Villanova University School of Law
School of Law Working Paper Series

Year 2005 *Paper* 34

Creative Commons and the New Intermediaries

Michael W. Carroll*

*Villanova University School of Law, mcarroll@wcl.american.edu

This working paper is hosted by The Berkeley Electronic Press (bepress) and may not be commercially reproduced without the permission of the copyright holder.

http://law.bepress.com/villanovalwps/art34

Copyright ©2005 by the author.

Creative Commons and the New Intermediaries

Michael W. Carroll

Abstract

This symposium contribution examines the disintermediating and reintermediating roles played by Creative Commons licenses on the Internet. Creative Commons licenses act as a disintermediating force because they enable end-to-end transactions in copyrighted works. The licenses have reintermediating force by enabling new services and new online communities to form around content licensed under a Creative Commons license. Intermediaries focused on the copyright dimension have begun to appear online as search engines, archives, libraries, publishers, community organizers, and educators. Moreover, the growth of machine-readable copyright licenses and the new intermediaries that they enable is part of a larger movement toward a Semantic Web. As that effort progresses, we should expect new kinds of intermediaries that rely on machine-readable law to emerge.

CREATIVE COMMONS AND THE NEW INTERMEDIARIES

MICHAEL W. CARROLL*

I. INTRODUCTION

In many early conversations about the Internet, the story that dominated was one of liberation from the intermediaries. Record companies, retailers of all stripes, and the mainstream media all were dinosaurs whose days were numbered. The Internet's end-to-end architecture enabled end-to-end commerce, end-to-end culture, and end-to-end news. Even the new intermediaries, like Internet service providers (ISPs), merely supplied infrastructure because end-to-end architecture greatly limited the kinds of control ISPs might try to assert.

After the revolutionary euphoria died down, however, many acknowledged that intermediaries are necessary to all kinds of transactions in commerce, culture, and news. Reintermediation soon follows from disintermediation, and the real question the Internet posed was not whether intermediaries are necessary but what kinds of intermediaries are necessary. When contemplating this question now, fifteen years after the invention of the World Wide Web, I want to highlight the disintermediating and reintermediating roles that Creative Commons licenses currently play on the Web and to also suggest that these licenses deserve lawyers' attention as a species of machine-readable law.[1]

Creative Commons licenses respond to the explosion of "copyright events" that digital technologies has let loose.[2] Explosions usually have violent consequences. The copyright explosion certainly has disrupted a

---

* Associate Professor of Law, Villanova University School of Law. Thanks to Barry Brewer for excellent research assistance. [Disclosure: I serve on the Board of Directors of Creative Commons, Inc. The views expressed herein are mine alone and do not necessarily reflect the views of Creative Commons or those associated with it.] Information current as of July 1, 2005. This Article is licensed under a Creative Commons Attribution-Non Commercial 2.5 License, at http://creativecommons.org/licenses/by-nc/2.5/legalcode. Attribution should be to Michael W. Carroll as author and to the Michigan State Law Review as original publisher.

[1] *See generally* Creative Commons *available at* www.creativecommons.org.

[2] I use the term "copyright events" to explain why the scope of copyright law's domain has expanded so dramatically with the growth of digital technology. A "copyright event" is any action in the world that entails the exercise of one or more of a copyright owner's exclusive rights to copy, distribute, perform, display or adapt information. Some copyright events are infringing and others are not. All implicate copyright law. The courts' response to digital technologies which require copying to function has been to permit copyright law to infiltrate almost every digital interaction. *See*, *e.g.*, MAI v. Peak, 991 F.2d 511 (9th Cir. 1993) (holding that every copy written to the Random Access Memory of a computer is a copy for the purposes of the Copyright Act).

number of industries and relationships that rely on copyright law. What is perhaps more interesting is how this radical expansion of copyright law's domain has not led to chaos. Instead, a number of implicit understandings have grown up around digital technologies, and these understandings have led to norms and implied licenses that serve important coordinating functions. As robust as these informal mechanisms are, however, greater clarity and coordination can often be had when copyright owners explicitly designate which copyright events they consider to be permissible. Enter Creative Commons licenses.

Moreover, the proliferation of Creative Commons licenses on the Web points up a new relevance dimension – the copyright status of information found on digital networks.[3] Imagine that you are an independent filmmaker in need of some music to accompany a montage in your film. You have no time or budget to clear the rights to the music. If you search for "Chopin," what is relevant is not simply whether there is information – such as a music file – that is accurately associated with your term, but also whether that information is available to you on terms that permit your desired use.[4] Recently, the creation of a Creative Commons search engine, followed by Yahoo!'s offering of a specialized Creative Commons search, enables searching along both the topical and copyright dimensions.

Creative Commons licenses act as a disintermediating force because they enable end-to-end transactions in copyrighted works. The licenses have reintermediating force by enabling new services and new online communities to form around content licensed under a Creative Commons license. Intermediaries focused on the copyright dimension have begun to appear online as search engines, archives, libraries, publishers, community organizers, and educators. Moreover, the growth of machine-readable copyright licenses and the new intermediaries that they enable is part of a larger movement toward a Semantic Web. As that effort progresses, we should expect new kinds of intermediaries that rely on machine-readable law to emerge.

II. CREATIVE COMMONS LICENSES AS INTERMEDIARIES

A Creative Commons license is a form copyright license that can be linked to via the Web. In addition to the legal code, the license is described by a "human-readable" Commons Deed, which identifies the key terms of

---

[3] The idea of "relevance dimensions" is familiar to many who undertake quantitative study. *See*, *e.g.* INEX Relevance Assessment Guide *available at* http://qmir.dcs.qmul.ac.uk/inex/Papers/ INEX02_Relevance_Assessment_Guide.pdf

[4] Of course, there also are many others searching the Net for music files for personal use who consider the copyright status of the works to be irrelevant.

the license and machine-readable metadata that associates the Internet location of the licensed resource with the Internet location of the license document. As of this writing, there are nearly 16,000,000 digital objects accessible over the Internet linked to a Creative Commons license.[5] These resources include scientific journal articles, music files, picture files, and weblog pages.

