AIRS 3
Advanced Information Research Skills

Modules:
7. Organising the literature
8. Managing data
9. Collaboration
Learning outcomes

To be able to:

1. Develop a systematic approach to coding the literature, note taking and managing notes.

2. Identify advantages and disadvantages of bibliographic management software.

3. Apply data management strategies to organise information proficiently, ethically and legally.

4. Commence your data management plan.

5. Identify a conference to attend with networking benefits.
Assessment: Resource Log (1500-2000 words)

1. Your research question
2. Preliminary search
3. Retrieving and evaluating information
4. Comprehensive search
5. Search alerts
6. Cited reference search
7. Organising the literature
8. Managing data
9. Collaboration
10. Getting published
11. Publication metrics
## Research support

<table>
<thead>
<tr>
<th>Processing and analysing your data</th>
<th>Search training events</th>
<th>Install software</th>
</tr>
</thead>
<tbody>
<tr>
<td>eResearch Office</td>
<td>Develop Research and Career Skills</td>
<td>ITassist</td>
</tr>
<tr>
<td>Help and advice</td>
<td>Training</td>
<td>Software</td>
</tr>
</tbody>
</table>
Bibliographic management

1. Install and use bibliographic management software
2. Code literature according to research question concepts and write useful notes
3. Organise your literature efficiently for later use
What is Endnote?

• Software that allows you to find, import, export, store, organise & use references to information sources

• A program that creates files of references called EndNote Libraries

• Your research database of academic and professional literature

• Storage for full text PDF, figures, pictures, tables, etc.
Endnote - Avoid pitfalls

• Cutting and pasting text containing formatted EndNote citations cannot be done safely without risk of document corruption.

Convert to unformatted citations **before** moving text around.

• 1 document, 1 Endnote Library

You cannot connect a single Word document to multiple Endnote Libraries.
Discussion

What do you use?
Endnote support

• Hands on workshops

• Installation assistance from the IT Helpdesk

• Help from Liaison Librarians

• Online tutorials from LinkedIn Learning

• Writing your thesis using Word reference guide
Software comparison table

Reference management tools
A number of reference management tools are in use by researchers at QUT. The main ones are EndNote, BibTeX, Mendeley and Paperpile.

EndNote
EndNote is free for QUT students and staff, and provides comprehensive functionality, especially when writing journal articles and theses. QUT Library provides workshops for researchers to help develop all the basic skills necessary to collect and organise references and PDFs.

BibTeX
BibTeX is a program and a file format used to describe and process lists of references, which primarily used for LaTeX documents.

Mendeley
Mendeley is a free reference manager and academic social network to help organise your research.
Coding the literature: skim, scan, select

Find relevant research  
Read effectively  
Write effective notes  

Leo Reynolds, ‘pile of books’  
‘http://www.flickr.com/photos/lwr/4077733012/
Coding the literature

Constrain your research with a research question.

Scan, skim and read literature and code it for relevance to your thesis.

Alter and narrow the scope of your research question as you progress.

The categories you use to code your literature come from the terms of your research question.
Sample bodies of literature

Sample research question:

“What factors characterize a successful mentoring relationship for minority students?”

1. Characteristics of successful mentoring
2. Characteristics of successful academic mentoring of minority students
3. Characteristics of successful academic mentoring of university students
4. Factors that affect the completion of graduate degrees
Create your codes

- a word or short phrase
- a numerical code
- an icon
- an acronym
- an abbreviation

Add tags to bibliographic software

Use codes to:

Write notes directly onto a .PDF or book

Structure writing sections
Activity: Create your codes (5mins)

Discuss your coding system with your neighbor.
Start to document your codes in your resource log.
Duffield, Lee R. (2011) Successes and stresses: A case study on relations between international higher degree students in Australia and their universities. In International Unity in Diversity Conference, Rydges Southbank Townsville Hotel, Townsville, North Queensland

Research support (code)

“in Australia, where general academic work is done only in English, supervising ISS who are NESB speakers can take more time, even for students with advanced English skills” (Duffield, 2011, p.8).

• This article argues that students with a non English speaking background can take more time to complete their doctoral studies

  • Yet, students can finish on time if adequate research support is offered by a University and the student is proactive about asking for help (see Peacock & Firth, 2008).
Note-taking software

**Evernote** for note-taking and archiving text, websites, pictures, voice memos and handwritten notes. You can photograph a concept map and store it with keywords, tags etc.

**Papers** stores .pdf documents, allows you to write notes on each page and bookmark notes. Your .pdfs become a database of references.

**Scrivener** works like an electronic ‘index card’ system. You can move notes and ideas around, trying different logical structures and connections.

