'Enroll in a Class' tab is highlighted with a red rectangular box. Below the navigation bar, the text 'NOW VIEWING: HOME' is displayed. Underneath, there is a section titled 'About this page' with the instruction: 'To enroll a class, enter the class ID and enrollment key and click submit. If you do not have a cl'. The main content area features a form titled 'Enroll in a class'. This form contains two input fields: 'Class/section ID: *' and 'Enrollment key: *', both of which are highlighted with a red rectangular box. Below these fields is a 'Submit' button.

Paper Submission


[Class Portfolio](#) [My Grades](#) [Discussion](#) [Calendar](#)

NOW VIEWING: HOME > TESTING 20211108

Class Homepage

This is your class homepage. To submit to an assignment click on the "Submit" button to the right of the assignment name. If the Submit button is grayed out, no submissions can be made to the assignment. If resubmissions are allowed the submit button will read "Resubmit" after you make your first submission to the assignment. To view the paper you have submitted, click the "View" button. Once the assignment's post date has passed, you will also be able to view the feedback left on your paper by clicking the "View" button.

Assignment Inbox: Testing 20211108

Assignment Title	Info	Dates	Similarity	Actions
Assignment 1		Start 08-Nov-2021 10:09AM Due 15-Nov-2021 11:59PM Post 16-Nov-2021 12:00AM		Submit View Download

Assignment Instructions

No additional instructions have been entered for this assignment.

- *Late paper submissions are not allowed for this assignment.*
- *Resubmissions are allowed for this assignment.*

Paper Submission

Submit: Single File Upload ▾

STEP ● ○ ○

Cut & Paste Upload

First name

Rebecca Shuk Ying

Last name

Tam

Submission title

[What can I submit?](#)

Requirements for single file upload

- File must be less than 100 MB ([read suggestions](#) to meet requirements)
If your file contains images, you may remove the images or save the file as a rich text or plain text file to make the file smaller. If your file is a PDF and exceeds the 100 MB restriction: (1) copy the text in the document, (2) change the upload process to 'copy and paste', and (3) paste the text in the body field.
[hide suggestions](#)
- The maximum paper length is 800 pages
- All file types are allowed and can receive marks in GradeMark and PeerMark; however, only Microsoft Word, PowerPoint, WordPerfect, PostScript, PDF, HTML, RTF, OpenOffice (ODT), Hangul (HWP), Google Docs, and plain text files with at least 20 words of text will be eligible to receive Similarity Reports

Choose the file you want to upload to Turnitin:

Choose from this computer



Choose from Dropbox



Choose from Google Drive

Digital Receipt

Congratulations - your submission is complete! This is your digital receipt. You can print a copy of this receipt from within the Document Viewer.

Author:
Rebecca Shuk Ying Tam

Assignment title:
Assignment 1

Submission title:
Assignment 1

File name:
definitions-plagiarism.docx

File size:
15.7K

Page count:
1

Word count:
316

Character count:
1137

Submission date:
08-Nov-2021 10:24AM (UTC+0800)

Submission ID:
1696038445

« Page 1 »

Definitions of plagiarism

According to Oxford English Dictionary (OED), it is the action or practice of taking someone else's work, ideas, etc., and passing it off as one's own. Sometimes, we simply refer this as literary theft.

More definitions include the following:

Plagiarism is defined in dictionaries as the "wrongful appropriation," "close imitation," or "purloining and publication" of another author's "language, thoughts, ideas, or expressions," and the representation of them as one's own original work, but the notion remains problematic with nebulous boundaries. The modern concept of plagiarism as immoral and originality as an ideal emerged in Europe only in the 18th century, particularly with the Romantic movement, while in the previous centuries authors and artists were encouraged to "copy the masters as closely as possible" and avoid "unnecessary invention".

抄襲 - 文藝創作學術領域、新聞學與公共政策、並以學術誠信和道德議題。是對對部原本或基本未經修改的抄襲。這是一種侵犯著作權的行為，但是一些時候是在轉述抄襲上較難以界定，例如像出一個位數的抄襲是否算抄襲就有很大爭議。一般人認為抄襲是剽竊的思想或「偷」的觀念，而不是剽竊的行為，因為法律只保護思想的表現方式，而不是思想本身。

References:
Oxford, O. E. "plagiarism, n." Oxford University Press.
Plagiarism. (2014, October 21). In Wikipedia, The Free Encyclopedia. Retrieved 08-05, October 29, 2014, from <http://en.wikipedia.org/w/index.php?title=Plagiarism&oldid=630419903>

- **Digital receipt** will be generated
(1) online immediately after the submission and
(2) a copy will be sent to your email.
- It provides the **Submission ID** which is unique for Turnitin tracking the paper in the system.

