Open Access to Legal Scholarship and Copyright Rules: A Law and Technology Perspective

Version 1.0 – April 2009

Roberto Caso
University of Trento-Faculty of Law
Open Access to Legal Scholarship and Copyright Rules:
A Law and Technology Perspective

Version 1.0 – April 2009*

Roberto Caso
University of Trento-Faculty of Law

Abstract. By applying copyright law, contracts, customs and technological standards it is possible to achieve two different kinds of control over digital information. In the first form, control is based on the closeness of information and it is rigid and centralized: see, e.g., the Digital Rights Management systems (DRMs). In the second form, control is based on the openness of information and it is flexible and decentralized: see, e.g., the GNU General Public License (GPL) and the Creative Commons Licenses (CCLs). Those two models of control correspond to two opposite trends in scientific community.

On one side, the risk is that a rigid and centralized control (such as the one based on DRMs), shaped by market considerations, invades the sector proper of the scientific community (which, on the contrary, is traditionally inspired by the logic of a flexible and decentralized control, based on customs and informal norms). This would strongly undermine the possibilities of access to scientific knowledge expressed in a digital format. This risk is prominent in the field of legal scholarship, where a vast amount of legal information (also covering the information that is, in theory, in public domain) is governed by rigid and centralized control.

On the other side, to counteract such a risk, part of the scientific community is promoting the logic of Open Access (mostly based on free licenses such as the GNU GPL or the CCLs) to scientific knowledge. The Open Access (OA) movement is quickly growing in importance for legal scholarship. Nonetheless, the institutional arrangements and the technological features of OA to legal scholarship are variegated and pose a vast array of problems.

Keywords: Open Access, Legal Information, Legal Scholarship, Copyright Law, Digital Rights Management, Creative Commons Licenses.

1. Introduction

We are confronting with a risk. A rigid and centralized control over digital information (such as that based on the Digital Rights Management systems), shaped by market considerations, may soon dominate the field of the scientific community at large (which, on the contrary, has been traditionally inspired by the logic of a flexible and decentralized control, based on customs and informal norms). This would strongly undermine the possibilities of access to scientific knowledge expressed in digital format. This risk is particularly acute in the field of legal scholarship, where a vast amount of legal information (also covering information that is, in theory, in public domain) is governed through rigid and centralized control. To counteract such a risk, part of the scientific community is promoting the logic of Open Access (mostly based on free licenses such as the GNU General Public License or the Creative Commons Licenses) to scientific knowledge.

Despite the initial delay, the Open Access (OA) movement is quickly growing in importance for legal scholarship. Nonetheless, the institutional arrangements and the technological features of OA to legal scholarship are variegated and pose a vast array of problems.

To understand the interaction between the law and technology in OA to legal scholarship Part 2 of this paper outlines the relationship between intellectual property and norms of science, Part 3 illustrates the two different kinds of control over digital information achieved by applying copyright law, contracts, customs and technological standards, Part 4 describes closed and OA models to scientific knowledge, Part 5 discusses the promises and perils of OA to legal scholarship. Eventually, Part 6 sketches some conclusions.

2. Intellectual property vs. [and] norms of science

Intellectual property has been shaped mainly by economic interests. In particular the ancestors of patents and copyrights emerged as privileges granted by the king to the representatives of merchant class, such as weavers and printers, to exercise their activity exclusively. The mechanism of privilege has then evolved to the exclusive right recognized by the law (David, 1993).

From an economic perspective, the exclusive right is a mechanism which is needed to balance the incentive to produce creative information with the possibility of accessing the same information. As distinct from material
goods, information is a non-rival and non-excludable good (Arrow, 1962). Because of non-excludability an information market cannot emerge. Intellectual property grants to the holder an artificial exclusive right, formally warranted by state law, mimicking the mechanism of ownership of material goods, thus laying the foundations of a market. The right holder can exclusively use the invention or the work, enjoying an advantage over competitors. In other words, the exclusive right is a sort of legal monopoly. Those enjoying the exclusive right can, indeed, charge a monopolistic price, i.e. a price higher than the marginal cost. This constitutes a benefit with regard to the incentive to produce information, but also a cost for society. The most important bundle of costs depends on the fact that anyone who is not willing to pay the monopolistic price will be cut out from the use of information protected by the exclusivity. Among these are also those who wish to re-elaborate the information in order to produce new inventions and works. It is, therefore, necessary that the social costs do not outweigh the social benefits. The limits of intellectual property rights are aimed at such goal and can consist in time limits (for example the patent for invention generally lasts twenty years; the copyright lasts usually for the lifetime of the author plus seventy years) and content limits (for example the patent for inventions concerns only new ideas; the copyright regards only the expressive form of an original work) (Menell and Scotchmer, 2005).

