

NIE Research Data Management Policy Guidelines



Title **NIE Research Data Management
Policy Guidelines**

Date **Apr 2019**

Review & Approval

Version: 1.10

Originator: Senior Officer, Director's Office

Approval: Director, NIE

National Institute of Education
1 Nanyang Walk
Singapore 637616

The Copyright in this work is vested in NIE and the document is issued in confidence for the purpose only for which it is supplied. It must not be reproduced in whole or in part or used for tendering or manufacturing purposes except under agreement or with the consent in writing of NIE and then only on condition that this notice is included in any such reproduction. No information as to the contents or subject matter of this document or any part thereof arising directly or indirectly there from shall be given orally or in writing or communicated in any manner whatsoever to any third party being an individual firm or company or employee thereof without prior consent in writing of NIE.

DOCUMENT HISTORY

Version No.	Issue Date	Review & Approval	Brief description of change (including affected page numbers)
1.0	21 Jan 2015	Reviewed by: Nenny Noorman, Senior Officer Approved by: Prof Paul Teng, Principal Officer	First draft of document
1.1	10 Feb 2015	Reviewed by: Nenny Noorman, Senior Officer Approved by: Prof Paul Teng, Principal Officer	Second draft of document
1.2	2 Apr 2015	Reviewed by: Nenny Noorman, Senior Officer Approved by: Prof Paul Teng, Principal Officer	Third draft of document; changes made to pages 9, 12
1.3	21 Apr 2015	Reviewed by: Nenny Noorman, Senior Officer Approved by: Prof Paul Teng, Principal Officer	Fourth draft of document; changes made to pages 4, 5, 6, 7, 8, 9, 12, 13, 15, 16, 20
1.4	10 Oct 2016	Reviewed by: Kumutha Krisnan, Senior Manager, on behalf of Nenny Noorman, Senior Officer Approved by: Prof Paul Teng, Principal Officer	Fifth draft of document; changes made to pages 4, 5, 6, 7, 8, 24
1.5	10 Jan 2017	Reviewed by: Nenny Noorman, Senior Officer Approved by: Prof Paul Teng, Principal Officer	Sixth draft of document; updated hyperlinks on pages 4, 10,11, 25; amended data security classification table on pages 21, 22

1.6	1 Feb 2018	Reviewed by: Nenny Noorman, Senior Officer Approved by: Prof Low Ee Ling, Chief Planning Officer	Updated NIE Research Integrity Point of Contact information on page 6 Updated the following hyperlinks: - NTU Research Integrity Policy on page 5 - link to definition of research data on page 7 - NTU DMP Template (version adapted by NIE) on page 9 - NTU guidelines regarding research ethics on page 12 - NTU Policy on Research Integrity and the Responsible Conduct of Research on page 26 - US Government Circular A-110 Revised 11/19/93 as further amended 9/30/99 on page 26
1.7	9 Mar 2018	Reviewed by: Nenny Noorman, Senior Officer Approved by: Prof Low Ee Ling, Chief Planning Officer	Updated declaration statement in Appendix A on page 18 to include link to NIE Data Repository.
1.8	7 Sep 2018	Reviewed by: Nenny Noorman, Senior Officer Approved by: Assoc Prof Chang Chew Hung, Chief Planning Officer	Updated the following sections: - Background on pages 5-6 - NTU Offline DMP Template for NIE on pages 8-9 - Upon Expiration of Data on page 13 - Data Retention, Archival and Disposal on page 14 - Appendix A pages 16 and 18 - References
1.9	15 Nov 2018	Reviewed by: Nenny Noorman, Senior Officer Approved by: Assoc Prof Chang Chew Hung, Chief Planning Officer	Updated the following sections: - Background on pages 5-6 - Upon completion of research project on page 10-11 - Data Retention, Archival and Disposal on page 13
1.10	11 Apr 2019	Reviewed by: Nenny Noorman, Senior Officer Approved by: Assoc Prof Chang Chew Hung, Chief Planning Officer	Updated the following sections: - Background on pages 5-6 - Table 1 on pages 16-17 - Research Data Declaration form on pages 18 and 20

Table of Contents

Background	5
1. Before commencement of research project	8
i. Introduction to Data Management Plan	8
ii. Benefits of having a Data Management Plan.....	9
iii. Evaluating data needs.....	9
iv. Data Management Plan Template.....	9
v. Guide and Samples.....	10
2. During conduct of research project	10
i. Documenting data	10
Metadata.....	8
Points for consideration	10
3. Upon completion of research project	12
Storing, backing up and keeping research data safe	12
Networked drives.....	12
Personal computers and laptops	12
External storage devices	12
Remote or online back-up services	13
Data security, protection & confidentiality	13
Data security.....	13
Data protection & confidentiality.....	14
Data Security classification	14
User Access Rights and Permissions	14
4. Upon expiration of the data retention period	14
Data Retention, Archival and Disposal	14
Appendix A	18
Research Data Declaration	18
Appendix B	23
File Naming Convention	23
Appendix C	26
Benefits of sharing data.....	26
Reasons to share	26
Reasons not to share	26
Open data	26
References:	28

Background

NIE is a national teacher training institute in Singapore committed to the vision of being an institute of distinction. Our mission is to be a world-class institute renowned for its excellence in teacher education and educational research.

As an institute of the Nanyang Technological University (NTU), NIE is obligated to adhere to the NTU Policy on Research Integrity and the Responsible Conduct of Research. This policy has been approved by the President's Coordinating Group and the Provost and Deans Group.

The Research Integrity policy and procedures' have been revised and harmonised (February 2014) to bring them in line with procedures under the faculty and staff disciplinary framework and the Whistleblowing Procedures.

