

The economics of open source hijacking and the declining quality of digital information resources: A case for copyleft**by Andrea Ciffolilli****Abstract**

The economics of open source hijacking and the declining quality of digital information resources: A case for copyleft by Andrea Ciffolilli

The economics of information goods suggest the need for institutional intervention to address the problem of revenue extraction from investments in those resources characterized by high fixed costs of production and low marginal costs of reproduction and distribution. Solutions to the appropriation issue, such as copyright, are supposed to guarantee an incentive for innovative activities at the price of few vices marring their rationale. In the case of digital information resources, apart from conventional inefficiencies, copyright shows an extra vice since it might be used perversely as a tool to "hijack" and privatise collectively provided open source and open content knowledge assemblages, even in the case in which the original information was not otherwise copyrightable. Whilst the impact of hijacking on open source software development may be uncertain or uneven, some risks are clear in the case of open content works. The paper presents some evidence of malicious effects of hijacking in the Internet search market by discussing the case of The Open Directory Project. Furthermore, it calls for a wider use of novel institutional remedies such as copyleft and Creative Commons licensing, built upon the paradigm of copyright customisation.

Contents[Introduction](#)[The economic nature of digital information resources: Virtues and risks of appropriation institutions](#)[Hijacking in cyberspace: The case of the Open Directory Project](#)[The promise of new IPR institutions: Copyleft and "customisation agencies"](#)[The Creative Commons approach: Tailoring a suitable license](#)[Conclusion](#)



Introduction

Digital information resources, expressing ideas such as creative works, education and scholarly materials, databases and software, are notoriously characterised by massive fixed costs of original production and low marginal costs of reproduction and distribution. As a consequence, the extraction of economic benefits from these goods is tricky and special institutional devices are needed to address the issue and encourage innovation. Awarding private ownership rights in order to allow commercial exploitation is a solution. While intellectual property right institutions, such as copyright, contribute to the solution of the problem of appropriation of rents, they are marred by few vices such as a monopoly deadweight, as well as inefficiencies regarding cumulative innovation, standardisation and modular development (David, 2000).

I argue that the digital age, in which those who arrange bits in certain order then own the arrangement [1], calls for an extra vice to be added to the list of copyright's undesirable effects. This is the possibility of subtracting other people's works from the public domain and embodying them in proprietary assemblages, even when the original content could not be otherwise copyrighted. I call this practice "hijacking" [2] and when undertaken, the very rationale behind copyright protection is abused.

Having spotted this extra weakness, far from considering obsolete the institution of copyright and its justification, I maintain that a degree of flexibility in its design and application would be beneficial in cases in which the social risks of hijacking are unquestionable. Hijacking may show up in different forms, from direct appropriation of content or code, to creation of a proprietary complementary product built upon a public domain work, whose potential developments result in being constrained (*e.g.* a proprietary application designed for an open source programme).

Moreover, hijacking is not alone but has an accomplice in the plot to privatise scientific information and data; it is the general regulatory trend towards extension of IPR protection [3].

The [second section](#) of this paper recalls briefly the peculiar economics of information goods, the problem of appropriation as well as the classical solutions applied to it. It is emphasised that currently hijacking carried out by certain commercial firms may possibly add to the list of vices marring the scheme of exclusive property rights granted, in order to foster creative productions.

The [third section](#) deals with a particular case of actual hijacking that is representative of the risks associated with proprietary appropriation of a collective good. The discussed case is that of The Open Directory Project, a pure public good exploited by commercial search engines and directories which incorporate its data in information arrangements whose quality appear, given the market dynamics, more and more dubious.

The [fourth section](#) discusses the emergence of new institutions, such as copyleft and think tank or "customisation agencies" (e.g. Creative Commons), capable of exorcising the problem. The [final section](#) draws conclusions.



The economic nature of digital information resources: Virtues and risks of appropriation institutions

Information as a transferable flow of facts and details, communicating concepts or ideas, constitutes an economic good that shows peculiar characteristics.

Information resources share with conventional public goods the properties of non-rivalry, non-excludability and high fixed costs of original production. Non-rivalry means that the amount of good available for consumption does not vary with the number of consumers drawing upon its stock. Non-excludability means that, given the low marginal cost of reproduction and distribution of a public good, it turns out to be tricky to charge a price for every taker. Massive fixed costs of original production imply the non-sustainability of a competitive market for this kind of goods.

Apart from these features, information is an experience good whose distribution is asymmetric. Assigning a value to an experience good is troublesome before consumption.

A bottle of wine is a typical experience good. Information experience goods such as newspapers, far from being necessarily mellifluous, are characterized by novelty each time they are consumed (Shapiro and Varian, 1999). Moreover, being asymmetrically distributed, it is not known when a piece of knowledge will be available in codified form (David, 2003).

Another unique feature of information resources is that knowledge, defined as a mix of information and other facts and details more difficult to codify and readily transfer (Polanyi, 1966; Cowan, *et al.*, 2000) which constitute cognitive human capabilities, is cumulative and interactive in a way in which advances in state of the art build upon previous findings in unpredictable manners.

The foregoing features render public goods in general and information resources in particular, susceptible to free riding and predisposed to being underprovided.