**OneNote** available with Microsoft Office
Information management workflow

Differs according to discipline & nature of research...

1. Use bibliographic software to build your library of references
2. Organise filing with naming system
3. Use the grouping function to organise references
4. Classify records/references with Labels & Research Notes
5. From the concept map, transfer headings into Word
6. Paper outline as list or concept map
7. Take notes - using program such as Evernote

Workshops: https://www.library.qut.edu.au/events/categories/research/
BREAK
Managing research data

1. Defining data
2. QUT Data Management Planning Tool (DMP)
3. Storage documentation & metadata
## Module 8: Managing data

<table>
<thead>
<tr>
<th>Question</th>
<th>Marks</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Describe your data. Include the format and scope of your research data</td>
<td>3</td>
</tr>
<tr>
<td>b) Outline your plans to securely store and back-up your data</td>
<td>3</td>
</tr>
<tr>
<td>c) Identify and discuss three types of contextual information (metadata) that help to make data F.A.I.R.</td>
<td>3</td>
</tr>
<tr>
<td>d) Discuss the commercial, legal and/or ethical considerations related to storing and sharing research data</td>
<td>3</td>
</tr>
<tr>
<td>e) Identify how long your data is legally required to be retained after completion of your research project/thesis. Refer to <a href="http://airs.library.qut.edu.au">Research Students – Managing Research Data</a> webpages and provide your reasoning as to how this applies to your research data.</td>
<td>3</td>
</tr>
</tbody>
</table>
Research data lifecycle

Adapted from UK Data Service Research Data Lifecycle https://www.ukdataservice.ac.uk/manage-data/lifecycle
QUT responsible research framework - data management guidance

“Research which is supported by public funding, and higher degree research student projects, must use the QUT research data management planning tool.”

QUT MOPP 2.8.7

The Research Integrity Online module explains your responsibilities.

<table>
<thead>
<tr>
<th>Code section</th>
<th>Relevant QUT policies</th>
<th>Supporting documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section 1: General principles of responsible research</td>
<td>• QUT Code of Conduct for Research [link]</td>
<td>• General principles of responsible research</td>
</tr>
<tr>
<td>Section 2: Management of research data and primary materials</td>
<td>• QUT Code of Conduct for Research [link]</td>
<td>• Management of research data and records</td>
</tr>
<tr>
<td></td>
<td>• Management of research data [link]</td>
<td>• Guidelines for the Management of Research Data at QUT [link]</td>
</tr>
<tr>
<td></td>
<td>• Records management policy [link]</td>
<td>• Data Management Planning Tool [link]</td>
</tr>
<tr>
<td></td>
<td>• Information privacy policy [link]</td>
<td>• University Sector Retention and Disposal Schedule information [link]</td>
</tr>
<tr>
<td></td>
<td>• Information security policy [link]</td>
<td>• Records on human research participants [link]</td>
</tr>
<tr>
<td></td>
<td>• Intellectual property policy [link]</td>
<td>• Guidelines on the use of the Data Management Planning Tool for Higher Degree Research Candidates (PDF, 35 KB) [link]</td>
</tr>
</tbody>
</table>

Section 2: Supervision of research trainees

<table>
<thead>
<tr>
<th>Code section</th>
<th>Relevant QUT policies</th>
<th>Supporting documents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• QUT Code of Conduct for Research [link]</td>
<td>• Supervision of research trainees webpage</td>
</tr>
<tr>
<td></td>
<td>• QUT Staff Code of Conduct [link]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Code of Good Practice for</td>
<td></td>
</tr>
</tbody>
</table>

Link - [http://qut.to/hgw1](http://qut.to/hgw1)
<table>
<thead>
<tr>
<th>Data Class</th>
<th>Process</th>
<th>Content examples</th>
<th>Data examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>Generated by lab equipment</td>
<td>Gene sequences; chromatograms</td>
<td>Laboratory notes; specimens; samples; methodology; slides; artefacts</td>
</tr>
<tr>
<td>Computational/Simulation</td>
<td>Generated from computational models - the actual model (and its metadata) may be more important than the output data</td>
<td>Climate models; economics prediction models</td>
<td>Methodology; data files; models; algorithms; scripts; workflows; standard operating procedures and protocols; simulation software</td>
</tr>
<tr>
<td>Observational</td>
<td>Recordings of specific phenomena at a specific time or location</td>
<td>Seismic data, medical imaging, opinion polls, climate data, interview or survey results</td>
<td>Transcripts; audio or video recordings; field notebooks; diaries; photographs; films; slides; questionaries; test responses; codebooks; text documents</td>
</tr>
<tr>
<td>Derived</td>
<td>Produced via processing or combining of other data</td>
<td>Data mining; compiled databases; GIS</td>
<td>Database contents; spreadsheet data; data files</td>
</tr>
<tr>
<td>Reference</td>
<td>Extracted from reference datasets</td>
<td>Genbank, HILDA, ABS CURF datasets</td>
<td>Spreadsheets; data files; contents of an application (schemas, input, output; log files for analysis software)</td>
</tr>
</tbody>
</table>
Activity: Identifying the data (5mins)