All those undertaking research in NTU, at whatever level, whether undergraduate or postgraduate students, technical and research staff or faculty, need to be fully aware of the policy. In particular, attention is drawn to the need to subscribe to the values set out in this policy, including the principles and responsibilities of a researcher in the Singapore Statement on Research Integrity, to abide by all the norms of good research practice and ethical behaviour and with the commitment to conduct research with care, rigour, honesty and respect for all and everything involved, which are the marks of good scholarship. This also covers the obligations to maintain full and accurate records of research and their storage in NTU, both in hard copy and as electronic records; to ensure that all those concerned or quoted are given appropriate credit and recognition; to respect all matters relating to privacy; and to comply fully with all relevant laws and codes of ethical behaviour pertaining to research.

Source: <http://research.ntu.edu.sg/rieo/RI/Pages/NTU-Research-Integrity-Policy.aspx>

In response to NTU's Policy on Research Integrity and the Responsible Conduct of Research, the NIE Research Data Management Policy (RDMP) was formulated in November 2014. The policy was updated in Apr 2019. By adopting this new RDMP (Version 1.5), NIE has committed itself to managing research data to the highest standards throughout the research data life cycle. This has implications for the responsibilities of the institute, Information Services, individual schools and research units. The RDMP covers the areas of data ownership, storage and archival, access rights and data retention and destruction. A portion of the RDMP covers the management of data of non-funded projects.

Also in line with the [NTU Policy for Intellectual Property \(IP\)](#), rights in intellectual property made or created by Staff Members and Students of NIE are owned by the University. These include all research data generated by funded and non-funded research projects conducted at or under the auspices of NIE. Unless stated otherwise by other specific terms of sponsorship, agreements or policy, the NIE RDMP stands as binding to the projects described herein. In the case of Students' IP, if the IP was developed using resources at the University that are ordinarily available to most Students (e.g. use of office, library, computers and storage servers which are ordinarily available to Students), the IP would then belong to the Student. Similarly, the Student shall own the copyright of his or her thesis subject to any commitments under any grant of agreement with external parties. The Student shall, however, grant to the University a royalty-free permission to use, publish, reproduce and/or distribute the thesis worldwide, in whole, in part or in whatever form, electronic or otherwise. As compliance to the NIE RDMP, the Student should

provide the University a copy of the final research data related to the thesis through his/her supervisor for archiving purpose. If there is publication resulted from his/her study, the final research data are to be archived in the NIE Data Repository (<https://researchdata.nie.edu.sg>).

The RDMP will govern the ownership, management and archival of research data for all research projects (funded and/or non-funded) conducted at or under the auspices of NIE.

For completed non-funded projects, which include students' projects, the final research data (including identifiers and consent forms) is only to be archived by the AG upon project completion, regardless of whether or not a publication is intended. Publishing in this case refers to any public dissemination of the results and although it may be interpreted in many ways, any medium in which the results may potentially result in any query or questioning on its integrity, should be viewed as deserving of archival. Hence, for completed non-funded projects, research data is to be archived. For higher degree student theses, the relevant final, anonymised research data* is to be archived by the data custodian (refer to data declaration form) in the NIE Data Repository**. Original hardcopies are preferred but softcopies are also accepted.

*The NIE Data Repository is for archival of softcopy of final research data only. The Library does not provide storage for physical or hard data format.

** Unanonymised data can be stored only if research participants have given explicit consent to include their data with identifiers. Any other data related to the non-funded projects (including higher degree students' work) but are not related to any form of publication (including higher degree students' theses) will be stored with the AG.

It is thus the responsibility of the faculty, staff or supervisor (in the case of higher degree students' work) to ensure that the data are kept throughout the project lifecycle. Further, for higher degree student work, it is the responsibility of the supervisor to ensure that the student understands that the final research data will have to be archived by the AG. For higher degree student theses, it is the responsibility of the data custodian (refer to data declaration form) to ensure that the final research data relevant to the higher degree student theses be kept in the NIE Data Repository for 10 years after the submission of the theses. At present, only Higher Degree students' projects, final year Bachelor degree students' projects and Dean's Commendation Undergraduate Educational Research Paper are included. This may, however, be subject to revision in the future.

This guide specifies the required processes for Principal Investigators to comply with the NIE RDMP. These processes apply to all research projects, regardless of whether the project is a funded or non-funded project. The Guide provides instructions for four phases of the research life cycle:

1. Before commencement of research project
2. During conduct of research project
3. Upon completion of research project
4. Upon expiration of the data retention period

These guidelines are designed to enable the NIE staff to abide by the NTU Research Integrity Policy and apply to data collected with effect from May 2016. Any further enquiries with regards to these guidelines should be forwarded to:

NIE Research Integrity Point of Contact,
Chief Planning Officer, Director's Office
1 Nanyang Walk
Singapore 637616
Email: chewhung.chang@nie.edu.sg

GUIDELINES TO HELP PRINCIPAL INVESTIGATORS (PI) COMPLY WITH THE NIE RDMP

1. Before commencement of research project

- a. Submission of Data Management Plan (also required by NTU IRB for all projects involving human subjects)

- i. Introduction to Data Management Plan

A data management plan is a formal document you develop at the start of your research project which outlines all aspects of your data (i.e., what you will do with your data during and after your research project). In a data management plan, you would systematically record how you intend to collect, process, organise, store and share your research data during and after your research project. This is necessary to fulfil research integrity requirements imposed by the University and external funding agencies.

As part of the NTU Research Data Policy requirements, you would need to complete and submit your data management plan at the start of your research project to show that you have a systematic plan in managing and using your research data.

Please update your data management plan as and when there is any major change or deviation from your earlier submitted data management plan.