Interpreting the Similarity Report

The similarity index does not reflect whether the paper carries plagiarized materials. It is an indication as to the **word-for-word match** with the Turnitin database.

- **Blue:** No matching text
- **Green:** One word to 24% matching text
- **Yellow:** 25-49% matching text
- **Orange:** 50-74% matching text
- **Red:** 75-100% matching text

Source: <https://help.turnitin.com/originality-check/turnitin-website/student/the-similarity-report/interpreting-the-similarity-report.htm>

Exclude Quotes and Bibliography

By removing the quotations and bibliography will reduce the similarity index temporarily. **It must be handled by the instructor.**

Bibliography and quote exclusion definitions - For papers written in English: <https://help.turnitin.com/feedback-studio/turnitin-website/student/the-similarity-report/bibliography-and-quote-exclusion-definitions.htm>

Filters and Settings [X]

Filters

Exclude Quotes ☒

Exclude Bibliography ☒

Exclude sources that are less than:

☐ words

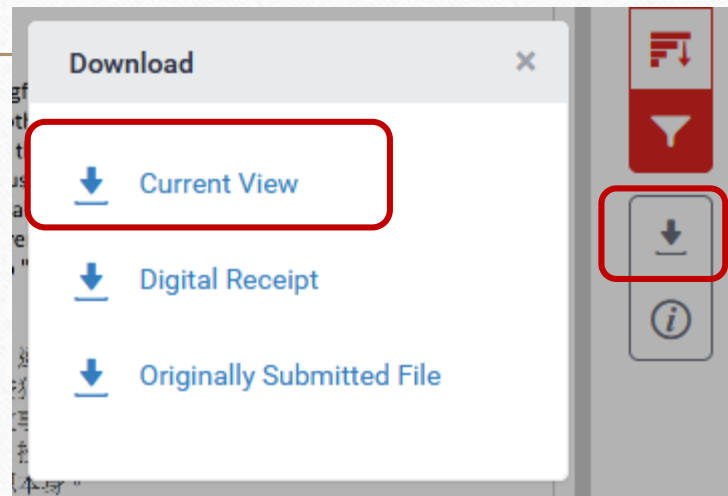
☐ %

☒ Don't exclude by size

Optional Settings

Multi-Color Highlighting ☒

Download Similarity Report / Text-Only Report



Assignment 1

ORIGINALITY REPORT

91%	82%	0%	91%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to University of Hong Kong	91%
	Student Paper	

Text-Only Report

Turnitin Originality Report

Processed on: 08-Nov-2021 10:25 HKT
ID: 1696038445
Word Count: 316
Submitted: 1

Assignment 1 By Rebecca Shuk Ying Tam

Similarity Index	Similarity by Source
91%	Internet Sources: 82% Publications: 0% Student Papers: 91%

include quoted include bibliography exclude small matches mode: quickview (classic) report Change mode print download

91% match (student papers from 26-Apr-2013)
[Submitted to University of Hong Kong on 2013-04-26](#)

Definitions of plagiarism According to Oxford English Dictionary (OED), it is the action or practice of taking someone else's work, idea, etc., and passing it off as one's own. Sometimes, we simply refer this as literary theft. More definitions include the following: Plagiarism is defined in dictionaries as the "wrongful appropriation," "close imitation," or "purloining and publication" of another author's "language, thoughts, ideas, or expressions," and the representation of them as one's own original work, but the notion remains problematic with nebulous boundaries. The modern concept of plagiarism as immoral and originality as an ideal emerged in Europe only in the 18th century, particularly with the Romantic movement, while in the previous centuries authors and artists were encouraged to "copy the masters as closely as possible" and avoid "unnecessary invention". 抄襲，亦稱作是偷竊、剽竊學術研究成果、違反學術誠信「來源標註」，是對於原著未經或基本未經修飾的抄錄，這是一種侵犯著作權的行為，但是一些時候是否構成抄襲比較難以界定，例如模仿一個故事的情節是否屬於抄襲就有很大爭議，一些人認為故事情節屬於思想範圍，而「抄襲思想」並不是犯罪的行為，因為法律只保護具體的表現方式，而不是思想本身。References: Dictionary, O. E. "plagiarism, n.": Oxford University Press. Plagiarism. (2014, October 21). In Wikipedia, The Free Encyclopedia. Retrieved 08:45, October 29, 2014, from <http://en.wikipedia.org/w/index.php?title=Plagiarism&oldid=630449801>