Creative Commons licenses permit certain royalty-free uses of the licensed copyrighted work. The most permissive license permits all uses so long as the copyright owner's directions concerning attribution are followed. Other conditions include a requirement that derivative works be licensed under the same terms, a limitation to non-commercial uses, and a prohibition on the creation of derivative works. These can be combined to create six permutations.[6] There are also some tailored licenses that respond to requests from particular communities. Musicians asked for a "sampling" license that permits commercial uses involving creation of derivative works through digital sampling. The sampling license comes in three flavors.[7] The Developing Nations license differentiates permission by geography, granting an Attribution license for uses in developing nations while reserving default copyright protection for uses in developed nations.[8] Creative Commons has also coupled its metadata with the pre-existing legal code from the Free Software Foundation for the use of creators of software who wish to license their creations under the GNU General Public License (GPL) or Lesser General Public License (LGPL).[9] In addition to these licenses, Creative

---

[5] *See* Mike Linksvayer, *CC in Yahoo! Advanced Search,* Creative Commons:About: Weblog, *at* http://creativecommons.org/weblog/entry/5456 (May 27, 2005).

[6] *See* Creative Commons: About: Choosing a license, *at* http://creativecommons.org/about/licenses (listing the various CC licenses) (last visited June 9, 2005). They are: "Attribution," "Attribution-NoDerivs"(No Derivative Works), "Attribution-NonCommercial-NoDerivs," "Attribution-NonCommercial," "Attribution-NonCommercial-ShareAlike," and "Attribution-ShareAlike." *See* Creative Commons: Creative Commons Licenses, *at* http://creativecommons.org/licenses (last visited June 9, 2005) (replacing the original four licenses, which could be combined into 11 permutations).

[7] *See* Creative Commons: Publish: Choose Your Sampling License Options, *at* http://creativecommons.org/license/sampling (last visited June 9, 2005) (including "Sampling,"which allows sampling for non-advertising purposes, "Sampling Plus," which is the same, but allows non-commercial copying of the entire work as well, and "Non-Commercial Sampling Plus," which allows only non-commercial sampling and copying).

[8] *See* Creative Commons: Commons Deed: Developing Nations 2.0, *at* http://creativecommons.org/licenses/devnations/2.0 (last visited June 9, 2005).

[9] *See* Creative Commons: Publish: Creative Commons GNU GPL, *at* http://creativecommons.org/license/cc-gpl?lang=en (last visited June 9, 2005), Creative Commons: Publish: Creative Commons GNU LGPL, *at* http://creativecommons.org/license/cc-lgpl?lang=en (last visited June 9, 2005).

Commons offers a service through which copyright owners can dedicate their works to the public domain.[10]

Creative Commons licenses facilitate cheap speech. For example, a teacher who wishes to find materials to copy for a course pack can see immediately that she can use content licensed under an Attribution license without asking for permission. In addition, by using Creative Commons licenses, millions of bloggers ensure that "news reader" programs may copy their respective RSS feeds and compile them into a derivative works. These speech transactions are made faster and cheaper by simple, machine-readable licenses.

Moreover, all of these licensed objects will function as a common pool. There will be new functions to be performed, similar to traditional functions related to traditional creative works, but within the context of the freedoms associated with digital objects licensed under Creative Commons licenses. In addition, Creative Commons licenses can be complemented by new licensing intermediaries who can facilitate transactions with respect to the rights reserved to the copyright owner under a Creative Commons license. For example, if a user finds a work licensed under a NonCommercial license, he or she can negotiate with the copyright owner for permission to use the content for commercial purposes. Existing intermediaries, such as the American Society of Composers, Artists and Publishers (ASCAP),[11] Broadcast Music, Incorporated (BMI)[12] and the Harry Fox Agency[13] in the music industry, for example, or new intermediaries may emerge to broker such negotiations.

III. NEW INTERMEDIARIES ENABLED BY CC LICENSES

The intermediaries enabled by Creative Commons licenses include search engines with added relevance dimensions, archives and libraries that include content tagged with CC licenses, new producers and publishers who facilitate uses of "some rights reserved" material made possible by Creative Commons, communities of Creative Commons creators, and even educational institutions. This section highlights primarily U.S.-based new intermediaries, but it is important to note that a whole range of such intermediaries also are emerging internationally.

---

[10] *See* Creative Commons: Public Domain Dedication, *at* http://creativecommons.org/licenses/publicdomain (last visited June 9, 2005).

[11] *See* ASCAP, *at* http://www.ascap.com (last visited June 9, 2005) (licensing music for public performances).

[12] *See* BMI, *at* http://www.bmi.com (June 9, 2005) (licensing music for public performances).

[13] *See* HFA, *at* http://www.harryfox.com (last visited June 9, 2005)(licensing copyrighted music for mechanical purposes such as Compact Discs).

A. Search

One of the earliest "reintermediaries" were the search engines. As the amount of information on the Net and on the Web continued to grow, connecting people with information they desired became increasingly difficult. The race was on to produce results that were most relevant to the terms used in a searcher's query. For the time being, Google's PageRank algorithm dominates along this dimension.

Most people measure relevance along more than one dimension, however, and the next stage in the search race will be to deliver multidimensional results. Creative Commons uses RDF for its metadata.[14] Potentially that metadata could be read by search engines to yield results that respond to results with both topical and copyright relevance. Recognizing the importance of finding licensed content, Creative Commons developed its own search engine.[15] The Firefox web browser now provides a toolbar link to this engine.[16]

Searching along the copyright dimension took a giant forward stride on May 23, 2005, when Yahoo! released the beta version of the Yahoo! Search for Creative Commons.[17] Searching Yahoo!'s far more comprehensive database, the search engine finds sites that have a Creative Commons license.[18] The site allows a searcher to choose among four

---

[14] *See* Creative Commons: Frequently Asked Questions, *at* http://creativecommons.org/faq#faq_entry_3329 (last visited June 9, 2005).

[15] Creative Commons Search, *at* http://search.creativecommons.org/index.jsp (last visited June 9, 2005). Creative Commons Executive Director Neeru Paharia and [who helped?] deserve credit for this advance.

[16] For those who use Firefox, the upper right corner defaults to a Google toolbar, but it is a pull-down menu that permits use of other search engines, including those provided by Yahoo!, Amazon, and Creative Commons.

