

Singapore Management University

## Institutional Knowledge at Singapore Management University

---

Research Collection Library

SMU Libraries

---

10-2019

### Taking a journey on research data management in Singapore

Pin Pin Yeo

*Singapore Management University*, [ppyeo@smu.edu.sg](mailto:ppyeo@smu.edu.sg)

Danping Dong

*Singapore Management University*, [dpdong@smu.edu.sg](mailto:dpdong@smu.edu.sg)

Follow this and additional works at: [https://ink.library.smu.edu.sg/library\\_research](https://ink.library.smu.edu.sg/library_research)



Part of the [Asian Studies Commons](#), [Scholarly Communication Commons](#), and the [Scholarly Publishing Commons](#)

---

#### Citation

Yeo, Pin Pin and Dong, Danping. Taking a journey on research data management in Singapore. (2019). *Pacific Neighborhood Consortium Annual Conference and Joint Meetings PNC 2019, Oct 15-17*. 1-26. Available at: [https://ink.library.smu.edu.sg/library\\_research/152](https://ink.library.smu.edu.sg/library_research/152)

This Presentation is brought to you for free and open access by the SMU Libraries at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection Library by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email [cherylds@smu.edu.sg](mailto:cherylds@smu.edu.sg).



**Taking a journey on research data management  
in Singapore**

**YEO Pin Pin**  
*Head, Scholarly Communication*

**DONG Danping**  
*Librarian, Scholarly Communication*

# Contents

- About SMU
- Changes in Singapore policy landscape
- Open access and open data
- Indicators for RDM at SMU
- Review of RDM policies
- SMU RDM Policy
- Data interviews
- SMU RDM journey
- Reflections

# SMU schools and research areas



**School of Accountancy**

- Corporate Reporting & Disclosure
- Corporate Governance, Auditing & Risk Management
- Financial Intermediation & Information
- Financial Performance Analysis
- Accounting Information System



**Lee Kong Chian School of Business**

- Organisational Behavior & Human Resources
- Strategy & Organization
- Finance
- Quantitative Finance
- Operations Management
- Marketing
- Corporate Communication



**School of Economics**

- Economics Theory
- Econometrics
- Applied Microeconomics
- International Economics
- Macroeconomics



**School of Social Sciences**

- Sociology
- Political Science
- Psychology
- Humanities



**School of Law**

- Asian and Comparative Legal Systems
- Public International Law, Regional and Trade Law
- Corporate Finance and Securities Law
- Private Law
- Innovation, Technology and the Law
- Dispute Resolution
- Legal Theory, Ethics and Legal Education
- Public Interest Law, Community and Social Justice
- Public Law



**School of Information Systems**

- Data Science & Engineering
- Cybersecurity
- Information Systems & Management
- Intelligent Systems & Optimization
- Software & Cyberphysical Systems

Home to 10,010 students (undergrad + post-grad), with 376 faculty members

# SMU Libraries



## Quick facts on SMU Libraries

- Over 462,000 books  
(Over 400,000 e-books)
- Access to over 80,000 e-journals
- Over 170 electronic databases
- More than 3 million downloads from institutional repository
- Over 1.27 million visitors in 2018
- 38 staff with 23 professional librarians



