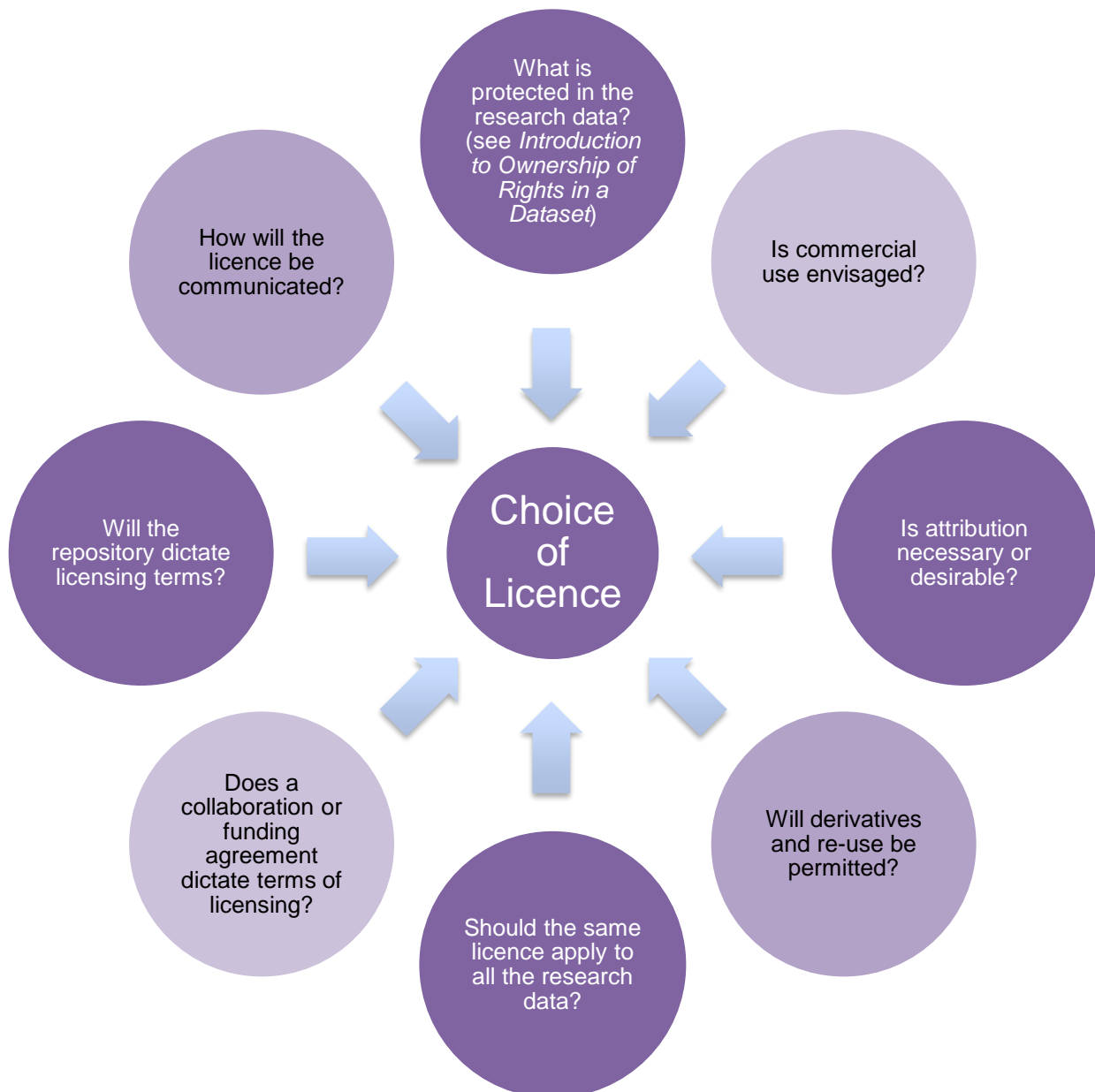# INFORMATION GUIDE (3)
CHOOSING A LICENCE FOR RESEARCH DATA

This information guide contains general responses to common concerns about licensing of research data. It is intended to aid researchers, teachers, librarians, administrations and many others using and encountering the need to make research data available under an Open Access licence. It may be necessary to consult specialist sources for certain research data and further sources are given at the end. There is a *Glossary* of terms used in this guide.

Factors to consider in choosing a licence for research data

What is protected in the research data? (see *Introduction to Ownership of Rights in a Dataset*)

How will the licence be communicated?

Is commercial use envisaged?

Will the repository dictate licensing terms?

Choice of Licence

Is attribution necessary or desirable?

