

Ivan Čukić

Mathematical institute of SANU

Faculty of Mathematics, University of Belgrade

KDE e.V.

INTELLECTUAL PROPERTY RIGHTS AND THE FREE/LIBRE LICENSES FOR PUBLISHING KNOWLEDGE AND DATA

Abstract. We present the legal framework for knowledge and data licensing with the focus on different kinds of free/libre licenses. The special attention is given to the Creative Commons and META-SHARE licenses since those define some concepts which are easy to understand and universally applicable to other licenses. We cover the most important licenses for documents and works of art by going through a several real-world examples. In the end, we give a short overview of the specificities of the software licensing along with the most commonly used ones like BSD and GNU licenses.

Keywords. Licensing, Freedom in Computing, Free Software, Open Source

1. Introduction

Copyright Law differs from one country to another, but the general concepts behind the different wordings are the same. It protects the original works of authorship (usually referred to as “IP – Intellectual Property”) including, but not limited to literary, musical and other works of art. The software copyright is a recent addition to the list and a special case since mathematical formulae (which computer programs in essence are) can not be copyrighted. The copyright is retained 70 years after the author's death [1].

The author of the work is granted the following exclusive rights: reproduction, communication, distribution, fixation, rental, broadcasting, creating derivative works and the moral rights, as stated in the Berne Convention for the Protection of Literary and Artistic Works. In the case of “work for hire”, the employer holds the copyright, and not the actual author. The author does retain the moral rights.

The copyrighted work can be sold, the IP rights transferred to another entity or, what is the usual case for the types of work we cover in this paper, licensed. The license is a contract that acts as a “permit” given by the author to the consumer that gives the consumer only the rights specified in it. This gives a whitelist-based framework where the consumer has only the minimal rights by default, and additional ones are given via the license.

The licenses vary in the amount of rights given to the consumer. Those that allow the minimal set of rights (usually just the rights required by the Copyright Law, and those deemed to be under “fair use”) are considered “restrictive”, as opposed to the “liberal” ones which grant all rights allowed by law.

2. Basic concepts

In this section, we cover the different concepts that are common to all licenses. We will use the concepts and symbols as provided by the Creative Commons [2] and META-SHARE [3] since those are quite clearly defined and easy to understand.

2.1 Creative Commons¹ (CC) license concepts. The idea behind the Creative Commons licensing is to isolate the different (mostly) orthogonal concepts that can be used as building blocks when deciding about the desired license features. By default, if none of the following concepts is chosen, the license (CC ZERO [L1]) equals to publishing the work under Public Domain (section 2.3).

Attribution (BY) places the requirement to retain the information about the original author or authors. When creating derivative works, it is an established practice to just append the name to the list authors of the original work.

NoDerivs (ND) prohibits creating derivative works. This alone does not forbid redistribution, copying or anything else, as long as the original work is unaltered.

NonCommercial (NC) forbids selling the original and the derivative work. Note that since the license applies to the consumer and not the original author, this does not imply that the author gives up the right to sell the work.

ShareAlike (SA) requires to distribute the derivative works under the exact same license the original work was published under. This means that if you have altered an article published under CC BY-SA license (which will be discussed in the section 3), and therefore made intellectual contribution to it, if you want to distribute the changed article, you need to do so under the same terms under which you got the original. This concept was instrumental in popularity of a few projects such as Wikipedia, Linux, etc.

2.2 META-SHARE² license concepts. While the concepts drawn from CC licenses cover most of the cases, META-SHARE introduces two new ones:

NoRedistribution (NoReD) does not allow the consumer to redistribute the work. The work can be used only for personal use, or in the case of META-SHARE, the work needs to stay inside the META network. Creative Commons licenses do not use this concept because the whole idea behind the project is to create a framework to allow sharing the knowledge and creativity.

For-a-Fee (FF) implies that in order to get the license, you need to pay the author. This is a strange concept since the fact that you have or have not paid a fee for the license does not need to be enclosed in the text of the license, so it does not need to be an integral part of it. Most of the licenses we will cover don't forbid the consumer to sell the original nor derivative works.

2.3 Basic license categories. The common licenses can be put into a several categories which will serve us later to easily distinguish between them, and discuss which license is best suited for particular kinds of works.

