# FACT SHEET ON CREATIVE COMMONS & OPEN SCIENCE V.0.1

This information guide contains questions and responses to common concerns surrounding open science and the implications of licensing data under Creative Commons licences. It is intended to aid researchers, teachers, librarians, administrators and many others using and encountering Creative Commons licences in their work.

What is Open Science?

Open Science is the movement to make scientific research and data accessible to all for knowledge dissemination and public reuse.

How should I licence my data for the purposes of Open Science?

We recommend you use the CCO Public Domain Dedication, which is first and foremost a waiver, but can act as a licence when a waiver is not possible.

CC ZERO LICENCE, 'NO RIGHTS RESERVED' LOGO
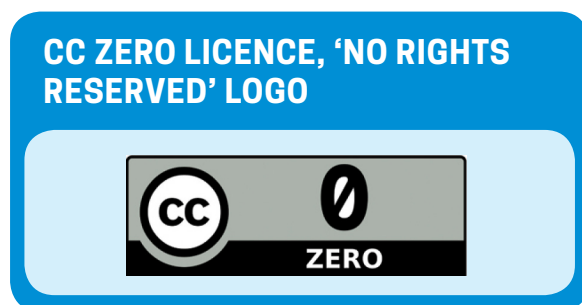
By applying CCO to your data you enable everyone to freely reuse your data as they see fit by waiving (giving up) your copyright and related rights in that data.

You should keep in mind that there are many situations in which data is *not* protected as a matter of law. Such data can include facts, names, numbers – things that are considered 'non-original' and part of the public domain thus not subject to copyright protections. Similarly, your database (which is a structured collection of data) might be considered 'non-original' and thus ineligible for copyright, and it might additionally be excluded
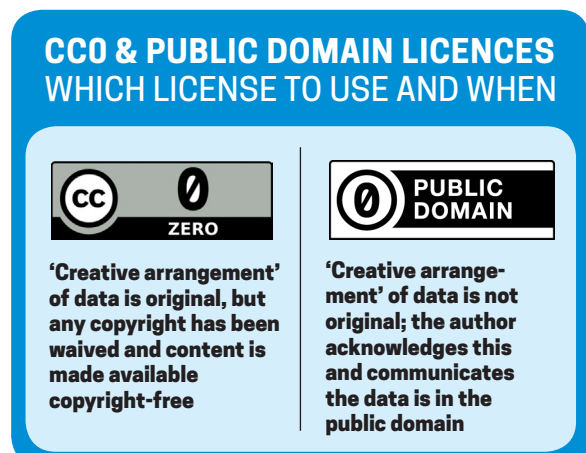
from other forms of protection (like the EU *sui generis* database right, also known as the 'SGDR', for non-original databases).

In these cases, using a Creative Commons licence such as a CC BY could signal to users that you claim a copyright in the non-original data despite the law, and perhaps despite your real intention.

Finally, if your data is in the public domain worldwide, you might state simply and obviously on the material that no restrictions attach to the reuse of your data and apply a Public Domain Mark.

PUBLIC DOMAIN MARK LOGO
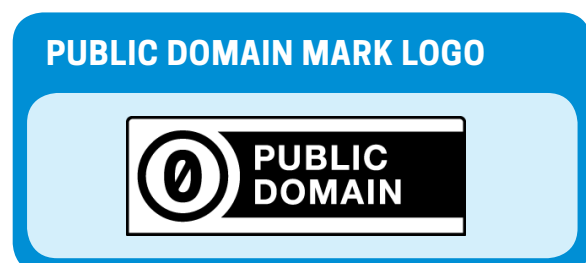
When in doubt, consider which use may be appropriate according to the chart below:

CCO & PUBLIC DOMAIN LICENCES
WHICH LICENSE TO USE AND WHEN

'Creative arrangement' of data is original, but any copyright has been waived and content is made available copyright-free	'Creative arrangement' of data is not original; the author acknowledges this and communicates the data is in the public domain

But I would like attribution when others use my dataset. In that case, shouldn't I use a CC BY licence?

We recommend that you avoid using a CC BY licence. Here's why:

While attribution is a genuine, recognisable concern, not only might using a CC BY licence be legally unenforceable when no underlying copyright or SGDR protects the work, but it may also communicate the wrong message to the world. A better solution is to use CC0 and simply ask for credit (rather than require attribution), and provide a citation for the dataset that others can copy and paste with ease. Such requests are consistent with scholarly norms for citing source materials.

Legally speaking, datasets that are *not* subject to copyright or related rights (and are thus in the public domain) cannot be the object of a copyright licence. Despite this, agreements based in contract law may be enforceable. Creative Commons licences, however, are copyright licences. Therefore, where the conditions for a copyright or related right are not triggered, copyright licences, such as the CC BY licence, are unenforceable.