Instead, the institutional features of the production of scientific knowledge have been mostly shaped by the practices and customs of the community of scientists, the so-called Republic of Science (Polanyi, 1962). An authoritative trend of the sociology of science has singled out the main informal norms governing the production of scientific knowledge: “universalism”, “communality”, “disinterestedness” and “organized skepticism” (Merton, 1973; Eisenberg, 1987; Rai, 1999; Burk, 2006).

Universalism means that the truthfulness of the results of the research is not bound to the scientist’s (national or institutional) identity.

Communality implies that the knowledge is the product of collaboration among colleagues and, therefore, it must be shared within scientific community. All the actual knowledge is built upon past knowledge and it is the basis for that of the future (as Isaac Newton said “if I have seen further it is by standing on the shoulders of Giants”).

Disinterestedness requires that scientists aspire to research the truth, not their personal interests.

Organized skepticism expects that scientists’ theories will be submitted to the critical evaluation of the community before being accepted.

These four norms are strengthened from the acknowledgement in terms of prestige (and of career progress) by the community. The scientific community, indeed, prizes those who make original contributions to
knowledge. The emphasis on originality generates the incentive to publish the works as soon as possible, trying to avoid being anticipated by others. But, after having published the work, the scientist does not have any more exclusivity over the knowledge she has produced.

Even though intellectual property and the norms of science have different paths, they also intersect. In the complex historic process that has led from the “secret science” to the “open science” (Rossi, 2007; David, 2004), the invention of scientific journals by the Royal Society of London in early 1665 – when Henry Oldenburg created the Philosophical Transactions (or Phil Trans) - was probably the first of these intersections. At its beginning, the scientific journal was, overall, the “public record of original contributions to knowledge [...]”. “[T]he Republic of Science claimed the right to grant intellectual property to scientific ‘authors’ and Phil Trans was its instrument of choice” (Guedon, 2001). Hence, it is paternity (i.e. the author’s right) at the center of this scenario, not the commercial aspect (the editor’s right) of copyright. In fact, for centuries scientific journals and scientific articles were not an editor’s business.

The landscape changed after the World War II, when supply and demand of scientific publications quickly rose. Scientific journals become a flourishing business. The rise of the concept of “core journals” has shaped the peculiar features of the market for journal publications and has led to so-called “serial pricing crisis” (Guedon, 2001). As a recent European study pointed out: the market of scientific journals is “an intermediated market, where libraries are the key buyers, which leads to lower reader price sensitivity. Moreover, it is a market where the best authors want to publish in highly-read journals and readers want to read journals which publish the best authors. This leads to ‘virtuous circles’ for journals, and to associated ‘natural barriers to entry’” (AA. VV., 2006). Moreover, “much of scientific activity is publicly funded: the output of research is typically not bought by journals but ‘donated’ by publicly-funded researchers” (AA. VV., 2006).

3. Digital technologies and access to information: closed models vs. open models

What role do digital technologies have in the access to scientific knowledge?

At a superficial glance, digital technologies multiply and speed up the possibility of access to scientific knowledge. Therefore, it is possible to think that their use within the scientific community has reinforced the trend of sharing the results of the research with the public. Nonetheless, the picture is much more complex.
The information and telecommunication technologies present revolutionary features (Pascuzzi, 2006). In the context of this discourse, we can single out two of them.

1) On one hand, it is possible to foreclose information totally (for example, by making the open source code of software secret or by encrypting a text file), making it understandable only to the machines or, better, making it accessible and usable (by a human user) through pre-determined modalities, machines or software. For example it is possible to program software for reading an e-book in a way which is compatible only with specific hardware.

2) On the other hand, it is possible to transmit information in one language understood by the computer (binary code) and in an open format (the so-called open source code), a format modifiable by those who know the programming language.

Following the scenario depicted so far, two models of production of digital information can be sketched (Caso, 2008).

a) The first model is based on the closeness of information and, therefore, on rigid and centralized control. Such a model creates hierarchical forms of production and distribution in which the information holders can choose where, how, when and who will be able to use information: this is the case for Digital Rights Management systems (DRMs), based on technological protection measures (TPMs).

b) The second model is based on the openness of information and, therefore, on a flexible and decentralized control. This model generates non-hierarchical forms of production and distribution (called “Peer to Peer”) where the actors, inspired sometimes by different incentives than payment in exchange for services, perform the functions typical of producers and consumers in a hybrid way: this is the case, for example, of the development of an open source software such as Linux or of the drawing of texts such as an on-line encyclopedia, where all the users can publish or amend the entries (Wikipedia).