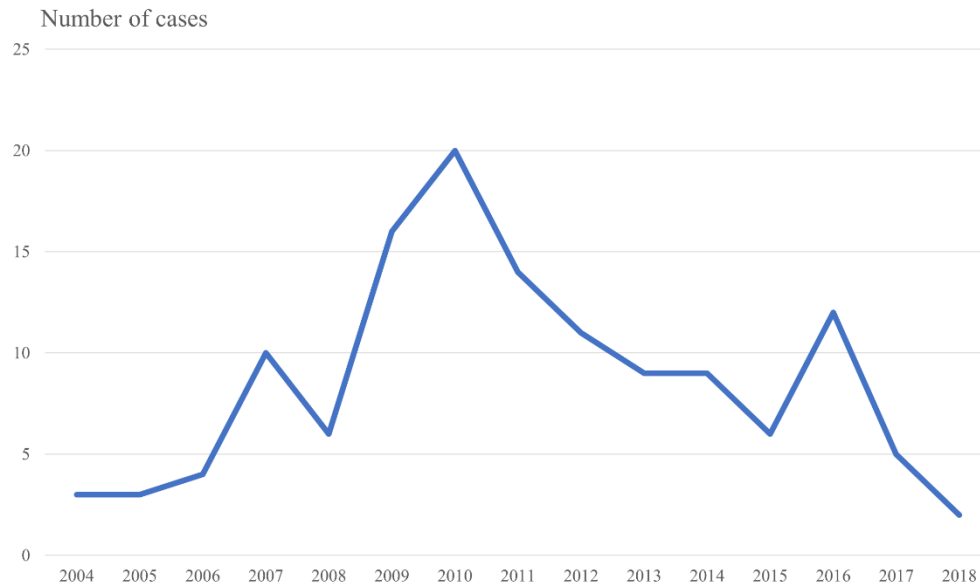
## Changes in Singapore policy landscape

- Universities adopt different approaches for research data policy and services
- No national mandate on open access and open data



# Changes in Singapore policy landscape

- Research integrity breaches which led to retraction of articles



Number of cases in Retraction Watch for Singapore



# Singapore Joint Statement on Research Integrity Relating to Scholarly Publications (2010, 2018)



**Reproducibility:** Research personnel must maintain accurate and detailed research records of procedures and results (for a minimum of 10 years<sup>3</sup>), to allow others to replicate the work, and ensure reproducibility of one's experimental results.

# Open Access and Open Data in Singapore

- 2015 May: Major research funders in Singapore added a normalized clause to require open access of publications within 12 months of publication
- 2015 Nov: National Medical Research Council added the requirement for sharing of data
- Social Science Research Council and the Institute for Adult Learning also added the requirement for sharing of data



