

Using Free and Open Online Resources: Licensing and Collections

By: Norm Friesen, PhD

January 2010

Using Free and Open Online Resources: Licensing and Collections

Abstract

The purpose of this document is to provide an overview of the “open” or “free” resources and materials available from online collections and providers that may be of significant strategic value to schools, universities and to other organizations with educational mandates, such as museums and archives. The initial sections of this paper describe some of the most basic characteristics of “open source,” or more accurately, “creative commons” licensing for cultural and educational resources. These sections also outline criteria for the inclusion of items in the annotated listing that is provided in the remainder of the document. This listing provides details for a number of recommended collections of resources with alternative licenses, indicating the terms governing their use in each case.

Introduction: Licensing Works on the Web

It goes without saying that the Internet and the Web have opened up a wide range of resource sharing and outreach possibilities for small-scale organizations working in educational and cultural sectors. However, technology on its own has not been enough. The free use, dissemination and exchange of information and contents that the Internet enables has required an appropriate legal framework to establish its legitimacy. Standard copyright law dictates that any creative or intellectual output, once it has a tangible or material form, is the property of its creator, and that it cannot be copied and distributed without his or her permission. The Internet, however, is essentially a technology for copying and distribution, and it is designed in such a way that reproduction of works, once made available, is difficult if not impossible to control. Those who want to freely create, distribute and develop using others’ material—and those who want to make original material freely available themselves—need a way of guaranteeing these freedoms.

“Open source” licensing has provided a way of guaranteeing these freedoms for the volunteer software development community, and “creative commons” licenses provide similar guarantees for those working with other kinds of intellectual and creative works. “Creative Commons” refers to a non-profit organization that has released licenses (in the USA, Canada and other jurisdictions) that allow copyright holders to specify in detail the terms under which material can be used. Each license specifies particular conditions related to the attribution, derivation, distribution and commercial use of the resource. Rights or terms related to use, distribution and revision that would automatically apply to a copyrighted resource under conventional law can be waived selectively for certain types of use. Licenses that apply to different jurisdictions vary in their precise wording and legal implications. Organizations engaged in education, cultural awareness and related activities have much to gain through access to materials licensed for use under these terms. And many organizations, large and small, already have been successful developing and providing content bearing only these alternative licenses. One prominent example is Flickr (www.flickr.com), which provides access to millions of alternatively-licensed digital photographs, which include those of the US National Archives). A second is Wikipedia (en.wikipedia.org), which provides access to millions of pages of texts and images that are similarly licensed. But many

challenges remain in understanding the applicability of different creative commons licenses in different circumstances.

Creative Commons and other alternatives to conventional copyright have been the subject of some confusion. In many cases, resources bearing a standard Creative Commons license are simply perceived to be “freely available,” whereas nothing could be further from the truth. For commercial users in particular, some creative commons licenses (specifically those with “non-commercial” and/or “share alike” clauses) can be more restrictive than conventional copyright licensing. In a recent report, the Hewlett Foundation focuses specifically on “open” educational resources, and comes to the following conclusion:

The principal finding of this study is that the terms “open” or “free” educational resources do not communicate much to educators, students, and other potential users with respect to actual uses that may be made. As a legal matter, use of these “open” resources is governed by a great diversity of license terms... Even standard terms may require a degree of interpretation because users are not familiar with copyright law or because there may be some ambiguity in the scope of uses permitted. (Hewlett Foundation, 2008, pp. 11-12)

As will be shown in the next section, one of the most “standard” or popular Creative Commons licenses—that is, the “Attribution-Share alike” license—is open to a number of interpretations, with each interpretation having rather different practical consequences.

License types and their Restrictions

Creative Commons licenses are labeled according to the types of uses they permit and disallow, with most licenses combining two or more specific conditions of use. The four conditions or restrictions combined in Creative Commons licenses are as follows:

1. All licenses stipulate that the use of the resource be **attributed**, that the author be given credit according to the manner requested.
2. One requirement specifies that only **non-derivative** or “verbatim copies” of the resource be used. In other words, the resource not be revised or combined with other resources in any way.
3. Some licenses stipulate that the resource be used only for **non-commercial** purposes, a condition that has been interpreted variously.
4. Other licenses require that the resource be used only under the condition that any copies or derivative versions be made available under the same terms as the original resource (i.e. as “share-alike” creative commons resources).

