



Virtual Physiological Human Network of Excellence

Grant Agreement: 223920

VPH ToolKit Guideline Document

Topic: Licensing

Version 1.0

04-Feb-2011

This page is intentionally blank

Document Information

IST Project Num	FP7 – 2007 - ICT - 223920	Acronym	VPH NoE
Full title	Virtual Physiological Human Network of Excellence		
Project URL	http://www.vph-noe.eu		

Document	Number	G07	Title	Guidance (Licensing)
-----------------	---------------	-----	--------------	----------------------

Status	Version 1.0	Final <input checked="" type="checkbox"/>
Dissemination Level	Public <input checked="" type="checkbox"/> Consortium <input type="checkbox"/>	

Authors (Partner)	UCL	Jacovella (now at QMU)		
	UOXF	Cooper		
	INRIA	Bleuzé		
	UPF	Martelli		
	USFD	Fenner, McCormack, Varma		
Responsible Author	Cooper		Email	jonathan.cooper@comlab.ox.ac.uk
	Partner	UOXF	Phone	

Abstract (for dissemination)	This document provides guidance on the licensing considerations for VPH ToolKit content.
-------------------------------------	--

The information in this document is provided as is and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. Its owner is not liable for damages resulting from the use of erroneous or incomplete confidential information.

Version Log			
Issue Date	Version	Author	Change
06-Jul-10	0.1	UOXF	First draft outline for comments
27-Jul-10	0.2	UOXF	Incorporate suggestions from UPF and material from UOXF and USFD
08-Dec-10	0.3	UOXF	Incorporate suggestions from INRIA and material on data and content licensing
14-Jan-11	0.4	UOXF	Final draft for internal review
27-Jan-11	0.5	UOXF	Incorporate material from UCL and final suggestions
04-Feb-2011	1.0	UOXF	First public release

This page is intentionally blank

Table of Contents

EXECUTIVE SUMMARY	8
Introduction	9
Guideline Topic: Licensing – an introduction	10
<i>Software licensing</i>	10
<i>Data licensing</i>	10
<i>Content licensing</i>	11
Underlying Concepts.....	12
<i>Software licensing</i>	12
<i>Data and content licensing</i>	13
Interactions and Dependencies.....	17
Applicable Legislation.....	18
Standards and Standards Bodies	19
Characteristics	20
<i>Recommended software licenses</i>	20
AGPL – The GNU Affero General Public License v3	21
Apache v2 – The Apache Licence (v2)	22
3-clause BSD – The 3-clause BSD Licence	22
4-clause BSD – The 4-clause BSD Licence	23
CeCILL v2 – The CeCILL Licence.....	23
EURL – The European Union Public Licence	24
GPL v2 – The GNU General Public Licence version 2.....	24
GPL v3 – The GNU General Public Licence version 3.....	25
LGPL v2.1 – The GNU Lesser General Public Licence version 2.1	25
MIT – The MIT Licence.....	26
MPL v1.1 – The Mozilla Public License 1.1.....	26
<i>Data and content licensing criteria</i>	27
<i>Recommended data licences</i>	28
Open Data Commons Public Domain Dedication and Licence (PDDL).....	28
Open Data Commons Attribution License (ODC-By).....	28
Open Data Commons Open Database License (ODbL)	29
Creative Commons CCZero (CC0)	29
<i>Recommended content licences</i>	30
Creative Commons Attribution licence family.....	30

Creative Commons CCZero (CC0)	30
GNU Free Documentation License (GFDL)	30
<i>Other “licences”</i>	31
Public domain	31
Implicit copyright	31
<i>Licence compatibility</i>	32
Methods of Verification	35
Training	36
Maintenance	37
Conclusions and Recommendations	38
Submitting Content to the ToolKit	39
Further Information	40
References	41

This page is intentionally blank

EXECUTIVE SUMMARY

This guideline is intended to assist potential contributors to the VPH ToolKit in preparing their content for submission. There are several such guideline documents in this series, covering the full range of issues affecting content providers. They are being developed over a period of time and, once finalised, these guidelines may be bound together into a single VPH-NoE resource.

A common issue for scientists in respect of interoperability is dealing with intellectual property considerations, particularly data and software licensing. The requirements of the host organisation(s) involved need to be taken into consideration, as well as the licences of any software libraries used in a software product. These topics need to be carefully considered when deciding how to release data or software, and under what licence. They are also worth considering when beginning any project, since decisions made earlier in development may have a significant impact on any later decision to release.

While much of the focus of this document is on software licensing, licensing of data and content (e.g. the documentation for a software tool) is no less important. However, historically, licensing associated with data sharing has received less attention, and so this field is less developed. The legal issues are also less clear. It is still crucial, however, that publicly released data and databases have suitable licence terms attached, particularly for the avoidance of doubt among users. It is a sad irony that, at a time when we have the technologies to permit global access to and distributed processing of scientific data, legal restrictions are making it harder to connect the dots.

(Note that licensing of VPH models encoded in a markup language is an emerging aspect of this topic. It is still being debated within the community, and so coverage of this issue has been deferred to a subsequent version of this document.)

This document provides some guidance on licensing issues, with links to further information. However, although recommendations are presented, this document was not produced by lawyers, and should not be considered as legal advice. It is merely a statement of our opinions. We make no warranties regarding the general legal information provided, and disclaim liability for damages resulting from its use. Please consult a lawyer to resolve your specific licensing and intellectual property issues.

The VPH-NoE advocates the use of a business-friendly open source licence whenever possible, in order to promote the wide usage and ongoing development of VPH tools. We further recommend selecting a licence from the list discussed in this document if possible. This is to keep the number of licences used by the VPH community manageable, and avoid the legal complexities that arise from building on software with an unusual licence. Multiple licences may be applied to permit alternative usage scenarios (e.g. to provide a commercial licence for a fee).

For data, we also advocate the use of an open license to encourage sharing and reuse, since innovation in science is best served by access to information. However it should be noted that considerations for clinical data in particular may preclude this in some cases.

This document itself is licensed under the open [Creative Commons Attribution-ShareAlike 3.0 Unported licence](#).

Introduction

This document is one of a series that together build to form a complete guide to the ideal content and presentation of materials offered for distribution via the Virtual Physiological Human Network of Excellence ToolKit Portal. The full set of Guideline Documents is summarised below.

Guidance Area	Description
Tool characterisation	The attributes important for inclusion in the documentation of Tools, including performance validation
Model characterisation	The attributes important for inclusion in the documentation of Models, including performance validation
Data characterisation	The attributes important for inclusion in the documentation of Data
Ontological annotation	Methods of knowledge representation, in particular the significance, benefits and methods of ontological annotation of ToolKit content
Interoperability	Key attributes and methods for enabling ToolKit content to be utilised in concert within a multistage workflow
Ethico-legal issues, provenance	The inherited responsibilities that are attached to any item of ToolKit content – perhaps particularly data – including legal, ethical and territorial restrictions
Licensing	The conditions that apply to the legitimate use of the content from an intellectual property standpoint
Usability and training	The factors that are important for the easy use and ready acceptance of ToolKit content, taking into account the environment, the likely users and the need for interoperability. Additionally, the nature of training facilities of all types appropriate to particular content categories.