Public subsidies to firms engaging in certain productions, direct public provision and regulated monopoly are classical solutions for the conventional public good problem. Analogously, there are three main remedies for the problem of appropriation of information rents. Sometimes recalled as "the three P's" (David, 2000), these are: *patronage*, that is awarding publicly funded grants based upon the submission of research

proposals; *procurement*, that is governmental provision or contracting for intellectual work; and, finally *property*. The last solution consists of the concession of exclusive property rights to new knowledge creators.

Information is an experience good whose distribution is asymmetric. Assigning a value to an experience good is troublesome before consumption.

Regarding the legal institution meant to solve the problem of appropriation of digital expression of ideas through concession of exclusive rights, namely copyright, some vices emerge to counterbalance the main advantage, say the incentive to innovation. The deadweight of monopoly and the network inefficiencies regarding standardization and modular development are among the main defects. The deadweight of monopoly implies that an information good available for a price higher than its marginal cost cannot be afforded and hence consumed by everyone even though additional supply practically does not cost anything. Inefficiencies with respect to standardization and modularity concern especially software products and indicate that exclusive property rights may oblige agents to pursue alternative, non-infringing innovation paths, with consequent proliferation of products and units characterized by incompatibility and technical inseparability.

Apart from these, there is the possibility to wickedly make the most of the appropriation institution and privatise public domain information resources.

This sort of predatory action, here defined as hijacking, implies taking control and possession of other's freely accessible works without leaving, in this case, any hope of deliverance. When an information resource is collectively provided and placed in the public domain, hijacking sounds even more censurable and in theory resembles a real theft.

But what can be the actual economic and social downfalls, if any, of this strategy? After all, we live in a world where *homo homini lupus est* (Plautus, 1994; Hobbes, 1641) and predatory actions are part of the competitive game, nourishing the Smithsonian *invisible hand*. Some reasons why it is argued here that hijacking is undesirable, reside in what will be discussed later on in this paper.

In a way it is paradoxical that the goods most susceptible to being hijacked, say libre software and open content works, are the very ones fuelled by a set of diverse motivations (*e.g.* Lerner and Tirole, 2000; Mateos Garcia and Steinmueller, 2003a; von Hippel and von Krogh, 2003; Bonaccorsi and Rossi, 2004) that render free-riding less relevant or even desirable for their development (Ghosh, 1998; Raymond, 1999; Weber,

2000). In other words, in their case, the positive externality, generated by increasing the size of the network outweighs the value of exclusivity as a reason to avoid free-riding (Ciffolilli, 2003a).

Free-riding of information does not imply depletion; hijacking is different since it means taking possession of and fencing otherwise freely accessible resources. Hence, hijacking translates in exhaustion with respect to all the individuals and bodies orphaned by the new unwarranted access barrier.

Although open source software endeavours can be definitely hijacked, there is no agreement on the fact that this necessarily constitutes a damaging circumstance. The diffused and rational worry is that the proprietary strategy to copyright a collective produced public good may "hold up" developers that lose the ability to customise a project to their needs (Lerner and Tirole, 2003). However, some commentators and project participants disagree and stress that hijacking is not likely to happen often and, when it does, it bears desirable spillovers (Reese and Stenberg, 2001). The latter claim is based on the belief that hijacking can only be avoided by using copyleft [4], but the highly restrictive nature of this kind of license may also hinder the development prospects of a project. The rationale behind this claim goes as follows.

While it is true that copyleft was originally designed by Richard Stallman [5] to prevent predatory and anti-cooperative behaviours leading to appropriation of public domain works, it is also a fact that commercial firms planning to exploit open source endeavours must face costs and barriers that make the actual encroachment difficult. In other words, it is not guaranteed that the savings in development and maintenance costs associated with an in-house product as well as other benefits of hijacking will exceed its costs.

In the case of software for instance, the costs and difficulties arising from the search for a useful product, the validation of the found object, its integration, assessment and maintenance may indeed make hijacking a non-profitable option for private firms. On the basis of similar claims, some copyleft opponents argue that there is little evidence of commercial exploitation in the real world and when it happens, consequences can even be positive for virtual communities [6].

If the foregoing considerations are definitely important for software, it may be claimed that in the case of open assemblages, characterized by complementary dependency (Mateos Garcia and Steinmueller, 2003b) and lower costs of exploitation, the argument seems somewhat more dubious.

Thus, even if the general frequency of hijacking as well as its effects on software development are debatable, I would dare to argue that, in the case of certain open content works, hijacking might clearly reveal itself as a real threat with iniquitous consequences.

Indeed, when endeavours aiming to contribute to public domain knowledge assemblages are appropriated, outcomes may be spoiled, with possible negative consequences on the

spontaneous provision of a public good. The case of The Open Directory Project aims to illustrate this idea.

Hijacking in cyberspace: The case of the Open Directory Project

The Open Directory Project or DMOZ (*i.e.* Directory Mozilla) is a human edited Web directory [7] constructed and maintained by a vast, global community of volunteer editors. It currently comprises over 4.5 million sites, more than 60 thousand editors and over 590,000 categories [8].

DMOZ was founded in the spirit of the open source/free software movement and is totally free. There is no cost either to submit a site to the directory or to use its data.

The ultimate vision of DMOZ is to build a definitive catalogue of the Web, therefore providing the means for the Internet to organize itself. At the root of this ambition is the possibility to exploit Linux's Law that can be interpreted here as: The more people there are editing the directory, the greater its comprehensiveness becomes and the higher its value in discriminating between the useless and the best Web content.