These restrictions are combined to produce six different types of licenses, each with its own requirements and shorthand label

1. Attribution (“BY”).

2. Attribution and use on a share-alike basis (“BY-SA”)
3. Attribution and use (or distribution) only in original form (“BY-ND”)
4. Attribution and non-commercial use only (“BY-NC”)
5. Attribution and non-commercial use only, with use only in original form (“BY-NC-ND”)
6. Attribution, non-commercial use only, on a share-alike basis (“BY-NC-SA”)

Many of these clauses, either individually or in combination, are relatively unrestrictive, and can be easily interpreted and implemented. Attribution requirements can generally be met through reference to the author’s name and/or Website. Forbidding the creation of derivative works means that the resource needs to be provided on its own, and without the alteration of its content.

The non-commercial restriction, on the other hand, restricts the use of the resource in commercial contexts, but can quickly become complicated. The text in the Canadian version of the license excludes use that “is primarily intended for or directed toward commercial advantage or private monetary compensation” (Creative Commons, 2010a). Some have understood this to mean that the provision of the work should not involve any kind of transaction involving (potential) profit. This would include giving access to the resource on a CD or in a book that is sold for profit, or on a Webpage that contains advertising (e.g., see: Möller, 2007). However, in the light of qualifying wording such as “*primarily* intended” and “*private* compensation,” such an interpretation seems overly restrictive. Correspondingly, it is generally understood that such a resource *can* be included in a course or other context where the user has paid for access and other services, as long as it is made available on a not-for-profit basis.

Although this “non-commercial” clause has been variously interpreted, it is the “share-alike” provision that seems at once to be the most popular and controversial, and it is of particular relevance for meriting detailed consideration. In the Canadian versions of these licenses, this clause permits users of a work to “to create and reproduce derivative works” but only under “the same terms and conditions as the license granted... under this license” (Creative Commons, 2010b). The American version of this provision is utilized by Wikipedia.org and its sister projects, as well by owners of about 50 million photos on Flickr.com. This particular type of restriction has been referred to as “copyleft,” implying a kind of inverse of the restrictions defined in copyright. Instead of forbidding the redistribution and alteration of a particular work, this restriction allows these actions, but *only* under the condition that resulting works continue to be distributed under the same terms. A photo from Flickr licensed in this way could be used “freely” in course documentation, but would require that documentation to similarly be “freely” available for use and adaptation.

This type of licensing has been derisively characterized as “viral” or even described as a “cancer” by commercial developers and vendors (most prominently, Steve Ballmer of Microsoft; see: Greene, 2001). Such characterizations arise from objections to the fact that the “copyleft” licensing requirement can be transmitted from a relatively small resource to an entire work, and can have the effect of changing a

whole program, book or Web resource work from being protected under “copyright” to being more freely open under “copyleft.” Through the addition of a relatively small component bearing a Creative Commons “share-alike” license, in other words, an entire aggregate resource would now have to be made available *without* any restrictions on its further distribution and revision. This also implies that the resource has to be freed of unnecessary technological barriers impeding access (e.g. being “locked” behind password protection, or being distributed in a PDF format that does not allow for editing, copying and other operations).¹ This presents obvious difficulties related to branding, quality assurance and possibly also the larger business model of some non-profit institutions.

However, this provision to “share-alike” is subject to interpretation and certain limitations, and the extended legal documentation that constitutes both the Canadian and US versions of this license identifies an important set of exceptions. These exceptions are outlined in the definition of “collective work” in the license:

“Collective Work” means a work, such as a dictionary, yearbook, encyclopedia, or a newspaper, review magazine or singular periodical and any work written in distinct parts by different authors, or in which works or parts of works of different authors are incorporated. A work that constitutes a Collective Work will not be considered a Derivative Work (as defined below) for the purposes of this license. (Creative Commons 2010b)

This definition presents the collective work, in other words, as kind of aggregate resource which is effectively exempt from the “viral” characteristics of copyleft. A copyleft resource (with a creative commons “share-alike” license) can be “incorporated” into such a collective work—for example, a “review magazine or singular periodical”—without the resulting aggregation being seen simply as a “derivative work” and as subject as a whole to copyleft. This exception is made in the case of works that are aggregations of multiple contributions, with each contribution (in the words of the American version of the license) “constituting [a] separate and independent” work in itself.