[17] Yahoo! Search Blog, *Larry Lessig on Searching Creative Commons* (Mar 23, 2005), *at* http://www.ysearchblog.com/archives/000092.html.

[18] Yahoo! Search: Creative Commons Search Beta, *at* http://search.yahoo.com/cc (last visited June 9, 2005). The Yahoo! search greatly increases the number of sites found with Creative Commons licenses. To make the copyright dimension visible the following test was run. Using the keywords "Eiffel Tower," a standard Yahoo! search yielded 1,970,000 results. *See* Yahoo! Search, *at* http://search.yahoo.com (last visited June 9, 2005). A test comparison between the Creative Commons engine and the Yahoo! Creative Commons Search yielded the following results: (a) with no restrictions other than a search for Creative Commons licensed content, Yahoo! produced 3,430 results and CC produced 32; (b) with the "Find content I can use for commercial purposes" option selected, Yahoo! produced 510 results and CC produced 1; (c) with the "Find content I can modify, adapt, or build upon" option selected, Yahoo! produced 2,250 results and CC produced 4; (d) with both options selected, Yahoo! produced 375 results and CC produced 1.

As with all new technologies, there is room for improvement. Web site owners

criteria.[19] The searcher can type in keywords to find any topically relevant Creative Commons licensed content, or the searcher can specify, "Find content I can use for commercial purposes," or "Find content I can modify, adapt, or build upon," or both.[20] This search works by adding a parameter for the Creative Commons license to the standard Yahoo! search.[21] The copyright relevance dimension has gone mainstream.[22]

B. Archives and Libraries

Traditionally, libraries have performed at least four basic functions. They collect and preserve information, disseminate information, index that information by creating and maintaining metadata about their collections in their card catalogs, and they enable searching of the index of metadata.[23] In the United States, copyright law traditionally facilitated libraries' performance of these functions.[24] Copyright in digital works is less hospitable to these traditional practices.[25] Creative Commons licenses facilitate a rebalancing that frees libraries to better perform their traditional roles as well as new ones called for by the digital environment.

Among online librarians seeking to perform these roles, Brewster Kahle stands out as a visionary.[26] Recognizing early on that the malleability

---

sometimes tag a web page with a Creative Commons license but do not also license audio or video files available through the site with a Creative Commons license. In such a case, the search engines will identify the site as relevant even though the content the searcher wants is not available under a Creative Commons license.

[19] See Id.

[20] *Id.*

[21] *See* Yahoo!: Developer Network Beta, *at* http://developer.yahoo.net/web/V1/webSearch.html (last visited June 9, 2005).

[22] Yahoo! Search Blog, *Larry Lessig on Searching Creative Commons* (Mar 23, 2005), *at* http://www.ysearchblog.com/archives/000092.html.

[23] Libraries also perform a latent authentication function. We do not think about authentication in a physical library. Generally, we assume that when a library has a book on the shelves called "Oliver Twist" by Charles Dickens, it really is that book. Manipulating a physical book is not easily done, and to the extent that there are multiple editions of this book, the differences among them are readily discernible. With digital objects, however, there are usually many versions and digital objects are easily manipulated. Authentication now emerges as a potential function for an online library. Online libraries will have to decide what will be archived, and which, if any, of the many manipulations or versions is canonical.

[24] *See*, *e.g.*, 17 U.S.C. § 109 (limiting exclusive right of distribution to not include lending of legally-acquired copy of a copyrighted work).

[25] *See*, *e.g.*, LAWRENCE LESSIG, FREE CULTURE: HOW BIG MEDIA USES TECHNOLOGY AND THE LAW TO LOCK DOWN CULTURE AND CONTROL CREATIVITY 226-28 (2004)

[26] *See* Lessig, *supra* note XX, at 110-15 (describing Kahle's vision for comprehensive online library).

of content on the Web presented an immediate challenge for preservation, he created the Internet Archive,[27] a non-profit organization that built and maintains an Internet Library.[28] The site provides access to historical material in digital format.[29] The Internet Archive stores texts, audio, moving images, and software as well as archived web pages.[30]

The Internet Archive has agreed to host content tagged with a Creative Commons license. Responding to this generosity, Creative Commons has written an easy-to-use piece of software, CC Publisher, which uses a drag-and-drop method for tagging content with a Creative Commons license and publishing the content to the Internet Archive.[31] This combination of a software tool for tagging and uploading content with a central repository for that content serves the intermediary function of enabling creators and users to more easily share. Creative Commons licensed content on the Internet Archive also appears in the results of a Yahoo! Creative Commons search.

C. Producers & Publishers

One large and important role for Creative Commons licenses is to facilitate amateur-to-amateur communication.[32] However, Creative Commons licenses also enable new intermediaries to create new business models for the distribution of creative works created by professional authors. For example, Magnatune, an on-line record label, was created to distribute music over the Internet and eliminate the problems inherent with traditional recording contracts.[33] Its business model is to target Internet radio listeners and "fans of music that gets little radio airplay or major record distribution, but has a fairly large audience."[34]

Magnatune is a new intermediary that incorporates Creative

---

[27] Internet Archive: Universal access to human knowledge, *at* http://www.archive.org (last visited June 9, 2005).

[28] *See* About the Internet Archive, *at* http://www.archive.org/about/about.php (last visited June 9, 2005).

[29] *Id.*; *see also* Kahle v. Ashcroft, 2004 WL 2663157 *2 (N.D. Cal. Nov. 19, 2004) (describing Internet Archive's collection).

[30] *Id.*

[31] *See* CC Publisher *at* http://creativecommons.org/tools/ccpublisher (last visited June 9, 2005).

[32] *See generally* Dan Hunter & Greg Lastowka, *Amateur-to-Amateur*, 46 WM. & MARY L. REV. __ (2004) (describing growth of amateur-to-amateur communication online and the obstacles that copyright law imposes on such communication).

[33] *See* John Buckman *Why I created Magnatune*, *at* http://www.magnatune.com/info/why (last visited June 9, 2005).

[34] *See* The Business Model: how we (and our artists) pay the rent., *at* http://www.magnatune.com/info/model (last visited June 9, 2005).

