RESEARCH DATA
COPYRIGHT – WHAT YOU NEED TO KNOW

This factsheet is intended for University of Newcastle staff members to use in the context of their employment only and provides general advice. University staff should contact the Copyright Advisor if they have specific questions or concerns.

Is my data (or dataset) copyrightable?

This can be a complex area with multiple layers. Much of what constitutes 'data' – facts, names, numbers, etc. – is not protected by copyright, regardless of the effort involved in gathering it. There may be protection in how you **express or use** this data, but generally if your dataset lacks a sufficient level of originality, any copyright protection may be quite 'thin'.

As 'data' is a broad term, there will be material that meets the minimum requirements for copyright protection. This could be quotes/transcripts/text, images, sound clips, etc. Where this data is taken from existing material, any copyright will most likely remain with the original owners and may limit the shareability of the data (see below for more on data scraping).

The more original the data(set), or creative the expression, the more likely you will have copyright protection. There are also usually other 'copyright-like' or 'copyright-adjacent' concerns with data, such as ownership and access, along with considerations for privacy/anonymity or commercial-in-confidence. 'Sui generis' rights may also apply in some countries.

What does this mean for open licensing?

Open licensing, such as Creative Commons (CC), functions within copyright. As such, if there is no (or at least very limited) copyright involved, a licence that specifically relies on copyright may be relatively unenforceable for some data.

Does this mean I can't make my data open access (OA) via CC?

It means that you may not be able to rely on a CC *licence* for some data. Creative Commons has a *public domain dedication tool* known as CC0 that allows creators to make materials OA and clears up any restrictions around copyright and usage. CC0 has a back-up function as a licence in countries where the public domain doesn't exist. Use of CC0 by some OA publishers for data that supports published articles has been in place for some time.

Alternative data licensing options include Open Data Commons (ODC) and the Community Data License Agreement (CDLA).

Does the CC0 tool include considerations such as 'Attribution', 'NonCommercial', etc., from the CC licences?

CC0 does not restrict any use of data. The table explains why CC-licence elements may not be appropriate for some data:

Attribution (BY)	Good scholarly practice is to always cite materials used. If concerned about this, just ask for credit.
NonCommercial (NC)	NC may restrict a lot of academic/scholarly use of data, including publishing in journals or books.
ShareAlike (SA)	SA may restrict how your data is applied, such as where the licensing is incompatible with others.
NoDerivatives (ND)	ND may restrict meaningful use of your data, such as combining it with other data or material.

What about data I have scraped or mined?

There is no specific copyright exception that allows for data scraping/mining in Australia. 'Fair dealing' may allow for certain uses with research but does not cover all reproductions of copyright material, limiting the ability to share this data.

Use of scraped online data for training generative AI is still, at time of writing, considered to be a potentially infringing use.

Best practice

- Check out the Library's Researcher Skills Toolkit for considerations around ownership, access, etc., for data(sets).
- Consider whether your data(set) is copyrightable and apply an open licence or the CC0 tool as appropriate.
- Where you have made use of, or mined/scraped from, third-party materials to create your data(set), it may not be shareable with others without permission from the copyright owners.

For more information see **Copyright Support** below.

Copyright Support

Ask a question: **newcastle.edu.au/copyright-request**
Phone: **(02) 492 16456**
Web: **newcastle.edu.au/copyright-support**

GETTING COPYRIGHT 'RIGHT'