What qualifies as research data

'*Research data*' is defined as the recorded factual material commonly accepted in the scientific community as necessary to validate research findings, but not any of the following: preliminary analyses, drafts of scientific papers, plans for future research, peer reviews, or communications with colleagues. This "recorded" material excludes physical objects (e.g., laboratory samples). Research data also do not include:

- (A) Trade secrets, commercial information, materials necessary to be held confidential by a researcher until they are published, or similar information which is protected under law; and
- (B) Personnel and medical information and similar information the disclosure of which would constitute a clearly unwarranted invasion of personal privacy, such as information that could be used to identify a particular person in a research study.'

Source: <https://www.whitehouse.gov/wp-content/uploads/2017/11/Circular-110.pdf>

Some examples of research data

- Documents (text, Word), spreadsheets
- Laboratory notebooks, field notebooks, diaries
- Questionnaires, transcripts, codebooks
- Audiotapes, videotapes
- Photographs, films
- Protein or genetic sequences
- Spectra
- Test responses
- Collection of digital objects acquired and generated during the process of research
- Database contents (video, audio, text, images)
- Models, algorithms, scripts
- Contents of an application (input, output, logfiles for analysis software, simulation software, schemas)
- Methodologies and workflows
- Standard operating procedures and protocols

Source: <http://www.lib.ncsu.edu/guides/datamanagement/define>

ii. Benefits of having a Data Management Plan

Data management is one of the essential areas of responsible conduct of research. Prior to the start of a new research project, PIs and research teams must address issues related to data management.

Some benefits of data management

- Meet funding body grant requirements.
- Ensure research integrity and replication.
- Ensure research data and records are accurate, complete, authentic and reliable.
- Enhance research efficiency.
- Save time and resources in the long run.
- Enhance data security and minimise the risk of data loss.
- Prevent duplication of effort by enabling others to use one's data.
- Comply with practices conducted in industry and commerce.

iii. Evaluating data needs

It is imperative for researchers to evaluate their data needs before the project commences. The following checklist could be used in their evaluation.

- What type of data will be produced? Will it be reproducible? What would happen if it got lost or became unusable later?
- Who will be the audience for your data and how will they use it now, and in the long run?
- Who controls it (Principal Investigator, computing officer in your school)?
- Are there any sharing requirements? For example, funding body's data sharing policy?
- Are there any other funding body requirements? For example, data management plan in proposal?
- How much data will be generated and how often will it change?
- How long should it be retained? For example, 5 years, up to 10 years, or permanently.
- Are there tools or software needed to create/process/visualise the data?
- What directory and file naming convention will be used?
- What file formats will be used? Are they long-lived?
- What project and data identifiers will be assigned?
- Is there good project and data documentation?
- Are there any special privacy or security requirements? For example, personal data, high-security data?
- What will be the storage and backup strategy?
- When and where the data will be published?
- Is there a community standard for data sharing/integration?

Source: <http://www.ed.ac.uk/schools-departments/information-services/research-support/data-library/research-data-mgmt/planning>

iv. Data Management Plan Template

NIE will adopt NTU's DMP tool in RIMS (the NTU Research Information Management System). NIE will use the offline fillable version of the NTU DMP using Microsoft Word and this will be used outside the RIMS environment. The completed DMP document must be submitted following the instructions with the award of the grant.

NIE's DMP template is enclosed below.



NTU_DMP_offline_f
orm_for_NIE_201801

v. Guide and Samples

You are strongly encouraged to refer to the respective **GUIDE and SAMPLES** first before answering each DMP question. Refer to the '[NTU Data Management Plan](#)' section of the [Research Data Management Guide](#) created by NTU Libraries for an online version of the guides, samples and FAQ.

2. During conduct of research project

- a. Ensure all research data are documented properly such as having a data classification system and following a Standard Naming Convention (Refer to Appendix B).

i. Documenting data

Data should be documented to ensure that it can be understood in the long term and usable. It is good practice in research to ensure that all data generated or collected through the course of your research are easy to understand and analyse.

Producing good documentation and metadata ('data about data') provides context for your data, tracks its provenance, and makes it easier to find and use in the long term. Documentation and metadata requirements should be identified from the start of your project and considered throughout the lifecycle of your data. This is the essence of good 'data curation'.

Metadata

Examples of specific metadata are File Name, creator/ author, date of indexing. Metadata are broadly defined into three categories:

Descriptive metadata - common fields such as title, author, abstract, keywords

Administrative metadata - preservation, rights management, and technical metadata about formats.

Structural metadata - how different components of a set of associated data relate to one another, such as tables in a database.

Points for consideration

Where a research protocol is used, much of the documentation needed will already exist. If instrumentation is used, calibration and other details need to be captured for the data to remain useful. Lab notebooks are perhaps the most rigorous form of documentation. Digitisation of lab notebooks would ensure its long term usability.

For qualitative data or small-scale surveys, the documentation should be written down while the data is fresh in your mind. This may include writing methodology reports, creating codebooks with full variable and value labels, documenting decisions about software, tracking changes to different versions of the dataset, recording assumptions made during analysis. PIs should be mindful of the need to protect the identities of participants as required by law.

Tip: A "readme.txt" file to describe the contents of files in a folder can be invaluable at a later date.

Ultimately the amount of effort put into documenting your data will depend on the intended lifespan and how broadly you intend to share it.

