

Description

Currently, researchers are finding that most funding agencies, organizations, and some professional journals, require data sharing as a condition of grants, awards, or for publication. Data sharing is beneficial because it allows researchers to promote their research and attain prompt dissemination of their results, which enables them to further their research and discovery efforts. In general, data is basically factual information that cannot be copyrighted. However, for data that is collected, collated, or manipulated using a significant investment of time or resources, copyrighting might be appropriate.

Researchers share data when they email it to colleagues or publish it on a project website. In addition, many researchers submit data to institutional repositories or data centers for archiving, curation, and access. However, simply releasing data without making clear the terms of use is counter-productive. This is where data licensing comes in.

A data license is a legal instrument that allows the researcher to authorize a second party to use the data in a manner that would otherwise infringe on the rights held. Thus, licenses typically grant permissions based on the idea that certain terms have to be met. The exact details of data licenses can vary, but in general, three conditions commonly found in licenses are attribution, copyleft, and non-commerciality. All of these conditions are meant to protect the researcher and their data.

Types of Licenses

Researchers are currently faced with the challenge of understanding the different licenses that are available. Here we provide details on some of the [common types of licenses](#) that can be used.

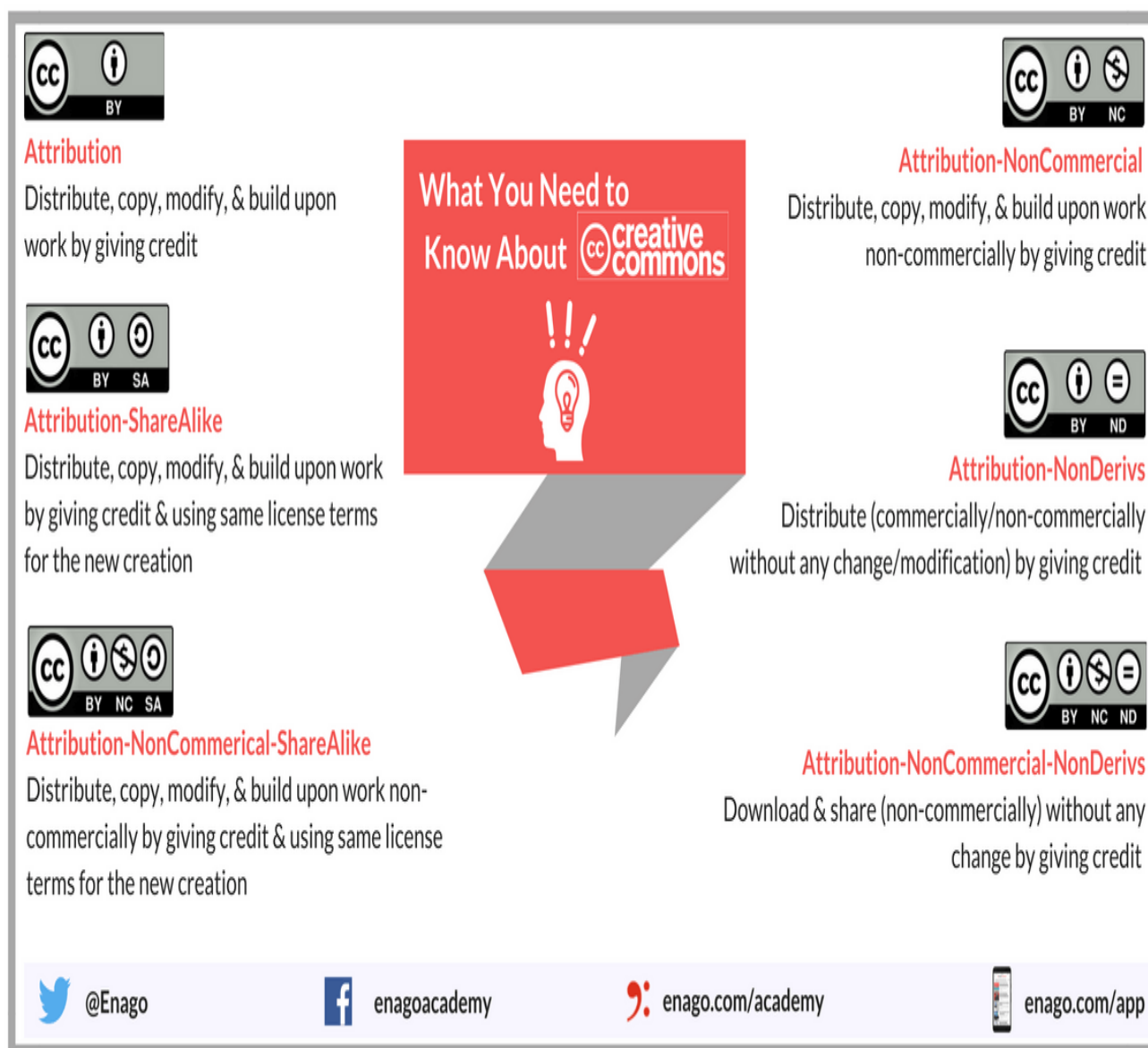
Prepared licenses: Often a researcher's department or institution already has a license prepared that researchers can apply to their data. These prepared licenses can be both at the institutional level and as public domain, such as in the case of genome data. Ultimately, when these are available to a researcher there is less work involved in obtaining the data license.

