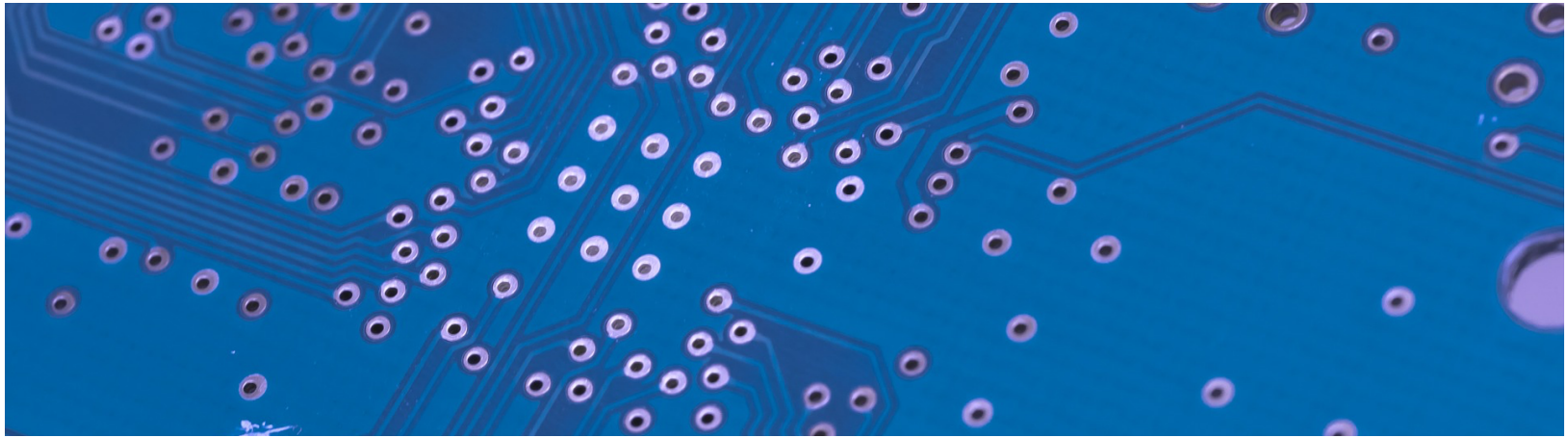
COPYRIGHT AND OPEN LICENSES FOR DATA



Anouk Santos, Data Manager Officer, University of Lausanne

COPYRIGHT: WHAT IS A PROTECTED WORK?

Three conditions must be reunited. A work is:

1. A creation of the mind,
2. Which has an individual character,
3. And is expressed in one form or another.

(Federal Act on Copyright and Related Rights and Ccdigitallaw 2020)

THE CASE OF FACTUAL SCIENTIFIC DATA

Copyright does not protect facts, information, ideas, formulas, algorithms, scientific measurements, etc.

Those type of data are not individual works of authorship but are discovered and compiled by a researcher's methods, something that copyright does not reward.

Therefore, factual scientific data are **not** protected by copyright.

(Pantaloni 2017)

LICENSES

- ✓ Compatible with data
- ✓ Quite simple to use
- ✓ Already widely used within the scientific field
- ✓ Well known (but not always fully understood)
- × not designed for software or computer code

THE OPEN DEFINITION

"Open data and content can be **freely used, modified, and shared** by **anyone** for **any purpose** (subject, at most, to requirements that preserve provenance and openness)."

(The Open Knowledge Foundation 2015)

ONLY 3 OPEN CC LICENSES

CC0

Data can be freely:
- ✓ used
- ✓ modified
- ✓ shared

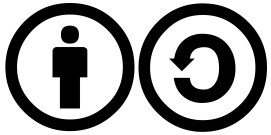
by anyone and for any purpose.

CC0 is in fact a Public Domain Dedication.

CC-BY

Data can be freely:
- ✓ used
- ✓ modified
- ✓ shared

by anyone and for any purpose, but:
- ❑ attribution is required

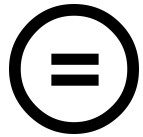


CC-BY-SA

Data can be freely:
- ✓ used
- ✓ modified
- ✓ shared

by anyone and for any purpose, but:
- ❑ attribution is required
- ❑ share alike is required

NO DERIVATIVES (ND): NOT OPEN



Data licensed with a ND element cannot be:

- Modified
- Combined or enriched with any other data
- Translated

→ A ND requirement forbids the creation of derivative works.

(Ball 2014 and Kreutzer 2014)

NONCOMMERCIAL (NC): NOT OPEN

What is a noncommercial use is not an easy notion to establish.