Commons licenses into a profit-driven business model. To market is music, Magnatune provides free radio stations that allow listeners to preview music from many different genres.[35] Royalty-free downloads are available under a Creative Commons Attribution Non Commercial ShareAlike license.[36] If listeners like what they hear, they can pay for downloadable albums or physical CDs.[37] The listener chooses what to pay, between $5 and $18 per album.[38] Purchasers can make non-commercial derivative works based on the works they purchase.[39] Magnatune also licenses music for commercial purposes.[40] The contract and the price are set by the type of use,[41] and the process is completely automated.[42] There is also no review of the use of the licensed music.[43]

The benefits of Magnatune for musicians are that the label splits revenue between itself and the artist on a 50/50 basis, which is much higher for the artist than a traditional record label.[44] Magnatune, unlike other online music sources, does not accept everyone.[45] It evaluates the artists like a traditional record label to maintain quality control.[46] According to Magnatune, "top artists make several thousand dollars per year."[47] With non-major artists on traditional labels, often no money is made by the artist.[48]

Creative Commons licensing is also being used in the publication of

---

[35] *See Id.*

[36] *See* What is "Open Music"?, *at* http://magnatune.com/info/openmusic (last visited June 9, 2005).

[37] *See Id.*

[38] *See Id.*

[39] *See Id.*

[40] *See* The Business Model: how we (and our artists) pay the rent., *at* http://www.magnatune.com/info/model (last visited June 9, 2005).

[41] *See Id.* (stating that the standard practice is for wealthier companies to be charged more for a license.)

[42] *See Id.*

[43] *See Id.*

[44] *See* Magnatune: Information: What's in it for musicians, *at* http://www.magnatune.com/info/musicians (last visited June 9, 2005). *See also* Magnatune: Information: Distribution contract terms, *at* http://www.magnatune.com/info/terms (last visited June 9, 2005) (Magnatune gives artists 50% of the *gross* on music downloads and licensing, but due to production costs, artists get 50% of the *profits* on physical items like T-shirts, posters, etc.)

[45] *See* Magnatune: Information: The Plan: problems with the music industry and how Magnatune is trying to fix them., *at* http://www.magnatune.com/info/plan (last visited June 9, 2005).

[46] *See Id.*

[47] *See Id.*

[48] *See* Magnatune: Information: What's in it for musicians, *at* http://www.magnatune.com/info/musicians (last visited June 9, 2005).

scientific research.[49] Two publishers, the Public Library of Science (hereinafter "PLoS") and Biomedcentral, use Creative Commons licenses to facilitate their respective missions to make the world's scientific and medical literature a public resource.[50] The vision is to give unlimited access to the latest scientific research, make it possible to search the text of every article to locate specific ideas, methods, experimental results, and observations,[51] and to facilitate innovative ways to explore and use the world's treasury of scientific ideas and discoveries.[52]

The journals published by both groups are peer-reviewed and feature established, well-regarded editorial boards. Rather than assign copyright to the publisher, authors grant the public a Creative Commons Attribution license, which enables these publishers to post articles on the public Web immediately upon publication. The immediate availability of this research has had noticeable effects. For example, PLoS began publishing two Journals, PLoS Biology and PLoS Medicine, and added PLoS Computational Biology in June 2005.[53] ISI Thomson, which assigns "impact factors" to scholarly journals based on the quantity and quality of citations received, assigned PLoS Biology an impact factor of 13.9, after only one year of publication.[54] BioMed Central's journals also have received impact factors that compare favorably with competing subscription-based journals, particularly in light of how young these journals are.[55] These "open access" publishers can use Creative Commons licenses in this way because they rely primarily on supply-side funding rather than the traditional demand-side funding through paid subscriptions.[56]

These two new business models show how Creative Commons can facilitate changes in the way we obtain both entertainment and important information. With Magnatune, the Creative Commons license helps listeners

---

[49] *See* Public Library of Science, *at* http://www.plos.org/index.html (last visited June 9, 2005).

[50] *See* About PLoS, *at* http://www.plos.org/about/index.html (last visited June 9, 2005); *What is BioMed Central? at* http://www.biomedcentral.com/info/ (last visited Jun. 9, 2005).

[51] *See id.*

[52] *See id.*

[53] *See* About the PLoS Journals, *at* http://www.plos.org/journals/index.html (last visited June 9, 2005).

[54] *See The First Impact Factor for PLoS Biology—13.9*, *at* http://www.plos.org/news/announce_pbioif.html#note. PLoS is launching two more journals in the near future. *See* PLoS Genetics, *at* http://www.plosgenetics.org (last visited June 9, 2005) (launching July 2005), PLoS Pathogens, *at* http://www.plospathogens.org (last visited June 9, 2005) (launching September 2005).

[55] *See* Frequently Asked Questions, *at* http://www.biomedcentral.com/info/about/faq?name=impactfactor.

[56] *See* PLoS Publishing Model, *at* http://www.plos.org/journals/model.html (last visited June 9, 2005).

and licensors find high-quality music that may not have mass appeal while creating revenue streams for artists who would have difficulty earning revenues under a traditional recording contract and would not be likely to reach as broad an audience. Open access publishers embrace the public goods nature of valuable information and use Creative Commons licenses in conjunction with a new financing model to make use of the Internet's disseminative power.

D. Creative Commons Communities

In some cases new intermediaries have adopted Creative Commons licenses as community norms. In other words, sharing is not just allowed, it is the point. A sampling of these includes the following:

Music. Opsound is an Internet record label,[57] but unlike Magnatune, this site contains an "open pool" into which all artists are invited to contribute music.[58] Opsound describes itself as "a kind of laboratory for looking at how artists can release music in a manner synergistic with the internet's capacity to encourage communication and sharing."[59] The site also describes itself as creating a "gift economy" among musicians.[60] The only requirement for adding music to the open pool is that the artist use the Creative Commons Attribution ShareAlike license, or place the music in the public domain.[61]

Creative Commons also launched its own musical sharing site, CC Mixter, with the help of a number of volunteers.[62] The site invites users to "sample, mash, & share music ... legally."[63] Everything on the site is licensed with a Creative Commons license.[64] Creators can sample and alter the music they find on the site to create their own works.[65] The only requirement is that the artist abides by the CC license used by the source artist.[66] CC Mixter also hosts contests in which artists can obtain material from the site and submit their creations.[67] In fact, CC Mixter recently held a contest in conjunction with Magnatune, the winner to receive a Magnatune contract.[68]

