Data Management 101

Well managed data supports data publishing and reuse.
What’s in data management for me?

Good data management is essential to my research career.
A checklist: to manage your data

- Policy ✔
- Storage ✔
- Metadata ✔
- Licencing ✔
- Ethics ✔
- Identifiers ✔
A checklist: to manage your data

● Policy/guide
● Storage
● Metadata
● Licencing
● Ethics
● Identifiers

● funder/journal /institution
● funder /institution
● research/er(s) /institution
● funder/partner /institution
● domain /institution
● domain/social /institution
The Case of Professor K

and a mixture of data.
A checklist: to manage your data

- Policy
- Storage
- Metadata
- Licencing
- Ethics
- Identifiers
I’m a newbie - help!?

A scholar moves to UNSW from a university overseas.
Her immediate concerns?

ARC LIEF grant application and data management
A checklist: to manage your data

● Policy/guide
● Funder/journal/institution
● ARC/UNSW
A11.5.2 Researchers and institutions have an obligation to care for and maintain research data in accordance with the *Australian Code for the Responsible Conduct of Research* (2007). The ARC considers data management planning an important part of the responsible conduct of research and strongly encourages the depositing of data arising from a Project in an appropriate publically accessible subject and/or institutional repository.
Her next question?

She’s thinking ahead: what’s this I hear about ARC Discovery rule changes around data management?
ARC Discovery Funding

An application now requires a data management plan.
A checklist: to manage your data

- Policy/guide
- funder/journal
- ARC
- /UNSW
3.12 What information am I required to provide in relation to Management of Data in the Project Description?

In line with responsibilities outlined in the *Australian Code for Responsible Conduct of Research (2007)* and international best practice, the ARC has updated wording in relation to the management of data.

The ARC does not mandate open data. However, researchers are encouraged to consider the ways in which they can best manage, store, disseminate and re-use data generated through ARC-funded research. The Project Description requires researchers to articulate briefly their plans for the management of data generated through the proposed Project. In answering this question researchers need not include extensive detail of the physical or technological infrastructure.

Answers should focus on plans to make data as openly accessible as possible for the purposes of verification and for the conduct of future research by others. Where it may not be appropriate for data to be disseminated or re-used, justification may be provided.

Further information and resources on managing data are available on the Australian National Data Service (ANDS) website at [http://www.andis.org.au/](http://www.andis.org.au/).

3.13 Is it sufficient to answer the Management of Data section in the Project Description by noting that I will comply with my institution’s requirements?

No. Whilst the ARC recognises that some institutions may have infrastructure and/or processes in place for storing, managing and sharing data and that these are valuable resources, to take into account the differences that may exist between institutions, disciplines and research projects, researchers are encouraged to highlight specific plans for the management of their research data in this section.

The Management of Data section in the Project Description aims to encourage consideration of ARC-funded research data at both an individual and institutional level, in accordance with the responsibilities outlined in the *Australian Code for Responsible Conduct of Research (2007)*. Researchers, in consultation with their institutions, are best placed to consider the management and future potential of their research data. This approach allows individuals to take into account the differences that may exist between disciplines and research projects as well utilise institutional resources and support available.

Details of compliance with institutional requirements should be included in this section, provided that they are supported by a description specific to the data arising from the individual research Project.
A checklist: to manage your data

- Policy
- Storage
- Metadata
- Licencing
- Ethics
- Identifiers
A checklist: to manage your data

- Storage
- funder
- /institution
Data Storage

What’s this I hear about research data storage (ReDS/RDSI) available?
More questions..

Does the research data we create and use qualify for ReDS/RDSI funding so it can go on the node or shall I get the data stored locally at the university?
Some answers..

- **The data stored will be safe.** Data will be stored in a certified tier 3 facility, subject to regular backup and replication.
- **The data stored will be secure.** Data will be stored behind security and access protocols designed in conjunction with dataset custodians.
- **The data stored will be accessible.** Data will be stored in a facility with very high speed access to AARNET, the NeCTAR research cloud, and high-performance computing facilities.
- **The data stored will be shareable.** Data can be made accessible, at researchers' discretion, via file sharing, databases or applications.
- **The data store will be robust.** The RDSI project includes provisioning for hardware refresh, ensuring the long-term sustainability of the data.
- **The data stored will incur no cost to researchers.** The Intersect RDSI node is funded by the Federal Government, the State Government and sustained by participating universities.
Data storage

- Safe?
- Secure?
- Accessible?
- Shareable?
- Robust?
Data is described in RDA

How do I make my data discoverable through the university and what is Research Data Australia?
A checklist: to manage your data

- Policy
- Storage
- Metadata
- Licencing
- Ethics
- Identifiers
A checklist: to manage your data

- Metadata
- research/er(s) /institution
Metadata.. that a researcher like you can use.
Metadata that...