Guideline Topic: Licensing – an introduction

Software licensing

A common and hugely important, although frequently misunderstood, issue for developers of scientific software is dealing with intellectual property considerations, in particular software licences. When you write software, you are creating a kind of property, which has a defined owner (typically yourself or your employer, although even this question can be complex). As property, computer software is automatically protected by copyright law, which gives the owner of the property (typically called a ‘work’ in this context) certain rights and makes it illegal for others to use the work as though they were its owner; action through the courts can be taken to enforce this. By default, only the owner can copy, adapt, or distribute the work. As soon as the software is ‘fixed’ – that is, recorded in some fashion, such as being saved to computer disk – copyright comes into being automatically; there is no need to register the work or state your copyright claim.

Hence, if anyone is to use, distribute, or extend the software that you create, explicit permission must be given to allow them to do so. Such permission usually takes the form of a software licence. In choosing an appropriate licence, many factors must be taken into consideration, including establishing the owner of the software, recognition of licences for any component parts used, and any potential exploitation plan for the software. The choice of licence should ideally be made at the outset of a project’s development, before other choices restrict the options available, or even make a release impossible without redevelopment work.

Data licensing

The situation as regards individual items of data, or collections thereof (e.g. in a database) is more complicated, and depends significantly on legal jurisdiction. Without any explicit license, it could be unclear what rights users actually have. Historically this has not been such an issue, since data was rarely shared except with collaborators. However a wider sharing of data is now becoming common, inspired by high visibility examples such as the Human Genome Project, and an increasing view that the results of publicly funded research, including the data collected or created, should be made openly available. Many funders and journals also now require that raw data is made available. This movement aims to make scientific data into a ‘commons’ – a resource collectively owned by and available to the scientific community, and indeed the general public. The benefits of this sharing, and of increased availability of medical research data, are well documented and include replication of previous findings, comparisons with independent datasets, testing of additional hypotheses, teaching, and patient safety.¹

Data licensing is important because it reduces uncertainty, explicitly clarifying the circumstances under which users can actually use the data, and for what purposes. Such clarification allows researchers to focus on more important matters.

¹ <http://www.bmj.com/content/340/bmj.c181.full>

Content licensing

Other content, for example documentation or videos, is also covered by copyright law, and thus its use is restricted by default. While it will often be distributed along with software, software licences were not written with such content in mind and so do not apply very well in these cases – different types of subject matter necessitate differences in licensing. A suitably open licence for documentation associated with open source software is particularly important, as it allows anyone who modifies the software to update the documentation accordingly.² Licences can also be applied more generally to any content, allowing authors to define how their works can be used.

This document provides some guidance on all these issues, primarily within the 'Characteristics' section (see p20), and links to further information. However, we reiterate that this document was not produced by lawyers, and should not be construed as legal advice. It is merely a statement of our opinions. Please consult a lawyer to resolve your specific licensing and intellectual property issues.

² See also the Free Software Foundation article at <http://www.gnu.org/philosophy/free-doc.html>.

Underlying Concepts

Software licensing

As described in the previous section, computer software is legally considered to be property, and is always protected by copyright law (unless the rights so granted are specifically disclaimed, for example by releasing the software into the public domain). This has important consequences for those wishing to use, modify, or distribute software – they need permission from the owner of the software in order to do so.

Identification of the owner is not always straightforward. If the software has been developed collaboratively by multiple individuals or organisations, there may be a complex shared ownership which can make it difficult or even essentially impossible to release the software. It is therefore crucial that agreements or contracts specify who will own the intellectual property that results from any collaboration, consortium or contract work.

Once ownership is determined, the domain of software use should be established. If it is only to be used within the context of current or future grant agreements, in which the relevant legal contracts cover permitted usage and development, then there is no need for a software licence. For a work to be included in the VPH ToolKit, however, it must be released to a wider audience, and hence must be appropriately licensed. A poor choice of licence can significantly restrict the utility of the software.

In order to promote the openness of the VPH community, and the widest possible use of VPH ToolKit content, the VPH NoE advocates the use of an Open Source licence whenever possible. The term “Open Source” describes a group of licences that all meet a certain set of conditions. The conditions are maintained by a group called the Open Source Initiative, and together they are referred to as the Open Source Definition (OSD). An open source licence must:

- grant the licensee the right to distribute the program themselves, including the right to charge money for it;³
- grant access to the program's source code;
- grant the right to modify the program;
- grant the right to distribute modified versions of the program;
- allow use of the program by all persons or groups in all fields of endeavour;
- apply to everyone who receives the program, without the need for any additional agreements;
- apply to the program it licenses, whether the program is obtained as part of a group of programs, or on its own;
- allow distribution with any other software; and

³ It is a common misconception that one cannot charge money for distributing open source software. This is not the case. In practice, it is rarely done, mainly because each one of your customers could give away the software in direct competition to your own sales.

- allow distribution in any form.

The desired effect of these conditions is to promote wide distribution of software, and to encourage people who receive the software to contribute to its functionality by modifying the source code. Although it might seem that granting these rights could lead to a proliferation of variants of the software, in practice successful open source software projects tend to absorb disparate modifications made by many contributors back into a single modified and improved version.

It cannot be stressed strongly enough that decisions about licensing should be taken at the very start of development, within the context of a wider discussion on exploitation. If this is not done, it may not be possible to release the software. Apart from potential difficulties in agreeing ownership, the licences used for components of the software may be incompatible – unfortunately, not all the source code that is available under an open source licence can be adapted and combined without restriction in order to produce new open source software. Two licences that each meet the requirements of the Open Source Definition individually may nevertheless contain terms that make them incompatible with each other. We give our opinion on which licences are compatible later in this document. For these reasons, all projects that produce software need to keep complete, detailed records of the licensing and ownership of contributions to that software, and all developers need to be aware of the licensing policy and its implications.

Further information, with excellent guidance on how practically to release software under an open source licence, listing the issues to consider, can be found at <http://www.oss-watch.ac.uk/resources/opensourceyourcode.xml>.

Finally, it is prudent to be aware of dual or multiple licensing. This involves releasing your software under more than one licence, and giving users the choice of which licence they adhere to. This can be a useful staging post towards an open source release, allowing software to be made available under an open licence, but also keeping a proprietary version that can be licensed commercially. It can also be useful in circumventing incompatibilities between open licences, allowing developers of derived works more choice in their licensing. OSS-Watch have [a useful guide explaining more about this topic](#).