---

[57] *See* Opsound, *at* http://www.opsound.org/opsound/about.html (last visited June 9, 2005).
[58] *See Id.*
[59] *See Id.*
[60] *See Id.*
[61] *See Id.*
[62] *See* CC Mixter, *at* http://ccmixter.org (last visited June 9, 2005).
[63] *See Id.*
[64] *See Id.*
[65] *See* Frequently Asked Questions, *at* http://ccmixter.org/isitlegal (last visited June 9, 2005).
[66] *See Id.*
[67] *See* CC Mixter, *at* http://ccmixter.org (last visited June 9, 2005).
[68] *See Id.*

Visual Art. A similar community is the Open Clip Art Library, which contains more than 3,400 clips contributed by more than 200 artists.[69] The Open Clip Art Library "aims to create an archive of clip art that can be used for free for any use."[70] It requires that all pieces of clip art submitted be placed into the public domain using the Creative Commons statement.[71] The clip art is then available to anyone to use for any purpose. The site's clip collection has been accumulated in little more than one year.[72] As the archive grows, this site will be a particularly useful resource for non-professionals with small budgets who create things like newsletters or promotional flyers. It also allows those who create small pieces of graphic artwork to disseminate their creations to the public.

Photographs. One of the fastest-growing communities that use Creative Commons licenses is flickr.[73] Flickr is a website that allows members to show photos either to everyone only to select friends and family.[74] Flickr enables but does not require users to post photographs under a Creative Commons license. Nonetheless, as of June 1, 2005, there were more than two million photographs hosted by flickr under a Creative Commons license.[75] Flickr serves as an intermediary both for those who only wish to view photos and for those who wish to use photos for their own creative works. The relevance dimensions added by the various searchable and browsable Creative Commons-licensed photos makes flickr a significant resource for creators who seek to share the works of others.

Blogs. Creative Commons is also an important part of the Blogosphere. Technorati, an online weblog ("blog") search engine, describes the blogosphere as a conversation in which millions of people express their ideas and millions respond to them.[76] About 32 million Americans are regular blog readers.[77] There are 38,000 new blogs a day, and roughly 500,000 posts daily.[78] A Yahoo! Creative Commons search for the term

---

[69] *See* Open Clip Art Library, *at* http://www.openclipart.org (last visited June 9, 2005).
[70] *See Id.*
[71] *See Id.*
[72] *See* News, *at* http://www.openclipart.org, (announcing their one year aniversary as of April 1, 2005).
[73] flickr, *at* www.flickr.com (last visited June 9, 2005); *see also supra* notes XX and accompanying text.
[74] *See* What is flickr?, *at* http://www.flickr.com/learn_more.gne (last visited June 9, 2005).
[75] *See Id.*(containing171,172 Attribution licensed photos, 58,567 Attribution-NoDerivs licensed photos, 582,019 Attribution-NonCommercial-NoDerivs licensed photos, 302,800 Attribution-NonCommercial licensed photos, 751,548 Attribution-NonCommercial-ShareAlike licensed photos, and 159,224 Attribution-ShareAlike licensed photos).
[76] *See* About Technorati, *at* http://www.technorati.com/about (last visited June 9, 2005).
[77] *See Id.* (citing the Pew Internet Study: The state of blogging, *available at* http://www.pewinternet.org/PPF/r/144/report_display.asp).
[78] *Id.*

"weblog" returns 3,230,000 hits, and a search for "blog" returns 5,960,000.[79] Creative Commons licenses facilitate the conversation. Since the blogosphere includes not just separate blogs, but blogs that respond to, cite and quote other blogs, the Creative Commons licenses allow bloggers to build the community conversation with the legal convenience provided by those licenses. Considering the large number of blogs searchable through the Yahoo! Creative Commons search, this facilitation may be one of the most powerful uses of the Creative Commons licenses.

E. Education

Last but not least, an important intermediary function facilitated by Creative Commons licenses is in the field of education. Colleges and universities serve the societal intermediary function of disseminating knowledge to members of society, preparing them for a productive life in their chosen field and enriching their lives. Creative Commons licenses enable institutions to disseminate information to an audience beyond the university community while retaining some control over copyrighted works.

At a time when numerous institutions of higher education looked at teaching materials produced on campus as a potential revenue source through distance education, the Massachusetts Institute of Technology ("MIT") launched OpenCourseWare, "a free and open educational resource for faculty, students, and self-learners around the world."[80] The program contains 900 courses[81] from thirty-four departments.[82] MIT will evaluate this experiment over the next five years, measuring its access, use and impact.[83] MIT uses a Creative Commons license for nearly all of its content.[84] The license has enabled people from all over the world, who have Internet access, to obtain, informally, many of the benefits of an MIT education. MIT not only hopes to spread its educational material but also to promote the concept

---

[79] *See* http://search.yahoo.com/cc, (all hits will not be separate, distinct weblogs, but it is indicative of the importance and popularity of the form).

[80] MITOPENCOURSEWARE, *at* http://ocw.mit.edu/index.html (last visited June 9, 2005).

[81] *See* MITOPENCOURSEWARE: About OCW, *at* http://ocw.mit.edu/OcwWeb/Global/AboutOCW/about-ocw.htm (last visited June 9, 2005) (stating that there were 900 courses as of September, 2004).

[82] *See* MITOPENCOURSEWARE: Master Course List, *at* http://ocw.mit.edu/OcwWeb/Global/all-courses.htm (last visited June 9, 2005).

[83] *See* MITOPENCOURSEWARE: About OCW, *at* http://ocw.mit.edu/OcwWeb/Global/AboutOCW/about-ocw.htm (last visited June 9, 2005).