Applying intellectual property law (patent, copyright, industrial secrets), contracts, customs and technological standards it is possible to obtain different forms of control over digital information.

Today, two forms of control are emerging, corresponding to the two models of production which have been previously indicated.

A) In the first form, the control is based on the closeness of information and it is rigid and centralized. This kind of control emerges from the market of the so-called “proprietary software” and finds its foundations on an (initially rudimentary) TPM: keeping the source code secret. The acknowledgment of the copyright protection and the diffusion of End User License Agreements (EULAs), have reinforced this control on a contractual basis.
This kind of control gives rise to DRM based on (cryptographic) TPMs (Bechtold, 2004; Caso, 2004). The control is extended from the source code of the software to any information which can be expressed in binary code (not only software, but also text, audio, video, etc.).

The goal of DRM is, indeed, that the conditions – written in the license - for access and use of information must be implemented by software and machines designed (on the basis of the standards proper of the DRM system) in order to enjoy the same information.

With regard to the control over information, the main components of the DRM systems are:
- the TPMs based mainly on the digital cryptography, but also on other technologies such as digital watermarking and fingerprinting;
- the metadata that describe restrictions in a language which is understandable by the computer:
- the content;
- the holder of the content;
- the user;
- the rules for enjoying the content (for example if it can be copied, printed, distributed, etc., where it can be enjoyed, through which machines it can be enjoyed), expressed in languages which are called Rights Expression Languages (RELs), such as the eXtensible rights Markup Language (XrML), which is one of the standard languages.

B) The second form of control is based on the openness of information and is flexible and decentralized. The first model of such control is represented by the GNU General Public License (GPL). These are general contractual conditions which, using copyright, are aimed at guaranteeing the right to copy (so-called “Copyleft”), modify and distribute software with open source code to anyone willing to accept such contractual conditions.

The mechanism of protection relies on the clause which declares that the software subject to the license is protected by copyright and that, at the same time, imposes on the users of the GPL to apply the same GPL, if the same software, or other derived software is distributed, to subsequent licensees. The GNU GPL model has inspired several different types of licenses. Among these, the Creative Commons Licenses (CCLs) are one of the most successful. Such licenses translate the GNU GPL model, successfully implemented in the field of software, to the broader field of all digital content and to all inventive works embodied in traditional media, such as books (Lessig, 2004).

4. Scientific knowledge: closed access models vs. open access models

The scenario described so far gives rise to the risk that rigid and centralized control (such as that based on DRM systems), shaped on market
considerations, invades the proper domain of the scientific community (which is, on the contrary, motivated by the logic of flexible and decentralized control, based on customs and informal norms), decreasing the possibility of access to scientific knowledge expressed in a digital format. Such a risk depends on many factors.

- Digitalization, along with other causes – such as the shortening of the distance between base and applied research – means that the scientific community can perceive knowledge as economic goods tradable on the market through intellectual property rights and TPMs (Nelson, 2003; David, 2003, Eisenberg, 1987). The phenomenon also concerns institutions financed by public funds, such as universities (Monotti and Ricketson, 2003).

- The contemporary scientific press is controlled by few private big editors, who apply market rules and intellectual property rights to the circulation of information concerning scientific knowledge. Despite the fact that digital technologies allow huge cuts in the costs of production and distribution of information, the price fixed by private editors to get access to digitized scientific information seems doomed to increase (Guedon, 2001).

- Western legal systems tend to strengthen and multiply intellectual property rights on digital (information) goods (David, 2003). New goods are subject to intellectual property (for example, software and databases); single goods can be subjected to many intellectual property rights (for example, in some cases software can be subject both to copyright and patent; databases in the European Union can be subject both to copyright and to a sui generis right); a plurality of subjects can claim intellectual property rights (for example, not only private individuals or enterprises, but also research institutes, universities, etc.). Rigid and centralized control, such as that implemented on DRMs, is part, therefore, of a picture in which intellectual property rights appear to be strengthened. Despite the enormous power of control and the many effects that this rigid and centralized control has on different legal aspects, western parliaments have created a discipline of legitimization and protection (only) for some of the components of the DRM systems, often by making room for DRM provisions in general copyright legislation (see the U.S. Digital Millennium Copyright Act (DMCA) of 1998 and European Union Copyright Directive (EUCD) of 2001). Lobbying by traditional (the entertainment industry) and emerging (the DRM industry) interests has prevailed on the public interest to regulate the control over digital information.