The NC requirement in a license could prevent:

- The use of data in a work for which the author receives a financial retribution (for example a published book)
- The use of data for the publication of an article in a journal owned by a commercial editor
- Public-private partnerships

(Ball 2014 and Kreutzer 2014)

SHARE ALIKE (SA):
OPEN BUT PROBLEMATIC

The problem is that copyleft licenses, like CC-BY-SA, are not compatible with other copyleft licenses:

A derivative work cannot respect both licenses

The SA element affects the interoperability of data and increases the incompatibility of licenses.

(Ball 2014 and Kreutzer 2014)

ATTRIBUTION (BY):
OPEN BUT PROBLEMATIC

The main problem is known as “attribution stacking”: citing all the authors of all the reused datasets can become progressively quite difficult.

Also, attribution cannot be legally binding by a license if the data are not protected by copyright… and therefore not eligible to licensing.

(Ball 2014)

CC0 (PUBLIC DOMAIN): THE BEST SOLUTION

- ✓ To solve the problem of licenses' incompatibility
- ✓ To achieve legal interoperability: CC0 is an answer to the ambiguity of data copyright
- ✓ Because there is a certain logic to put publicly funded data into the public domain
- ✓ **Because Open Science is easier to achieve with the least restrictions to impact data reuse**

(Fortney 2016, Lämmerhirt 2017, Labastida & Margoni 2020, Murray-Rust et al. 2010)

Thank you for your attention!

My master's thesis (in French) is available from: https://doi.org/10.5281/zenodo.3967402

You can reach me at: anouk.santos@unil.ch

REFERENCES 1/2

BALL, Alex, 2014. *How to License Research Data*. DCC How-to Guides. Edinburgh: Digital Curation Centre. Available from: http://www.dcc.ac.uk/resources/how-guides/license-research-data.

CCDIGITALLAW, 2020. 2.1 Protected work. *Ccdigitallaw*. 1st April 2020. Available from: https://ccdigitallaw.ch/index.php/english/copyright/2-what-work/21-urheberrechtlich-geschuetztes-werk

FORTNEY, Katie, 2016. *CC BY and data: Not always a good fit.* Office of Scholarly Communication (University of California). 15 September 2016. Available from: https://osc.universityofcalifornia.edu/2016/09/cc-by-and-data-not-always-a-good-fit/

KREUTZER, Till, 2014. *Open Content: A Practical Guide to Using Creative Commons Licences*. Bonn: German Comm. for UNESCO, 2014. Available from: https://meta.wikimedia.org/wiki/File:Open_Content_-_A_Practical_Guide_to_Using_Creative_Commons_Licences.pdf.

LABASTIDA, Ignasi and MARGONI, Thomas, 2020. Licensing FAIR Data for Reuse. *Data Intelligence*. January 2020. Vol. 2, n° 1-2, pp. 199-207. Available from: https://doi.org/10.1162/dint_a_00042

LÄMMERHIRT, Danny, 2017. Avoiding data use silos: How governments can simplify the open licensing landscape. *research.okfn.org*. December 2017. Available from: https://research.okfn.org/avoiding-data-use-silos/

REFERENCES 2/2

MURRAY-RUST, Peter, NEYLON, Cameron, POLLOCK, Rufus et WILBANKS, John, 2010. Panton Principles: Principles for Open Data in Science. *Panton Principles*. 19 February 2010. Available from: https://pantonprinciples.org/

OPEN KNOWLEDGE FOUNDATION, 2015. Open Definition 2.1. Available from: http://opendefinition.org/od/2.1/en/.

PANTALONI, Nazareth, 2017. Copyright and data curation. *Indiana University Libraries Blogs*. 12th December 2017. Available from: https://blogs.libraries.indiana.edu/scholcomm/2017/12/12/copyright-and-data-curation/.

SANTOS, Anouk, 2020. *Données de la recherche : cadre juridique et licences*. Geneva: Haute école de gestion de Genève. Master's thesis. Available from: https://doi.org/10.5281/zenodo.3967402

SWISS CONFEDERATION, 2020. Federal Act on Copyright and Related Rights (Copyright Act, CopA) of 9 October 1992. Status as of 1st April 2020. Available from: https://www.admin.ch/opc/en/classified-compilation/19920251/index.html

CREDITS

designed by fjstudio from Flaticon

designed by Freepik from Flaticon

designed by Smartline from Flaticon

designed by ultimatearm from Flaticon

Creative Commons logo, buttons and icons taken from:
https://creativecommons.org/about/downloads/