Agency for  
Science, Technology  
and Research



Ministry of Education  
SINGAPORE

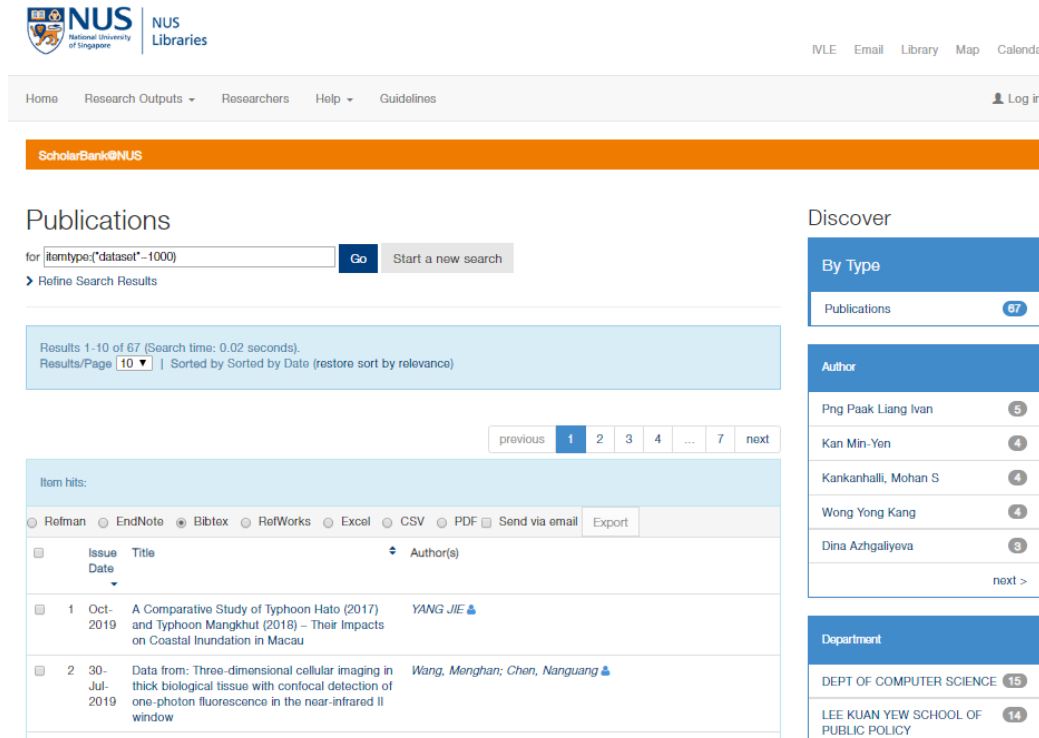


MINISTRY OF HEALTH  
SINGAPORE

**NATIONAL RESEARCH FOUNDATION**  
PRIME MINISTER'S OFFICE  
SINGAPORE

# Response by NUS

- 2012 National University of Singapore (NUS) implemented research data policy
- 2017 launched data repository using Dspace (same platform for pubs repository)



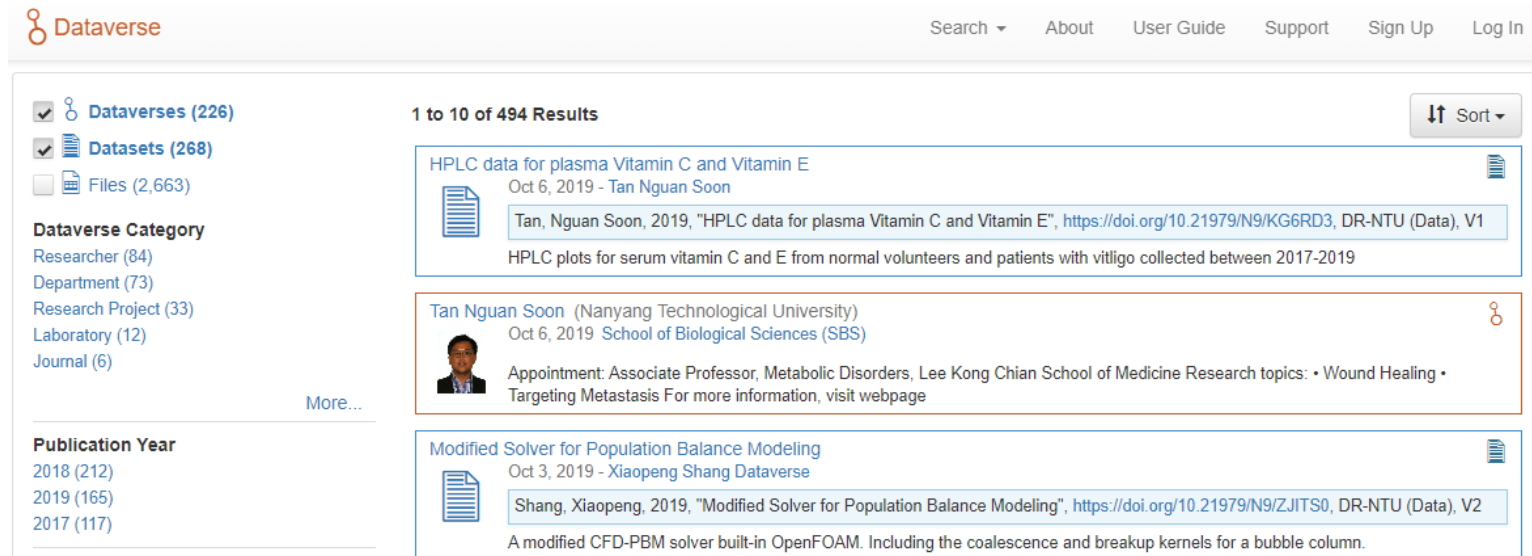
The screenshot shows the NUS ScholarBank@NUS Publications page. The search query is 'itemtype:"dataset"--1000'. The results show 67 items, with the first two items displayed:

Item hits:	Issue Date	Title	Author(s)
1	Oct-2019	A Comparative Study of Typhoon Hato (2017) and Typhoon Mangkhut (2018) – Their Impacts on Coastal Inundation in Macau	YANG JIE
2	30-Jul-2019	Data from: Three-dimensional cellular imaging in thick biological tissue with confocal detection of one-photon fluorescence in the near-infrared II window	Wang, Menghan; Chen, Nanguang

