

Copyright & Research Data

What is copyright?

Copyright covers “original works of authorship fixed in any tangible medium of expression” (17 U.S. Code § 102). Copyright gives authors the exclusive rights to control:

- reproduction of the copyrighted work
- preparation of derivative works
- selling copies or transferring copyright ownership
- performing or displaying the copyrighted work publicly (17 U.S. Code § 106)

The current term of copyright is the life of the author plus 70 years.

Does my research data have copyright?

Probably not. Copyright does not cover facts and most research data is considered to be factual.

The key distinction comes from *Feist v Rural*, where the U.S. Supreme Court ruled that a telephone book is not original enough to have copyright. Hard work alone is not enough to guarantee copyright – there needs to be some creativity involved.

For example, if another researcher can follow scientific protocols reproduce your raw data, the data is not copyrightable. Additionally, where data analysis is scientifically driven – and not creatively driven – analyzed data also is not likely to qualify for copyright either.

Where data is made into figures for publication, the figures may have copyright even if the underlying data does not. Data compilations can also be creative enough to merit copyright even when member datasets are uncopyrightable.

How do I share my research data?

Since data rarely has copyright, the best way to share research data is with a public domain waiver such as Creative Commons Zero (CC0) (“CC0 use for data,” 2013). This recommendation is in accordance with the Panton Principles (Murray-Rust, Neylon, Pollock, & Wilbanks, 2010) and makes the copyright status clear, thus enabling easier reuse.

What rights do I have over my data?

If data is creative enough to merit copyright, the data creator has the exclusive rights to reproduce, sell, transfer copyright on, publically perform, and prepare derivative works of the data. This is the same as for any copyrighted work.

If your data does not qualify for copyright (and is thus public domain), you cannot exclusively control what others can do with your data if you publicly share it. You can, however, limit reuse by controlling how and if you share the data.

Patents more commonly apply in research than copyright. Patents give exclusive control over inventions whereas copyright gives control over how others can reuse your content (data). Data is still important to patents in that relevant data should be tightly controlled prior to filing for the patent.

Can I use someone else’s data?

Where data is licensed, preferably with CC0, you are free to use another researcher’s data in accordance with the license.

You can use public domain (uncopyrightable), factual data in your research, even if you did not generate the data. Note that contracts or licenses can supersede this allowance.

Fair use also allows for some reuse of copyrighted content, particularly for research purposes.

Be aware that good research practice requires data citation with reuse, no matter the status of the data’s copyright. More information on data citation is available from Data Pub (California Digital Library, 2014).

Resources

California Digital Library. (2014). Data Citation.
<http://datapub.cdlib.org/datacitation/>

CC0 use for data. (2013).
http://wiki.creativecommons.org/CC0_use_for_data

Murray-Rust, P., Neylon, C., Pollock, R., & Wilbanks, J. (2010). Panton Principles.
<http://pantonprinciples.org/>