All of these considerations are important, given that many organizations (such as colleges or museums) typically retain copyright of resources created or compiled by employees and apply such licensing to content they provide over the Web. This reality needs to be understood in combination with further factors: namely, that many online works are licensed on a “share-alike” or copyleft basis (e.g. all of Wikipedia), and that educational organizations lack the resources to pay for the use of other copyrighted works. Taken all together, these factors have a wide range of specific implications for different types of content, different kinds of use and different forms of licensing. Many of these implications hinge on the definition and interpretation of the notion of “collected works” as discussed above. The tables below attempt to systematize some of these implications, showing how combinations of copyleft and (conventional) copyright resources—and the kind of collective work that they would constitute together—may or may not meet the requirements of share-alike licensing. Also, in as many cases as

¹ “ You may not Use the Work with any technological measures that control access or use of the Work in a manner inconsistent with the terms of this Licence Agreement” (Creative Commons, 2010b)

possible, examples of collections providing the type of resource referenced are collections that are ranked and reviewed in this report.

Table 1: Combinations of cultural and educational works likely to constitute a collective work

In this case, *it seems quite clear* that a number of kinds of share alike or copyleft resources could be used together with copyright resources to constitute a collective work (and thus be adaptable to the conventional copyright practices of many organizations, despite share-alike requirements).

Copyleft Resource	Copyright Resources	Collective Work incorporating both
Textbook (e.g., from textbookrevolution.org, see below)	Online course resources or a “course manual”	“Course” constituted through the provision of the two resources.
Article (e.g., from DOAJ)	Other articles as readings	Edited collection of readings associated with another offering (e.g. course, exhibit)
Video or Audio podcast ²	Other multimedia resources	Multimedia resources part of a display, Website or course

*Table 2: Combinations of works that **may** constitute collective works*

Given the terms of the above wording for a collective work, it is *difficult to be certain* about the status of other resource types and their combinations on other kinds of collective works.

Copyleft Resource	Copyright Resource	Collective Work incorporating both
Podcast (accessed via iTunes)	Integrated resource online (e.g. a resource instructing user to access iTunes podcasts at a specific point).	Integrated resource linking to and requiring the use of other resources.
Collection of resources (e.g. merlot.org; Wikipedia.org)	Integrated course resource online referencing specific resources in the copyleft collection.	Integrated resource requiring the use of other resources.

*Table 3: Combinations of works that **do not** constitute collective works*

Finally, given the terms of the wording, above, it is *nearly certain* that collective works developed from the following types of resources would *not* be allowed.

Copyleft Resource	Copyright Resource	Collective Work incorporating both
Interactive resource (e.g. simulation from merlot.org)	Specific exercises for the use of the simulation.	Fully integrated course manual or other kind of resource.
Photograph from	Text referencing photo or diagram and	Fully integrated course manual or

² The effective limitation of collective work to “written works” in the Canadian version of the license casts the viability of this last example into some doubt (where the use of a Canadian resource is concerned). This would apply where “Canadian” resources referencing this license are used. Note that such a limitation (to text and writing) is *not* present in the American license, and thus the use of resources available under the American version of this license does not suffer from this same limitation.

Flickr or diagram from Wikipedia	its particular contents.	other kind of resource.
Video or Audio podcast (e.g. clip from YouTube EDU)	Web page from course materials that directly embeds YouTube player and clip, and adds contextualizing text.	Password protected Web (HTML) documents constituting course materials.

Based on these resource combinations listed above, it is important to note that *types of resources* and *forms of integration* are of paramount importance in considering the legal ramifications of the use of many creative commons resources. As long as resources licensed under the “share-alike” provision are unmodified and otherwise retain their integrity as individual works, it seems that they can be brought together with other materials to constitute Collective Works, especially if these collective works are loosely rather than tightly integrated.

Specific Criteria for Selection and Ranking

The reason for the extended discussion of licensing terms, above, and the examples of their application in different situations derives from the fact that licensing is the most important and complex criterion involved in the selection and ranking of collections of resources. This criterion as well as the others used in creating and ranking the listing provided in this report are as follows:

1. **Licensing flexibility (rated out of 5):** those resource collections whose resources are licensed in a highly flexible manner, and/or can be readily used in a Collective Work, as discussed above. The criterion for this involves not only the specific licensing of the work, but also the kind of aggregation and integration that is or is not allowed using the types of resources featured in a given collection.
2. **Quality (rated out of 3):** The quality of the resource collection is generally assessed through reference to independent reviews of the resource, or the existence of peer reviews for the resources collected. Also considered are the reputation of sponsor or affiliate, and the history of the resource collection.
3. **Stability (rated out of 3):** The stability of the resource collection is assessed either in terms of its longevity (i.e. number of years in existence), or its affiliation to an institution.
4. **Fitness for purpose (rated out of 5):** The suitability of the resources collected for open course designs, based on the subject matter and type of resource involved. This criterion is applied and explained on an item-by-item basis.