[84] *See* MITOPENCOURSEWARE: Legal Notices, *at* http://ocw.mit.edu/OcwWeb/Global/terms-of-use.htm (last visited June 9, 2005) (noting exceptions for some licensed 3d party material included in some courses).

of opencourseware in general.[85] At least ten other universities from the United States, Japan, and Viet Nam have launched opencourseware programs, indicating the concept's attraction.[86]

Connexions from Rice University represents another educational use of Creative Commons licenses.[87] Connexions disaggregates learning materials by using small "chunks" known as "modules" as the basic unit of course material.[88] These modules can be organized and linked into courses.[89] Learning need not be linear, and the use of modules can show "relationships both within and between topics," and show that "knowledge is naturally interconnected."[90] The goal of Connexions is to create a commons of high-quality diverse content through grassroots collaboration,[91] facilitated by use of a Creative Commons Attribution license.[92] According to the site, "[m]ore than one million people from 157 countries are tapping into over 2,400 modules and 90 courses developed by a worldwide community of authors in fields ranging from computer science to music and from mathematics to biodiversity."[93] Because of the open nature of Connexions, quality control is handled by allowing third parties to review the content, presented in the form of "lenses" that include ratings based on popularity, feedback by universities and other reliable sources, and peer assessments.[94] The modules are also being translated into several languages.[95] This shows how Creative Commons facilitates not only dissemination, but also collaboration and community building in the educational context.

Finally, Berklee Shares is a collection of music lessons prepared by the faculty of the Berklee College of Music licensed under Creative

---

[85] *See* MITOPENCOURSEWARE: About OCW, *at* http://ocw.mit.edu/OcwWeb/Global/AboutOCW/about-ocw.htm (last visited June 9, 2005).

[86] *See* MITOPENCOURSEWARE: Other Opencourseware Programs, *at* http://ocw.mit.edu/OcwWeb/Global/AboutOCW/otherocws.htm (last visited June 9, 2005) (including Johns Hopkins University School of Public Health, the Fulbright Economics Teaching Program in Viet Nam, and the Tokyo Institute of Technology).

[87] *See* Connexions, *at* http://cnx.rice.edu (last visited June 9, 2005).

[88] *See Id.*

[89] *See* Connecxions: Philosophy, *at* http://cnx.rice.edu/aboutus/philosophy (last visited June 9, 2005).

[90] *See Id.*

[91] *See* Connexions Web Tour: Vision, *at* http://cnx.rice.edu/aboutus/tour/8.html (last visited June 9, 2005).

[92] *See* Connections Content Commons, *at* http://cnx.rice.edu/content/browse_content (last visited June 9, 2005) (linking to Creative Commons Deed Attribution 2.0, *at* http://creativecommons.org/licenses/by/2.0).

[93] Connexions *at* http://cnx.rice.edu/.

[94] *See* Connexions: Tour: Quality, *at* http://cnx.rice.edu/aboutus/tour/10.html (last visited June 9, 2005).

[95] *See Id.*

Commons licenses.[96] The goal here is to provide free music lessons for the musical community around the world and to promote the Berklee College of Music.[97] While not as broad in scope as MIT OpenCourseWare or Connexions, the Berklee use stems from the same philosophy that learning should be more widely available. It also demonstrates the potential to use content offered under a Creative Commons license for promotional purposes. Berklee Shares specifically states that one of its reasons for making its content available is "to reach interested students and make them aware of the possibility and potential of a Berklee education."[98]

* * * * *

The rapid adoption of Creative Commons licenses by individual copyright owners and by a variety of new intermediaries demonstrates the utility of standardized understandings that enable some sharing of copyrighted works while reserving other rights to the copyright owner. To date, this utility has been derived primarily from the simplicity of the human-readable Commons Deed and associated icons, which quickly communicate the essential permissions and restrictions for each Creative Commons license. Soon, however, chances are that the machine-readable description of these licenses is likely to become paramount as efforts to build a Semantic Web progress.

IV. THE SEMANTIC WEB

Frustrated by technological inabilities to share documents across computing platforms, Tim Berners-Lee invented the hypertext mark-up language (HTML) and other protocols that are the foundation for the World Wide Web.[99] Having achieved document interoperability, Berners-Lee and his colleagues at the World Wide Web consortium (W3C) share a vision of a next-generation Web that takes interoperability to a higher level, a Web in which machines mine mountains of metadata in order to automate a wide variety of transactions. They call this the Semantic Web.

The idea is to add logic to the Web, meaning to use "rules to make inferences, choose courses of action and answer questions."[100] Two

---

[96] *See* Berklee Shares.com: Welcome, *at* http://www.berkleeshares.com (last visited June 9, 2005).
[97] *See Id.*
[98] Berklee Shares.com: Frequently Asked Questions: Why is Berklee doing this?, *at* http://www.berkleeshares.com/faq (last visited June 9, 2005).
[99] *See* Lawrence Lessig, The Future of Ideas:
[100] Tim Berners-Lee et al., *The Semantic Web: A New Form of Web Content That Is Meaningful To*

technologies exist for developing the Semantic Web, "eXtensible Markup Language (XML) and the Resource Description Framework (RDF).[101] XML allows creators to extend the standardized tags used in HTML to tag their content however they like, and RDF gives meaning to that content.[102] The goal of RDF is to enable machines to identify relationships among data at a conceptual level by using XML tags to create "triples," much like subject, verb, object in a normal sentence.[103] Each part of the triple is identified by a Universal Resource Identifier (URI), rather than a normal phrase.[104] This allows similar but different concepts, universally defined, to be distinguished by machines.[105]

RDF uses "ontologies" to describe relations of terms.[106] Ontologies enable machines not only to distinguish between similar but different concepts, but also, through the use of "equivalence relations," to understand that some things are the same though they are described using different terms.[107] In the field of Internet search, for example, ontologies can improve the accuracy of Web searches along the familiar topical dimension by looking for only those pages that refer to a precise concept, ignoring those that use

---

Computers Will Unleash a Revolution of New Possibilities, SCIENTIFIC AMERICAN 3 (May 2001), *available at* http://www.scientificamerican.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21&catID=2 (last visited June 9, 2005) (contemplating a future web where web connected devices will use "agents" to communicate and perform tasks). The site that claims to be the first site on the Semantic Web, see Mindswap *at* http://www.mindswap.org/.

[101] *Id.*

[102] *Id*, (tagging means to attach a hidden label to content that can be used by programs).

[103] *Id.* (noting that this allows web pages to assert that "things . . . have properties . . . with certain values").

[104] *Id.* (explaining that the most common example of a URI is a Uniform Resource Locator or URL, which is the format for the location of all pages on the current Web).

[105] *See Id.* (using as an example the difference between an address that is a post office box, an address that is a street address, and a speech).