Anybody can sign up and contribute to DMOZ by choosing a category of interest and applying. The project is also characterized by a system of distributed authority (Mateos Garcia and Steinmueller, 2003a). As editors gain experience with specialized subjects, they can move up in the hierarchy and edit more general categories.

The copyright of the catalogue is owned by the Netscape Communications Corporation. The directory is made available to the public under the terms of the Open Directory License, a non-exclusive license that allows free use and download of DMOZ content as long as recognition is given to Netscape [9].

The Open Directory Project was born mainly in response to the problem of long delays with which the well-known directory Yahoo! processes applications and lists Web sites. Its current dimension and relative success notwithstanding, DMOZ hardly joins the list of the most popular Internet search sites (see [Figure 1](#)). This might be due in part to technical troubles [10] thought to plague the directory (Olsen and Hu, 2003), in part to the fact that DMOZ did not actually manage to solve delay problems affecting commercial players. In this respect, congestion costs play an important role. Indeed, some editors [11] observe that fifty percent of the sites submitted for review are spam links. The huge backlog from bad submissions has led to a delay in the process of site reviewing of up to two years.

Moreover, the very vision of building a definitive catalogue of the Web appears intrinsically somewhat problematic. In the case of knowledge assemblages characterized by complementary dependence, such as DMOZ, subjective value judgements are heavily involved in the process of submission of contributions. There are also problems of agreement in both the directory structure and listing policy.

Despite these downfalls, the Open Directory Project database constitutes a massive and valuable resource, regularly exploited by commercial search engines and directories [12]. Google and AOL (which owns Netscape) are usual "shoppers" and even Yahoo! uses DMOZ data to enhance its relevant search results [13]. All this would not be a big deal, if the search engine market was not going through serious and important changes.

In general, Web directories are dropping behind search engines. The latter automatically crawl the Internet and record sites found on the basis of certain search algorithms that, at first glance, seems to guarantee better results, either in terms of the reach or the quality of the searched information.

The number of search engines has reduced substantially over the last few years, probably to an extent as a consequence of the new economy crisis that opened the millennium. In general, there is less advertising funds keeping them afloat (Vaughan, 2003). For instance, Open Text started in 1995 and terminated its Web search services in 1997; both Magellan and Infoseek, born in 1995, closed in early 2001; Snap ended its internal search technology in 2001, after four years of activity; Direct Hit was born in 1998 and deceased in 2002. Some very popular engines such as WebCrawler, Lycos, Excite and HotBot started outsourcing search technology (Sullivan, 2003). Others, such as AltaVista, have been acquired and even if they did not disappear completely, they eventually lost their appeal and their market share.

Some search engines are characterized by technical problems that affect the quality of information retrieval. First of all, there are a variety of markup "tricks" [14] that confuse indexing procedures, thus making it difficult for a specific search engine to discriminate between a catalogue of quality and a collection of rubbish.

Secondly, a problem of reduced quality may arise naturally from the search algorithm itself. In Google, for example, PageRank, assumes that the most relevant pages are those that attract the greatest number of links. Accordingly, the top results of a search are often online shops (if the searched item can be sold); as stressed by Johnson (2003), the reason is probably two-fold. When a product is mentioned on the Web, a link to an online shop is also conventionally inserted. In other cases, there are some sites engaged in tracking prices and online availability of items, creating a great deal of links to stores in databases of search engines. A further problem is "googlewashing," a phenomenon that happens when a group of prolific linkers can drive the online identification with a certain word. For instance, the search for the word "apache" on Google, produces over 20 million results [15]. Most of them are related to the Apache open source Web server. Some results link to Web sites of organisations whose names contain the word "apache."

Where are our American Indian friends then? Actually, we must wait for about the third page of results to obtain some information we have been looking for, as well as some news concerning the deadly helicopter called, alas, Apache. The point is that pages dealing with the Apache Web server gather a high concentration of links, simply because the majority of very active Internet users, as well as bloggers, are more interested in the Web server than American Indians. There might be a great deal of pages dealing with tribes, and swarms of interested surfers seeking them, but none of those pages would ever generate the same amount of hypertext links that only one major Apache Web server portal is able to gather.

Even if the existence of these flaws is unquestionable, at the same time it cannot be denied that a wise use of search techniques makes it possible overcoming most biases. The case of "apache" is deliberately an exaggeration since one can use the keyword "tribe" in conjunction with "apache," for example, in order to retrieve relevant outcomes.

A possibly related, but much more relevant, problem is that free and paid rankings are beginning to be mixed together without users aware of these alterations. This development is one of the latest to appear in the search engine market.

In July 2003, Yahoo! launched a 1,6 billion (U.S.) dollar operation to acquire Overture. Overture is a leader in the market of paid rankings. The results of a search on Overture are listed according to fees paid by the owners of specific sites, rather than on the basis of different criteria such as relevance to a particular search [16].

With this acquisition and the development of its own technologies [17], Yahoo! clearly intends to displace Google's leadership in the search engine market. This development would have been considered incredible just a few months ago, considering that Yahoo! played a crucial role in the growth of Google, adopting it as an official search engine on its Web site since October 2002.

It is not surprising given that Google is entering the market of fee-influenced searching [18]. Overture fought back in a way by acquiring AltaVista and AllTheWeb, with the purpose of undermining Google's leadership in the field of free searching.