The work that is marked as **Public Domain** doesn't require anything from the consumer, and gives her all the rights. This means that the work can be modified and redistributed in any manner desired. It even allows removing the information about the original authors. The works without a known author, or the works whose copyright has expired automatically become Public Domain.

1 Creative Commons is a non-for-profit organization that enables the sharing and use of creativity and knowledge through free legal tools.

2 META – Multilingual Europe Technology Alliance is a joint international effort towards furthering Language Technology as a means towards realising the vision of a Europe united as one single digital market and information space.

Since the notion of Public Domain is not recognized in all countries, there are licenses crafted specifically to give the consumer the same rights. In the countries which do not allow the transfer or abandonment of authorship, these licenses might be deemed invalid.

Copyfree [4] licenses grant all the rights that the Public Domain does with the sole obligation to retain the copyright notice of the original work (the *Attribution* clause). Because of the copyright retention, these are recognized in most legal systems in the world.

Copyleft [5] licenses build upon the previous group and add the requirement to publish the derivative works under the same or equivalent license the original work was published under (the *ShareAlike* clause).

Unlike the previous categories, the **proprietary** licenses usually place restrictions even on redistribution (the *NoRedistribution* clause). While there are licenses with different amount of rights given, these usually give only the rights required by law.

2.4 Free/Libre licenses³. The desire to define the requirements for licenses that grant the consumer freedom came from the Free Software Foundation and the GNU project. The *freedoms* required for a license to be considered free are as follows [6]:

1. The freedom to run the program, for any purpose.
2. The freedom to study how the program works, and change it so it does your computing as you wish. Access to the source code is a precondition for this.
3. The freedom to redistribute copies so you can help your neighbour.
4. The freedom to distribute copies of your modified versions to others. By doing this you can give the whole community a chance to benefit from your changes. Access to the source code is a precondition for this.

While these rules were tailored for software, it is easy to distil them to be applicable to other kinds of author works [7]:

1. The freedom to **use** the work and enjoy the benefits of using it
2. The freedom to **study** the work and to apply knowledge acquired from it
3. The freedom to **make and redistribute** copies, in whole or in part, of the information or expression
4. The freedom to **make changes and improvements**, and to distribute derivative works

Additionally, there must be no restrictions placed on the above freedoms. This means that the licenses that have the *NonCommercial*, *NoDerivs* or *NoRedistribution* clauses are not considered to be *free*. Non-free licenses are usually referred to as *proprietary*.

When considering the previously defined categorization (section 2.3), the first three categories define the free licenses. It can be said that the *public domain* and *copyfree* licenses give the freedom to the consumer, while the *copyleft* gives the freedom to the work itself, by ensuring it cannot be locked down and made proprietary.

3. Documents and works of art

While the licenses used for software (section 4) can be used for documents and works of art, and vice versa, it is not a recommended practise in most cases because the software has some fundamentally different concepts. We will cover the most used licenses by going through a few distinctive examples.

3 Due to the ambiguity of the term *free* – free as in freedom, versus the free as in price – it is not uncommon to write it as *free/libre* when referring to the former.

Example 1: E. A. Blair wrote a dystopian novel that he wants to be distributed across the world so that the people can get educated about the political systems. He wants to allow free copying, but doesn't want anyone to be able to change the original text and doesn't want others to be able to charge for its distribution.

The license requirements are clear – the authorship needs to be kept. BY
 Distribution should be allowed but restricted to be non-commercial, and altering ND
 the work should be forbidden. NC

The best fit would be the Creative Commons Attribution-NoDerivs-NonCommercial License [L2] (or CC BY-ND-NC for short). Mind that this license is not considered free because of the ND and NC restrictions.

Example 2: R. Johnson recorded a blues tune. He wants to share it with the world in a way which will allow other musicians to build upon his work, to be able to learn improvisation and even to improve parts of the song. He wishes to ensure that the other artists will play fairly, and will publish their versions in the same terms he did, but that nobody should be able to sell the original nor the modified versions of the song.

The license should require keeping the authorship. Distribution of the BY
 modified work should be allowed, but only under the same terms as the original. SA
 Redistribution should be non-commercial only. NC

The license which covers this case is the Creative Commons Attribution-NonCommercial-ShareAlike [L3] (CC BY-NC-SA). The other notable license that fits this case is the Open Content License. These licenses are not free because of the NC clause.