Data and content licensing

Parallel to the open source movement for software, the concept of open data is emerging. This is part of a wider *open knowledge* field. The term knowledge is taken to include content such as music, films, or books, as well as data be it scientific, historical, geographic or otherwise.

Openness and licensing for data are crucial to enabling progress in the VPH field, and to scientific progress more generally, by allowing the interoperability of data.^{4,5} The volume of scientific data, and the interconnectedness of the systems under study, make integration of data a necessity. Open data available under a permissive licence is much easier to break-up and recombine, to use and reuse, and hence much more valuable for the common good. The technical challenge of such integration is itself significant, although emerging technologies

⁴ <http://blog.okfn.org/2009/02/02/open-data-openness-and-licensing/>

⁵ <http://sciencecommons.org/projects/publishing/open-access-data-protocol/>

are helping. But the forest of terms and conditions around data make integration difficult to perform legally in many cases, and the law varies much more widely between different parts of the world than is the case for software. In many areas, a database may not be openly reusable without an explicit licence granting permission.

When talking about databases we first need to distinguish between the structure and the contents of a database. Copyright law will generally cover the structure of the database, but may not cover the contents as a collection. Individual contents may be covered depending on their nature. For example, the contents of a database listing the melting points of various substances would not be copyrightable, since they are “facts”. Forms of protection for the contents as a collection fall under copyright law in some jurisdictions, and/or a “sui generis” right for collections of data may exist. Both apply within the EU. Contract law may also be used to protect (closed) databases by providing access only to registered users who have agreed to particular terms and conditions. For further information see <http://www.opendefinition.org/guide/data/>.

These variations can create uncertainty or practical difficulty for those wishing to share databases and their underlying data, but retain a limited amount of rights under a “some rights reserved” approach to licensing as outlined in the Science Commons Protocol for Implementing Open Access Data.⁶

Differences in the natures of these various types of knowledge (software code, content, or data) necessitate differences in licensing, taking account of the differing usages and legal frameworks. Specific differences between data and content include:⁷

- differing legal rights as described above;
- the need to distinguish between a database and its contents, and to ensure both are appropriately licensed;
- the distinction between a database and content generated from it (for example, a geospatial database used to create a map), which is different from the relationship between a database and a derivative database (e.g. a database adding post office locations to the original); and
- the prominence of such derivative works – unlike content (but like code) databases have a high level of reuse.

In order to enhance the utility of data and content shared within the VPH, the VPH-NoE advocates the use of a licence compliant with The Open Knowledge Definition (OKD; <http://www.opendefinition.org/okd/>). Developed by the Open Knowledge Foundation (OKF; <http://okfn.org/>), this outlines the principles that define open knowledge, and by which to judge whether a knowledge license is open. Many of the principles were inspired by the OSD.

Under the OKD, an open data or content licence must, in brief:

- ensure that the work is made available as a whole in a convenient and modifiable

⁶ <http://sciencecommons.org/projects/publishing/open-access-data-protocol/>

⁷ See <http://www.opendatacommons.org/faq/licenses/#why-not-use-a-creative-commons-or-freeopen-source-software-license-for-databases> from which this list is taken.

form;

- allow others to redistribute the work either on its own or as part of a collection;
- allow for modifications and derivative works to be made and distributed;
- ensure the work is provided in such a form that there are no technological obstacles to the performance of the above activities;
- not discriminate against any person or group of persons;
- not restrict anyone from making use of the work in a specific field of endeavour (e.g. commercially);
- apply to all who receive the work without the need for any additional agreement;
- not depend on the work being part of a particular collection of works; and
- not place restrictions on other works that are distributed along with the licensed work.

In addition, an OKD compliant licence may:

- require as a condition for redistribution and re-use the attribution of the contributors and creators to the work, but not in an onerous fashion; and
- require as a condition for the work being distributed in modified form that the resulting work carry a different name or version number from the original work.

In addition, and following the Science Commons Protocol, it is considered that the most suitable approach to achieving both simplicity and low transaction costs for users of databases, and legal certainty, may be to release data into the public domain, while also specifying non-enforceable community norms to express the wishes of the data provider. This is not a ubiquitous view however.⁸

Clinical data, of particular importance to the VPH, poses additional questions that must be considered when looking to release and licence data. Many of the issues are primarily related to ethical considerations, for example the need to de-identify datasets and obtain appropriate ethical approval, ideally with informed consent from participants. These topics are discussed in the accompanying Ethico-legal Guideline. [Hrynaszkiewicz et al. \(2010\)](#) also provide practical guidelines and suggest a minimum standard for de-identification. They note that “restrictions on access to certain aspects of data may be warranted, such as when removal of information that could identify the data would negate its scientific value. In circumstances where data must be behind a barrier to universal access, the data could be made accessible only to those who agree to certain conditions of use, and to individuals who meet certain professional criteria. Embargoes on access to data could also be applied.” Such considerations require more specific licensing arrangements than can be covered in this document, and legal advice should be sought. Other further information on open access to biomedical data can be found via BioMed Central.⁹

There are several good guides available that give further information about data licensing, and how to license works. The Creative Commons provide an article

⁸ See e.g. <http://blog.okfn.org/2009/02/02/open-data-openness-and-licensing/>

⁹ <http://blogs.openaccesscentral.com/blogs/bmcblog/category/Open+Data>

(<http://creativecommons.org/about/think>) which describes how to use one of their licences and what needs to be thought about beforehand. The Open Knowledge Foundation also provides a guide to open licensing (<http://www.opendefinition.org/guide/>).

Interactions and Dependencies

Three other documents in this series of ToolKit Guidelines are particularly closely connected with this Guideline.

The issues around data licensing are clearly tightly coupled with the nature of the data itself. Preparing data for sharing is discussed in detail in the Data Characterisation Guideline.

Copyright law is not the only legislation that may apply to computer software or data. Other areas of ethical and legal concern are addressed in the Ethico-legal Guidelines document.

Finally, the software licence used is a key characteristic of each VPH tool. Such matters are discussed further in the Tool Characterisation Guideline.

Applicable Legislation

The VPH ToolKit Guidelines cut across legal issues that are subject to continuous review and modification at national, European and global level. It is impractical to provide an exhaustive and definitive reference here. Rather, we seek to provide non-legal researchers an indication of possible legal issues affecting their work, with references to the most important current legislation. We also focus on legislation affecting the European Union, since this is the sphere most closely associated with the VPH.

The most represented law in this document is copyright law, the key concepts of which have been described above. Contract law and international treaties also play a role to regulate interactions between researchers and other parties. Examples include the World Trade Organisation's Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS)¹⁰ and the Berne Convention for the Protection of Literary and Artistic Works¹¹.