Does a collaboration or funding agreement dictate terms of licensing?

Should the same licence apply to all the research data?

Will derivatives and re-use be permitted?
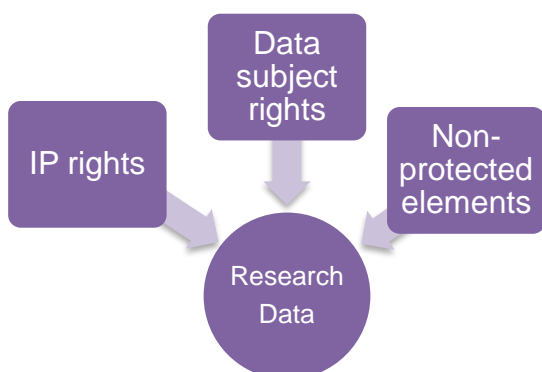
What is a licence?

A licence is a statement in which a legal rightsholder gives permission to use something protected by law under certain conditions. It does not diminish the ownership of the rightsholder, but may give others the ability to exploit and use the thing licensed. Licences can be negotiated directly and be bespoke, or a unilateral standard licence might be used that all potential users have to accept to use the research data.

There are many types of licence – permissive, restrictive, open access. Without a licence, the default "All rights reserved", i.e. that it is very difficult to make use of a protected work or database because of the automatic **copyright** and/or **sui generis**

> Copyright and database rights that arise automatically in research data are explained in our information Guide *Introduction to Ownership of Rights in Research Data.*

database rights (SGDR) in the research data.

There are other rights that may arise in research data (or sub-sets of research data) which may influence licensing decisions, or which may require alternate terms to be imposed to protect these other rights.

Examples of licensing
The licence for some of the most common operating systems on a tablet or phone restricts copying, redistribution, or modification of the licensed work. On the contrary, Wikipedia uses the licence (CC BY) that tells users that they can copy, modify and redistribute the original work under certain obligations such as the attribution of the original author.

A licence allows the owner of a database or work to set the terms on which the user may use the subject matter. These might be very broad and essentially allow any re-use. Or these might impose certain conditions such as the need to credit (**attribute**) the author of the work(s), restrict the ability to repurpose, **remix** or create new research data (**derivatives**) or control redistribution of the work for commercial purposes.

Within the **open access** framework there are a number of standard licences that are commonly used, which will be discussed in this guide. (More restrictive licensing is also used, although this tends to be bespoke to the particular research data.)

> **What is Open Access?**
> Open Access is the movement to make research and knowledge freely available through the Internet for public reuse, with the only constraints the right to preserve the integrity of the research and the right to be properly acknowledged. For more, see the *Budapest Declaration on Open Access*, 2002.

Publicly funded research institutions (and funders) are now committed to open access, and this has been included in the next UK Research Excellence Framework Exercise, creating an obligation on researchers in these settings to make their research data publicly available, unless legal or ethical issues prevent this.

Comparing Open Access licences

(An Open Access licence is one which meets the Budapest Declaration standard, making the licensed work available only subject to restrictions to preserve the integrity of the work and an obligation of attribution)

	Use permitted?	Derivatives permitted?	Attribution required?	Derivatives required to use same licence?
CC.0	✓	✓		
PPDL	✓	✓		
OGL 3.0	✓	✓		
CC-BY v4.0	✓	✓	✓	
ODC-BY 1.0	✓	✓	✓	
CC-BY-SA	✓	✓	✓	✓
ODC-ODbL	✓	✓	✓	✓

Creative Commons licences

Creative Commons, a global not-for-profit organisation which provides legal tools to promote the legal sharing and reuse of intellectual endeavours, has produced a number of licences which meet the criteria for Open Access. These offer different levels of permission.

Interoperability and stacking
When choosing a licence for protected research data, it is necessary to consider how this might interact across the whole research data (including non-protected elements) and how it might interact with future derivative works produced from re-use or re-purposing. OpenMinTeD has a useful matrix presenting the compatibility of different licences. Using two or more licences may require **stacking** of attribution of rights in the licensed work.

Not all licences can accommodate different layers of protected work. Creative Commons licences are better at this than OGL 3.0. If research data is multi-faceted choice of licence needs carefully considered.

Creative Commons offers licences readable at three different levels: legal, machine (the metadata that the data is tagged with to convey the licence) and human (non-legal descriptions). Creative Commons has a useful tool to help you determine the licence

CC0.
This is a waiver, in which the creator or author waives the copyright, SGDR and some other rights in the work so others have the ability to re-use, re-mix and share the work without any restrictions. Attribution of the research data is not required.