Useful Tips

- You are allowed **three resubmission attempts** where the Similarity Report will **generate immediately**. After three attempts, you'll have to **wait 24 hours** before a new Similarity Report can be generated.
- **Re-submission** will **overwrite** the original file submission in the same assignment box.
- **Never use other student's account** to submit your paper as it will be stored in Turnitin database permanently.
- Students **cannot delete a paper permanently** in the database. Explain this to your instructor who, if convinced, can inform HKU Turnitin Administrator to remove the match from the current report.

Further Assistance...

- Turnitin@HKU:
<https://lib.hku.hk/turnitin/index.html>
- Student Guides on Feedback Studio:
<https://help.turnitin.com/feedback-studio/turnitin-website/student/student-category.htm>

Final Thesis Submission: RPg and TPg theses

Theses@HKU <http://etd.lib.hku.hk/>



Theses @ HKU

HKU Policy

Format of HKU ETDs

How to create an ETD

How to submit HKU Theses

Theses Binding Service

FAQs

Relevant web sites

Theses @ HKU

The Libraries acquire printed and electronic HKU theses for PhD, MPhil and SJD programmes. The printed theses are transferred and stored in The HKU Archives while the ETDs (Electronic Theses & Dissertations) are on open access in the [Scholars Hub](#) under the licensing of the [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](#). For all other theses producing degrees, the Libraries acquire and store electronic copies only, which are also accessible in the Scholars Hub.

For enquiries, please email to libetd@hku.hk or call 3921-2902. Enquiries on thesis binding, please email libadmin@hku.hk or call 3917-2249.

HKU Policy

RPG students

- RPg students entering MPhil and PhD programmes on or after 1 January 2001 are required to submit an **electronic copy** (ETD) of their theses along with the **print copy**.

TPG students

- Since 2013 the Libraries only accession TPg **ETDs** and no longer accession print TPg theses.
- TPg students should check with their **Faculty Office** on the requirement for ETD submission.

Access to HKU theses

Print

- HKU Archives

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

ETDs (Electronic Theses & Dissertations)

- Open access in the Scholars Hub
- Attribution-NonCommercial-NoDerivatives 4.0 International



How to submit HKU theses?

- Visit <https://etd.lib.hku.hk/submit.html>
- Watch the video <https://www.youtube.com/watch?v=me8HOJstKXw>

Theses @ HKU

HKU Policy

Format of HKU ETDs

How to create an ETD

How to submit HKU Theses ▶

Theses Binding Service

FAQs

Relevant web sites

How to submit HKU Theses

<https://etd.lib.hku.hk/submit.html>

Research Postgraduates (RPg)

Starting from 1 January 2017, the [Thesis Submission E-Form: MPhil & PhD](#) will replace the previous print form to facilitate a more efficient and convenient workflow for RPg theses submission.

Before submitting, please check '[Format of HKU ETDs](#)' for proper computer format and number of files to submit, and '[Create an ETD](#)' on instructions on how to convert your thesis to a PDF file.

When you are ready to submit, login to the [E-Form](#) with your HKU Portal account or via a combination of your student ID, telephone number, and HKID. Instructions are available as you go along. For an overview of the submission process, please refer to the [online submission workflow](#).

Taught Postgraduates (TPg)

TPg students should first check with their Faculty Office on the requirement for ETD submission. For submission, they should fill out the [Other HKU Degrees Thesis Submission Form](#) and save their thesis on a CDROM or DVDROM. Please enclose the CDROM or DVDROM in a paper or plastic envelop. Do not apply sticky tapes to the CDROM or DVDROM surface. Please also note that the CDROM or DVDROM will not be returned to the students.

Thesis Submission: MPhil & PhD Thesis

Do you still have access to HKU Portal?