Two European Directives¹² are of particular relevance. The first is Directive 2001/29/EC (22 May 2001) on the 'harmonisation of certain aspects of copyright and related rights in the information society'. This aims to bring consistency to national copyright legislation across the EU.

There is also a database-specific 'Database Directive': Directive 96/9/EC on the 'legal protection of databases'. It provides for both copyright and the sui-generis right though with some restrictions on when you can use the copyright (old common-law jurisdictions and many others allowed copyright in simple data no matter how 'unoriginal'; the EU directive requires authorship involving personal intellectual creativity).

¹⁰ See http://www.wto.org/english/thewto_e/whatis_e/tif_e/agrm7_e.htm

¹¹ See http://en.wikipedia.org/wiki/Berne_Convention_for_the_Protection_of_Literary_and_Artistic_Works

¹² Such directives are subject to national ratification before becoming law. Those mentioned here are widely ratified however.

Standards and Standards Bodies

The primary standards body for software licensing is the Open Source Initiative (OSI; <http://www.opensource.org/>). The OSI is a not-for-profit corporation formed to educate about and advocate the benefits of open source and to build bridges among different constituencies in the open-source community. Its role as a standards body is to maintain the Open Source Definition for the good of the community. The OSI Approved License trademark and program is designed to create a nexus of trust around which developers, users, corporations and governments can organize open-source cooperation.

Free and open source software has also been greatly championed by the Free Software Foundation (FSF; <http://www.fsf.org/>), which has a separate Free Software Definition, and promotes “copyleft” licensing in particular. A copyleft licence requires that derivative works, if released, are made available under an open source licence, typically the same as the original.¹³ It thus propagates open source, reusable software. The FSF sponsors the GNU project, and is responsible for the GNU family of open source licences.

The Open Knowledge Foundation (OKF; <http://okfn.org/>) is a not-for-profit organization founded in 2004 and dedicated to promoting open knowledge in all its forms. It developed the Open Knowledge Definition, and was initially responsible for the Open Data Definition. Its major function is acting as a hub for work on open knowledge, drawing together representations from across the community in various working groups and projects.

Creative Commons (<http://creativecommons.org/>) is a not-for-profit corporation dedicated to making it easier for people to share and build upon the work of others, consistent with the rules of copyright. It has developed the Creative Commons family of open content licences.

The Science Commons initiative (<http://sciencecommons.org/>) has grown out of Creative Commons, and designs strategies and tools for faster, more efficient web-enabled scientific research. It identifies unnecessary barriers to research, crafts policy guidelines and legal agreements to lower those barriers, and develops technology to make research, data and materials easier to find and use. Its goal is to speed the translation of data into discovery — unlocking the value of research so more people can benefit from the work scientists are doing.

¹³ See <http://www.gnu.org/copyleft/>

Characteristics

In this section we discuss what characteristics high quality ToolKit content should have in the realm of licensing. **The key point to note is that each entry submitted to the ToolKit must indicate the name(s) and version(s) of the licence(s) under which it is released.** Characterisation then consists of characterising the features of the various licences.

This document contains a summary of each of a selection of the most popular open licences for software, data, and content. In the case of software licences, these are already available as options in the 'licence' field of the VPH ToolKit Portal. We recommend choosing a licence from this list, in order to keep the number of licences used by the VPH community manageable, and avoid the legal complexities that arise from building on work with an unusual licence.

If ToolKit content is released under a licence not in our list, further details of that licence must be provided when submitting the entry. At a minimum this should consist of a link to the licence document. However, we recommend also providing an informative summary of the licence in the same manner as given here. Furthermore, any supporting documentation should be indicated, for example details of the licences of all third-party components used.

The recommended licenses are presented below, under the headings:

- Recommended Software Licences
- Recommended Data Licences
- Recommended Content Licences
- "Other" Licences

Recommended software licenses

The first distinguishing characteristic of software licences we consider is whether or not the licence is open source. This is preferred, for the reasons discussed above. If software cannot be made available under an open source licence, being freely available for academic use is highly recommended.

Each open source licence in this document is characterised according to five key criteria:

- Whether or not it has been approved by the Open Source Initiative.
- Whether the licence is "business friendly" - that is, whether software released under the licence can be incorporated into a closed source software product which is then released under a commercial licence.
- Whether the licence is compatible with the GNU General Public Licence (GPL). This is important because the GPL is a viral "copyleft" licence, and so if GPL-licensed code is incorporated within your software, your software can then only be released under the GPL. If other components are released under licences incompatible with the GPL, it then becomes impossible to release your software. The GPL is only compatible with licences whose restrictions are subsets of those in the GPL. Note that this means that a GPL-compatible licence will not be "business friendly". At the end of this section we have a table indicating which licences we believe to be compatible with each other.

- Attribution – whether you need to acknowledge the work in any derivative of it that you release. Most open source licences require preservation of copyright, permission, and warranty disclaimer notices.
- The legal jurisdiction specified in the licence, if any. Some people are reluctant to use the GPL because it is unclear whether all its conditions are enforceable outside the USA. Also, at least one UK university is known to be averse to using the Mozilla Public Licence (MPL), since it specifies that it shall be governed by California law provisions, and they are reluctant to adopt the legal risk associated with this.

The following software licences are recommended and described in this document:

- The GNU Affero General Public License v3
- The Apache Licence (v2)
- The 3-clause BSD Licence
- The 4-clause BSD Licence
- The CeCILL Licence
- The European Union Public Licence
- The GNU General Public Licence versions 2 and 3
- The GNU Lesser General Public Licence version 2.1
- The MIT Licence
- The Mozilla Public License 1.1

AGPL – The GNU Affero General Public License v3

- [OSS Watch information page](#)
- [Licence text](#)
- An OSI-approved Open Source licence
- Not business friendly - the GPL is a viral copyleft licence
- GPL v3 compatible
- Attribution: yes
- Legal jurisdiction: unclear; would probably need to be agreed between the parties in any dispute

This is an adapted version of the GNU General Public License v3, designed to better protect software whose chief use is providing useful services over a network.

As the provision of services over the web has become a bigger and bigger sector of the economy, some observers have become concerned that such service providers are reaping huge benefits from the use of free and open source software without returning anything to the community that created it. After all, such companies are not actually distributing their versions of the software in question, just exposing its functionality to users over networks like the web. As they are not distributing the software, they have no responsibility to make the source code to their version available. This has become known as the 'ASP (Application Service Provider) loophole'.

To address this, the AGPL stipulates that any adaptation of software that it covers must prominently offer its source code for download to users who interact with it over a network. It also explicitly permits software to be created by combining code which is under the GPL v3 and the AGPL v3. In this case the resulting software would be under a combination of both licences, with each section of code retaining its original licence. In all other respects the AGPL v3 is identical to the GNU GPL v3.