Before these recent developments, both Google and Yahoo! had already begun to sell "real estate" to online stores, a strategy already pursued in the past by Lycos and Infoseek. Yahoo! started in 1999 to require fees from Web sites in order to retain their listings. Later, Google inaugurated its sponsored links. Negative consequences for information reliability are indisputable, especially for those who use search engines for educational and research purposes.

Now, the latest move of Yahoo! (*i.e.* the acquisition of Overture) outlines a new scenario, characterized by an extremely concentrated market and the contraposition of a few big actors competing in both free and paid searches. [Figure 1](#) provides a snapshot of the Internet search market.

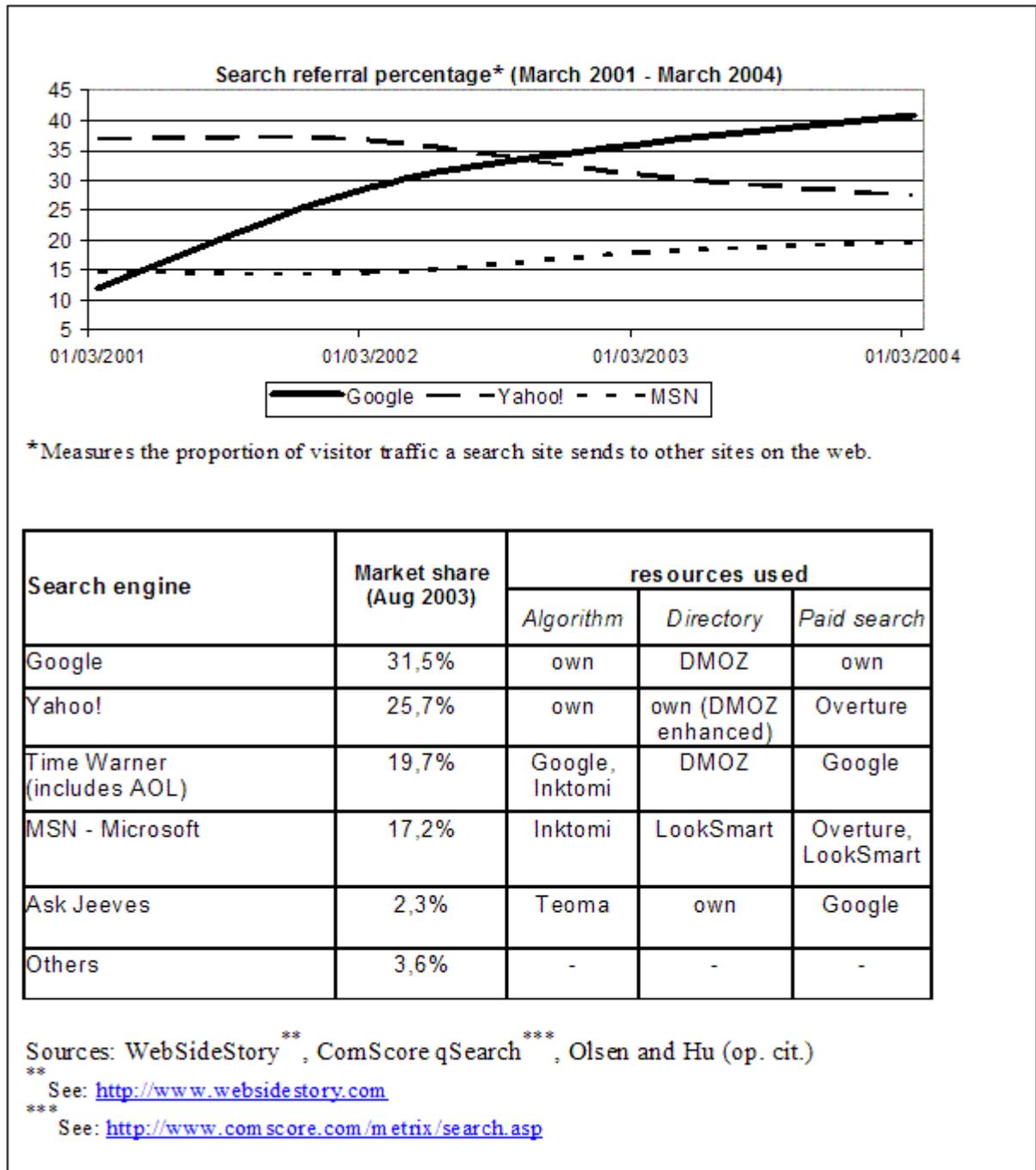


Figure 1: Some data on the Internet search market.

There is a risk that the distance between paid- and relevance-driven searches is going to fade. In general, the non-commercial roots of Internet-based directories are drying up and with them the opportunity to distinguish between useful and worthless information.

The value of a knowledge assemblage characterized by complementary dependence is reduced by the subtraction of a piece of information. Even if this is not detrimental to the

usability of the collection, the systematic incorporation of low quality contributions (*e.g.* the mix between paid and free searches) may render the assemblage useless, in the medium term.

The systematic encroachment of the Open Directory Project data, given market dynamics and the Open Directory attribution license, does not constitute a piece of welcome news. Indeed, the very vision of creating the ultimate catalogue of an over congested Web is at risk when information may be blended, according to the monetary weight of sponsors rather than to some objective standard of relevance, and eventually included in a "new" copyrighted digital arrangement. All of these developments may likely undermine an editorial process that fights the commercialisation of Web searching, attempting to preserve the *raison d'être* of a directory as an information source. Hence, it seems important that DMOZ defend its uniqueness and integrity, as well as its commitment to the construction of a useful and reliable open resource.