[106] *See Id.* (ontology includes taxonomy, meaning the definition of "classes of objects and the relations among them" and inference rules, which allow the computer to manipulate the terms ... in ways that are useful and meaningful to the human user"); *see also* Scientific American.com: Sidebar: May 18, 2001: Glossary: Ontologies, *at* http://www.scientificamerican.com/article.cfm?articleid=00019A83-2B28-1CBF-B4A8809EC588EEDF (defining ontologies as " Collections of statements written in a language such as RDF that define the relations between concepts and specify logical rules for reasoning about them. Computers will "understand" the meaning of semantic data on a Web page by following links to specified ontologies") (last visited June 9, 2005).

[107] *See* Tim Berners-Lee et al., *The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities*, SCIENTIFIC AMERICAN (May 2001), *available at* http://www.scientificamerican.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21&catID=2, at 6 (last visited June 9, 2005) (using the example that zip code and postal code are different phrases that describe the same thing).

ambiguous keywords.[108] Moreover, ontologies theoretically could facilitate multidimensional searches.[109]

The dream of the Semantic Web has been elusive, but progress is being made. Last year, the W3C approved RDF and the Web Ontology Language (OWL), as standards.[110] Berners-Lee has encouraged developers to create applications to "justify the Semantic Web in the short term."[111] An application called Haystack, developed at MIT, reportedly "knocks down the partitions that separate e-mail clients, file systems, calendars, address books, the Web and other repositories so that information can be worked with regardless of its origin."[112]

If realized, the Semantic Web vision has profound consequences for law – deeper than the now-familiar concerns about electronic agents[113] and machine-enforceable rules.[114] Creating machine-interpretable and machine-

---

[108] *Id.* at 4.

[109] If successfully deployed, the Semantic Web also would greatly increase the role of electronic agents. *Id.* at 5. Berners-Lee and his colleagues offer a hypothetical in which a pair of siblings make a doctor's appointment for their mother using their respective web agents. *See Id.* at 1 (theorizing that the agent could find a doctor in their mother's insurance plan, the best office location with considerations for traffic and the best time to schedule the appointment to avoid major conflicts with existing obligations, all automated). These agents communicate, verify the identity of other agents, ask for "proofs" of the data they receive to ensure accuracy, and locate agents across the web that provide desired services through a directory. *See id.* Agents will understand each other through the exchange of ontologies, and indeed agents will be able to acquire new "reasoning capabilities" as they find new ontologies. *Id.* at 6. Eventually, this will extend from the web to the physical world when other items become web enabled. *See id.* (theorizing that in the future, devices will be able to communicate and control each other, like a phone call triggering a reduction in the volume of a stereo or television). According to the authors, even microwaves may be able to contact the manufacturer of a frozen meal to learn the perfect way to cook that meal. *See id.*

[110] *See* Darryl K. Taft, *W3C Approves Pair of Semantic Web Specs,* eWEEK: Web Services, *at* http://www.eweek.com/article2/0,1759,1524304,00.asp (Feb 11, 2004) (citing support from twenty-four organizations involved in technology including IBM, Adobe, and the U.S. Department of Defense).

[111] *See* Anne Chen, *Semantic Web Is 2 Steps Closer,* DevSource: Add Ons, *at* http://www.devsource.com/article2/0,1759,1621521,00.asp (July 6, 2004) (quoting Tim Berners-Lee).

[112] Lisa Vaas, *Berners-Lee Maps Vision of a Web Without Walls,* eWEEK.com: Infrastructure, *at* http://www.eweek.com/article2/0,1759,1734926,00.asp (Dec. 2, 2004) (describing Berners-Lee's vision and describing Haystack as a prototype that utilizes the ideas of the Semantic Web); Haystack *at* http://simile.mit.edu/hayloft/index.html

[113] *See*, *e.g.*, Anthony J. Bellia, Jr., Contracting With Electronic Agents, 50 EMORY L.J. 1047 (2001); Stephen T. Middlebrook & John Muller, *Thoughts On Bots: The Emerging Law Of Electronic Agents*, 56 BUS. LAW. 341 (2000); *see also* Margaret Jane Radin, *Online Standardization and the Integration of Text and Machine*, 70 FORDHAM L. REV. 1125 (2002); Margaret Jane Radin, *Humans, Computers, and Binding Commitment*, 75 IND. L.J. 1125 (1999).

[114] *See generally* Symposium Issue, *The Law & Technology of Digital Rights Management*, 18 Berkeley Tech. L. J. 487 (2003); Dan L. Burk & Julie E. Cohen, *Fair Use Infrastructure For*

actionable concept maps of the law will enable more radical departures from the default rules the law supplies than we have previously experienced.[115] Further, the process of building machine-interpretable concept maps is likely to alter our understandings of the concepts being mapped. Creative Commons licenses, which use RDF at the machine-readable layer, are just the tip of this particular iceberg. Efforts to create a "policy aware" Web, appear to be a next step that lawyers should keep an eye on.[116] Although developers imagine the policies of which the Web should be aware to be private policies adopted by those who provide Web resources, the technologies also could be adapted to reflect public policies as well.

Some see the Semantic Web project as fundamentally flawed.[117] These critics charge that the vision requires too much complexity and demands that users adapt to the needs of machines instead of adapting machines to the needs of users.[118] Machines use rules to process information, and rules require classification of information to be useful. People may use rules to classify information, but we often use different rules depending on context, and we may not agree about which rule to apply in any given situation. In a well-argued essay Clay Shirky writes that semantics are in the users not the system and that ontological classifications, such as those required for the Semantic Web, work in certain limited domains but will not

---

Rights Management Systems, 15 HARV. J.L. & TECH. 41 (2001); Julie E. Cohen, *Lochner In Cyberspace: The New Economic Orthodoxy Of "Rights Management"*, 97 MICH. L. REV. 462 (1998).

[115] This point should not be confused with the argument that "code is law." *See generally* LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE (1999); Update *at* http://codebook.jot.com/System/Templates/single_tab/Search?key=code+is+law. That argument is that software regulates behavior separate and apart from the way that law does and that on the Internet these regulatory modalities are interchangeable, with code being the more effective in many cases. For further discussion, see, e.g., Tim Wu, *When Code Isn't Law*, 89 VA. L. REV. 679 (2003); R. Polk Wagner, *On Software Regulation*, 78 S. CAL. L. REV. 457 (2005); James Grimmelmann, Note, *Regulation By Software*, 114 YALE L. J. 1719 (2005).