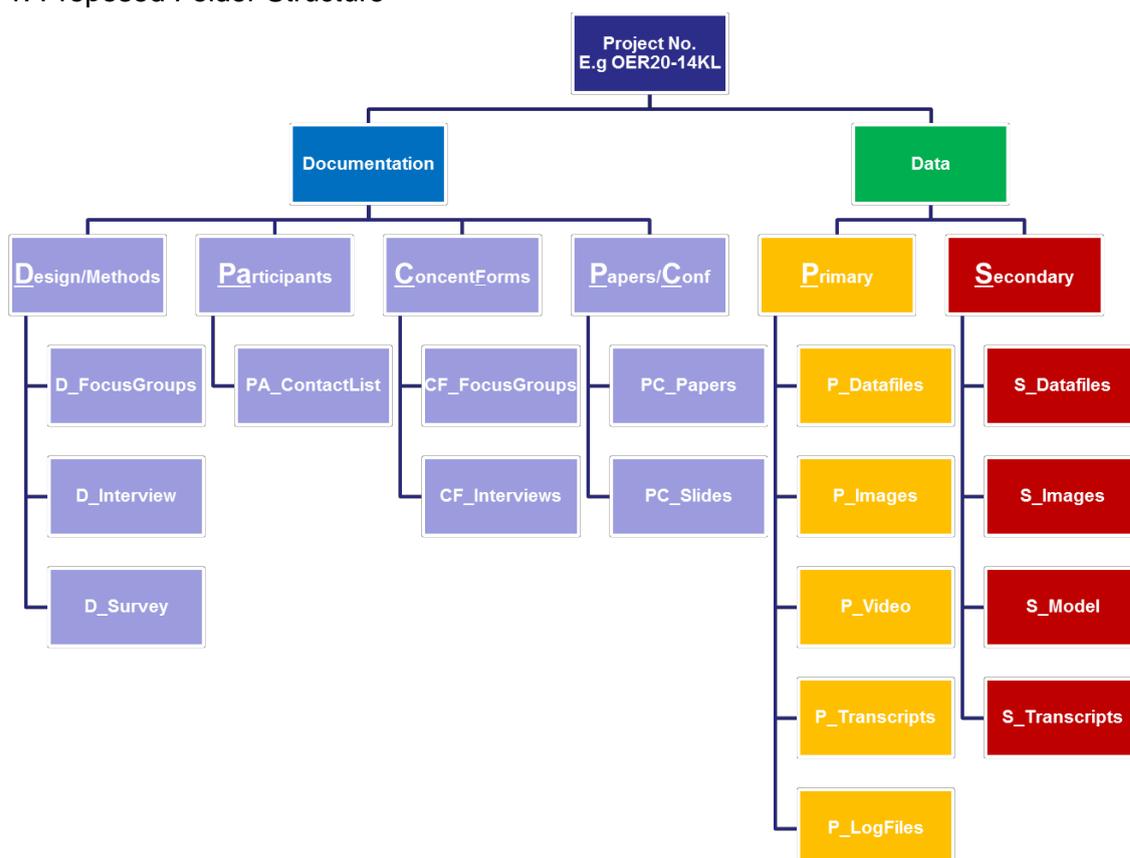
Ideally these decisions should be made at the outset to avoid having to carry out a rescue mission on the data, sometimes known as 'digital archaeology', when a key member of staff leaves, or renewed interest in a topic suddenly puts a dataset in demand.

Data classification

Classification is the process of systematically grouping data according to function or subject matter. Data are classified so that related information is kept in the proper context and can be retrieved easily.

Diagram 1 shows the folder hierarchical structure defined to cater to the researcher's needs. The design of the folder structure is based on input by the OER Knowledge Management (KM) Team and the UK Data Archiving institute. This folder structure should be implemented together with the Standard Naming Convention (refer to Appendix B).

Diagram 1: Proposed Folder Structure



3. Upon completion of research project or when the PI/Supervisor leaves, whichever occurs first

- a. Ensure all research data are properly stored and archived for 10 years from the end of project or 10 years after the publication of a paper, whichever is later.
- b. Submission of Research Data Declaration (Refer to Appendix A) to the research funding administration office (OER or GPL) for funded projects and to the AG administration office for non-funded projects, for filing.
- c. A copy of the data is to be handed over to OER/GPL/OTE/AGs.

Storing, backing up and keeping research data safe

Through the course of your research you must ensure that all research data, regardless of format, are stored securely and backed up or copied regularly. It is strongly recommended to keep at least 3 copies of your data, for example, original, external/local, and external/remote, and have a policy for maintaining regular backups. For photo-sensitive printouts produced in the course of the research, it is up to your discretion to determine the best way to conserve them.

Generally, research data can be saved and backed up on:

Networked drives

These are managed by IT staff centrally or within your Academic Group/ Department. It is highly recommended to store your research data on regularly backed-up networked drives such as:

- File servers managed by your research group or school.
- File servers managed by the Academic Computing & Information Services, such as NIE MySpace (<http://myspace.nie.edu.sg>).

This way you will ensure that your data will be:

- Stored in a single place and backed up regularly.
- Available to you as and when required.
- Stored securely minimising the risk of loss, theft or unauthorised use.

Personal computers and laptops

Local drives may fail or PCs and laptops may be lost or stolen leading to an inevitable loss of your data. These are convenient for storing your data temporarily but should not be used for storing master copies of your data.

External storage devices

External storage devices such as hard drives, USB flash drives, CDs and DVDs are not recommended for the long term storage of your data, particularly your master copies as:

- Their longevity is not guaranteed, especially if they are not stored correctly. For example, CDs degrade and tapes shrink in the long term.
- They can be easily damaged, misplaced or lost.
- Errors with writing to CDs and DVDs are common.
- They may not be big enough for all the research data, so multiple disks or drives may be needed.
- They pose a security risk.

If you choose to use CDs, DVDs and USB flash drives (for example, for working data or extra backup copies), you should:

- Choose high quality products from reputable manufacturers.