A copyleft, rather than a simple attribution licence, might be in this case a wise choice. It does not forbid commercial exploitation, but only shields the coherence of the collectively constructed public good. In this case, it reduces the exploitation of genuine information and its loss from the public domain by its incorporation into pay-per-play catalogues.

Think of DMOZ like a river whose water is clean in the proximity of its source but then gets polluted along the way to the sea. Providing healthy mineral water to a thirsty market implies bottling nearby the source. Those who draw upon DMOZ database are otherwise delivering information "bottled" near the estuary, where a thriving industrialising town is located.

However, the "source" may dry up if, in the long term, contributors are discouraged from participating in the project. If those who control proprietary information assemblages embodying DMOZ data, predominate the search market and its distribution channels, DMOZ resources might eventually become accessible only in formats in which their blenders are willing to provide them (*e.g.* a mixture of paid- and relevance-driven information).

There are certainly long-term effects. Other endeavours aiming to contribute to public domain knowledge can be appropriated, outcomes may be spoiled, with possible negative consequences on the spontaneous provision of a public good. This case clearly demonstrates some of the risks in the exploitation of the public digital domain.



The promise of new IPR institutions: Copyleft and "customisation agencies"

Copyleft is a novel license provision which, thanks to a creative and wise use of copyright law, seems able to permanently affect the development path of digital knowledge assemblages released under its terms [19]. Indeed, if a work is copylefted, everyone can copy, use and modify it, and then distribute the modified versions without asking the copyright holder for permission, as long as the derivative works are also released under the same license terms. These conditions, sometimes derogatively referred to as "viral" in nature, assures that a work is freely available, remains as such, and ensures the same conditions for any improvements or enhancements of the original. These licenses represent a new paradigm in the design and interpretation of intellectual property rights. In the market of digital information goods, this paradigm is competing with the traditional one, copyright. Both aim to solve a certain set of legal and socio-economic issues, namely the appropriation of economic benefits and the promise of a certain life expectancy to collectively created digital goods.

Copyleft does not preclude commercial exploitation of a piece of work. Complementary services and improvements or modifications of the work itself can be sold but the copylefted content will not be subtracted from the conservancy in which it was placed and raised [20].

All these features render this path-breaking legal innovation a restrictive [21] provision (Lerner and Tirole, 2003), not completely free from shortcomings (Ciffolilli, 2003b). In general, a high degree of restrictiveness [22] can be smoothed by pursuing further IPR customisation. This can be interpreted as either *ad hoc* design of license provisions or application of dynamic licensing (Bezroukov, 2002). The former case resembles the strategy of the Creative Commons, the latter implies designing licenses in a way that their terms change according to the life cycle of an information resource.



The Creative Commons approach: Tailoring a suitable license

The Creative Commons approach towards licensing proves how the use of restrictive provisions can be made flexible and hence solves the seeming oxymoron of "flexible copyright." Even if all of the issues of a restrictive licence cannot be solved once and for all, the friction among positive and negative effects can definitely be smoothed with an innovative approach to licensing that makes flexibility and customisation its main virtues.

The vision

The Creative Commons (CC) was founded in 2001; it is housed at Stanford Law School, from which it receives support and shares staff and premises [23]. The organization is conducted by a pool of cyberspace, "cyberlaw" and intellectual property rights experts.

CC makes direct reference to U.S. law, but its intellectual property strategies are, in principle applicable anywhere.

U.S. legislative changes of 1976 (Copyright Act) and 1988 (Berne Convention) [24] introduced automatic copyright for creative works. These changes have certainly affected artists; borrowing artwork may now invoke a court date. If the current copyright laws had been in effect earlier whole genres, such as collage, hip-hop, and Pop, might have never existed [25].

CC supporters believe that, without the legal provision of "copyright by default," many authors would have been willing to choose a different degree of protection for their works. In a way, CC is representative of a way in which some individuals would either like to dedicate their creations to the public domain or to exercise some, but not all, of their legal intellectual property rights [26].

Open content is a philosophical context in which the Creative Commons develops its menu of licenses.

CC's main goal is to provide an easy mechanism that allows authors to customize copyright law creatively according to their notions of flexibility. The project's vision makes direct reference to the legal concepts of public domain, the idea of a commons, open content and intellectual property conservancies [27].

Public domain is defined as a body of intellectual endeavours unfettered by law. Innovation and creativity rely on this heritage. It is particularly important in this digital age of collaborative creative activities when it is not threatened by the expansion of intellectual property protection. This expansion contributes to the implementation of a commodity transaction model of information creation and distribution, endorsing the interests of some economic agents or constituencies, whilst utterly disregarding others (Mansell and Steinmueller, 2000).

If public domain can be considered a container, then the commons represents its content of inexhaustible resources, jointly held and accessible without permission. Open content is a philosophical context in which CC develops its menu of licenses, a set of legal provisions that allows anyone to use certain works without specific permission or royalties. The final goal of CC is the manufacture of an intellectual property conservancy, where works of particular public importance are not allowed to become exclusively owned by any one party and are protected from obsolescence.

The licenses' menu

In December 2002, Creative Commons started its activity with the release of a set of copyright licenses free for public use.

Although CC was inspired openly by the GNU GPL, the organization does not deal with software, but designs licenses for other kinds of creative works, such as Web sites, scholarship, photography, films, literature, and music [28].

