Data Ownership, Legal and Ethical Guide

This guide is intended to cover common questions researchers may have around ownership, legal and ethical issues for managing and sharing their research data. These issues should be dealt with early in any research data management planning.

1. Data Ownership and Intellectual Property
 Ownership and intellectual property of research data can be quite complex and take into account whether the research is funded by a grant-making body, whether the research is a collaboration with other institutions and if it has commercial potential or applications.

Ownership and rights issues for research data should be confirmed and laid out at the beginning of a research project, as such issues will affect any future storage, sharing or re-use of the data.

All Pawsey Supercomputing Centre institutions have intellectual property policies which should be consulted when preparing your data management plan. The Pawsey Supercomputing Centre does not hold any IP over the data it stores on behalf of researchers.

Who owns the Copyright of the data?
The creator of the data usually owns the copyright. However when the work is made by an employee as part of employment, the employer (university) is usually the first owner of copyright. There is some variation between institutional IP policies, so familiarise yourself with your local policy and seek advice if you are unclear what this means for your data specifically. Any data stored on the Pawsey Supercomputing Centre must not breach copyright laws.

What if the data has two or more creators?
If the data has multiple creators then there is joint copyright for all creators. Similarly, where a research project has multiple researchers or institutions then there is joint copyright for all researchers/institutions. It is best practice to codify this in a contract before commencing research.

Do data centres, such as Pawsey Supercomputing Centre, hold copyright?
No, generally data centres store and publish data on behalf of the copyright holders. Rather, copyright holders give permission to data centres to store datasets and make them accessible to others.

Do I need to license my data?
All data intended for publication and reuse should have a licence. A license is a legal document that sets out how the data may (and may not) be re-used and attributed to the original researcher. It also removes a barrier to potential re-users.
Writing your own licence is not recommended and unnecessary given the availability of standard licences (see below). If you find that standard licences are not suitable for your work, please consult your institution’s research office and/or legal department for guidance.

Is there a standard licence I can use for my data in Australia?
The Australian Government Open Access and Licensing Framework (AusGOAL) provides standard licences suitable for Australian research data and other publicly funded work. AusGOAL guides you through a set of simple questions to select an appropriate licence for your data. The AusGOAL suite of licenses includes:

- Creative commons (CC) licenses
- AusGOAL Restrictive License template (RL)
- The BSD 3-Clause Software Licence

Creative Commons licences are standardised copyright licences that grant permission to use copyright works, in accordance with the particular standard set of conditions selected by the copyright owner.

Creative Commons Licences

<table>
<thead>
<tr>
<th>Icon</th>
<th>Licence</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="https://creativecommons.org/favicon.ico" alt="CC BY" /></td>
<td>Attribution (CC BY)</td>
</tr>
<tr>
<td><img src="https://creativecommons.org/favicon.ico" alt="CC BY-NC" /></td>
<td>Attribution Non Commercial (CC BY-NC)</td>
</tr>
<tr>
<td><img src="https://creativecommons.org/favicon.ico" alt="CC BY-SA" /></td>
<td>Attribution Share Alike (CC BY-SA)</td>
</tr>
<tr>
<td><img src="https://creativecommons.org/favicon.ico" alt="CC BY-NC-SA" /></td>
<td>Attribution Non Commercial Share Alike (CC BY-NC-SA)</td>
</tr>
<tr>
<td><img src="https://creativecommons.org/favicon.ico" alt="CC BY-ND" /></td>
<td>Attribution No Derivatives (CC BY-ND)</td>
</tr>
<tr>
<td><img src="https://creativecommons.org/favicon.ico" alt="CC BY-NC-ND" /></td>
<td>Attribution Non Commercial No Derivatives (CC BY-NC-ND)</td>
</tr>
</tbody>
</table>

The AusGOAL Restrictive License template (RL) was developed specifically to licence material that contains personal or other confidential information that has a high security risk associated with its release.

The BSD 3-Clause Software Licence can be applied to software that has been wholly created by your organisation, where the code does not include code licensed under the GNU GPL Licence.
Other licenses to choose from include the Open Data Commons run by the Open Knowledge Foundation and provides three further licences similar to Creative Commons Licences, but designed specifically for data.