In summary, the AGPL

- closes the 'ASP Loophole' by mandating the delivery of source code by service providers
- ensures that modified versions of the code it covers remain free and open source.

Apache v2 – The Apache Licence (v2)

- [OSS Watch information page](#)
- [Licence text](#)
- An OSI-approved Open Source licence
- Business-friendly
- Compatible with GPL v3 (but probably not with GPL v2)
- Attribution: yes; see below
- Legal jurisdiction: unclear; would probably need to be agreed between the parties in any dispute

This is an extension of the 3-clause BSD licence. In brief a licensee of Apache Licensed v2 software can:

- copy, modify and distribute the covered software in source and/or binary forms
- exercise patent rights that would normally only extend to the licensor

provided that:

- all copies, modified or unmodified, are accompanied by a copy of the licence
- all modifications are clearly marked as being the work of the modifier
- all notices of copyright, trademark and patent rights are reproduced accurately in distributed copies
- the licensee does not use any trademarks that belong to the licensor

Furthermore, the grant of patent rights specifically is withdrawn if:

- the licensee starts legal action against the licensor(s) over patent infringements within the covered software

The licence also defines the notion of a *Contributor*, as distinct from someone who just modifies the software, who by default also grants a licence to their modification back to the original authors.

In summary, the Apache Licence v2

- explicitly grants patent rights where necessary to operate, modify and distribute the software
- permits code that it covers to be subsumed into closed source projects.

3-clause BSD – The 3-clause BSD Licence

This is also known as the modified or new BSD licence, to contrast it with the original 4-clause BSD licence.

- [OSS Watch information page](#)

- [Licence text](#)
- An OSI-approved Open Source licence
- Business friendly
- GPL compatible
- Attribution: yes; see below
- Legal jurisdiction: unclear; would probably need to be agreed between the parties in any dispute

This is one of the simplest open source licences, and is very permissive regarding what licensees are allowed to do. There are just 3 restrictions:

- redistributions of source code must retain the copyright notice, list of conditions and disclaimer
- redistributions in binary form must reproduce the copyright notice, list of conditions and disclaimer in the documentation and/or other materials provided with the distribution
- the names of the previous contributors are not used to promote any modified versions without their written consent

A variation on the 3 clause "modified" licence was recently approved by the OSI, containing only the first 2 clauses; this is referred to as the "simplified" BSD licence.

4-clause BSD – The 4-clause BSD Licence

- [OSS Watch information page](#)
- [Licence text](#)
- Not OSI-approved
- Business friendly
- Not GPL compatible
- Attribution: yes; original author must be acknowledged in all advertising
- Legal jurisdiction: unclear; would probably need to be agreed between the parties in any dispute

This original BSD license contained a clause not found in later versions of the licence, known as the "advertising clause":

- All advertising materials mentioning features or use of this software must display the following acknowledgement: This product includes software developed by (developer).

As reasonable as this might seem, it threatened to make the practice of aggregating open source software extremely impractical. It was officially [rescinded](#) by the Director of the Office of Technology Licensing of the University of California on July 22nd, 1999. There is still other software licensed under the old version of the BSD licence, and effort continues to contact the authors and persuade them to reissue their work under the revised licence.

Apart from the advertising clause, this licence is identical to the 3-clause BSD licence.

CeCILL v2 – The CeCILL Licence

- [Wikipedia information page](#)
- [Licence text](#)
- Not OSI-approved (has not been submitted for approval) but is approved by the Free Software Foundation

- Business friendly: depends on the variant
- GPL compatible: depends on the variant
- Attribution: yes (details depend on the variant)
- Legal jurisdiction: French Law, although another jurisdiction may be chosen by mutual agreement

Due to legal uncertainties about using the GPL outside the USA, three French public research organisations, the [CEA](#), the [CNRS](#), and [INRIA](#), launched the CeCILL project (from "**CEA CNRS INRIA Logiciel Libre**"), to develop open source licences in conformance with French law. They are also suited to wider international use.

The CeCILL licence is modelled after the GPL v2, and hence has essentially the same conditions. There are also variants of the CeCILL licence modelled after the 4-clause BSD and LGPL licences: CeCILL-B and CeCILL-C respectively.

EUPL – The European Union Public Licence

- [Wikipedia information page](#)
- [Licence text](#)
- An OSI-approved Open Source licence
- Not business friendly - the EUPL is a copyleft licence
- GPL v2 compatible
- Attribution: maintain all notices intact, and prominently display notification and dates of any modifications
- Legal jurisdiction: that of the EU country where the licensor resides (or is registered), or Belgian Law if the licensor is not in the EU

The EUPL is a software licence that has been created and approved by the European Commission. It is available in 22 official languages of the European Union, all with identical legal value. Its main goal is focusing on being consistent with the copyright law in the 27 Member States of the European Union, while retaining compatibility with popular open-source software licences such as the GPL v2. To achieve the latter aim, the EUPL specifically authorizes derived works including material under one of the following licences to be released under that licence:

- GPL v2
- [Open Software License](#) v2.1, 3.0
- [Common Public License](#) v1.0
- [Eclipse Public License](#) v1.0
- CeCILL v2

GPL v2 – The GNU General Public Licence version 2

- [OSS Watch information page](#)
- [Licence text](#)
- An OSI-approved Open Source licence
- Not business friendly - the GPL is a viral copyleft licence
- GPL compatible (naturally)
- Attribution: yes; all distributed copies (modified or not) must carry a copyright notice and exclusion of warranty
- Legal jurisdiction: unclear; would probably need to be agreed between the parties in

any dispute

The GPL v2, like nearly any licence, grants rights under certain provisos. These are briefly listed here. A licensee of GPL v2-licensed software can:

- copy and distribute the program's unmodified source code (Section 1)
- modify the program's source code and distribute the modified source (Section 2)
- distribute compiled versions of the program, both modified and unmodified (Section 3)

provided that:

- all distributed copies (modified or not) carry a copyright notice and exclusion of warranty (Section 1 and 2)
- all modified copies are distributed under the GPL v2 (Section 2)
- all compiled versions of the program are accompanied by the relevant source code, or a viable offer to make the relevant source code available (Section 3)

In summary, the GPL

- ensures that modified versions of the code it covers remain free and open source
- attempts to spread *copyleftism* by mandating the use of the GPL v2 for distributed adaptations of GPL v2-licensed code

GPL v3 – The GNU General Public Licence version 3

- [OSS Watch information page](#)
- [Licence text](#)
- An OSI-approved Open Source licence
- Not business friendly - the GPL is a viral copyleft licence
- GPL compatible (naturally)
- Attribution: yes; all distributed copies (modified or not) must carry a copyright notice and exclusion of warranty
- Legal jurisdiction: unclear; would probably need to be agreed between the parties in any dispute

More than sixteen years separate the GPL v3 and its predecessor, GPL v2. In this period, several perceived threats to the ideals of software freedom were identified, and the GPL v3 was drawn up to build on the enormous success of the GPL v2 while combating these.