CC does not provide legal advice, but a Web application that allows a copyright holder to choose between several license options [29]. Not all CC license provisions are copyleft, indeed, their menu can be adjusted according to several degrees of restrictiveness. The option that gives a copyleft flavour to a CC license is the "share alike" type. It allows others to distribute derivative works only under a license identical to the original work. [Figure 2](#) summarizes the basic features that can be combined, matched and mixed together to obtain a customized *ad hoc* provision.

<i>name of option</i>	<i>rights given to any taker</i>	<i>rights retained</i>	<i>icon</i>
Attribution	copying, distribution, displaying, performing of original and derivative works	credit	
Noncommercial	copying, distribution, displaying, performing of original and derivative works	commercial exploitation	
No Derivative Works*	verbatim copying, distribution, displaying, performing of the original work	integrity of original	
Share Alike*	distribution of derivative works	invariability of license provision over derivative works	

*The combination of options Share Alike & No Derivative Works is not possible since the Share Alike condition applies only to derivative works.

Source: Creative Commons website: <http://creativecommons.org/learn/licenses/> (last accessed 12 April 2004)

Figure 2: Some rights reserved: Options that can be combined in a Creative Commons license.

When the choice has been made, the applicant obtains the following: a summary of the chosen license with the appropriate icon [30] — Commons Deed; a fine print that can stand up in court — Legal Code; and finally, a machine readable version able to help search engines to identify the work by its terms of use — Digital Code [31].

CC provides a blog for general discussion and a "discuss page" in which several groups, each coordinated by a Project Lead, engage in more in-depth research on an issue meant to deserve investigation. The purpose of each discussion group is to produce a workable proposal to be implemented by CC, when needed.

All of these innovative features characterizing the activity of CC allow designs of *ad hoc* licenses that may retain their copyleft strength against hijacking and, at the same time, may soften the side effects of a GPL-style restrictive provision.

Since a willingness to be flexible relative to copyright must take into account legal precepts and international trends towards IPR extension, the case of CC represents a great learning exercise for new knowledge creators as well as policy-makers.

It is worthwhile stressing here that probably no available CC license nor copyleft directly suit the case of DMOZ. However, it is important to note that provisions can be designed — if a flexible and customisation-driven approach is taken.



Conclusion

The possibility of hijacking public domain digital information and including it into proprietary assemblages — which might happen even if the original resources were not otherwise copyrightable — is just another serious flaw characterizing the concession of exclusive property rights to new knowledge creators. The possible side effects of hijacking are particularly sinister in the case of collectively produced open content works such as the Open Directory Project which is, in fact, a pure public good.

The main goal of DMOZ — building a comprehensive catalogue of the Web — seems at risk. Systematic hijacking of DMOZ data by commercial search engines and directories, as well as their embodiment in digital arrangements that tend to mix paid ranks and relevance-driven information results, constitute an example of privatisation and depletion of public domain knowledge.

Far from arguing that the institution of copyright is obsolete, this essay suggests that IPR can be customised and adapted to circumstances in which hijacking is likely to reveal itself as disruptive.

The principle of copyleft constitutes a powerful tool available for digital content creators and policy-makers, implying that information arrangements built upon freely accessible resources should be distributed under licensing terms similar to those covering the original resources. In cases where copyleft appears so restrictive that participation in a collective project may be discouraged, further customisation is always a feasible strategy. The case of Creative Commons illustrates this point and represents a critical learning exercise. 

About the author

Andrea Ciffolilli is a PhD student at the Department of Economics, Università Politecnica delle Marche, Ancona, Italy. He is interested in the social and economic implications of IPR extension and in the best ways to use ICT in order to improve daily and business life. Besides, together with his dad, he enjoys producing Montepulciano d'Abruzzo, a medium body, rich color Italian wine.

E-mail: andreciffo@libero.it

Acknowledgements

I wish to thank Judita Jankovic and Jenna Bailey for proofreading, Laura Vici and Patrick Wagstrom for useful suggestions. Many thanks are due to Prof. Ed Steinmueller for comments on the essay from which this article has been substantially extracted. The author is responsible for any error or omission which survived in the text. I am grateful to the Department of Economics, Ancona, for financial support.

Notes

1. For instance, within the European regulative framework, the European Commission's Directive on the legal protection of databases (issued 11 March 1996) prohibits copying of a substantial part of a digital content arrangement, regardless of the fact that the originality of arrangement is also copied. This provision, in fact, removes the distinction between protection of ideas and their expressions. Directive available at <http://europa.eu.int/ISPO/infosoc/legreg/docs/969ec.html>, accessed 9 April 2004.
2. The definition of hijacking employed hereby does not necessarily imply stealing software code or obscuring a domain, as the more conventional wisdom would suggest.
3. This article does not focus on international extension of IPR regulation and advances in technologies of protection. For an insightful analysis of these issues and their

consequences, see, for example, Drahos and Braithwaite (2002), Aigrain (2003) and David (2003).

4. The characteristics of these licenses are discussed in a [later section](#).

5. For a biography of R. Stallman and a history of copyleft see, for example, Wikipedia, the free encyclopaedia at http://en.wikipedia.org/wiki/Richard_Stallman, text available under the terms of the GNU Free Documentation License, http://en.wikipedia.org/wiki/Wikipedia:Text_of_the_GNU_Free_Documentation_License, accessed 9 April 2004.