In some cases, however, rights may exist (like the *sui generis* database right previously mentioned), and permission for others to use your dataset may be legally required. These rights are meant to protect the maker's investment, rather than originality. As such, database rights do not include the moral right of attribution. So by using a CC BY licence, you signal to users that you restrict access to your dataset beyond the protections provided by the law. We are not saying that this cannot be done, we are just saying that if you choose to do this, you should make sure you fully understand what it entails.

I'm uncomfortable with others using my research for commercial purposes. Should I use a non-commercial licence for my dataset?

We recommend you avoid using a non-commercial licence. Here's why:

For legal purposes, drawing a line between what is and is not 'commercial' can be tricky; it's not as black and white as you might think. For example, if you release a dataset under a non-commercial licence, it would clearly prohibit an organisation from selling your dataset to others for a profit. However, it might also prohibit someone using the dataset in their research if they intend to eventually publish that research. This is because most academic journals are commercial businesses that charge some sort of fee for access to their content, hence, such use could qualify as 'commercial'. Consequently, using a non-commercial licence prevents researchers from using your data in work destined for publication. This can subsequently affect the dissemination, recognition, and impact of your dataset.

BECAUSE WHO DOESN'T LOVE A GOOD VENN DIAGRAM?

NON-COMMERCIAL USE · COMMERCIAL USE · ACADEMIA
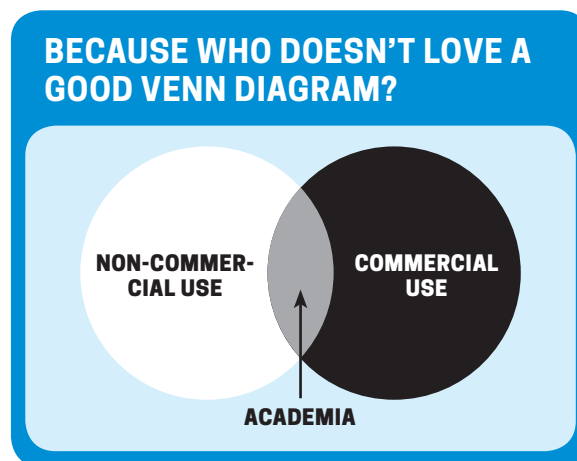
Please also consider that the current definition of 'Open Access' in the relevant international declarations states that limiting reuse to non- commercial activities does *not* comply with 'Open Access' (see the Berlin Declaration, Bethesda Statement on Open Access Publishing, and Budapest Open Access Initiative).

Ultimately, the decision is yours. However, the better open science practice is to avoid restricting use of your dataset to only non-commercial use.

I'm uncomfortable permitting use of my research for any and all purposes. Should I use a 'No Derivatives' (ND) licence for my dataset?

We recommend you avoid using a 'No Derivatives' licence. Here's why:

Similar to how a non-commercial licence might restrict meaningful reuse of your dataset, a ND licence can have the same effect: it may prevent someone from recombining and reusing your data for new research. For data to be truly Open Access, it must permit these important types of reuse.

What happens if I use 'Share Alike' (SA) licensed material in my work? Does that mean I have to make my work available under the same SA licence?
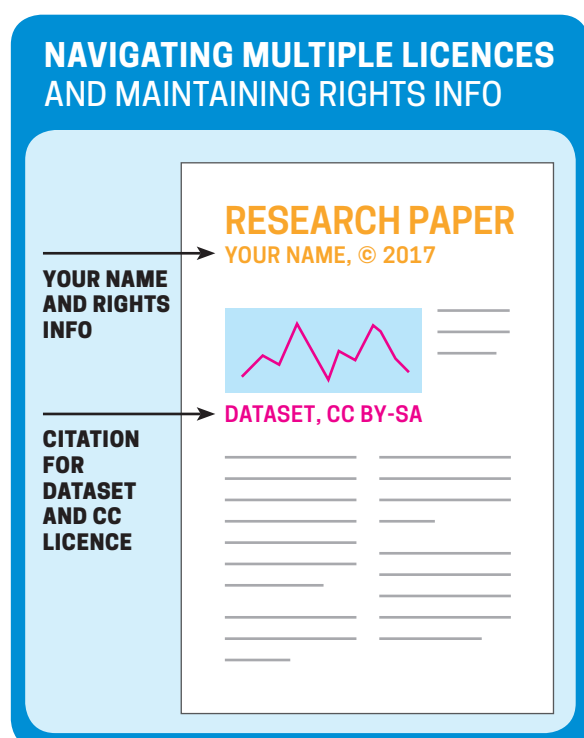
Not necessarily, but it depends on how you use the SA licensed content.

A 'Share Alike' CC licence applies only to the content licensed as SA that you have used. It does not require you to also make your work available under a SA licence, so long as you have not combined the independent works into one new work (known as a 'derivative' work).

When using SA content in your work, be sure to maintain the SA licensing information in regards to the content used. This can be done by providing the SA licensing information next to the content in your work and by designating it as SA when listing the other restricted content in your rights statement.

