LIBRARIES & COLLECTIONS

KING'S College LONDON

Licensing your data

Why should I license my data?

Applying a licence to your data allows you to set terms and conditions governing what others can do with your data. There are a variety of licences available, ranging from public domain licences which waive copyright ownership and give permission for others to access and reuse the data without restriction, to licences that retain copyright ownership for the data creator and allow access and reuse under certain conditions.

If you are depositing your data in a repository or data centre you will probably be required to select a licence for your data as part of the deposit process, so it is worthwhile familiarising yourself with the various kinds of licenses and their conditions so that you can feel confident you are applying the licence that best meets your requirements.

What licence should I use?

Public domain dedications:

A public domain dedication allows others to access and reuse your data without restriction. Although attribution is not required under these licences, it is expected that the usual academic protocols of attribution and referencing will be followed.

Creative Commons Zero Public Domain Dedication (CC0)

Open Data Commons Public Domain Dedication and Licence (PDDL)

Creative Commons licences:

Alternatively, you can license your data for public access and reuse but only under certain conditions e.g. the data creator should be acknowledged, the data cannot be used for commercial purposes, the data cannot be altered, or the data must be redistributed using the same licence:

Creative Commons Attribution (CC BY) - lets others distribute, remix, tweak, and build upon the work, even commercially, as long as they credit the licensor for the original creation.

More information about the full range of creative commons licences is available on the Creative Commons web site and our copyright for academics and staff FAQ.

Open Data Licences:

A standard Creative Commons licence might not always suitable for databases; licences specifically designed for databases include:

Open Data Commons Attribution License (ODC-By) allows licensees to copy, distribute and use the database, to produce works from it and to modify, transform and build upon it for any purpose.

Open Data Commons Open Database License (ODC-ODbL) is the same as above but has a share-alike condition so that any adaptations must be offered under the same license.

Open Government License allows users to copy, publish, distribute, adapt, and exploit the data for commercial purposes but attribution is required. A non-commercial variation of the license is also available.

Licences for source code and software:

The GNU General Public License 3.0 and the MIT Licence are popular open licences for software.

Help with choosing a software licence is available from GitHub.

Licensing considerations

Irrevocability:

You cannot withdraw a Creative Commons licence once it has been issued.

Third Party Rights:

Before choosing a licence ensure that ownership of the data has been clarified and that any third party rights are not being infringed. Only the rights holder or someone with permission or authority to act on their behalf can grant a licence. It is therefore essential that any intellectual property rights relating to the data are established before a licence is issued.

Creative Commons licences might not always be suitable for datasets that contain third party rights materials as additional clearances might be required.

Creative Commons licences do not provide any guarantees regarding information about third party materials contained in the dataset, nor any indemnities for the user in the case that they do - leaving the licensee to take all the risk.

Attribution Stacking:

If a dataset is at the end of a long chain of derivations, or if large teams of contributors were involved, the list of credits might well be considered too unwieldy.

Although using a public domain dedication (e.g. CC0) can be used to prevent attribution stacking, attribution can be important as a means to recognise both the source and the authority of the data.

Sources: Licensing Open Data: A Practical Guide (Naomi Korn, and Charles Oppenheim, HEFCE, 2011); How to License Research Data (Alex Ball, Digital Curation Centre, 2014).

Help with choosing a licence

Creative Commons have created an online tool to help choose a licence. The Ufal Public License Selector is another useful online tool that can help you select the most appropriate licence for your data.

How do I apply a licence to my data?

Assigning a licence to your data is generally very straightforward, and guidance is usually available on the licence provider web site.

For example:

Guidance on marking works with a Creative Commons licence, including datasets, is available from the Creative Commons Wiki.

Help with applying the Open Data Commons Attribution License is available from the Open Data Commons.

Further reading:

Digital Curation Centre - Creative Commons Licensing

Creative Commons - Data