On the right side, the 'Discover' section shows filters for 'By Type' (Publications: 67) and 'Author' (Phng Paak Liang Ivan: 5, Kan Min-Yen: 4, Kankanhalli, Mohan S: 4, Wong Yong Kang: 4, Dina Azhgaliyeva: 3). The 'Department' section shows 'DEPT OF COMPUTER SCIENCE' (15) and 'LEE KUAN YEW SCHOOL OF PUBLIC POLICY' (14).

# Response by NTU

- 2015: Nanyang Technological University (NTU) implemented research data policy
- 2016: Data management plans must be submitted to access research funds
- 2017: Launched DR-NTU (Data) using DataVerse platform



The screenshot shows the DataVerse website interface. At the top, there is a navigation bar with the DataVerse logo, a search bar, and links for About, User Guide, Support, Sign Up, and Log In. Below the navigation bar, there are filters on the left for Dataverses (226), Datasets (268), and Files (2,663). There are also filters for Dataverse Category (Researcher, Department, Research Project, Laboratory, Journal) and Publication Year (2018, 2019, 2017). The main content area displays search results for '1 to 10 of 494 Results'. The first result is 'HPLC data for plasma Vitamin C and Vitamin E' by Tan Nguan Soon, dated Oct 6, 2019. The second result is 'Tan Nguan Soon (Nanyang Technological University)' with a profile picture and appointment information. The third result is 'Modified Solver for Population Balance Modeling' by Xiaopeng Shang, dated Oct 3, 2019.

## Early indicators for RDM at SMU

- 2010: Report on research data services by T. Bennett recommended the initiation of a distinct research data services function that provides, at minimum, identification and dissemination of data resources, instruction in finding and using data, and some technical assistance with the use of specialized software
- 2012: Added a part-time position and hired a postgraduate student
- 2014: Converted to full-time position to focus on research data services
- 2016: New hire with broader responsibilities
- Librarians have always fielded questions on data and datasets from SMU community

## Starting a RDM policy at SMU

- Building on existing relationship between SMU Libraries and Office of Research & Tech Transfer (ORTT)
- 2017: Started a project on Research Data Policy
  - Studied local and overseas institutional RDM policies
  - Studied data policies of journals where SMU faculty tend to publish in
  - Feasibility study of data repository solutions

# Data repository platforms

- Environmental scan of data repository solutions
- Prefer hosted over open source
- Solutions adopted by other local universities
- Identified solutions and analyzed their strengths & weaknesses
- Understand technology aspects of RDM and prepare for next stage

## Engagement with stakeholders

- Draft policy proposed to Council of Deans in July 2018 for endorsement
- Met with the School Deans and Vice-Deans of Research
- Consulted the Academic Faculty Senate
- Interviewed faculty about their research data by SMU Libraries
- Update at the Provost Council in Sep 2019





# SMU RDM Policy

- Policy has evolved from the draft version in 2018 to the current version in 2019
- Terminology changed to provide more clarity
- Moved from centralised approach to decentralised approach for data storage and retention
- Data ownership and rights of use aligned with existing Intellectual Property Policy
- Planned to be effective in Jan 2020 together with Research Data Repository



# Interviews with faculty on their research data

## ■ Objectives

- Understand the research workflow and nature of research data produced from the major disciplines of SMU (business, finance and management, computer science, social sciences, economics, accounting and law)
- Understand current data management practices
- Get feedback on draft policy

## ■ Interviewed 18 faculty from all 6 Schools

- Jointly conducted by liaison librarian & RDM librarian
- Transcripts were coded for analysis

# Themes and categories

Data  
Characteristics

Documentation

Storage,  
Backup and  
Archival

Ownership

Data  
Repository

Reproducibility

Data Sharing

## Documentation

The spectrum of current practices in data documentation...