Example 3: L. Sanger decided to create an encyclopedia that can be edited by everyone willing to contribute. Since this is supposed to create a live system where one article is edited multiple times, any subsequent edit needs to retain the original license. Other than that, the readers and researchers using the encyclopedia should not have any other restrictions.

The license should require keeping the authorship. Distribution of both the BY
 original and the modified work should be allowed, but only under the same SA
 licensing terms.

The most used licenses for this are Creative Commons Attribution-ShareAlike [L4] (CC BY-SA) and GNU Free Documentation License [L5] (GNU FDL). The later was used for Wikipedia before 2009, when it was replaced by CC BY-SA. These are the recommended free licenses to be used for documentation or works of art because they forbid re-licensing the original work as proprietary.

Example 4: Prof. M. Stonebraker created a database system and started writing the user manual for it. He decided that everyone should be able to contribute to the document. He understood that some companies that use his system will want to modify the manual to include some of their specific work-flows which are either useless for the general public, or are trade secrets. So, it was impossible to require that all derivative works be published under the license he chose.

The license should require keeping the authorship. Everything else should BY
 be up to the consumer.

Licenses that can be used for this case are Creative Commons Attribution License [L6] (CC BY) and FreeBSD Documentation License [L7]. These licenses are also free, but allow making the work proprietary.

4. Software

Software has some additional concepts that need to be taken into consideration when creating licenses for it. Usually, the concept of the source code, library linking and such do not exist in the previously covered examples⁴.

4.1 Freedom for software. For software, freedom to study (*freedom 2, section 2.4*) and to make changes and improvements (*freedom 4, section 2.4*) imply the necessity to have the source code of the program, along with the freely available tools for generating binaries (executable programs, dynamic libraries) from it. It also implies that providing obfuscated source code that is incomprehensible to understand doesn't meet the needed criteria. The need for the source code is one of the reasons behind the term *Open Source* as a replacement for the original name *Free Software*.

4.2 Copyleft for software. In the case of documents or works of art published under a copyleft (or ShareAlike) license, it is sometimes hard to find the border between which part is a derivative of the original work, and the separate original work. For example, if a part of an article from Wikipedia is copied to a document, then does the whole document need to follow the original license? Or should the whole collection to which the document belongs do as well?

For software, this distinction is clear, and is based on interlinking between the files. If there is any code in the file that was released under a copyleft license, or the file is statically linked to such code, the whole file needs to be published under the same license. This is called the “weak” copyleft. This allows for the *free* code to be used in proprietary projects by using it as a dynamically linked library.

The “strong” or “viral” copyleft imposes the aforementioned restriction even on dynamically linked code which stops any proprietary software from using it. The term “viral” comes from the fact that if a project uses even one line of the strong copyleft code, even dynamically loaded at runtime, the license spreads out to the whole project.

4.3 Free Software licenses. We are going to present the Free Software licenses grouped by whether they are copyleft (and which type) or not. We will skip the examples since the decision which license fits a particular project boils down to just two questions – is the source code supposed to remain free forever; and if the project develops a dynamic library, should it be possible for it to be used in proprietary projects or not. Historically, this choice was more often an ideological than a pragmatic one, based on the project's affinities and interpretation of the *free software* term.

If it should be possible to lock down the source code, then a non-copyleft license should be used. Good examples for these are the Revised BSD License [L8], the X11/MIT License [L9] and the Apache License [L10]. Usually, these are all based on the original BSD license, with some smaller changes. The main reason for choosing this type of licensing is the desire of the original authors to provide a commercial proprietary version of the software at a later stage without the need to get the permission from all contributors. Examples of the projects using these are FreeBSD, X.org, PostgreSQL, Boost etc.

If the software is supposed to remain free, then a copyleft license is the obvious choice. The default is the GNU General Public License (GPL), which is a strong copyleft license used in the projects like Linux, MySQL, etc. If the software is a web application,

4 In the cases that do have these notions such as LaTeX or XML-based documents, XML-based graphics formats etc., the licenses described in this section can be used if desired. It is not uncommon to find the software documentation published under the same license as the code.

GPL's restrictions can be circumvented because of the fact that the actual software is not being redistributed to the user – it is running on a server, and the user only accesses it remotely (the so called “Application service provider – ASP loophole”). In this case, it is wiser to use the Affero GPL [L12] which requires that the complete source code be made available to any network user of the work licensed under AGPL.