Broadly speaking the threats were:

- 'Tivoisation' and Technological Protection Methods
- Unintentional incompatibility with some open source licences
- US-specific legal terminology in the GPL v2
- The rise of the web application as a means of realising value from software
- Software patent non-enforcement covenants as a means of dividing the free and open source software community

The essential aims of the GPL v3 are, however, unchanged from the GPL v2. It

- ensures that modified versions of the code it covers remain free and open source
- attempts to spread *copyleftism* by mandating the use of the GPL v3 for distributed adaptations of GPL v3-licensed code

LGPL v2.1 – The GNU Lesser General Public Licence version 2.1

This is a variation of the regular GNU General Public License (GPL). Originally known as the

GNU Library General Public License, it was drafted by the Free Software Foundation to provide a weaker (or Lesser) form of copyleft for use in certain specific circumstances.

- [OSS Watch information page](#)
- [Licence text](#)
- An OSI-approved Open Source licence
- Partially business friendly - software which uses a library licensed under the LGPL can be distributed with any licence, subject to a few restrictions
- [GPL compatible](#)
- Attribution: yes; all distributed copies (modified or not) must carry a copyright notice and exclusion of warranty
- Legal jurisdiction: unclear; would probably need to be agreed between the parties in any dispute

The LGPL v2.1 is identical to the GPL v2 in many of its provisions. The main difference is that where the GPL mandates that all derivative works be distributed under the GPL, if at all, the LGPL v2.1 defines a separate class of works which may be derivative but which nevertheless can be licensed in any way. These are referred to as *works that use the library*. In summary, the LGPL

- keeps modified versions of the library itself open source
- allows non-open source software to use the library, and be distributed with it

MIT – The MIT Licence

- [Wikipedia information page](#)
- [Licence text](#)
- An OSI-approved Open Source licence
- Business friendly
- GPL compatible
- Attribution: yes
- Legal jurisdiction: unclear; would probably need to be agreed between the parties in any dispute

The MIT Licence is a free software licence originating at the Massachusetts Institute of Technology (MIT). It is similar to the 3-clause BSD licence, except that the BSD licence contains a notice prohibiting the use of the name of the copyright holder in promotional materials.

MPL v1.1 – The Mozilla Public License 1.1

- [OSS Watch information page](#)
- [Licence text](#)
- An OSI-approved Open Source licence
- Business friendly
- Not GPL compatible
- Attribution: yes
- Legal jurisdiction: California law

The MPL differs most significantly from the GPL in the way that it governs how works containing the MPL-licensed code can be distributed and licensed. Put simply, someone can take an MPL-licensed work and build upon it with new components. The resulting work can

be distributed with the MPL covering the use of the original work and any licence covering the rest. Clearly in this way a company could add closed source components to an MPL-licensed work and thus build a proprietary product.

The MPL also divides the granting of rights into two sections - one concerning the grant of rights by the code's initial author, and another concerning the grant of rights by other people who have added code to the initial author's work. The initial author grants these rights:

- to use, reproduce, modify, display, perform, sublicense and distribute the source, and modified versions of the source;
- patent rights to use and make available the original code (where relevant);
- to distribute works which contain the code in combination with new code, and to license the new code in any way the distributor wishes.

The contributor grants these subtly different rights:

- to use, reproduce, modify, display, perform, sublicense and distribute the source of their modifications;
- patent rights to use and make available both the modifications and the entire work (original code plus modifications);
- to distribute works which contain the code in combination with new code, and to license the new code in any way the distributor wishes.

Both these grants are subject to the following conditions:

- all distributed copies (original or modified) must include the source code or advice on how to obtain the source code;
- all modifications are described in accompanying documentation;
- any patent rights necessary to operate the software are clearly described in accompanying documentation;
- all copies of the code, original or modified, have a statement of copyright and an exclusion of warranties attached;
- all modified files must be distributed under the MPL. New files containing entirely new code need not be distributed under the MPL.

In summary, the MPL

- explicitly grants patent rights where necessary to operate the software;
- keeps the covered code itself open source;
- allows extensions of the code to be licensed in non-open ways.

Data and content licensing criteria

The 5 key criteria described above for software licences have some parallels in the data and content world.

- Whether or not it conforms to the Open Knowledge Definition.
- Whether it allows commercial use.
- Whether it is a viral copyleft licence, i.e. it includes a 'Share Alike' clause restricting creators of derivative works to the same licence. Such clauses also affect licence compatibility as detailed at the end of this section.
- Attribution – whether you need to acknowledge the work in any derivative of it that

you release.

- The legal jurisdiction specified in the licence, if any.

Recommended data licences

The following data licences are recommended and described in this document:

- Open Data Commons Public Domain Dedication and Licence
- Open Data Commons Attribution License
- Open Data Commons Open Database License
- Creative Commons CCZero

Open Data Commons Public Domain Dedication and Licence (PDDL)

- [Licence text](#) and [information](#)
- Conforms to the Open Knowledge Definition
- Allows commercial use
- Is not viral
- Does not require attribution
- Legal jurisdiction is intended to be universal

The PDDL allows others to use the data or database in any way they wish, with no restrictions, insofar as this is permitted by law. It is intended to cover all copyright and database laws, so that however database rights are protected (under copyright or otherwise), those rights are all waived. However, it does not affect trademark and patent rights. If waiver of rights is not possible within a particular jurisdiction, the PDDL acts as a licence giving all applicable rights to the user.

A particular feature of the ODC licences is that they allow a Community Norms document¹⁴ to be associated with the database. This has no legal standing, but sets out explicitly principles (norms) for those that share data. It allows the sharer of the database to specify the desired code of conduct for the community.

Open Data Commons Attribution License (ODC-By)

- [Licence text](#) and [information](#)
- Conforms to the Open Knowledge Definition
- Allows commercial use
- Is not viral
- Requires attribution
- Legal jurisdiction is that in which the License terms are sought to be enforced

This ODC licence permits any use of the database as a whole, providing suitable attribution is made for any public use of the database, works produced from the database, or redistribution.

¹⁴ Such as this one drawn up by the ODC: <http://www.opendatacommons.org/norms/odc-by-sa/>

Databases can contain a wide variety of types of content (images, audiovisual material, and sounds all in the same database, for example), and so this license only governs the rights over the database itself, and not the contents of the database individually. Licensors may therefore wish to use this license together with another license for the contents.

Open Data Commons Open Database License (ODbL)

- [Licence text](#) and [information](#)
- Conforms to the Open Knowledge Definition
- Allows commercial use
- Is a viral copyleft licence
- Requires attribution
- Legal jurisdiction is that in which the License terms are sought to be enforced

This shares the same provisions as the ODC-By licence, except that it adds two further restrictions. The first is a ‘share-alike’ clause – if a user publicly uses any adapted version of the database, or a work produced from an adapted database, then they must release that adapted database also under the ODbL. Also, while they are free to use technological protection mechanisms, they must also release a version of the adapted database that does not use them.