- Follow the instructions provided by the manufacturer for care and handling, including environmental conditions and labelling.
- Regularly check the media to make sure that they are not failing, and periodically 'refresh' the data (that is, copy to a new disk or new USB flash drive).
- Ensure that any private or confidential data are password-protected and/or encrypted.

Remote or online back-up services

These provide users with an online system for storing and backing-up computer files e.g. Dropbox, GoogleDrive. Typically, they:

- Allow users to store and synchronise data files online and between computers.
- Employ cloud computing storage facilities.
- Provide the first few gigabytes free and users pay for more facilities including space.

Note: Due to the Personal Data Protection Act, researchers are advised against storing institutional data on a server that resides outside of Singapore.

When considering your backup strategy you need to know:

- Whether all data, or only changed data, will be backed up. A backup of changed data is known as an "incremental backup", while a backup of all data is known as a "full backup".
- How often full and incremental backups will be made.
- How long backups will be stored.
- How much hard-drive space or number of DVDs will be required to maintain this backup schedule.
- If the dataset is sensitive, how will it be secured and (possibly) destroyed.
- What backup services are available that meet these needs and, if none, what will be done about it.
- Who will be responsible for ensuring backups are available.

Data security, protection & confidentiality

What forms of control do I have over access to my data? Do I need to restrict access in any way?

Data security

Securing digital research data is part of the issue of information technology security. You should always have up-to-date anti-virus software installed on your office and home computer.

If you have sensitive data that are covered by privacy laws or confidentiality agreements, it is best to store them on a computer that is not connected to any network. If this is not possible, then you can also consider encrypting your data.

Physical security is also an issue to be considered. A computer that is not connected to a network is still vulnerable to theft and malicious damage/modification to data.

More information can be found in these sites:

- [Acceptable Use Policy for Computing](#) [NIE Portal > Services > IT & Media > IT Security, Policies and Guidelines > ICT Policies > Computing Acceptable Use Policy]
- Encryption, whereby data is transformed into code, is a good way of ensuring its confidentiality and security. You will find the [NIE's recommendation for encryption](#) useful.
- Guide on the disposal of confidential and/or sensitive waste (held on paper, CDs, DVDs, tapes, discs and disposal of records [NIE Portal > Services > IT & Media > IT Security, Policies and Guidelines > Information Security Policy]

Data protection & confidentiality

If your research involves human subjects, you will need to consider both legal and ethical obligations regarding sharing your data. Please refer to [NTU's guidelines regarding research ethics](#).

Data Security classification

Access to research data will be determined using two levels: **data security classification** and **access rights of the user**.

All NIE research data will be categorised using a standardized Security Classification system, comprising five levels based on the criticality of the information to NIE (refer to Appendix C).

These are:

- Level 1: Public
- Level 2: Restricted
- Level 3: Confidential
- Level 4: Highly Confidential
- Level 5: Secret

For highly sensitive information such as personnel data, grant matters, patents, grievance issues, disciplinary matters and Director's correspondences, a new category of "Highly Confidential" has been created to facilitate secure management of such records. To aid users, a list of information categories that would qualify for the "Highly Confidential" classification will be provided for reference. All data classified under "Highly Confidential" will be blocked from general search requests. Access to such data by individuals other than the data creator/owner will be based on specific permissions.

Staff will be required to identify the security classification of a data at the point of creation. To provide further reference to users, the classification label should be included on each page of the document (e.g. 'Confidential'). Strong justification will have to be given for classifying information types as Highly Confidential which are not on the reference list.

User Access Rights and Permissions

PIs should ensure that the research team members keep the research data confidential and store it securely. Appropriate user access rights and permissions can be accorded based on the data security classification designated to their research data.

4. Upon expiration of the data retention period

The PI or the data custodian will inform LIBRIS, the research funding administration office or AG administration office which is archiving their data that the data is due to be destroyed.

The task to destroy is only undertaken by the PI/data custodian upon approval from the office. PIs may wish to consider the option of sharing their data (refer to Appendix D).

Data Retention, Archival and Disposal

In accordance to the NTU Research Integrity Policy, the PI is responsible for how the data is to be managed and subsequently archived, according to that which is stated within the Research Data Declaration form.

The final research data must be archived for a minimum retention period of 10 years with effect from the completion date of the study or 10 years after the publication of a paper, whichever is later. The Research Data Declaration will be filed and kept by the research funding administration office (OER or GPL). For non-funded projects, the Research Data Declaration will be kept by the AG administration office. Similarly, a copy of the data is to be handed over to OER/GPL/OTE/AGs upon project completion or when the PI/Supervisor leaves, whichever occurs first.

Upon expiration of the data retention period, the PI or the data custodian will inform LIBRIS, the research funding administration office or AG administration office that the data is due to be destroyed. The task to destroy is only undertaken by the PI/data custodian upon approval from the office.

