SHARING YOUR RESEARCH DATA VIA SCHOLARBANK@NUS
MEET OUR TEAM

Gerrie Kow
Head, Scholarly Communication
NUS Libraries
✉ gerrie@nus.edu.sg

Estella Ye
Research Data Management Librarian
NUS Libraries
✉ estella.ye@nus.edu.sg
What we do

1. PLAN & CONDUCT
Learn about literature review, databases, academic writing via workshops and personalised advisories

2. PUBLISH RESEARCH
Find out more about publishing in impactful journals and with scholarly presses through talks and workshops

3. ENHANCE VISIBILITY
Enhance your research visibility through the use of ORCID IDs, the adoption of green open access and more

4. ARCHIVE & PRESERVE
Preserve publications and research data in our ScholarBank@NUS to facilitate new knowledge creation and fulfil funders’ mandates

5. MEASURE IMPACT
Measure and benchmark the quality and impact of your research output using citation metrics using tools like Scopus, SciVal, Web of Science
WHAT WILL BE COVERED

• Why Share Research Data
• What is ScholarBank@NUS
  o Why deposit research data into ScholarBank@NUS
  o Research data content policy
• Submitting Research Data into ScholarBank@NUS
  o Before Your Submission
  o How to Submit
  o Persistent Identifiers & Data Citation
WHY SHARE RESEARCH DATA

FUNDERS’ MANDATE

8. DATA SHARING POLICY

8.1. Subject to restrictions related to research ethics, confidentiality and intellectual property, all data generated from research funded by the Thematic Grant should be made available to user communities at the earliest feasible opportunity. This would generally be no later than the release through publication of the study’s main findings, or in line with established best practices in the respective fields.

8.2. All peer-reviewed publications arising from research funded by the Thematic Grant must be made publicly available no later than 12 months after the official date of publication. A copy of the publication may be deposited in the Host Institution’s Open Access repository or any other subject Open Access repository, in accordance with the Host Institution’s Open Access policy.
Data Sharing

The implementation of the Framework of Research Data Sharing and Governance will be in two phases. The first phase, which will commence in May 2015, will encompass the open access of publications for all existing and new projects. **The sharing of datasets will be implemented in Phase 2 which is targeted to be in May 2016.**
WHY SHARE RESEARCH DATA

PUBLISHERS’ REQUIREMENTS

- Science
- AAAS
- Data Deposition Policy

- Nature
- Springer
- Data Sharing and Citation

- Wiley
- Data Sharing and Citation

Research Data Policy Types
WHY SHARE RESEARCH DATA
GOOD FOR SCIENCE, GOOD FOR YOU

- Provides higher ROI for public investment
- Encourages scientific enquiry and debate
- Facilitates innovations and technological advancement
- Provides you credit - research data as a research output in its own right
- Leads to new collaborations between data users and data creators
- Increases the impact and visibility of research

Source: UK Data Archive
Clinical and imaging data collected in the Alzheimer's Disease Neuroimaging Initiative (ADNI) were made available immediately for scientists to download and analyze.

This type of open clinical data sharing policy was unprecedented at that time – “exceptional.”

2010 to date – 771 papers have been published.

“We make data transparent and available so that other people can analyze the data and discover different things, this is going to accelerate all of science and it's a relatively inexpensive way to get more value out of all of the work that we do.”

Michael Weiner, principal investigator of ADNI

This is now a possible model for looking at Parkinson's disease.
CASE STUDY:
The English Lexicon Project by Washington University

- 77,976 Unique visits to the Dataset
- 893 Citation Counts of Related Article
- 13.05 Field-Weighted Citation Impact of Related Article

Date: Sept 22, 2017

Scan Here to View the Dataset
The English Lexicon Project publication

**Scopus Metrics**

- **Citation Count**: 893
- **Field-Weighted Citation Impact**: 13.05
- **Citation Benchmarking**: 99th percentile

**Cited by**

- **893 Citations**

**Benchmarking**

- Measures of activity relative to specific research domains, based on cited by in Scopus
- Compared to Psychology articles of the same age and document type

**Date range**:
- **2013** to **2017**
- Exclude self citations
- Exclude citations from books

**Update**
SCOPUS: The British Lexicon Project publication


Scopus Metrics

Citation Count
84
Cited by in Scopus

Field-Weighted Citation Impact
7.52

Citation Benchmarking
99th percentile
Compared to Experimental and Cognitive Psychology articles of same age and document type

Cited by

84 Citations
Date range: 2013 to 2017

Benchmarking
Measures of activity relative to specific research domains, based on cited by in Scopus
Compared to Experimental and Cognitive Psychology articles of the same age and document type

All Citations 99th percentile
The Open Data Citation Advantage

- 85 studies, 69% increase
- Gene expression microarray data
- 430 studies, 10 years, doubling of citations
- Astrophysics & astronomy
- Citation advantage: 20-50%
- Political science

Read more here
WHAT IS SCHOLARBANK@NUS

ScholarBank@NUS is the Institutional Repository (IR) of NUS

OUR GOALS

Collect and showcase the research output of NUS researchers in order to

• Support NUS researchers in increasing their research visibility

• Demonstrate the research excellence of NUS to the world

Starting from Oct 2017, ScholarBank@NUS accepts research data produced by NUS researchers.
WHY SCHOLARBANK@NUS?