A number of unexpected results were produced by applying these rankings to a range of collections. First, there are unfortunately *no* resource collections that combine peer review or other standard quality assurance processes to resources that are, at the same time, exempt from copyleft limitations. There *are* collections that provide resources of varying quality which can be adapted to conventional copyright approaches, but the precise quality of these offerings needs to be evaluated on a resource-by-

resource basis. Second, there are many collections with resources of uniformly high quality. However, because these collections feature resources that would be valuable for re-use and adaptation in the form of smaller works or units, and because they use copyleft licensing, they are ranked relatively low in their potential value. So there is an unfortunate pattern of mutual exclusion: collections allowing for flexible use lack quality assurance, and collections of uniformly high quality provide resources that are not readily adaptable to conventionally copyrighted uses.

However, three resource collections present encouraging—albeit limited—exceptions to this pattern. These are Wikimedia Commons, Flickr and Connexions. These three collections are large, well-known, and while they do not have peer review, they have few or no copyleft limitations, giving them a high degree of licensing flexibility.

Top-Ranked Collections (rated 9 or higher)

rated 14	Connexions	URL: http://cnx.org	No. of Resources: about 15,000 instructional “modules”
Stability: High (3). In existence since 1999; identified closely with and supported by Rice University.		Quality: Varies (2). Connexions is <i>working towards</i> a peer-review approach.	
Description: “a place to view and share educational material made of small knowledge chunks called modules that can be organized as courses, books, reports, etc.” Covers all levels of education, but has a post-secondary emphasis.			
Licensing: (5) All resources are available for use <i>requiring attribution only</i> (i.e. an indication of their authorship). In other words, all resources <i>are exempted from copyleft licensing</i> . As a result, the instructional units can be mixed with others to create derivative resources that can be subject to any kind of licensing.			
Fitness for purpose: (4) Despite the varying quality of the resources in this collection, the licensing used makes this collection of significant potential strategic value to Open Learning. The emphasis on post-secondary and North American education, as well as the availability of some resource sets as textbooks further adds to this value.			

rated 14	DOAJ Directory of Open Access Journals	URL: http://www.doaj.org	No. of Resources: Over 4000 journals; likely well over 30,000 articles
Stability: High (3). In existence since 2002; selected contents are archived digitally.		Quality: High (3). All resources (articles) are subject to peer review.	
Description: “This service covers free, full text, quality-controlled scientific and scholarly journals. We aim to cover all subjects and languages.”			
Licensing: (5) Articles of the kind collected by the DOAJ can be aggregated to form collective works; many of the articles provided are licensed using Creative Commons licenses.			
Fitness for purpose: (3) Not, strictly speaking, a collection of educational or cultural resources, but potentially of great value for post-secondary usage and other contexts in which it is appropriate to expose users to academic writing and research. Using articles from this collection obviates the need for library licensing of specific articles, and for users to navigate access procedures usually associated with this kind of institutional licensing. All subject areas are covered in this collection.			

rated 13	Wikimedia Commons	URL: commons.wikimedia.org	No. of Resources: more than 5,000,000
Stability: High (3) Wikimedia Commons went online in 2004, is closely affiliated with the Wikipedia project, and is a part of the Wikimedia project. It consequently has a relatively robust funding model.		Quality: Varies (2). Contents are not peer-reviewed in a strictly systematic manner. Quality needs to be evaluated on a case-by-case basis.	
Description: “a repository of images, sound and other multimedia files. It is a project of the Wikimedia Foundation, from which uploaded files can be used across all Wikimedia projects in all languages, including Wikipedia, Wikibooks, Wikisource and Wikinews, or downloaded for offsite use.”			
Licensing: (4) Licensing conditions vary, with many resources being associated with copyleft licenses. However, there are also many resources (above all images and diagrams) that are available as a part of the public domain, meaning that their use is unrestricted.			
Fitness for purpose: (4) Despite the need to evaluate quality on a case-by-case basis, this collection has significant potential for many organizations, especially since it can be used to provide images (often of historical, cultural or scientific value) that are explicitly in the public domain.			