A focus group
- Participants provide demographic and personal information via an online survey
- They sign a release and consent form
- They engage in semi-structured conversation, responding to pre-defined discussion topics
- They are video-recorded and audio-recorded
- They follow a useability test on a website, using eye tracking software
- They complete a feedback questionnaire

What datasets might be generated?  
What file types or formats may be generated?  
How will each dataset need to be managed?
Activity: Identifying the data (5mins)

<table>
<thead>
<tr>
<th>Datasets generated</th>
<th>File types or formats</th>
<th>Data management considerations</th>
</tr>
</thead>
</table>
| • Audio files and transcripts  
• Video files  
• Survey forms  
• Hand written feedback forms - scanned, transcribed  
• Statistical information - transcribed into a spreadsheet  
• Software report/heatmap  
• Nvivo report | Paper, .xls, .doc, .mov, .aup, .mp3, .prproj, .mp4, .pdf, .csv, .nvp, .twb, .tde, .spo | • Saved in a non-proprietary format  
• Filed in an archive box, labelled and locked. Location recorded. Scanned copies  
• De-identified  
• Saved to encrypted hard drive; saved to a password protected networked drive  
• Compressed  
• Accompanied by an index; a file hierarchy or relationship schema; a copy of the software; the survey instrument; the questionnaire; the discussion topics; a glossary |
Collecting data - research methodologies
Collecting data – using existing data

**Find existing research datasets**

To support your project, you can access licensed research data sets through the QUT Library.

- **Access research datasets**

The library also curates published datasets and repositories that support the social sciences and research in the science, technology, engineering, mathematics and medical fields.

- **Datasets for research in social science**

- **Datasets for Health and STEM**

To request access to additional datasets, contact the Research Data Librarian.

[https://qutvirtual4.qut.edu.au/group/staff/research/conducting/managing-research-data/find-existing-research-datasets](https://qutvirtual4.qut.edu.au/group/staff/research/conducting/managing-research-data/find-existing-research-datasets)
Stage 2/Research Proposal

- Includes a data management plan in Appendices

- Requirements of the first milestone are available from the Higher Degree Research page QUT Research Students Centre wiki:

The document requires indication of the following:

- I have not yet considered data management issues but will do so before confirmation
- I have read the data management planning information and discussed the need to complete the Data Management Plan (online) with my supervisor
- I have read the data management planning information, discussed it with my supervisor, and have completed the Data Management Plan (online)
Activity (10mins)

Create a data management plan using the Data Management Planning (DMP) tool

https://dmp.qut.edu.au/

Work on Section 1
# Data storage at QUT

<table>
<thead>
<tr>
<th>Storage option</th>
<th>For research data</th>
<th>For students</th>
<th>Very large data (&gt;500GB)</th>
<th>Sensitive data</th>
<th>Share data with QUT</th>
<th>Share data with non-QUT</th>
<th>Remote access to data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Data Storage Service (RDSS)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>AARNet’s CloudStor</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>QCIF’s QMIRdata</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

- **H: drive / U: drive folders (except U:\Research)**
  - ![X](http://qut.to/8c9ob)
- **OneDrive / Google Drive / DropBox**
  - ![X](http://qut.to/8c9ob)
- **Portable storage (e.g. USB / CD / DVD / external hard drive)**
  - ![X](http://qut.to/8c9ob)

[http://qut.to/8c9ob](http://qut.to/8c9ob)
3-2-1 storage rule

Tracy uses QLD Transport data.

A month of data files produced hourly in .csv file format is provided by encrypted USB.

Raw data must be kept on QUT network storage, for security, processing and analysing purposes.

Derived data results are shared with her industry partner, QLD Transport.

3-2-1 Solutions

**Master:** Tracy saves the .csv files from USB via QUT computer to her RDSS raw data folder. She doesn’t use these files but makes copies of them for processing.

**Working:** Tracy saves a copy of the master .csv files to QUT RDSS, in a working data folder. Tracy uses these copies for processing and analysis. Derived data results are shared securely with industry partner via AARNeT’s Cloudstor FileSender.