[116] *See*, *e.g.*, Policy Aware Web *at* http://www.policyawareweb.org/; Daniel J. Weitzner, Jim Hendler, Tim Berners-Lee & Dan Connolly, *Creating a Policy-Aware Web: Discretionary, Rule-Based Access for the World Wide Web*, *in* WEB AND INFORMATION SECURITY ___ (E. Ferrari & B. Thuraisingham, eds. forthcoming), *also at* http://www.w3.org/2004/09/Policy-Aware-Web-acl.pdf. I thank Hal Abelson for this point.

[117] Berners-Lee reminds that many also saw the vision for the World Wide Web as fundamentally flawed. *See An [sic] parenthetical discussion to the Web Architecture at 50,000 feet. [sic] and the Semantic Web Roadmap, at* http://www.w3.org/DesignIssues/RDFnot.html (last visited June 9, 2005).

[118] *See*, *e.g.*, Eric Nee *Web Future is Not Semantic, Or Overly Orderly,* CIO|INSIGHT, *at* http://www.cioinsight.com/article2/0,1397,1817758,00.asp (May 5, 2005) (quoting Google cofounder Sergey Brin as saying "I'd rather make progress by having computers understand [sic] what humans write, than by forcing –humans [sic] to write in ways that computers can understand").

work for the Web at large.[119]

Shirky and other critics, influenced by recent thinking about complex systems, argue that simple technologies like Really Simple Syndication (RSS) and "social bookmarking" better enable user-defined complex organization and classification.[120] RSS enables users to automatically check to see if web pages marked with the appropriate XML tags and to aggregate results. Used by millions of bloggers and now most mainstream news sites, RSS has been one of the most rapidly-adopted Internet technologies in recent years.[121]

Social bookmarking is a development destined to warm every postmodernist's heart. Social bookmarking and tagging enable quick publication and aggregation of metadata about resources, such as web pages, available on the Internet. Tagging theoretically enables us to forgo hierarchical classifications – such as "organizing your favorites" into folders – and the habits of mind associated with such classification.[122] Moreover, the technology enables probabilistic classifications that democratize and make explicit the social construction of meaning.[123] By publishing the list of web pages that you have bookmarked in your Web browser, you implicitly make a statement that of all the resources available on the Web, these are relevant to you in some way. Social bookmarking sites, such as www.deli.icio.us, offer to host a user's bookmark file – thereby making the file available to the user on any Net-connected computer – and to publish the file, or parts of it, to all, or selected, Web users.[124]

These sites also permit users to also to associate "tags," i.e., keywords, with these Web addresses and make more explicit the ways in which these sites are relevant. Flickr's photo hosting site, discussed above, is one of the fastest-growing uses of social tagging, enabling searches for

---

[119] *See*, *e.g.*, Clay Shirky, *Ontology is Overrated: Categories, Links, and Tags at* [LINK]

[120] *See id.*; Nee, *supra* note XX, (promoting the technology of Google and Really Simple Syndication (RSS) as pragmatic alternatives to the Semantic Web theory).

[121] CITES Made popular by bloggers, nearly all major news sites provide RSS feeds now. CITES

[122] I say "theoretically" because even though tagging does away with the need for visual representations of conceptual hierarchies – such as a file folder organization scheme – most people use conceptual hierarchies to make sense of the world and we should expect to see those hierarchies reflected in their tags.

[123] A probabilistic classification asserts that *Z*% of users think that *X* is relevant to *Y* rather than asserting that *X is* relevant to *Y*. *See* Shirky. It is much easier to give a computer the authority to make the former statement than the latter. *Id.*

[124] As scholarly research continues to migrate to the Web, some scholarly publishers see the value of social bookmarking for communities of researchers as well. The Nature Publishing Group's Connotea site targets scientific researchers to signal to each other which Web resources, such as online articles, they deem to be most important or relenvant. *See generally* www.connotea.org; www.flickr.com.

photographs along the topical dimension by popular tags,[125] or along the copyright dimension for photographs available under a Creative Commons license.[126]

Much of the opposition to the Semantic Web is misdirected. The spread of these simple technologies is not antithetical to the Semantic Web. Indeed, the Semantic Web vision requires that there be rich metadata associated with information available on the Web. The creation of metadata is costly. It may well be that simple technologies that supply incentives for the creation of such metadata are prerequisites to realization of a Semantic Web. RSS tags give you the news of the day, social bookmarks can influence what you read, and Creative Commons metadata tells you about the copyright status of the information you encounter.

Moreover, RDF's first mission is to enable interoperability. As various social bookmarking and tagging systems emerge, RDF can serve as a bridge between these systems. Similarly, as machine-readable licensing becomes more common, RDF can be used to identify equivalence relations between licenses and/or license terms. When applied to public law, RDF could also be used to identify equivalence relations between the legal codes of various jurisdictions – taking international legal harmonization in a new direction. In many ways, Creative Commons licenses are a test case for the possibilities of machine-readable law, and this development is worth following.

V. CONCLUSION

The number of copyright events occurring in our daily lives continues to grow as our collective use of digital media continues to expand. Creative Commons licenses facilitate coordination and regulation of these events by enabling end-to-end copyright transactions and by fueling the growth of new intermediaries that rely on the common pool of Creative Commons-licensed content. As this pool expands, it is possible to imagine the growth of a cultural counter-canon along the copyright dimension, but Creative Commons-licensed content need not compete with content available on more traditional copyright. Indeed, Creative Commons licenses are being integrated into traditional commercial licensing practices, although this remains an underexploited growth opportunity for new and old intermediaries. Finally, intermediaries increasingly will begin to use and rely upon the machine-readable descriptions of Creative Commons licenses, expressed in RDF, as the importance of the copyright relevance dimension

---

[125] *See* Hot Tags, *at* http://www.flickr.com/photos/tags/

[126] *See* Creative Commons, *at* http://www.flickr.com/creativecommons/ (grouping photographs by license terms); *see also infra* Section III.F. (discussing Flickr further).

increases and as the idea of machine-readable law becomes better understood.