Harvard Dataverse

DRYAD

figshare

Mendeley DATA

GitHub

re3data.org
WHY SCHOLARBANK@NUS?

NUS’ Research Story

- Free forever
- Longevity
- Personalized service
SCHOLARBANK@NUS
RESEARCH DATA CONTENT POLICY

• Who can contribute?
  o NUS staff with assigned staff numbers and NUSNET IDs
  o NUS departments (including NUS affiliated Research Institutes/Centres)

• What type of datasets do we accept?
  o Research data that are tied to publications
  o Research data may be of use to researchers and the public
  o The submission will be reviewed on a case-by-case basis
SUBMITTING YOUR RESEARCH DATA INTO SCHOLARBANK@NUS

1. Start A New Submission
2. Describe the Dataset
3. Apply Data License
4. Upload Data Files
5. Describe Data Files
6. Define Access Setting
7. Verify Submission Details
Describe Your Dataset

**Metadata** is defined as "structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource".

**Dublin Core Metadata Schema**

**Descriptive**
- Title
- Contributor
- Subject
- Description
- Identifiers (e.g. DOI, URI)
- Type

**Structural**
- Table of contents
- Related publications

**Preservation**
- File formats
- Provenance

**Administrative**
- Department
- License information
- Access control
- Versions
- WBS code
- Grant ID

*Source: Understanding Metadata, National Information Standards Organization United States, 2004*
Why you should give your research output a DOI

What is a DOI?
A DOI is a unique alphanumerical string assigned to identify content and provide a persistent link to its location on the Internet. It can be assigned to intellectual property, such as:

- Datasets
- AV materials
- Event
- Image
- Software
- Sound
- Text
- Website
- Workflow

Benefits
- Persistence
  - Accessible for long-term use
- High level of confidence in the quality and accuracy of DOIs: supported by the International DOI Federation (IDF) and Registration Agencies infrastructure
  - Unambiguous data citation
  - DOIs enable citation tracking
- Accessibility
  - DOI resolves to an online metadata record which provide easy access

Source: Australian National Data Service; American Psychological Association
**Title:** Dataset testing 20170918

**Authors:** YECH SAW ALBRENDA, Chan, Alexandre

**Subject:** Testing

**DOI:** 10.5072/85VH-QH21

**Related Publications:** 10835/106686

**Citation:** YECH SAW ALBRENDA, Chan, Alexandre (2017-09-18). Dataset testing 20170918. 6. ScholarBank@NUS Repository. [Dataset]. https://doi.org/10.5072/85VH-QH21

**License:** Attribution-NonCommercial 4.0 International

http://creativecommons.org/licenses/by-nc/4.0/

**Appears in Collections:** Staff Dataset
Mapping Early European Writings on Southeast Asia

Villamar Cuauhtemoc Tonatiuh, Ho Richard Cheng Yong & Feng Yikang

Website published 2017 via National University of Singapore Libraries

This is an interactive visualisation of the geographic and temporal aspects of about 57 early writings on Southeast Asia, a small proof of concept and voluntary effort initiated by a historian and two librarians, in hope of stimulating conversations in the digital humanities within the local historian community. It enables exploration of where early texts were published, their subsequent reprints or translations over time, and the relationships between the influences on the work which may...

https://doi.org/10.25541/v5af-1bb
<description xml:lang="en-us" descriptionType="Abstract">This is an interactive visualisation of the geographic and temporal aspects of about 57 early writings on Southeast Asia, a small proof of concept and voluntary effort initiated by a historian and two librarians, in hope of stimulating conversations in the digital humanities within the local historian community. It enables exploration of where early texts were published, their subsequent reprints or translations over time, and the relationships between the influences on the work which may comprise of writers, editors, publishers, sponsors and authorities. The visualisation is based on information transcribed from the books, which were captured as an event-based graph model that takes reference from CIDOC CRM. The interactive visualisation augments the exploration of system of influences behind early European writings on Southeast Asia. Details of the bricolage, historical context and implementation will be presented in a workshop in HistoInformatics2017 - the 4th International Workshop on Computational History, in conjunction with the 26th ACM International Conference on Information and Knowledge Management (CIKM 2017), Singapore. </description>
Mapping Early European Writings on Southeast Asia
https://doi.org/10.2504/kin-18-1.4
This essay is an invitation to read and also to see the Southeast Asia region through a selection of 57 books that were written in Europe in the sixteenth century.

Under Western Eyes: European Writings on South East Asia
https://www.soas.ac.uk/courseunits/155607001.html
The module will encompass writings by Europeans on South East Asia from the early colonial period to the present day. Individual topic are likely to change as ...

Routledge Handbook of Southeast Asian History - Google Books Result
https://books.google.com/sg/books?id=151764375X
Norman G. Owen - 2014 - History
These opinions became more prevalent in European writings with the widespread colonization of Southeast Asia from the mid-nineteenth to early twentieth ...