**Backup:** RDSS research data is backed up to removable tapes, which are duplicated and housed in two security compliant geographically dispersed data centres.
Process and analyse – software at QUT

• Software to download at home: https://secure.qut.edu.au/ithelpdesk/qut/softwaredownloads/downloads.jsp

• Self-install software on a QUT computer via IT Assist icon

• Full catalogue via IT Helpdesk website: https://secure.qut.edu.au/ithelpdesk/qut/softwaredownloads/downloads.jsp

• QUT hands-on training: https://unihub.qut.edu.au/students/events/

• Online training via LinkedIn.com https://libguides.library.qut.edu.au/databases/lynda
Make your data F.A.I.R.

<table>
<thead>
<tr>
<th>F</th>
<th>Findable</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Accessible</td>
</tr>
<tr>
<td>I</td>
<td>Interoperable</td>
</tr>
<tr>
<td>R</td>
<td>Re-useable</td>
</tr>
</tbody>
</table>

Publishing data

Data repositories

Multidisciplinary

- FigShare
- Dryad

 Discipline specific

- DataONE
- GenBank
- PANGAEA
- Scientific Data (Nature's data journal)
- Australian Data Archive (social sciences)
- TERN (terrestrial ecosystem data)

QUT’s data repository

https://researchdatafinder.qut.edu.au/
Metadata

Metadata is data about data.

It contextualises the data.

It makes it findable and usable.

= ?

+ = Useable

Metadata
Examples of metadata

- Endnote labels to categorise resources
- #hashtags on Twitter
- iTunes music genres
- Flickr aperture settings
- Grocery shop aisle signs
- Library subject headings ...
Metadata types

**Descriptive metadata** – the who, what, why, when and how.

**Provenance metadata** - data source, version tracking & transformations.

**Technical & Structural metadata** - file types, software, file size & contents; how the components relate.

**Rights and Access metadata** – preservation & confidentiality; ownership & licensing; access restrictions & timelines.
Activity: Types of Metadata (5mins)

Explore a Research Data Finder record with your neighbour.

Does it meet the F.A.I.R. data principles using the metadata types below?

- Descriptive metadata
- Provenance metadata
- Technical & Structural metadata
- Rights and Access metadata

https://researchdatafinder.qut.edu.au/
# Data retention

<table>
<thead>
<tr>
<th>Nature of research</th>
<th>Retention Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research data which does not result in a patent</td>
<td>Retain for 5 years after last action.</td>
</tr>
<tr>
<td>Research data which results in a patent</td>
<td>Retain for 7 years after expiry of patent.</td>
</tr>
<tr>
<td>Clinical trials</td>
<td>Retain for 15 years after completion of clinical research/trial AND 10 years after last patient service provision or medico-legal action.</td>
</tr>
<tr>
<td>Significant</td>
<td>Permanent</td>
</tr>
<tr>
<td>Research data which is of high public interest or significance to the discipline, such that it has or will change a commonly held view or approach.</td>
<td>Permanent</td>
</tr>
<tr>
<td>Projects which:</td>
<td>Permanent</td>
</tr>
<tr>
<td>• are controversial</td>
<td>Permanent</td>
</tr>
<tr>
<td>• are subject of extensive debate</td>
<td>Permanent</td>
</tr>
<tr>
<td>• rouse widespread scientific or other interest</td>
<td>Permanent</td>
</tr>
<tr>
<td>• have the potential to cause major adverse impacts on the environment, society or human health</td>
<td>Permanent</td>
</tr>
</tbody>
</table>

The Queensland State Archives University Sector Retention and Disposal Schedule - statutory retention periods for research data (section 601.2/A50).
Additional training

- Datasets for Research
- File and Data Management 101
- Research Data Management Drop-In
- Hacky Hour

https://unihub.qut.edu.au/students/events/search
Collaborating with other researchers

Finding potential collaborators at QUT, Australia and worldwide

Online tools for data sharing and collaboration
Building your academic profile

• Create a profile

• Create content (e.g. social media, digital repositories of work, open education resources, content curation)

• Become a player in your field – professional associations, conferences, webinars

• Get published & get ranked (see AIRS 4)
Activity (5mins)

Search **QUT Staff Profiles** OR **QUT ePrints** to identify potential collaborators at QUT
Networking – in person & online

Supervisor

Faculty events and conferences

Peers

Industry engagement

hdr.intern@qut.edu.au

QUT ePrints

Linkedin

twitter
Activity: Find a conference (5mins)

Search **Conference Alerts** for suitable academic conferences. Set up an alert for conferences in your field.

http://conferencealerts.com/

https://thinkcheckattend.org/
AIRS 4 workshop

You will:

1. Make decisions about where to publish.
2. Use bibliometric and altmetric tools.
3. Access training to prepare for your post PhD career.