A data retention schedule would guide PIs on the decision on data management before, during and after the research is completed. A data retention schedule (Refer to Table 2 as a guide) ensures that:

- Data are not retained any longer than mandated regulatory frameworks
- Data truly contributing to institutional memory are permanently retained
- Data fulfilling their retention period are destroyed/disposed in an effective manner

Table 1: Data Retention and archival

Type of Data	Recommended Data Archive format
<p>Quantitative tabular data with extensive metadata a dataset with variable labels, code labels, and defined missing values, in addition to the matrix of data</p>	<p>proprietary formats of statistical packages e.g. SPSS (.sav), Stata (.dta) MS Access (.mdb/.accdb)</p>
<p>Quantitative tabular data with minimal metadata a matrix of data with or without column headings or variable names, but no other metadata or labelling</p>	<p>comma-separated values (CSV) file (.csv) tab-delimited file (.tab) delimited text of given character set - only characters not present in the data should be used as delimiters (.txt) widely-used formats, e.g. MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb), dBase (.dbf) and OpenDocument Spreadsheet (.ods)</p>
<p>Qualitative data Textual</p>	<p>Rich Text Format (.rtf) plain text data, ASCII (.txt) Hypertext Mark-up Language (HTML) (.html) widely-used proprietary formats, e.g. MS Word (.doc/.docx) some proprietary/software-specific formats, e.g. NUD*IST, NVivo and ATLAS.ti</p>
<p>Digital image data</p>	<p>JPEG (.jpeg, .jpg) but only if original created in this format TIFF (other versions) (.tif, .tiff) Adobe Portable Document Format (PDF/A, PDF) (.pdf) standard applicable RAW image format (.raw) Photoshop files (.psd) BMP (.bmp) PNG (.png)</p>
<p>Digital audio data</p>	<p>Free Lossless Audio Codec (FLAC) (.flac) MPEG-1 Audio Layer 3 (.mp3) but only if created in this format Audio Interchange File Format (AIFF) (.aif) Waveform Audio Format (WAV) (.wav)</p>
<p>Digital video data</p>	<p>MPEG-4 (.mp4) OGG video (.ogv, .ogg) Motion JPEG 2000 (.mj2)</p>

Documentation and scripts	Rich Text Format (.rtf) PDF/A or PDF (.pdf) HTML (.htm) OpenDocument Text (.odt) plain text (.txt) some widely-used proprietary formats, e.g. MS Word (.doc/.docx) or MS Excel (.xls/.xlsx) XML marked-up text (.xml) according to an appropriate DTD or schema, e.g. XHMTL 1.0
Geospatial data Vector and raster data	ESRI Shapefile (.shp, .shx, .dbf, .prj, .sbx, .sbn optional) ESRI Geodatabase format (.mdb) Geo-referenced TIFF (.tif, .tfw) CAD data (.dwg, .dxf or .svg) Tabular GIS attribute data Binary formats of GIS and CAD packages Geography Markup Language (.gml) MapInfo Interchange Format (.mif) for vector data Keyhole Mark-up Language (.kml) Adobe Illustrator (.ai)

Reference from: <http://www.data-archive.ac.uk/create-manage/format/formats-table>

Research Data Declaration

With reference to NTU Research Integrity Policy, as a member of the academic staff, the Principal Investigator (PI) or researcher is obliged to maintain full and accurate records of the research and their storage. The researcher is obliged to ensure that the data will be stored and maintained at NTU for a minimum period of 10 years post publication. This form is to be completed by the staff for each project (funded or unfunded), and to be submitted to the research administration office (OER, GPL, OTE or AG).

Project Reference Number	
Name of Principal Investigator (PI) / Researcher	
Department (AG/RC/PO)	
Date of Project Completion	

RESEARCH PRIMARY DATA (raw unprocessed data)

For each of the data types, please select "A", "B" or "C" where

A: Data exists and will be archived B: Data exists but will not be archived C: Not Applicable

S/N	Data Type	A	B	C	Remarks (e.g. indicate reason for not archiving etc.)
1	Data files	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
2	Images (photos, images)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
3	Video files	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
4	Audio files	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
5	Transcripts (raw)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
6	Fieldnotes (Observations)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
7	Artefacts	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
8	Log files (for web based)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
9	Others, please specify	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

RESEARCH SECONDARY DATA (processed, analysed or anonymised data)

For each of the data types, please select “A”, “B” or “C” where

A: Data exists and will be archived B: Data exists but will not be archived C: Not Applicable

S/N	Data Type	A	B	C	Remarks (e.g. indicate reason for not archiving etc.)
1	Data files (SPSS, HLM, Mplus, SAS, others, please specify: _____)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
2	Images (photos, images)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
3	Models	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
4	Transcripts	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
5	Analysis of the instruments outcomes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
6	Documentation of project best practices (if any)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
7	Others, please specify	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

RESEARCH RELATED DOCUMENTS

For each of the data types, please select “A”, “B” or “C” where

A: Data exists and will be archived B: Data exists but will not be archived C: Not Applicable

S/N	Data Type (* mandatory if data is collected)	A	B	C	Remarks (e.g., indicate reason for not archiving etc.)
1	Consent forms (FGD, interview, survey)*	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
2	Participants contacts list	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
3	Ethics clearance letter (e.g. from NTU-IRB)*	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
4	MOE (or other institution) approval letter for data collection	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
5	Design documents (FGD, interview, survey, questionnaires, modelling, programming)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
6	Coding Schemes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
7	Instruments (survey forms, test forms)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
8	Programs (compiled exe, source codes)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
9	List of schools, name of liaison teacher and principal involved	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

10	Acknowledgement of vouchers by participants	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
11	Others, please specify	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

I hereby acknowledge and confirm that I have secured, stored, and archived the project's research data (primary and secondary), as indicated in this form, in accordance to the guidelines stipulated by the NIE Research Data Management policy.

I understand that I am the custodian of the research data so long as I remain a staff of NIE. As the research data is an intellectual property owned by the institute, it is the responsibility of the data custodian to ensure that the data is made available upon request from the institute.

I understand that I shall also upload the final set of publication-related research data to the NIE Data Repository (<https://researchdata.nie.edu.sg>).

I agree to submit a copy of the research data related to the project (including rights to access via the online archiving system, if any) should I decide to terminate my employment with the institute. In such case, I shall first hand over the research data to the project co-PI, if the subject in question is still a staff of the institute. The relevant research administration office (OER, GPL, OTE or AG) will be duly informed upon completion of the handover. If the project does not have an existing co-PI, the research data related to the project will be handed over to the relevant research administration office (OER, GPL, OTE or AG).