rated 13	Wikibooks	URL: http://wikibooks.org	No. of Resources: 37,978 pages or content modules
Stability: (3). Wikibooks went online in 2003, is closely affiliated with the Wikipedia project, and is a part of the Wikimedia project. It consequently has a relatively robust funding model. A particular, stable version of any given book or resource can be readily accessed using a stable URL, avoiding the problems presented by ongoing editorial work.		Quality: Varies (2). Contents are not peer-reviewed in a strictly systematic manner. Quality needs to be evaluated on a case-by-case basis.	
Description: “Wikibooks is a Wikimedia project for collaboratively writing educational content on various subjects.”			
Licensing: (4) All contents are subject to copyleft licensing. However, because books are generally not considered as constituting parts of derivative <i>or</i> collective works, the resources here can be readily used by organizations as whole books or book chapters.			
Fitness for purpose: (4) Despite the need to evaluate quality on a case-by-case basis, the right resource for the right purpose could be of great value.			

rated 10	TED Talks	URL: www.ted.com	No. of Resources: over 450 20 min. talks
Stability: (2). Supported by the Sapling Foundation and by corporate sponsors, but in existence on the Web only since 2007.		Quality: High (3). Process for selecting speakers is unclear, but only high-profile academics and others are selected.	
Description: "We're building here a clearinghouse that offers free knowledge and inspiration from the world's most inspired thinkers, and also a community of curious souls to engage with ideas and each other."			
Licensing: (2) The collection makes available relatively short or small resources using CC licensing, but also requires them to be shared alike.			
Fitness for purpose (3) It is permissible to link to individual resources conventionally-licensed material, but integrating the material in any other way is not advisable.			

rated 9	MIT Open Courseware	URL: http://ocw.mit.edu	No. of Resources: 1900 full courses, "virtually all" of MITs courses
Stability: (3) In existence since 1999; identified closely with and supported by the Massachusetts Institute of Technology. World famous.		Quality: High (3). Quality of courses and events generally assured through the reputation of the host university. Criteria for technical quality are also in place.	
Description: "MIT OpenCourseWare (OCW) is a web-based publication of virtually all MIT course content. OCW is open and available to the world and is a permanent MIT activity."			
Licensing: (2) All materials are made available through a CC license requiring them to be attributed and "shared alike." Materials are likely to be used in part, and not as a whole (e.g. in the way a textbook or research article is used). This means that it is that the material in this collection generally <i>cannot</i> be adapted in the development of copyright material.			
Fitness for purpose (2) It should be possible to link to individual resources from courses and other contexts, but integrating the material in any other way will be problematic from a licensing perspective.			

9	Flickr	URL: www.flickr.com	No. of Resources: 133 million available via Creative Commons licenses
Stability: (3) Owned by Yahoo Inc., in operation since 2004.		Quality: Varies (2). Quality needs to be	

	evaluated on a case-by-case basis.
Description: An “online photo management and sharing application” that offers free accounts to users. Its collection is comprised almost entirely of users’ photographs, but videos are also collected.	
Licensing: (4) The type of license varies from one resource to another, but a very large number of resources (80 million) are available without share-alike restrictions.	
Fitness for purpose (2) Despite the need to evaluate quality on a case-by-case basis, this collection has significant potential for many organizations, especially since it can be used to provide images (often of historical, cultural or scientific value) that are explicitly in the public domain. 60 museums, art galleries, archives and libraries have added substantial collections to Flickr’s creative commons works.	

References:

Creative Commons. (2010a). Creative Commons Legal Code: Attribution-NonCommercial 2.5 Canada. Accessed January 29, 2010 from: <http://creativecommons.org/licenses/by-nc/2.5/ca/legalcode.en>

Creative Commons. (2010b). Creative Commons Legal Code: Attribution-ShareAlike 2.5 Canada. Accessed January 29, 2010 from: <http://creativecommons.org/licenses/by-sa/2.5/ca/legalcode.en>

Greene, T.C. (2001, June 2). Ballmer: “Linux is a cancer” Contaminates all other software with Hippie GPL rubbish. *The Register*. Accessed January 29, 2010 from: http://www.theregister.co.uk/2001/06/02/ballmer_linux_is_a_cancer/

Hewlett Foundation. (2008). What status for “open”? An examination of the licensing policies of open educational organizations and projects. Accessed January 29, 2010 from: <http://learn.creativecommons.org/what-status-for-open>

Möller, E. (2007). The Case for Free Use: Reasons Not to Use a Creative Commons-NC License. Accessed January 29, 2010 from: <http://freedomdefined.org/Licenses/NC>