There is an associated “Database Contents License” which can be applied to the contents of the database if desired (the ODbL only applies to the database itself), or another licence (or licences) may be used for the contents.

The conditions imposed by this licence depend on the licensor either having IP rights over the material (whether in copyright, or database right), or the license being considered a contract between the licensor and the user of the database. In some jurisdictions, e.g. the EU, the relevant “DB” rights certainly do exist, while in others, e.g. the USA, the situation is less clear. Where neither “DB” rights nor a contract exists no license will be enforceable. If this is of concern to you your only real alternative is not to make the database available.

Creative Commons CCZero (CC0)

- [Licence text](#) and [information](#)
- Conforms to the Open Knowledge Definition
- Allows commercial use
- Is not viral
- Does not require attribution
- Legal jurisdiction is intended to be universal

A person using CC0 (called the “affirmer” in the legal code) waives all of his or her copyright and neighbouring and related rights in a work, to the fullest extent permitted by law. If the waiver isn’t effective for any reason, then CC0 acts as a license from the affirmer granting the public an unconditional, irrevocable, non exclusive, royalty free license to use the work for any purpose. The work is thus effectively placed in the public domain, insofar as this is possible for a given jurisdiction. Note that CC0 is intended to cover all copyright and database laws, so that however database rights are protected (under copyright or otherwise), those rights are all waived. However, it does not affect trademark and patent rights.

Recommended content licences

The following content licences are recommended and described in this document:

- Creative Commons Attribution licence family
- Creative Commons CCZero
- GNU Free Documentation License

Creative Commons Attribution licence family

- [Licence texts](#) and information
- Conforms to the Open Knowledge Definition, unless “Non-commercial” or “No derivatives” are selected
- Allows commercial use, unless “Non-commercial” is selected
- Is not viral, unless “Share-Alike” is selected
- Requires attribution
- Legal jurisdiction may be left unspecified, or chosen from a long list

Creative Commons have developed a range of licences allowing authors to choose how many restrictions they apply, while still maintaining some copyright over the work. Copyright owners may select from 4 features they wish to include or not, giving a total of 6 different licences. All these licences require attribution – giving the original author credit for the work. The other 3 features are as follows.

- Share-Alike – others may distribute derived works, but only under identical licence terms.
- Non-commercial – others may copy, distribute, display, and perform the work, or make derivative works, but only for non-commercial purposes.
- No derivatives – others may copy, distribute, display, and perform the work, but not create derivative works based on it. (This option is thus incompatible with Share-Alike.)

All the licences include a human-readable summary of the main points of the licence, as well as the full legal document. Creative Commons also works to “port” the core Creative Commons Licenses to different copyright legislations around the world – both linguistically translating the licenses and legally adapting them to particular jurisdictions.¹⁵

At least two previous open content licences have now been deprecated in favour of one of the CC licences.

Creative Commons CCZero (CC0)

See above under “Recommended data licences” for details.

GNU Free Documentation License (GFDL)

- [Licence text](#) and information
- Conforms to the Open Knowledge Definition, with some provisos.
- Is a viral copyleft licence

¹⁵ See <http://creativecommons.org/international/>

- Requires attribution
- Legal jurisdiction is unspecified

The creators of the license recommend that it is used “principally for works whose purpose is instruction or reference,” such as software manuals and documentation or textbooks. Due to the requirement to include the entire licence text within the licensed document, it is not very suitable for short works.¹⁶

The GFDL is only considered to confirm to the OKD if the following provisos are met:¹⁷

- no invariant sections or cover texts;
- no “Acknowledgements” or “Dedications” section;
- the DRM restriction (section 2) is amended to be less broad.

Other “licences”

Public domain

- [Wikipedia information page](#)
- Not a licence as such (but see also the [CC0](#) and [PDDL](#) licences)
- Business friendly, suitable for commercial use
- GPL compatible, not a viral copyleft licence
- Attribution: no
- Legal jurisdiction: probably that in which the work was created

A work released to the public domain is essentially no longer subject to copyright law, although exact details vary from country to country. In some cases works automatically enter the public domain after a period of time has elapsed (e.g. 70 years). Authors may also explicitly release a work into the public domain (for example by using the CC0 licence), and so allow free use of the work in any context.

Implicit copyright

- Not business friendly
- Not GPL compatible
- Attribution: yes
- Legal jurisdiction: probably that in which the work was created

Any work released without an explicit license is automatically afforded implicit protection under copyright law. This applies even without an explicit copyright notice anywhere within the work. Without specific permission from the author, the user cannot utilise the unlicensed work within any project. The user must ask permission from the copyright owner to release the work under a suitable license or provide a written statement of declaration of right specifically to the user.

¹⁶ See for example <http://www.oss-watch.ac.uk/resources/relicensing.xml>

¹⁷ See <http://www.opendefinition.org/licenses/gfdl/>

Licence compatibility

As described above, two licences *A* and *B* may contain clauses which render them incompatible with each other, in the sense that you cannot combine (portions of) two works, *a* and *b*, licensed under *A* and *B* respectively, within the same derivative work, and then release the derivative work. This situation most commonly occurs when making use of software libraries. A similar issue may apply when combining data from multiple sources to produce a new resource (e.g. a “mash-up”).

The following table gives our interpretation of open source licence compatibility, for the software licences described on this site. The entries state whether you can release a derived work combining two works licensed under the licences in question. The items in brackets indicate the licence under which a combined work must be released, if there is any restriction.

For data and content licences, we believe the situation to be more straightforward. The only conflict which may arise (at least for the licences described in this document) is if different licences which both include ‘share-alike’ terms, requiring the use of the same licence for derivative works, are used. In this case, but no other, creating a work which derives from both sources is impossible. Note that in the case of the ODC-ODbL, the ‘share-alike’ terms include provision for the use of ‘compatible licences’ for derived works, in which the decision of which licences are compatible rests with the licensor. Hence the incompatibility may not be insurmountable in this case.