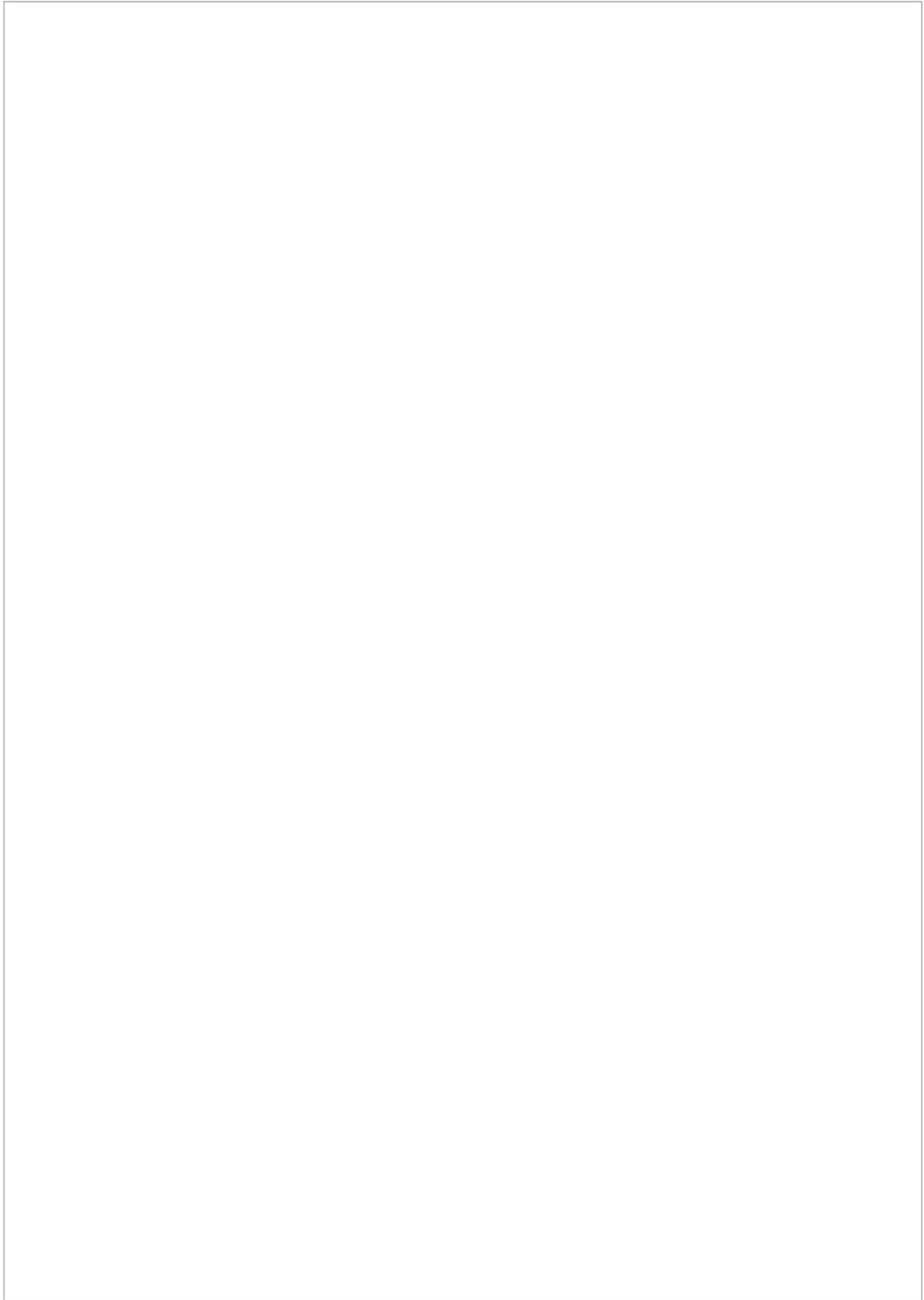
I understand that I may be contacted in the future should there be any query pertaining to the project's research data, even after having left the institute.

SIGN OFF

Signature _____
Principal Investigator / Researcher

Date _____

Screen Capture of the Folder Directory



Screen Capture of the Folder Directory



File Naming Convention

The identity of a file is important for document retrieval, and as such, it is imperative that there is some form of best practice that the researcher can follow to standardise their file name.

The proposed file naming convention and standards are based on the UK Data Archiving Standard, "<http://www.data-archive.ac.uk/create-manage/format/organising-data>".

The Proposed File Naming Convention should contain:

- a. Department Abbreviation (3 to 4 characters), department abbreviation can be omitted if it is part of the project number
- b. Project Number (10 to 15 characters)
- c. Description of the file (20 to 30 characters), please use '-' as space for the description
- d. Date, YYYYMMDD (8 characters)
- e. Version Number if there are 2 or more versions, e.g., v02 (3 characters)
- f. Only use _ to join the above 5 elements
- g. Do not use spaces and special characters
- h. File name should not contain more than 124 characters

Here is an example on how a researcher can name a video recording on classroom observation at Yishun Sec School for an English class. The video recording is done on Sec 1E2 on 26 Aug 2014. The project doing this is OER 20/14KL.