6. Reese and Stenberg (2001), two software developers that first used copyleft and then turned away from it because of its limitations, revealed that they have never been interested in making money off their hijacked projects, and therefore, did not suffer from free-riding. On the contrary, learning has been their main goal and achievement. Moreover, they emphasize that, when they changed the license of their software to be compatible with the GNU GPL, they did not receive any contribution from the open source individuals who had requested this change. Instead, contributions were received from commercial corporations, since the license adopted by the two developers allowed making changes at one's discretion. Clearly, open source contributors may be moved by a variety of motivations from learning to signalling, to inherent communitarianism. Therefore, the evidence given by Reese and Stenberg is only part of the story.

7. A directory is a Web catalogue resembling the table of contents of a book. It is characterized by a specific structure and edited by individuals. A search engine does not have a contents outline and is not usually human edited but instead uses an automatic programme to crawl cyberspace in search of keywords or keyword lists as defined by Webmasters.

8. See <http://www.dmoz.org/about.html>, accessed 9 April 2004.

9. The Open Directory License is therefore an attribution license, not a copyleft provision. See <http://dmoz.org/license.html>, accessed 10 April 2004.

10. For example, hardware failures over the 2002 winter holidays caused the directory to be out of commission for several months.

11. The problem is directly addressed by DMOZ volunteer editor Elisabeth Osmeloski, in Olsen and Hu (2003).

12. For a list of the sites using DMOZ data, see http://dmoz.org/Computers/Internet/Searching/Directories/Open_Directory_Project/Sites_Using_ODP_Data/, accessed 13 April 2004.

13. See the Overture Web site at <http://www.content.overture.com/d/USm/ps/wspi.html>, accessed 12 April 2004.

14. Some of these tricks are described in Vaughan (2003). For instance, the possibility of embedding the name of an opponent political coalition either in the html code, using the background colour, or in the keyword list of a certain political Web site, resembles a particular case of diverting searchers towards a specific destination.

15. See <http://www.google.com/search?hl=en&ie=ISO-8859-1&q=apache>, accessed 13 April 2004.

16.

<http://www.content.overture.com/d/advertisers/p/bjump/?o=UCJ&b=10&AID=9442328&PID=1466800>, accessed 14 April 2004.

17. Yahoo! also developed a new algorithm for ranking Web sites, Web Rank, that equips its toolbar. Web Rank is covered by a patent allowing the use of a weighted average which combines both content and sales revenue to rank Web sites in results pages. Sponsors might as well bid in order to alter the weights of search results.

18. Paid ranks or sponsored links are listed on any page of Google's search results, on the right side of the screen.

19. See, for example, <http://www.gnu.org/copyleft/copyleft.html>, accessed 12 April 2004.

20. If, of course, copyleft proves to be able to survive the test of trial in court. A recent German court ruling stated that the main clauses of the GNU General Public License are valid under German copyright and contract law. However, it should be considered that validity of copyleft may depend on a particular legal system. See <http://yro.slashdot.org/article.pl?sid=04/07/23/1558219&tid=117&tid=123>, accessed 2 September 2004.

21. Restrictiveness is intended to prevent licensees of a copylefted software, for example, from distributing a proprietary modified version of the product without releasing the source code.

22. A high degree of restrictiveness can be intended as the condition, characterising the GNU General Public License, which prevents open source licensees from mingling copylefted source code with non-copylefted code.

23. See <http://creativecommons.org/learn/aboutus/>, accessed 13 April 2004.

24. Copyright Law of the United States of America and related laws contained in title 17 of the United States Code, available at <http://www.copyright.gov/title17/circ92.pdf>, accessed 9 April 2004.

25. See <http://www.illegal-art.org/index2.html>, accessed 13 April 2004.

26. <http://creativecommons.org/learn/legal/>, accessed 13 April 2004.

27. <http://creativecommons.org/learn/legal/>, accessed 14 April 2004.
28. <http://creativecommons.org/learn/aboutus/>, accessed 13 April 2004.
29. Currently CC provides a selection of eleven licenses. In addition, it provides the Commons Deed and metadata that can be added to GNU GPL, GNU LGPL, public domain dedications, sampling licenses and founder's copyright (a license granting exclusive rights for a shorter period than usual copyright). See <http://creativecommons.org/learn/licenses/> and <http://creativecommons.org/license/>, accessed 13 April 2004.
30. An online available work, for example, should include a button or icon linking back to the Commons Deed that acts as a notification for the public.
31. Metadata included in this Digital Code improves access to some sites on the Internet. See <http://creativecommons.org/learn/licenses/>, accessed 12 April 2004.

References

- P. Aigrain, 2003. "Positive intellectual rights and information exchanges," European Commission, Information Society Technologies R&D Programme, at <http://opensource.mit.edu/papers/aigrain.pdf>.
- N. Bezroukov, 2002, "BSD vs. GPL, Part 2: The dynamic properties of BSD and GPL licenses in the context of the program life cycle," at http://www.softpanorama.org/Copyright/License_classification/social_dynamics_of_BSD_and_GPL.shtml.
- A. Bonaccorsi and C. Rossi, 2004. "Altruistic individuals, selfish firms? The structure of motivation in Open Source software," *First Monday*, volume 9, number 1 (January), at http://www.firstmonday.org/issues/issue9_1/bonaccorsi/.
- A. Ciffolilli, 2003a. "Phantom authority, self-selective recruitment and retention of members in virtual communities: The case of Wikipedia," *First Monday*, volume 8, number 12 (December), at http://www.firstmonday.org/issues/issue8_12/ciffolilli/.
- A. Ciffolilli, 2003b. "Will a serf defeat an empire? An economic analysis of copyleft as a tool to block hijacking of Open Source projects," MSc dissertation, SPRU — Science and Technology Policy Research, University of Sussex, Brighton, UK. (September).
- R. Cowan, P.A. David, and D. Foray, 2000. "The explicit economics of codification and tacit knowledge," *Industrial and Corporate Change*, volume 9, number 2 (Summer), pp. 211–253.