Licence	GPL v2	GPL v3	LGPL v2.1	3-clause BSD	4-clause BSD	MPL v1.1	CeCILL v2	CeCILL-B v2	CeCILL-C v2	EUPL	AGPL	Apache v2
GPL v2	n/a	Yes (GPL) {3}	Yes (GPL) {3}	Yes (GPL)	No	No	Yes (GPL)	No	No	Yes (GPL)	No	No
GPL v3		n/a	Yes (GPL) {3}	Yes (GPL)	No	No	Yes (GPL)	No	No	No	Yes (Combined AGPL & GPL v3)	Yes (GPL)
LGPL v2.1			n/a	Yes (LGPL) {1}	No {1}		Yes (CeCILL) {1}	No {1}	Yes (LGPL) {1}	No {1}	Yes (AGPL) {1}	No {1}
3-clause BSD				n/a	Yes	Yes {2}	Yes (CeCILL)	Yes	Yes (CeCILL-C)	Yes (EUPL)	Yes (AGPL)	Yes
4-clause BSD					n/a	Yes {2}	No	Yes	No {1}	Yes (EUPL)	No	Yes
MPL v1.1						n/a	No	Yes {2}	No {1}	No	No	Yes {2}
CeCILL							n/a	Yes (CeCILL)	Yes (CeCILL)	Yes (CeCILL)	No	Yes (CeCILL)
CeCILL-B								n/a	Yes (CeCILL-C) {1}	No	No	Yes
CeCILL-C									n/a	No {1}	No {1}	Yes (CeCILL-C) {1}
EUPL										n/a	No	No
AGPL											n/a	Yes (AGPL)

Notes:

1. The LGPL and CeCILL-C licences have a special provision for "works that use the library" which can be released under any licence (subject to any other restrictions which may apply). Other derivative works must be released under the LGPL or CeCILL-C licence, respectively.
2. Someone can take an MPL-licensed work and build upon it with new components. The resulting work can be distributed with the MPL covering the use of the original work and any licence covering the rest.
3. [The Fedora project wiki](#) has a nice table showing how to combine code under different versions of the (L)GPL.

Methods of Verification

Each tool that is released as part of the VPH ToolKit must be released under some form of licence. Otherwise it is unusable. That licence must be specified in the appropriate field when submitting the tool to the ToolKit portal.

Verifying that the content is allowed to be released under this licence can be very complex, and is beyond the scope of the VPH-NoE. We thus rely on self-certification.

High quality ToolKit content will, however, include sufficient documentation on the licensing of the software and any constituent parts to enable users to determine permitted activities. Furthermore, the authors of the software should maintain an audit trail for contributions, in order that a formal verification of licence validity could in principle be performed. See for instance <http://www.oss-watch.ac.uk/resources/versioncontrol.xml>.

At present, the ToolKit portal does not provide for the submission of data or content other than tools. However, when this functionality is eventually added, equivalent requirements and concerns will apply.

Training

There is extensive educational and training-related material available online, referenced by this document, on this topic. Indeed, this whole document is written largely from a training perspective. In future versions of this document it is planned to include here some case studies on the licensing decisions made for some ToolKit content.

Maintenance

While we have made every effort to ensure that the information presented in this document is accurate and timely, we are not legal experts, and there may be inaccuracies. Copyright legislation does change from time to time, and new licences are created more often than is desirable. There is therefore a need for ongoing maintenance of this Guideline. We encourage contributions from the community in correcting our errors and updating out-of-date material. A formal mechanism for this may be introduced if desired; in the meantime please email the corresponding author.

Some additions are already planned for future versions of this document. In particular, it is intending to incorporate some case studies with the Training section, and provide further details on how to report the licence(s) used by data once a data repository for the VPH is available.

Conclusions and Recommendations

The VPH ToolKit welcomes all kinds of content, released under any licence. Commercial tools can have great utility, as can tools available only through collaborations with the authors. However, in order to enhance the growth of the ToolKit, and promote the widest possible use, testing, and further development of the tools contained therein, we do make the following recommendations for best practice in VPH software licensing.

The precise choice of licence will depend on your circumstances, as described above. We recommend choosing an OSI-approved business-friendly GPL-compatible open source licence, to allow your users maximum flexibility. If this is not possible in your circumstances, we recommend choosing an open source licence from those listed in this document.

Remember that a user cannot copy (or download), modify, or distribute your software without an explicit licence or grant of right from you. Also remember that it is your responsibility to ensure that you use an appropriate licence, compatible with the licences of all components used in your software, and to ensure that all the stakeholders of your project are aware of the implications of its use with the project. Failure to do so leads to the potential of litigation. We therefore recommend that licensing and exploitation are considered carefully at the start of any project, and also whenever adding a new dependency.

Submitting Content to the ToolKit

When submitting tools to the ToolKit portal, the licence used should be indicated by filling in the Licence field in the submission form. A similar mechanism will be used for data when this service is available.

In addition, there are some best practices in documenting the licence(s) used which should be followed within the tool's source code and documentation.

- State on your tool's website which licence is applied, with a link to the licence text.
- Provide the full text of the licence in a file distributed with the tool (both source and binary distributions, if these are distinct).
- Provide a boilerplate notice at the top of each source file. The exact form of this boilerplate varies from licence to licence, and further details can normally be found alongside the licence itself in its original home. Generally it indicates who owns the copyright, and which licence is used.
- Ensure you have attributed third-party code that you use in compliance with its licence.

Further Information

There is a wealth of information online discussing copyright law, and the various open source copyright licences.

The Open Source Initiative (OSI; <http://www.opensource.org/>) is a not-for-profit corporation formed to educate about and advocate the benefits of open source, and to build bridges among different constituencies in the open-source community. It maintains the Open Source Definition, and a list of approved Open Source licences.

Similarly, the Open Knowledge Foundation (OKF; <http://okfn.org/>) provides a range of resources about the more general field of Open Knowledge, including issues surrounding data licensing.

OSS Watch (<http://www.oss-watch.ac.uk/>) is a public (and free) service for higher and further education institutions in the UK, which provides services to help institutions and projects who are using or developing free and open source software. It also provides extensive guidance and informative material about software licensing on its website. The FAQ (<http://www.oss-watch.ac.uk/about/faq.xml>) is a good starting point. Various portions of the text in this document have been drawn originally from their resources.

The IPR Helpdesk (<http://www.ipr-helpdesk.org/>) provides advice on intellectual property matters to EU-funded projects.

The Software Freedom Law Centre (SFLC; <http://www.softwarefreedom.org/>) provides legal representation and other law-related services to protect and advance Free, Libre and Open Source Software (FLOSS). It also frequently publishes the results of its legal analysis of various FLOSS-related legal issues. An article of particular relevance to the VPH discusses software used on implanted medical devices.¹⁸

¹⁸ <http://www.softwarefreedom.org/resources/2010/transparent-medical-devices.html>

References

Hrynaszkiewicz, I., Norton, M. L., Vickers, A. J., Altman, D. G., January 2010. Preparing raw clinical data for publication: guidance for journal editors, authors, and peer reviewers. *BMJ* 340 (jan28_1), c181+. URL <http://dx.doi.org/10.1136/bmj.c181>

Wilbanks, J., 2008 The Control Fallacy: Why OA Out-Innovates the Alternative. *Nature Precedings* (713). URL <http://dx.doi.org/10.1038/npre.2008.1808.1>