Yes

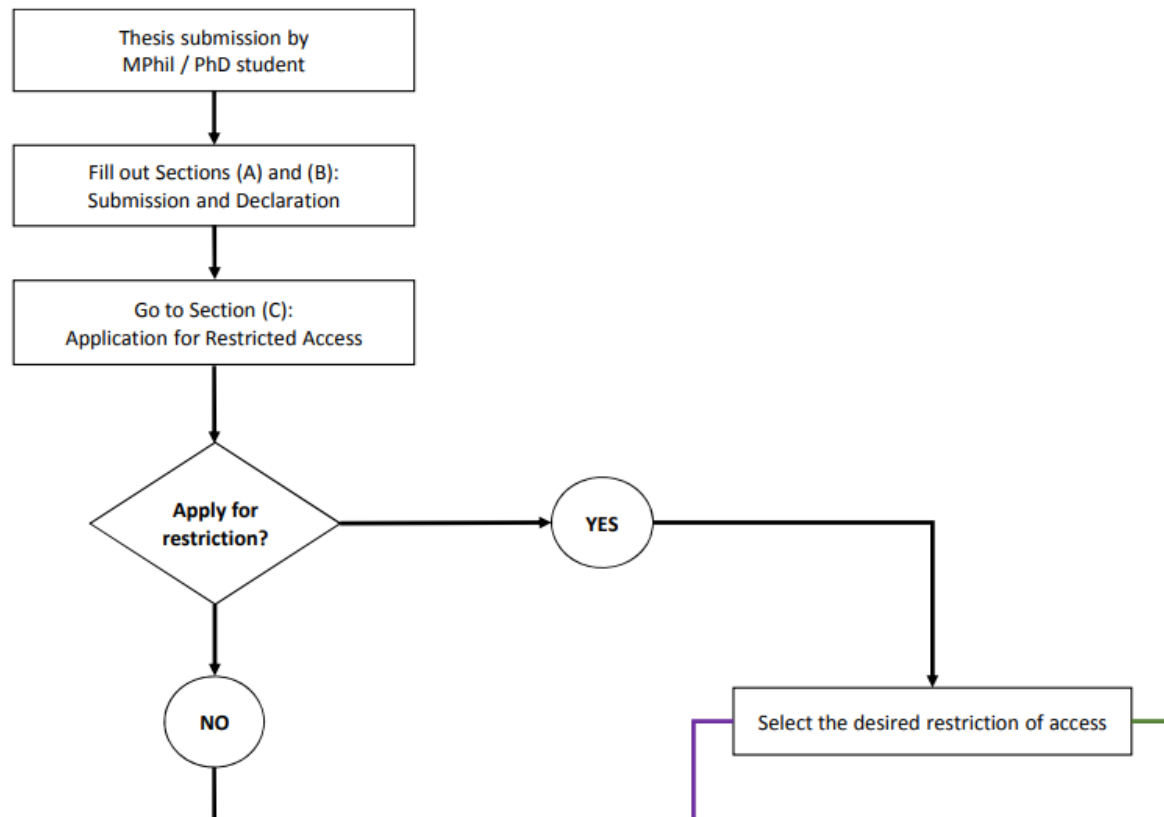
No

- You have active student portal account
- Don't use staff portal account
- If you are in the grace period for accessing portal, it doesn't work

- The system will ask for UID
- Authenticate with phone no and HKID no.

Thesis Submission Workflow

<https://etd.lib.hku.hk/thesis-form/MPhil%20&%20PhD%20Thesis%20Submission%20Workflow.pdf>



Apply for restriction

- RPg students
 - Select the restriction options when doing the online submission
 - Restriction for **up to 5 years** -> approval from **Faculty Board**
 - **Beyond 5 years** -> recommendation from Faculty Board and approval from **Graduate School**
- TPg students
 - Write to your Faculty Board to apply for up to 5 years restriction

Important Notes

- Should the Libraries not receive the properly signed copy of this Thesis Submission & Declaration Form or confirmation of approval for restricted access from the proper authority within **one month of the date of the Congregation** at which his/her degree is conferred, the Libraries will assume that the intent of the author is to make both the print and the electronic copy freely available immediately.

Important Notes

- Submit the **finalized copy** only
- Contents of the submitted **electronic** thesis must be the **same** as in the **printed** thesis.
- Remove all **copyrighted reproductions** from the electronic thesis if without receiving the permission of the copyright holders
- Submit the thesis **supplementary material and data** to the HKU DataHub.

Questions?

Research Data Services

2241-5841

researchdata@hku.hk

Theses Submission

3921-2902

libetd@hku.hk

Turnitin

3917-2202

turnitin@hku.hk