### “It’s all in the paper!”

- reading the paper could help understand the dataset
- Expect other researchers to figure out the meaning on their own
- Too much effort with little gain

### “I keep some notes for myself...”

- Some simple notes within software for self
- Some codes may be available, but may not be easily understood
- No additional effort to describe the data for others to understand

### “I’ve got everything!”

- Codes are available and comprehensive which captures almost all steps of research
- Data is understandable or codebook available
- Reproducible research

## Documentation

How much effort does it take?

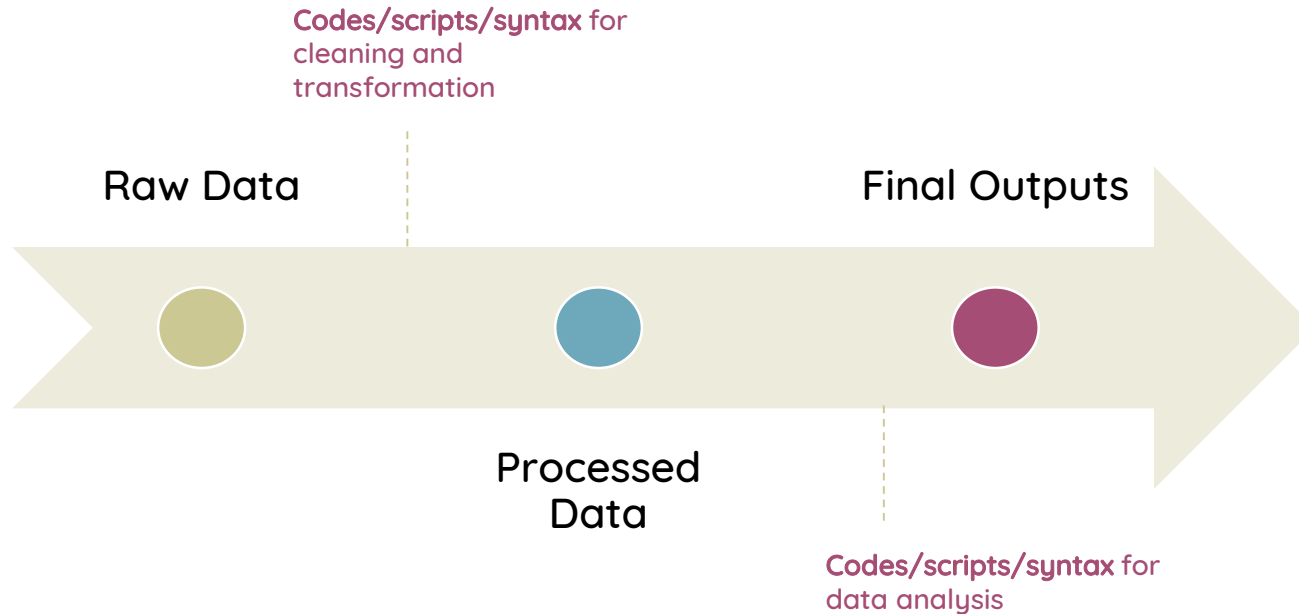
- Considerable overhead to describe all data and codes for sharing
- Sometimes impossible to track all manual steps or minor changes

- Relatively easy to provide the final dataset, codes and analysis outputs
- too much effort to document everything from raw to final

- Not much effort if you do it consistently from the start
- In the long run it helps with efficiency
- Some see it as necessary for good science

## Reproducibility

What should be archived for reproducibility?