This content downloaded from 66.249.79.158 on Sat, 23 Sep 2017 20 ...
https://www.jstor.org/stable/25441892
Whatever the biological and cultural origin, Southeast Asian history is independent ... history. Hall's historiography, especially for the pre-European period, is still mainly ... For example, in writing of Anglo-Dutch seventeenth-century rivalry, Hall.
Apply Open Data License

**Default License:** Creative Commons Attribution-NonCommercial 4.0 International ([CC BY-NC 4.0](https://creativecommons.org/licenses/by-nc/4.0/))
This grants any user the permission to remix, tweak, and build upon your work non-commercially.

If your research is funded by NUS, the University assigns automatic rights to NUS Libraries to publish the research data under CC BY-NC 4.0.

<table>
<thead>
<tr>
<th>License Name</th>
<th>Commercial Use of Work</th>
<th>Modification of Work</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attribution 4.0 International (CC BY)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Attribution-NonCommercial 4.0 International (CC BY-NC)</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Attribution-NoDerivatives 4.0 International (CC-BY-ND)</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Attribution-Non-Commercial-NoDerivatives 4.0 International (CC BY-NC-ND)</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Attribution-ShareAlike 4.0 International (CC BY-SA)</td>
<td>Yes</td>
<td>ShareAlike</td>
</tr>
<tr>
<td>Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA)</td>
<td>No</td>
<td>ShareAlike</td>
</tr>
</tbody>
</table>
Upload Data Files

1. **Group files into a dataset**
   - What to share
   - Up to **1GB** (individual file) for manual submission

2. **Organise your data**
   - Consistent file naming
   - Convert to recommended formats
   - Deal with sensitive data

3. **Prepare your documentation**
   - README file
   - Supplementary information (if any)
Prepare your data

File Naming

1. Include important contextual information
2. Try not to have very long file names
3. Use date format YYYYMMDD or YYMMDD
4. Avoid special characters (i.e. ~ ! @ # $ % ^ & * ( ) ` ; < > ? . [ ] { } ' " |) and space in a file name
5. Include the versioning information (if needed)

File Formats

1. Non-proprietary formats is preferred
2. "Lossless" formats
3. Commonly used by the research community

Sensitive Data

1. Sufficiently anonymise your data
2. Explicit consent needed for any research subjects whose identity could be revealed from the data (including images)
A README file is intended to help ensure that your research data can be correctly interpreted and re-used by others.

- Write your README document as a plain text file
- Name your README file as "readme.txt"

Contents to be considered:

- **General information**, e.g. title, creator, date/date range of data collection
- **Data and file overview**, e.g. brief description to the dataset or each data file
- **Sharing and access information**, e.g. licensing and rights
- **Methodological information**
- **Technical information**, e.g. system requirements/software to access or analyse the data
Define Access Setting

- “Open” by default
- Embargo setting is available (no more than 2 year)
- For other access settings (Restricted/Closed), please contact us and it will be considered on a case-by-case basis
WHY DEPOSIT RESEARCH DATA INTO SCHOLARBANK@NUS

Your research data will stay with NUS IT within the campus network.

We keep your data in perpetuity with frequent backups.

DOIs are assigned to archived datasets. Through this, formal citations and usage statistics are trackable.

Usage and citation statistics tracking is enabled.

With authors' ORCiD added to Dataset description, records of dataset will be automatically updated on your ORCiD profile.
RESEARCH DATA SERVICES@NUS LIBRARIES

- Data Management Planning
- Data Organisation Strategies
- Metadata & Documentation
- Data Sharing & Publishing
- DOI Minting
- Workshops & Trainings
- And More

Contact Us
nusl.scomm@nus.edu.sg
gerrie@nus.edu.sg
estella.ye@nus.edu.sg
SCHOLARBANK@NUS
Future Plans

• Engagement with Departments/Funders
  o To facilitate a one time/one source deposit to save researchers’ time

• Explore APIs that would allow harvesting of NUS researchers’ output
  o To facilitate a one time/one source deposit to save researchers’ time

• Actively gather feedback from end-users
  o To develop and enhance existing system and services

• Design discipline specific workshops on RDM
  o To upskill researchers in RDM best practices
USEFUL RESOURCES

• ScholarBank@NUS
  Scholarbank.nus.edu.sg
• ScholarBank@NUS – Research Data Content Policy
  Scholarbank.nus.edu.sg/policy-for-datasets.jsp
• NUS Libraries LibGuide Research Data Management
  libguides.nus.edu.sg/rdm
• NUS Libraries LibFAQ
  libfaq.nus.edu.sg/search/?t=0&adv=1&topics=ScholarBank%40NUS
• NUS Research Data Management Policy
  www.nus.edu.sg/dpr/InfoForResearchers/research-dmp.html
• Creative Commons Licenses
 creativecommons.org/licenses/
YOUR FEEDBACK IS GREATLY APPRECIATED

https://nus.edu/2gLm0fo
FAQ
THE END

THANK YOU