Below is how the filename can be written:

OER20-14KL_Observation-ELL-YSS-S1E2_20140826.mp4

Data Security Classification in NIE

Classification	Description	Impact if lost	Examples	Who can read the information
Public	<ul style="list-style-type: none"> • Information for all types of audiences • Published on NIE website, publications, & other public domains 	<ul style="list-style-type: none"> • No implications 	<ul style="list-style-type: none"> • Ethics clearance letter • MOE (or other institution) approval letter for data collection 	<ul style="list-style-type: none"> • Anyone
Restricted	<ul style="list-style-type: none"> • Information only for NIE staff and predetermined / designated outsiders 	<ul style="list-style-type: none"> • Some inconvenience • Minimal disruption to operations • Modifications can be tracked back 	<ul style="list-style-type: none"> • General correspondences emails • Design documents • Coding schemes • Instruments • Programs • Consent forms • Participants contacts list • List of schools, names of liaison teachers and principals involved 	<ul style="list-style-type: none"> • All NIE staff • Selected external parties given access rights /permissions
Confidential	<ul style="list-style-type: none"> • Information strictly for NIE staff with access rights • Senior management issues • Highly sensitive research information 	<ul style="list-style-type: none"> • Loss of credibility and integrity • Legal ramifications 	<ul style="list-style-type: none"> • Data files • Images • Models • Transcripts • Analysis of instruments outcomes • Personal data • Financial accounts 	<ul style="list-style-type: none"> • All PIs, Research Associates • Selected Research Assistants given access rights /permissions • Director, NIE • Designated research staff given access rights / permissions

Secret

- Only to be used in line with Government's 'Secret' classified documents

- Breach of government security provisions

- papers
- Parliamentary correspondence
- Diplomatic correspondences

- Correspondences with Minister

- Designated research staff from project originating documents
- Designated appointment holders and other officers assigned access rights or permissions

- Only Director, NIE and assigned staff with security clearance

Benefits of sharing data

Why should I consider sharing research data that I collect, create or collate in the course of my research? Are there legitimate reasons not to share?

Reasons to share

- Scientific integrity - publishing your data and citing its location in published research papers can allow others to replicate, validate, or correct your results, thereby improving the scientific record.
- Funding mandates – Organisations which are granting research grants are increasingly mandating data sharing so as to avoid duplication of effort and save costs.
- Increase the impact of your research - those who make use of your data and cite it in their own research will help to increase your impact within your field and beyond it. Users of your data may include those in other disciplines, sectors, and countries.
- Preserve your data for your own future use - by preparing your data for sharing with others, you will benefit by being able to identify, retrieve, and understand the data yourself after you have lost familiarity with it, perhaps several years hence.
- Teaching purposes - your data may be ideal for students to learn how to collect and analyse similar types of data themselves.
- Making publicly funded research available publicly - there is a growing movement for making publicly funded research available to the public

Reasons not to share

- If your data has financial value or is the basis for potentially valuable patents that could be exploited by the University, it may be unwise to share it, even with a data licence or terms and conditions attached.
- If the data contains sensitive, personal information about human subjects, it may violate the Data Protection Act, ethics codes, or your own written consent forms to share it, even with other researchers. Often there are ways to anonymise the data to remove the personally identifying information from it, thus making it sharable as a public use dataset.
- If parts of the data are owned by others, such as commercial entities or authors, then even if you have derived wholly new data from the original sources you may not have the rights to share the data with others.

By writing a data management plan near the beginning of your research project, you can work through these issues and determine if you will be able to produce a version of your data that can be shared with others.

Open data

Some PIs may wish to classify their research data as open data. More information can be found at <http://opendatahandbook.org>.

Source: <http://www.ed.ac.uk/schools-departments/information-services/research-support/data-library/research-data-mgmt/benefits-sharing-data>

Other non-NTU Research Data Repositories

- **Subject Specific** data repositories

- Biological and Life Sciences

- Protein DataBank

- Chemistry

- Cambridge Structural Database

- Social Sciences

- Open ICPSR (Inter-University Consortium for Political and Social Research)

- **General** data repositories

- Figshare

- Dryad

- Visit **Data Repository Registries** to find more repositories

- Re3data.org

References:

- NIE Research Data Management Policy
- NTU policy on research integrity and the responsible conduct of research (<http://research.ntu.edu.sg/rieo/RI/Pages/NTU-Research-Integrity-Policy.aspx>)
- NTU Policy on Intellectual Property ([http://www.ntu.edu.sg/docs/Revised%20NTU%20IP%20Policy%20\(effective%20Apr%202014\).pdf](http://www.ntu.edu.sg/docs/Revised%20NTU%20IP%20Policy%20(effective%20Apr%202014).pdf)) – login may be required
- US Government Circular A-110 Revised 11/19/93 as further amended 9/30/99 (<https://www.whitehouse.gov/wp-content/uploads/2017/11/Circular-110.pdf>)
- North Carolina University Library - Defining research data: data management (<http://www.lib.ncsu.edu/guides/datamanagement/define>)
- UK Data Archive – File formats table (<http://www.data-archive.ac.uk/create-manage/format/formats-table>)
- UK Data Archive – Organising data (<http://www.data-archive.ac.uk/create-manage/format/organising-data>)
- North Carolina University Library - Examples of Data Management Plan (http://www.lib.ncsu.edu/guides/datamanagement/dmp_examples)
- NIE Records Management Policy Framework (NIE Portal - Services > Records Management > Records Management Policy)
- The University of Edinburgh Research Data Management Guidance (http://www.ed.ac.uk/files/atoms/files/uoerdm-roadmap_-_v2_0_0.pdf)
- The University of Edinburgh Research Data Management Planning (<http://www.ed.ac.uk/information-services/research-support/research-data-service/planning-your-data>)
- The University of Edinburgh Research Data Management Guidance: Benefits of Sharing Data (<http://www.ed.ac.uk/schools-departments/information-services/research-support/data-library/research-data-mgmt/benefits-sharing-data>)