P.A. David, 2003. "Koyaanisqatsi in cyberspace," Stanford Institute for Economic Policy Research (SIEPR), Discussion Paper, number 02–29 (March), at <http://siepr.stanford.edu/papers/pdf/02-29.pdf>.

P.A. David, 2000. "The digital technology boomerang: New intellectual property rights threaten global "Open Science"," World Bank Conference Volume ABCDE–2000, at <http://www-econ.stanford.edu/faculty/workp/swp00016.pdf>.

P. Drahos and J. Braithwaite, 2002. *Information feudalism: Who owns the knowledge economy?* London: Earthscan.

R.A. Ghosh, 1998. "Cooking pot markets: An economic model for the trade in free goods and services on the Internet," *First Monday*, volume 3, number 3 (March), at http://www.firstmonday.org/issues/issue3_3/ghosh/.

T. Hobbes, 1641. *De Cive: The English version entitled, in the first edition, Philosophicall rudiments concerning government and society*. Critical edition by H. Warrender. Oxford: Clarendon Press.

S. Johnson, 2003. "Digging for Googleholes: Google might be our new God but it's not omnipotent," at <http://www.slate.msn.com/id/2085668/>.

J. Lerner and J. Tirole, 2003. "The scope of Open Source licensing," National Bureau of Economic Research, Working Paper, number w9363 (December), at <http://www.nber.org/papers/w9363>.

J. Lerner and J. Tirole, 2000. "The simple economics of Open Source," National Bureau of Economic Research, Working Paper, number 7600 (March), at <http://www.hbs.edu/research/facpubs/workingpapers/papers2/9900/00-059.pdf>.

R. Mansell and W.E. Steinmueller, 2000. *Mobilizing the information society: Strategies for growth and opportunity*. Oxford: Oxford University Press.

J. Mateos Garcia and W.E. Steinmueller, 2003a. "The Open Source way of working: A new paradigm for the division of labour in software development?" Brighton: Science and Technology Policy Research, Open Source Movement Research, INK, Working Paper, number 1 (June), at <http://www.sussex.ac.uk/spru/publications/imprint/sewps/sewp92/sewp92.pdf>.

J. Mateos Garcia and W.E. Steinmueller, 2003b. "Applying the Open Source development model to knowledge work," Brighton: Science and Technology Policy Research, Open Source Movement Research, INK, Working Paper, number 2 (January), at <http://www.sussex.ac.uk/spru/publications/imprint/sewps/sewp94/sewp94.html>.

S. Olsen and J. Hu, 2003. "The changing face of search engines," at <http://www.news.com.com/2100-1032-993677.html>.

Plautus, Titus Maccius, 1994. *Asinaria*. Translated by M. Scandola. Milan: Rizzoli.

M. Polanyi, 1966. *The tacit dimension*. London: Routledge & Kegan Paul.

E.S. Raymond, 1999. "The magic cauldron," In: E.S. Raymond. *The cathedral and the bazaar: Musings on Linux and Open Source by an accidental revolutionary*. Sebastopol, Calif.: O'Reilly and Associates, at <http://www.catb.org/~esr/writings/magic-cauldron/magic-cauldron.html>.

B. Reese and D. Stenberg, 2001. "Working Without Copyleft," at <http://www.oreillynet.com/lpt/a/1403>.

C. Shapiro and H. Varian, 1999. *Information rules: A strategic guide to the network economy*. Boston: Harvard Business School Press.

D. Sullivan, 2003. "Where are they now? Search engines we've known and loved," *SearchEngineWatch.com* (4 March), at <http://searchenginewatch.com/sereport/article.php/2175241>.

F. Vaughan, 2003. "The computer curmudgeon: Give a little, get a little ...," *Computer Bits Online*, volume 13, number 4 (April), at <http://www.computerbits.com/archive/2003/0600/vaughan0306.html>.

E. von Hippel and G. von Krogh, 2003. "Open Source software and the "private–collective" innovation model: Issues for organization science," *Organization Science*, volume 14, number 2 (March–April), pp. 209–223, and at <http://opensource.mit.edu/papers/hippelkrogh.pdf>.

S. Weber, 2000. "The political economy of Open Source software," Berkeley Roundtable on the International Economy (BRIE), Working Paper, number 140 (June), at <http://repositories.cdlib.org/cgi/viewcontent.cgi?article=1011&context=brie>.

Editorial history

Paper received 27 July 2004; revised 2 September 2004; accepted 3 September 2004.

Contents Index

Copyright ©2004, *First Monday*

Copyright ©2004, Andrea Ciffolilli

The economics of open source hijacking and the declining quality of digital information resources: A case for copyleft by Andrea Ciffolilli
First Monday, volume 9, number 9 (September 2004),
URL: http://firstmonday.org/issues/issue9_9/ciffolilli/index.html