## Takeaways from the interviews

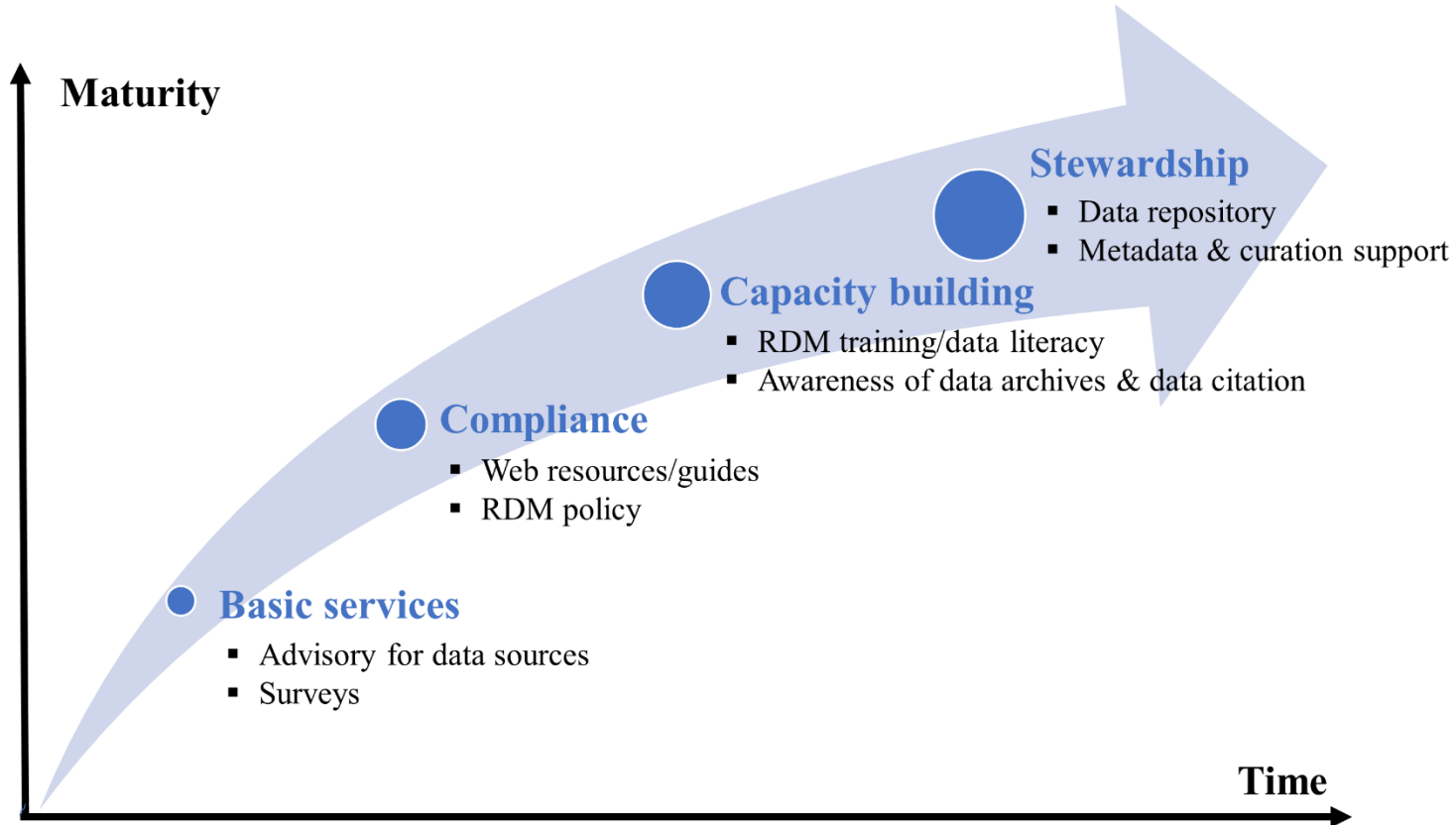
- Research data management is highly discipline-specific - no uniform approach to achieve reproducibility
- Useful insights on revisions to the draft policy
- Gain a better understanding of a typical 'data workflow', and nature & characteristics of SMU's research data
- Valuable information for future RDM work

# Work-in-progress & looking forward

- In-progress: SMU Data Repository Project
- 2020 Jan: Policy to be effective
- 2020 Q1: Repository to be launched
- Focus on communication and engagement



# SMU journey



## Reflections on RDM journey

- RDM needs to be a strategic and university-wide initiative
- Critical to get the support and buy-in from senior management
- Policy needs to come first
- Good collaboration between research office, IT department and the library
- Engagement with stakeholders is important