For example, if you include a CC BY-SA dataset in your research, you do not have to licence the entire body of work under a CC BY-SA, but the CC BY-SA dataset must retain the original licence. However, if you create a new dataset by combining two existing datasets, one of which belongs to you and the other is licensed under a CC BY-SA, then the new work (a derivative work) must be licensed CC BY-SA.

We understand that might be confusing, so here's an illustration to help:

NAVIGATING MULTIPLE LICENCES
AND MAINTAINING RIGHTS INFO

YOUR NAME AND RIGHTS INFO

RESEARCH PAPER
YOUR NAME, © 2017

DATASET, CC BY-SA

CITATION FOR DATASET AND CC LICENCE

It sounds like you're really pushing for the use of CC0 for open science datasets.

Exactly. Data is only open if anyone is free to use, reuse, and distribute it. This means it must be made available for both commercial and non-commercial purposes under non-discriminatory conditions that allow for it to be modified.

When data is made available for all reuse, others can create new knowledge from combining it. This leads to the enrichment of open datasets and further dissemination of knowledge. Accordingly, CC0 is ideal for open science as it both protects and promotes the unrestricted circulation of data.

And remember, it's bad science not to cite the source of data you use. To help others cite your data include a citation that users can copy and paste to give you credit for your hard work.

For example, the citation for this document is:

'Fact Sheet on Creative Commons and Open Science', Creative Commons UK, DOI: 10.5281/zenodo.840652, CC BY 4.0, https://creativecommons.org/licenses/by/4.0/

After reading this document, should you still wish to use CC BY make sure to include the citation for your dataset so others may cite your work with ease.

The font used is Cooper Hewitt, an open source typeface designed by Chester Jenkins and commissioned by the Cooper Hewitt museum.

Citations and other useful links

» 'About the Licenses', Creative Commons, https://creativecommons.org/licenses/, CC BY 4.0

» 'Berlin Declaration on Open Access to Knowledge in the Science and Humanities', Open Access Max-Plank-Gellschaft, 22 October 2003, https://openaccess.mpg.de/Berlin-Declaration

» 'Best practices for attribution', Creative Commons, https://wiki.creativecommons.org/wiki/Best_practices_for_attribution, CC BY 4.0

» 'Bethesda Statement on Open Access Publishing', Scientists and Scientific Societies Working Group, 20 June 2003, http://legacy.earlham.edu/%7Epeters/fos/bethesda.htm

» 'Budapest Open Access Initiative Declaration', Budapest Open Access Initiative, 14 February 2002, http://www.budapestopenaccessinitiative.org/, CC BY 3.0

» CC0 1.0 Universal Public Domain Dedication, Creative Commons, https://creativecommons.org/publicdomain/zero/1.0/

» 'CC Factsheet', Creative Commons, https://wiki.creativecommons.org/wiki/CC_Factsheet, CC BY 4.0

» Definition of Budapest compliant open access, Open Access Working Group, https://access.okfn.org/definition/2/, CC BY 3.0

» 'Europeana Publication Policy v1.0: A guide for aggregators and data providers on metadata and content requirements', Europeana, 20 November 2014, http://pro.europeana.eu/files/Europeana_Professional/Aggregation/Europeana%20-%20Publication%20Policy.pdf, CC BY-SA

» 'Frequently Asked Questions', Creative Commons, https://creativecommons.org/faq/, CC BY

» 'Global Open Access Portal', United Nations Educational, Scientific and Cultural Organization, http://www.unesco.org/new/en/communication-and-information/portals-and-platforms/goap/open-science-movement/

» 'How to Cite Datasets and Link to Publications', Alex Ball & Monica Duke, DDC How-to Guides, Edinburgh: Digital Curation Centre, 2015, http://www.dcc.ac.uk/resources/how-guides/cite-datasets

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

» 'Licensing Open Data: A Practical Guide', Naomi Korn and Prof Charles Oppenheim, June 2011 version 2.0, http://discovery.ac.uk/files/pdf/Licensing_Open_Data_A_Practical_Guide.pdf, CC BY-NC-SA

» 'The Problem of the Yellow Milkmaid: A Business Model Perspective on Open Metadata', Harry Verwayen, Martijn Arnoldus and Peter B. Kaufman, November 2011, http://pro.europeana.eu/files/Europeana_Professional/Publications/Whitepaper_2-The_Yellow_Milkmaid.pdf, CC BY-SA

» 'Promoting your articles to increase your digital identity and research impact', Jon Tennant, March 2017, http://blog.scienceopen.com/2017/03/promoting-your-articles-to-increase-your-digital-identity-and-research-impact/

» 'Public Domain Mark', Creative Commons, https://creativecommons.org/choose/mark/, CC BY

» 'Safe to be Open: Study on the Protection of Research Data and Recommendations for Access and Usage', Nil Dietr, Lucie Guibault, Thomas Margoni, Krzysztof Siewicz, Gerald Spindler, Andreas Wiebe, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2391280, CC BY 4.0