Supports:
- Discovery
- Reuse
- Citation

Describes:
- Research activity
- Researcher(s)
- Institution
- Grant
- Publication
Institutional support

e.g. UNSW ResData
Library Service

Goal: efficient discovery, sharing and reuse of UNSW research data collections.

Institutional systems connected.
Auto-filling (to reduce rekeying).
Reuse of institutional information.
A checklist: to manage your data

- Policy
- Storage
- Metadata
- **Licencing**
- Ethics
- Identifiers
Some of the data is mine

...and some of it is other researchers or owned by a government agency.
What data do I licence and how?

- This researcher creates her own metadata (about the data she is looking at). ✔
- This researcher uses data and metadata from a state government agency. ✗
- This researcher collaborates and may have access to data and metadata provided by other researchers. ✗
Licence options:

Institutional licence e.g. [CSIRO Data Licence](https://www.csiro.au/en/Licensing/CSIRO-Data-Licence)
Creative Commons licence e.g. [CC-BY 3.0 AU](https://creativecommons.org/licenses/by/3.0/au/)
One-off licence i.e. a contractual agreement

- Be aware of funding or institutional policy on making data available
- Take a look at [AusGOAL](https://www.biodiversity.edu.au/ausgoal)
A checklist: to manage your data

- Policy
- Storage
- Metadata
- Licencing
- Ethics
- Identifiers
A checklist: to manage your data

- Ethics
- funder/partner /institution
A question of ethics

- What ethical concerns may arise through data collection, research or publishing?
- What processes of de-identification or anonymisation need to be in place?
- How does making data available impact upon seeking research participants?
- Who is the author/owner of the data?
- What constraints are placed by the funder?
In the case of Professor K

Cannot publish data and metadata that has been made accessible to her by a government agency that has a CC-BY-NC licence and or through a research collaboration and legal agreement.
Professor K

The university owns the IP. She will need to check, to publish and licence her (meta)data. There is UNSWorks for open access scholarly publications (provided to support researchers working within ARC/NHMRC funding requirements) and ResData to support discovery of UNSW research data.
A checklist: to manage your data

- Policy
- Storage
- Metadata
- Licencing
- Ethics
- Identifiers
A checklist: to manage your data

- Identifiers
- domain/social
- /institution
Identifying Professor K and her work

Identifiers for her, the universities she has worked at, and her outputs.
# Identifiers for Professor K

<table>
<thead>
<tr>
<th></th>
<th>Personal</th>
<th>Organisational</th>
<th>Grants</th>
<th>Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORCID?</td>
<td>Could be: <a href="http://orcid.org/#########">http://orcid.org/#########</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VIAF</td>
<td><a href="http://viaf.org/viaf/#############">http://viaf.org/viaf/#############</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LinkedIn</td>
<td><a href="http://www.linkedin.com/firstname-lastname">http://www.linkedin.com/firstname-lastname</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Academia.edu</td>
<td><a href="http://www.academia.edu/firstname-lastname/">http://www.academia.edu/firstname-lastname/</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARC</td>
<td><a href="http://purl.org/au-research/grants/arc/LP########">http://purl.org/au-research/grants/arc/LP########</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inderscience</td>
<td><a href="http://dx.doi.org/10.1504/IJDCET.2008.######">http://dx.doi.org/10.1504/IJDCET.2008.######</a></td>
<td><a href="http://dx.doi.org/10.1007/################">http://dx.doi.org/10.1007/################</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Springer</td>
<td><a href="http://dx.doi.org/10.4225/13/###">http://dx.doi.org/10.4225/13/###</a></td>
<td><a href="http://dx.doi.org/10.4225/13/###">http://dx.doi.org/10.4225/13/###</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANDS?</td>
<td>Could be: <a href="http://dx.doi.org/10.4225/13/###">http://dx.doi.org/10.4225/13/###</a></td>
<td><a href="http://dx.doi.org/10.4225/13/###">http://dx.doi.org/10.4225/13/###</a></td>
<td><a href="http://dx.doi.org/10.4225/13/###">http://dx.doi.org/10.4225/13/###</a></td>
<td></td>
</tr>
</tbody>
</table>

Cite person (national) | Cite person (international) | Cite person (international) | Cite person (social) | Cite person (social) | Cite organisation (national) | Cite grant (national) | Cite output (book) | Cite output (article) | Cite output (conference paper) | Cite output (data?) |
Any questions?

Ingrid Mason  ingrid.mason@ands.org.au