For the projects that develop dynamic libraries, and want to enable proprietary usage, or usage from the code published under a different free license, but want to disallow making the library itself proprietary, the best choice is the GNU Lesser General Public License [L13] (LGPL, sometimes referred to as Library GPL) which is a weak copyleft license used in such projects as Hibernate, Qt, KDElibs etc. The alternative is to use GPL with the classpath exception [L14] as used in the OpenJDK implementation of the Java programming language.

5 Conclusion

This paper has developed a model for understanding the relationship between the author of the particular work, and its consumers. We have presented concepts that stand behind most of today's Free licenses, along with a few most commonly used licenses for works of art, documentation, data and software. A quick reference of the aforementioned licenses, along with the concepts they have is given in the *Appendix A*.

The benefit of choosing one of the licenses covered in this paper is that they are widely deployed, well-known, multinational and, to a large degree, tested in court. Even more, all of these licenses are easily understood even by a legal layperson which is a benefit for both sides – for the authors and consumers alike.

Acknowledgements. The work presented here was partially supported by the Serbian Ministry of Education and Science with the projects III 178006 and III 044006.

References

- [1a] Art. 1 (1), Directive 2006/116/EC, European Union
- [1b] 17 U.S.C. § 302(a), United States of America
- [1c] Art. 100, Law on Copyright and Related Rights, Serbia
- [2] Creative Commons, <http://creativecommons.org/licenses/>
- [3] META Network, <http://www.meta-net.eu/meta-share/licenses>
- [4] Copyfree, <http://copyfree.org/>
- [5] Copyleft definition, <http://www.gnu.org/copyleft/>
- [6] Free software definition, <http://www.gnu.org/philosophy/free-sw.html>
- [7] Free cultural works definition, <http://freedomdefined.org/Definition>

Licenses

- [L1] <http://creativecommons.org/publicdomain/zero/1.0/>
- [L2] <http://creativecommons.org/licenses/by-nc-nd/3.0>
- [L3] <http://creativecommons.org/licenses/by-nc-sa/3.0>
- [L4] <http://creativecommons.org/licenses/by-sa/3.0>
- [L5] <https://www.gnu.org/copyleft/fdl.html>
- [L6] <http://creativecommons.org/licenses/by/3.0>
- [L7] <http://www.freebsd.org/copyright/freebsd-doc-license.html>
- [L8] <http://www.freebsd.org/copyright/freebsd-license.html>
- [L9] <http://opensource.org/licenses/MIT>
- [L10] <http://directory.fsf.org/wiki/License:Apache2.0>
- [L11] <https://www.gnu.org/licenses/gpl.html>
- [L12] <https://www.gnu.org/licenses/agpl.html>
- [L13] <https://www.gnu.org/licenses/lgpl.html>
- [L14] <http://openjdk.java.net/legal/gplv2+ce.html>

Appendix A: Licenses overview

Table 1: Overview of the most commonly used licenses

License	Attribution	Commercial usage	Derivative works	Copyleft	Free/Libre
CC ZERO	no	yes	yes	no	yes
WTFPL	no	yes	yes	no	yes
CC BY	yes	yes	yes	no	yes
CC BY-SA	yes	yes	yes	yes	yes
CC BY-SA-NC	yes	no	yes	yes	no
CC BY-SA-ND	yes	yes	no	yes	no
Rev. BSD	yes	yes	yes	no	yes
X11/MIT	yes	yes	yes	no	yes
Apache	yes	yes	yes	no	yes
GNU FDL	yes	yes	yes	yes	yes
GNU GPL	yes	yes	yes	yes*	yes
GNU LGPL	yes	yes	yes	yes	yes
GNU AGPL	yes	yes	yes	yes*	yes

* “strong” copyleft

Table 2. Overview of the META-SHARE NoRedistribution* licenses

License	Attribution	Commercial usage	Derivative works	Free/Libre
C-NoReD-FF	yes	yes	yes	no
C-NoReD	yes	yes	yes	no
C-NoReD-ND-FF	yes	yes	no	no
C-NoReD-ND	yes	yes	no	no
NC-NoReD-FF	yes	no	yes	no
NC-NoReD	yes	no	yes	no
NC-NoReD-ND-FF	yes	no	no	no
NC-NoReD-ND	yes	no	no	no

* META-SHARE also provides COMMONS set of licenses that are equivalent to their Creative Commons counterparts

ivan@math.rs