CC BY 4.0
This is another Creative Commons licence that permits the work to be shared and adapted in a permissive way but with the condition that the author is always attributed. CC BY 4.0 complies with Open Access, and is useful if attribution is important to the research data owner(s). Version 4.0 of CC BY is best suited to research data.

CC BY-SA 4.0
This is more restrictive as it requires **attribution** of the author & imposes a requirement that if any **derivatives** are created from the original they are also shared under a CC BY-SA licence. A CC BY-SA grants rights to share, use and adapt the work, subject to these conditions. Share Alike requirements can inhibit the **interoperability** of data.
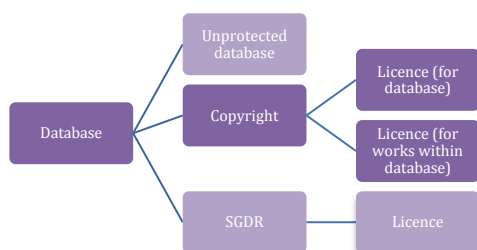
best for you. More restrictive CC licences are unlikely to meet Open Access requirements (e.g. because they impose restrictions on creating derivatives or on commercial use).

Other Open Access licences for research data

For researchers in the **public sector** the **Open Government Licence (OGL) v3.0** is a licence developed by the UK Government for **public sector information**. This encourages use and re-use of the licensed work in a free and flexible manner, similar to Creative Commons licence. It requires users to acknowledge the source of the licensed work, using either a format specified in the licence of the generic description "contains public sector information licensed under the Open Government Licence 3.0." This licence does not allow for granularity within the research data. As a researcher you are welcome to use data under the OGL, but you should not license your data under the OGL as it is specifically designed for Government activities.

Open Data Commons is another not-for-profit that has produced standard licences designed specifically for research data.

Open Access licences for databases

- Referring to the licence on the landing page or host site for a digital research data;
- Attaching the licence to the metadata of the research data.

Creative Commons licences are drafted to contemplate different layers of protection within the licensed work, which is useful for research data, which can be a 'bundle' of rights.

Creative Commons Licences compatible with Open Access

ODC provides a **Public Domain Dedication and Licence (PPDL)** for those seeking the most permissive solutions. This dedicates works to the public domain, which means they can be copied, shared, re-used and adapted, for commercial and non-commercial purposes, with no conditions attached.

If you are seeking to licence open source software, GitHub provide a comprehensive guide to open source licensing options.

How are licences applied to research data?

Licences are not automatic. The owner of rights in research data must make it clear that a licence is imposed, or that it is dedicated to the public domain. Repositories may specify the licence applied to data deposited in their repository. Since the RCUK Concordat suggests the default licence should be an open access licence, choice of repository should be considered carefully if more restrictive licence is required.

Applying a licence can happen by: -

- Stipulating the licence terms within the research data;

- Setting up a Read Me file for the data.

For more see the Digital Curation Centre Guide to mechanisms for licensing data. If a standard licence from Creative Commons or

Open Commons is used, they will have [tool](#) to help attach the licence effectively.

For more information on attaching licences, and on metadata, see the Research Data Alliance [Metadata Standards Directory](#) or UK Data Service information on attaching **metadata** [here](#).

Disclaimers

An area of concern for many researchers is whether they will incur liability for re-use and re-purposing of their data. For example, if research data contains data relating to architecture or construction, could the researcher be liable if that data was re-purposed to construct a building that later collapsed?

Standard licences usually offer a clause about limitations and disclaimers, which make it clear that the research data is shared 'as is' and that the original researcher is not responsible for the outcomes of re-use and re-purposing. If a bespoke licence is used, limitation of liability clauses must be drafted to cover this.

Other considerations in licensing

Research ethics – what is expected in relation to making data available?
Dynamic data – is an interim licence needed for initial results?
Termination – most standard licences are irrevocable.
Enforcement - It is usually up to the owner of the protected rights in research data to be vigilant in monitoring use and taking action to enforce the terms of the licence if infringement is discovered.

More Information Guides
- *Introduction to Ownership of Rights in Research Data*
- *Making Research Data Available*
- *FAQ: Using Research Data*

Legal sources
- Copyright, Designs and Patents Act 1988
- Directive 96/9/EC of 11 March 1996 on the legal protection of databases

Sources on licensing
- [Creative Commons](#)
- [Open Data Commons](#)
- [Github – Choosealicense.com](#)
- [CopyrightUser.org](#)
- [UK Data Service](#)
- [OpenMinTeD](#)
- [OpenAIRE](#)
- [EUDAT](#)