The most permissive way of releasing data is under a dedication to the public domain. This is where all copyright interests and database rights are waived, allowing the data to be used as freely as possible. Creative Commons has a CC0 licence for releasing into the public domain. Using CC0, you can waive all copyrights and related or neighbouring rights that you have over your work, such as your moral rights (to the extent waivable), your publicity or privacy rights, rights you have protecting against unfair competition, and database rights and rights protecting the extraction, dissemination and re-use of data.

2. Ethical and Legal Issues
Researchers collecting and using data from humans and animals are expected to maintain high ethical standards and comply with relevant national and state legislation. Ethical guidance is provided by professional bodies, academic institutions and funding organisations. This research is subject to ethical reviews and researchers are obliged to address ethical requirements through ethics applications. When depositing data with the Pawsey Supercomputing Centre researchers must follow the ethical guidance provided by these bodies.

In Australia, research on human subjects is covered by the National Statement on Ethical Conduct in Human Research (2007). Researchers collecting data from humans need to consider and explain the processes used to protect research participants including the steps taken to ensure anonymity of data. This could also include an explanation of constraints to data sharing. It may not be appropriate to share sensitive, confidential data with other researchers depending on the level of consent that has been received from participants.

Are there any privacy regulation that affect my data?
Chapter 18 of the National Statement on Ethical Conduct in Research Involving Humans outlines the legal and ethical obligations of researchers to participants of research. The Privacy Act 1988 requires Commonwealth agencies to comply with the Information Privacy Principles (IPPs) regarding personal information. Researchers must consider relevant privacy legislation when storing data at the Pawsey Supercomputing Centre and ensure the collection, storage and use is permitted by law.

Do principles of confidentiality and consent affect my data?
The National Statement on Ethical Conduct in Research Involving Humans, particularly Chapter 3.2 on Databanks, outlines researchers' responsibilities in the areas of collecting, storing, using and disclosing research data.

Researchers must respect any confidentiality agreement regarding stored data that has been made with participants. Consent processes must include information about
the form in which the data will be stored (specifically about identities of subjects) and the purposes for which the data will be used and/or disclosed. You must retain records of these agreements and ensure that data will not become available for uses to which participants did not consent.

Do I have to consider cultural sensitivities surrounding my data?
Research involving different cultures of people may contain more ethical challenges. The research could include anthropological studies of other cultures or research involving indigenous people. This research may have special data management requirements: decision making and communicating about methods of collecting information, including where data is stored and how it is accessed, may need special attention.

Data management planning may also highlight a need to take into account indigenous or other cultures’ intellectual and cultural property rights, in addition to other ownership and use of copyright and intellectual property.

The Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS) provides Guidelines for Ethical Research in Indigenous Studies and other resources for learning more about these issues and how to address them.

Can I use sensitive administrative data in my research?
Through the Population Health Research Network (PHRN) there are procedures in place to make health and health related data available to approved Australian researchers. This data is mainly administrative data, which is generated during the provision of a service, such as registering a birth, death or marriage or visiting a hospital.

If you need to use sensitive administrative data about health and related information for the WA population in your research the WA Data Linkage system has been created for ethically approved research, planning and evaluation projects which aim to improve the health of Western Australians. There are strict protocols on the use of this data to protect confidentiality and security.

Data provided through this system can only be used for the specific project that has been approved. It must not be combined with other datasets or used for other projects, therefore it must not be shared outside the project.

Is it possible to share sensitive data?
In sharing any sensitive information through the Pawsey Supercomputing Centre or anywhere else you must make sure that you are in compliance with your ethics body’s requirements. This means that you need to start planning to share at a very early stage in the project.

In order to share any data about people it is optimal to have informed consent from the research participants about the proposed future uses of the data. In many cases you
can protect participants identities through data anonymisation or avoid this by not collecting any personal data.

It may not be necessary to share all parts of your data with everyone. Managing and regulating access to approved researchers is an important part of protecting sensitive data. The data could also be embargoed for a period of time deemed long enough to nullify any confidentiality concerns.

A researcher must not store clinically identifiable patient data on the Pawsey Supercomputing Centre. All such data must be subject to anonymisation before being stored. Please refer to the Pawsey Supercomputing Centre Data Storage and Management Policy for further details.

Acknowledgement to Sarah Olesen (Australian National Data Service) for her time in reviewing this guide.