Bespoke licenses: These are custom licenses that are not easy to prepare. These are not commonly used, but when there is a significant commercial value associated with the data or the researcher needs to clarify his or her responsibilities and those of the re-users with respect to the data, then a custom license is often necessary.

Standard licenses: These are the most commonly used licenses, as most research projects are better served by using one of the standard licenses. In the next section, we describe these in more detail.

Types of Standard Licenses

1. Creative Commons was established in 2001 as a non-profit corporation with the sole purpose of producing simple yet robust licenses for creative works, including research data licensing. These licenses allow the researchers to have detailed control over how they can be used, as opposed to simply declaring them in the public domain or reserving all rights. These licenses are easy to decipher because they all have brief clear summaries and a canonical URL for use in HTML, RDF, and other code. There are six types of [Creative Commons](#) licenses. Each license includes the Attribution condition, which makes it unique. They are as follows: Attribution (CC BY); Attribution Share Alike (CC BY-SA); Attribution No Derivatives (CC BY-ND); Attribution Non-Commercial (CC BY-NC); Attribution Non-Commercial Share Alike (CC BY-NC-SA) and Attribution Non-Commercial No Derivatives (CC BY-NC-ND).
2. Open Data Commons is a project that was established in 2007. The first license it produced was a public domain dedication for databases. The project was transferred to the Open Knowledge Foundation in 2009 and now has two additional licenses that are somewhat similar to the nature of the Creative Commons licenses, but are designed specifically for databases. All three, similar to the Creative Commons model, provide a clear summary and canonical URL alongside the full legal text.
3. Open/Non-Commercial Government License is part of the UK Government Licensing Framework and was released in September 2010. It is made for the use of the UK public sector and government resources, and is particularly for datasets, source code, and collected or original information. For this license, attribution is required, derivative works and commercial uses are clearly allowed, and there is no copyleft condition. This license does not allow the use of personal information, unpublished information, public sector logos, armorial bearings, military insignia, identity documents, information subject to patents, trademarks, design rights, third party copyright (unless authorized), etc.
4. Public domain is probably the most liberal way of releasing data. In this manner, all copyright interests and database rights are waived, allowing free use of the data. Using a public domain for licensing is not simple, and thus, Creative Commons and [Open Data Commons](#) provide special tools for this purpose. Releasing data to the public domain means that the researcher permanently relinquishes numerous rights and protections, including protection against unfair competition. Thus, for researchers still interested in exploiting their data academically or commercially, this might not be the best option.



Tools for Selecting the Appropriate License

Before a researcher considers the licensing options that are available, first, they should determine whether there is an obligation or a strong encouragement to use a certain license as a condition of funding or as a matter of local policy. This can save a great deal of time. If a researcher is in a position where they have to decide which license to use, then we propose these tools for doing so:

- The [License Chooser](#) tool at Creative Commons
- Gain a full understanding of all types of licenses.
- The Open Data Commons

Category

1. Publishing Research
2. Understanding Ethics

Date Created

2017/06/20

Author

daveishan