- Intellectual property legislation on digital goods vary according to the legal system considered. For example, the US legal framework is different from that of the EU, but also within EU one can detect important differences between countries. This is a further hurdle to access to and circulation of scientific knowledge expressed in digital form.
To counteract this risk, part of the scientific community is promoting the logic of OA to scientific knowledge (Willinsky, 2006; Suber, 2004-2006; Guedon, 2001).

Many scientific communities publish their results on websites freely accessible to anyone through the Internet. This can be the case for the publication of drafts, articles already published in “for payment” journals or, again, it may be the sole manner of publishing the research (Guedon, 2004).

The logic of OA is also promoted in solemn declarations, such as the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities in 2003.

Nonetheless, the institutional arrangements and the technological features of OA are variegated and pose an array of problems (Burk, 2006).

For example, sharing the source code of software contemplated by the GNU GPL is based on contract and informal norms of the community of software programmers. On one hand, respect for the GPL is warranted by the threat to apply the law of contracts. This seems to be a different aspect than that of the traditional informal norms of the scientific community. On the other hand, respect is warranted by the reputation created by the norms of the software programmers. But this does not mean that such norms can be adapted to other communities (such as the communities of the biologists or lawyers).

To make a further example, the CCLs project is developing a program specifically dedicated to scientific knowledge called “Science Commons”. However, the transplant of the CCLs logic in the scientific context gives rise to many questions.

CCLs are very recent tools, while the institutional arrangements of the scientific community are quite old. The CCLs are contracts which can be used for all the typologies of authors. The scientific community is formed by many scientific communities. Each of them has, along with the general informal norms already mentioned in Part 2 of this paper, specific informal norms. Moreover, the CCLs raises, like all standard contracts, the problem of the licensee’s protection. In particular, an extremely interesting field of investigation is that aimed at improving the cognitive and informational position of the licensee, making it possible for him to choose preferred contractual arrangements with increased awareness (Hillman and Rachlinski, 2001). Such a perspective requires a major investigation also on the programming side, particularly with regard at the aim of improving the technologies for the digital management of the contracts (Mulligan and Burstein 2003).

5. Open Access to legal scholarship: promises and perils
The risk of rigid and centralized control based on market considerations invading the proper domain of the scientific community is acute if one considers the field of legal scholarship. A vast amount of legal information - also covering information that is, in theory, in the public domain - is presently accessible only through closed and proprietary databases such as Lexis and Westlaw. The increasing concentration of the market of legal databases has led to escalating prices for legal information (Arewa, 2006). The contractual and market power of databases holders is strengthened by new copyright laws (such as the EU directive 96/9 on the database protection and the DMCA as well as the EUCD on TPMs).

But, despite the initial delay, the OA movement is quickly growing in legal scholarship (Carroll, 2006; Hunter, 2006; Solum, 2007). The OA model to legal scholarship has the potential to subvert the present dominant publishing model. Perhaps this is true as well as in other scientific fields. This impression depends on the peculiar features of the legal field.

- The peer review system (traditionally managed by commercial editors) is much less important in law than in the hard sciences.
- Some types of legal publications are for non academic lawyers (judges, attorneys, etc.). Besides, legal scholarship is more and more interdisciplinary and globalized (Carroll, 2006). Hence, the public interested in legal publications is very large and heterogeneous.
- The raw legal data (i.e., the primary sources like acts and judicial opinions) are – thanks to clear copyright rules - in many western legal systems in the public domain.

The emerging OA model the legal scholarship has the following institutional structure.

- The major functions of publications (selecting the best works, making the works accessible, publicizing the works, and archiving the works) is based on the old participants (commercial editors, university press, law reviews student-edited, etc.) and new intermediaries (legal scholarship repositories like Social Science Research Network’s Legal Scholarship Network and Berkeley Electronic Press Legal Repository, Wikipedia, Internet search engines like Google Books and Google Scholar, social software, etc.) (Solum, 2007).
- Production costs are – as in the past – borne by the authors and their institutions (universities and law faculties). Dissemination costs – lower than in the past - are shared among the authors, their institutions and the old and new intermediaries. The incentive system is based on the “reader’s attention”. E.g., in the publishing of a post-print on the OA repository such as SSRN, the authors and
law reviews increase the probability of reading and citation, while the repository multiplies the circulation of its own brand, and the Internet search engines increases the number of users and visits. Moreover, OA dramatically reduces the delay in publication (making the information more useful to potential readers, if it is true that time is an important variable in the contemporary legal arena, especially for attorneys and for judges) and the costs of reading/accessing that material.

- The author retains copyright (in particular, the right of attribution) over the publication and grants - through open licenses such as CCLs (see the Open Access Law Program of the Science Commons Project) (Carroll, 2006) – to public and intermediaries a limited set of rights.

This is a revolutionary approach. OA to legal scholarship changes the form of the legal publication - as a scholar has pointed out, we are facing new kind of publications such as the “idea-paper”, the blog post, the Wikipedia article (Solum, 2007) - and shifts the “quality selection” function from the traditional intermediaries to the new intermediaries (e.g. search engines and social software) and the readers.

Nonetheless, the revolution has just started, and we are in the middle of it. The future of OA to legal scholarship is not clear. Some scholars have argued that the OA movement complements the old publishing model and is not a substitute. In particular, an author has pointed out that the success of the OA depends on the power to supply the so called “economy of prestige” that has been managed for centuries by legal scholarship through the old publishing model (Madison, 2006).

If it is true that the managing the economy of prestige is still (and will be) a cornerstone of (legal) scholarship, nevertheless the revolutionary power of digital technologies is modifying the features not only of the publishing model but also of the academic lawyer itself. As argued above, thanks to information and communication technologies, legal scholarship is more and more interdisciplinary and globalized.

Rather, the actual problems of the OA to legal scholarship seem the same as those of the larger OA movement. These problems can be synthesized in the following list.

- Until now OA has been a bottom up and decentralized movement, based on different policies, solemn declarations and contractual arrangements such as CCLs. In fact, there is no unified definition of OA. Yet, the formal law shows a growing attention to OA (see, in the USA, the National Institute of Health Public Access Policy implements the Division G, Title II, Section 218 of PL 110-161, Consolidated Appropriations Act, 2008). Hence, the future of OA depends on the intersection between formal law, OA policies and
social norms of science. E.g., the CCLs are formal contracts (even if contracts with special features) and customs of (legal) science are informal norms. The future will tell us if the application of CCLs in the courts will be compatible with social norms of (legal) science.
- OA is a powerful instrument making work accessible. The success of the OA in the other three functions of publishing (selecting the best works, publicizing the work, and archiving the work) will depend not only on institutional arrangements but also on the development of a trustworthy technological system based on standardization of metadata (Carroll, 2006; Madison, 2006), search engines with sophisticated Boolean operators, and digital formats which guarantee long term preservation of the works. From this perspective the openness of software and formats will play a fundamental role.
- OA is not Nirvana. The dislocation of publishing functions to new intermediaries raises a number of new risks. E.g., Google’s algorithm is not neutral. “Google ranks the relevance of any given Web site by determining the number of other sites that are linked to it” (Hunter, 2006). This is a value choice. Moreover, “the precise method for producing the rank order is a trade secret” (Solum, 2007). Last but not least, many have raised concerns about the Google’s power to process an infinitive quantity of personal data (Guarda, 2009). In other words, dislocating publishing functions means also dislocating power, with new problems related to that power.

6. Conclusions

OA is an extremely powerful tool for the dissemination of knowledge made possible by digital technologies. It has some real advantages as regards the traditional publishing model. Moreover, through OA it is possible to counteract the risk that the rigid and centralized control (such as that based on the DRM systems), shaped on market considerations, will invade the proper domain of the scientific community. But the definitive success of OA (in legal scholarship as well as in other scientific fields) depends on our capacity to understand and manage the complex intersection among intellectual property law, contracts, norms of science and technological standards.
One may say that the OA is the revenge of “author’s right” on the “editor’s right”. Nevertheless we must be conscious that - in the digital age - copyright law is only one (and not the most important) among many other instruments which may govern the production and distribution of information.
The dislocation of the publishing functions to new intermediaries raises a number of new risks. To counteract these new risks we have to take Internet governance seriously and to see beyond copyright law.

Acknowledgements

I wish to thank Pippo Bellantuono, Matteo Ferrari, Paolo Guarda, Umberto Izzo, Federica Lorenzato and Tina Piper for helpful comments on earlier drafts of this paper.

References


David P. A. (2003), Can ‘Open Science’ be Protected from the Evolving Regime of IPR Protections?, available at:


