Fitzgerald, Brian F. and Pappalardo, Kylie M. (2007) The Law as Cyberinfrastructure. *CT Watch Quarterly* 3(3):pp. 51-57.

CTWatch QUARTERLY
August 2007

The Law as Cyberinfrastructure

Brian Fitzgerald,
Queensland University of Technology, Australia
Kylie Pappalardo,
Queensland University of Technology, Australia

In almost everything we do, the law is present. However, we know that strict adherence to the law is not always observed for a variety of pragmatic reasons. Nevertheless, we also understand that we ignore the law at our own risk and sometimes we will suffer a consequence.

In the realm of collaborative endeavour through networked cyberinfrastructure we know the law is not too far away. But we also know that a paranoid obsession with it will cause inefficiency and stifle the true spirit of research. The key for the lawyers is to understand and implement a legal framework that can work with the power of the technology to disseminate knowledge in such a way that it does not seem a barrier. This is difficult in any universal sense but not totally impossible. In this article, we will show how the law is responding as a positive agent to facilitate the sharing of knowledge in the cyberinfrastructure world.

One general approach is to develop legal tools that can provide a generic permission or clearance of legal rights (e.g., copyright or patent) in advance (usually subject to conditions) that can be implemented before or at the point of use. This has become known as open licensing and will be discussed below in terms of copyright and patented subject matter. [1]

However, open licensing will not be adopted by everyone nor in every situation is it suitable. A generalisation is that it will be advocated in the context of publicly funded research producing tools and knowledge upon which platform technologies are built where considerations such as privacy are not an issue.

Where open licensing is not being used, the many parties to a collaborative endeavour will normally be required to map the scope and risk of their mutual endeavour through a contract. Contracts can take time to negotiate and, in many instances, promise to frustrate the fast paced and serendipitous nature of research fuelled by high powered cyberinfrastructure. To this end a number of projects throughout the world, for example The Lambert Project in the UK,[2] the University Industry Demonstration

Project (UIDP) in the USA,[3] and (amongst other projects) the 7th Framework Project in the EU,[4] have begun asking how we might be able to improve this situation. Suggestions include standard form or off the shelf contracts covering a variety of situations, a database of key clauses and, in the case of the UIDP project, a software based negotiation tool called the Turbo-Negotiator. Legal instruments that can match the dynamic of the technology and appear seamless and non-invasive are the goal. More work in this area is needed (and happening) and is critical to ensuring we have the law and technology of cyberinfrastrcuture working to complement each other.

In the remainder of this article we will focus on the open licensing model.

Open Licensing

1. Open Content Licensing

From a legal perspective, one of the most significant responses to the technological advances that have revolutionized the creation and distribution of copyright materials during the last decade has been the development of new systems for licensing (or authorising) others to obtain access to and make use of the protected material. These new forms of licences – usually referred to as "open content" – are founded upon an acknowledgement of the existence of copyright in materials embodying knowledge and information, but differ from licences commonly used before the advent of the digital era in key respects. As well as being relatively short, simple and easy to read, they are standardised, conceptually interoperable with other open content licences, machine (computer) enabled and have the advantage that, since they are automated and do not require negotiation, they eliminate (or at least minimise) transaction costs. Running with the copyright material to which they are attached (thereby avoiding the privity issue where rights are conferred contractually), open content licences identify materials that are available for reuse and grant permissive rights to users, thereby facilitating access and dissemination. [5]
The most widely used of the open content licences are the Creative Commons licences.[6] These licences attach to the copyright material and provide that anyone can reuse the material subject to giving attribution to the author of the material and subject to any of the optional conditions as selected by the licensor. The optional conditions are:

- non-commercial use;
- no derivative materials based on the licensed material are to be made; or
- share alike – others may distribute derivative materials based on the licensed material, but only under a licence identical to that covering the licensed material.

Creative Commons licences have more commonly been applied to publications than to research data. They have been particularly useful for academic authors depositing their publications in university or scholarly digital repositories or databases. Repositories help to make publications more accessible to the research and general communities. The advantage of a Creative Commons licence is that it tells people accessing the publication what they can and cannot do with the material, without the copyright owner having to deal with permissions on a case-by-case basis.

Below are two examples of scientific research publication projects that promote open access and reuse of material by utilising open content licensing models.

Example One – PLoS ONE

The Public Library of Science (PLoS) is a non-profit, open access, scientific publishing project that aims to create a library of peer-reviewed scientific and medical journals that are made available online without restrictions under open content licences.[7] PLoS ONE is a peer-reviewed, scientific literature journal that enables scientific research to be published and disseminated within weeks, avoiding delays associated with traditional means of publication.[8]

The features of PLoS ONE include:

- *rapid publication* – realising that the rapid publication and dissemination of research is one of the highest priorities, PLoS ONE ensures a streamlined electronic production workflow that ensures papers are published within weeks of submission;
- *freedom of use and ownership* – in accordance with the CC attribution licence, PLoS ONE enables users to read, copy, distribute and share papers freely without restrictions and formal permission, provided that the original author and source are cited; and
- *high impact* – PLoS ONE has been designed in light of the fact that papers published in OA journals are more likely to be read and cited given the lack of barriers to access.

Example Two – Nature Precedings

Nature Precedings is an online database designed to allow scientific researchers to share pre-publication research, unpublished manuscripts, presentations, white papers, technical papers, supplementary findings and other scientific documents.[9] Contributions are taken from biology, medicine (except clinical trials), chemistry and earth sciences. The database is free of charge to access and use, and is intended to provide a rapid means of disseminating emerging results and new theories, soliciting opinions and recording the provenance of ideas.

Nature Precedings aims to make scientific documents citable, globally available and stably archived. To this end, it can also be used as an archiving tool for scientists to store their work for their own future convenience.

Submissions made to Nature Precedings are screened by a professional curation team for relevance and quality, but are not subject to peer review. The database is designed to complement scientific journals by providing a more rapid and informal communication system, but submissions to Nature Precedings are not subject to the same rigorous and time-consuming reviews as submissions made to scientific journals.

The Nature Precedings website states that scientists should own copyright in a document and have permission from other copyright holders (e.g., co-authors), before

they submit the document to Nature Precedings.[10] Copyright then remains with the author. However, the website encourages scientists to release their work under a Creative Commons Attribution Licence so that content can be quoted, copied and disseminated, provided that the original source is correctly cited.[10]

Authors who own copyright in their publication will be able to place a Creative Commons licence on their work, but if they have assigned copyright to their publisher or another party, they will need to ask permission from that party before they can attach a Creative Commons licence. A problem that often arises in this situation is that authors are unsure of whether they own copyright or their publisher owns copyright. Even when authors know that they have transferred copyright to their publisher, they may be reluctant to ask their publisher if they can attach a Creative Commons licence to their work for fear of jeopardising their relationship with the publisher.[11]

These issues are best dealt with through established policies. Every research and academic institution should have in place policies relating to copyright management, including the licensing of copyright works. These policies should deal with the legal impediments to making copyright material openly accessible, including determining who owns copyright, how to obtain necessary permissions from copyright owners and how to licence material in a way that grants the appropriate rights but retains the appropriate controls. The policies may also deal with non-legal issues, including how to get authors interested in open access repositories and how to assist authors in maintaining a positive relationship with their publisher while asserting additional rights.[11]

The Creative Commons open content principles have been extended to the sharing of scientific data and publications through the Science Commons Project.[12] As explained on the Science Commons website, Creative Commons licences can be used in relation to databases that attract copyright protection.[13] An example of a database that uses a Creative Commons licence appears below.

Example Three – UniProtKB/Swiss-Prot Protein Knowledgebase

UniProtKB/Swiss-Prot is a protein knowledgebase established in 1986 and maintained since 2003 by the UniProt Consortium. The UniProt Consortium is a collaboration between the Swiss Institute of Bioinformatics and the Department of Bioinformatics and Structural Biology of the Geneva University, the European Bioinformatics Institute (EBI) and the Georgetown University Medical Centre's Protein Information Resource.

The data held within UniProtKB includes protein sequences, current knowledge on each protein, core data (sequence data; bibliographical references and taxonomic data) and further annotation. The database is organised through a web interface that displays the data associated with each protein sequence.

The UniProt Consortium states that the public databases maintained by UniProt

Consortium members are freely available to any individual and for any purpose.

A copyright statement on the UniProtKB website states:

We have chosen to apply the Creative Commons Attribution-NoDerivs Licence to all copyrightable parts of our databases. This means that you are free to copy, distribute, display and make commercial use of these databases, provided you give us credit. However, if you intend to distribute a modified version of one of our databases, you must ask us for permission first.

[14]

The UniProtKB open access system has been described as operating on an "honour system" on the basis that the user community is small and so accurately monitored by electronic tracking that non-compliance with the copyright licence would risk unacceptable costs in loss of reputation, peer pressure and possible denial of privileges.

2. Open Patent Licensing

Increased interest in sharing data also raises issues in relation to patents. Patents protect products and processes that are novel, useful and involve an inventive or innovative step. Patents must be registered and confer on the patentee the exclusive right to use or sell the patented product during a certain period of time (usually 20 years).

For researchers intending to seek patent protection for inventions derived from their research, a primary concern is whether they will be able to obtain a patent and whether disclosure of their data to other researchers could prevent them from obtaining a patent (because the product would no longer be "novel"). For researchers who do not intend to patent, a concern is whether another person could secure a patent over an invention that encompasses the researcher's data.

Some researchers will be more interested in making their data openly available to advance research than in commercialising patented products or processes derived from their research. These researchers will not be concerned that public disclosure of their research data could prevent them from obtaining a patent because the invention is no longer novel or is obvious. However, disclosure of data, in itself, will not always be enough to prevent patenting. The problem arising from the public release of data is that it leaves the way open for another party to make improvements to the disclosed data and then make those improvements proprietary.

Claire Driscoll of the NIH describes the dilemma as follows:

It would be theoretically possible for an unscrupulous company or entity to add on a trivial amount of information to the published…data and then attempt to secure 'parasitic' patent claims such that all others would be prohibited from using the original public data.[15]

Where information or data is used to develop a patentable invention, the subsequent patent rights may be broad enough to cover use of the actual data forming part of the invention. As Eisenberg and Rai explain:

Although raw genomic data would not undermine claims to specific genes of identified function, annotated data might do so. A major goal of annotation is to identity coding regions in the genome and add information about the function of the protein for which the region codes.[16]

Consequently, some research projects have relied on licensing methods, similar to the open content copyright licences described above, in an attempt to keep the data "open," rather than simply releasing the data into the public domain.

One example is the HapMap Project, which required anyone seeking to use research data in the HapMap database to first register online and enter into a click-wrap licence for use of the data. The licence prohibited licensees from filing patent applications that contained claims to particular uses of data obtained from the HapMap databases, unless that claim did not restrict the ability of others to freely use the data.[17]

Another approach – currently being practised by the CAMBIA project - is to obtain a patent and then open licence the use of the patented invention on certain conditions. Some argue that, in specific areas, effective open access will only be achieved by allowing a certain level of use of the copyright and patented material.

2.a The CAMBIA Approach

CAMBIA is an international, independent, non-profit research institute led by well known scientist, Richard Jefferson. CAMBIA was designed to "foster innovation and a spirit of collaboration in the life sciences."[18] This goal is achieved through four interconnected work products:

- *Patent Lens*, which provides tools to make patents and patent landscapes more transparent;
- *Biological Open Source Initiative (BiOS)*, which advocates for the sharing of life sciences technology and data through a series of licences;
- *BioForge*, a research portal (or repository) that makes data and technologies openly available for others to use in new innovations, whether for research, commercial use, or humanitarian use; and
- *CAMBIA's Materials*, new technologies developed by CAMBIA, particularly in the field of genetics, which CAMBIA makes openly available under a BiOS licence.

CAMBIA has also applied for and obtained twelve patents of biological material in different patent offices around the world. CAMBIA's approach involves obtaining patents over products or processes, but then licensing the use of those inventions under open terms. A primary object of this is to ensure that the biological material is not patented by others under restrictive terms, which do not allow for open access and use by others. Another object is to encourage innovation. CAMBIA

Strives to create new norms and practices for dynamically designing and creating the tools of biological innovation, with binding covenants to protect and preserve their usefulness, while allowing diverse business models for wealth creation, using these tools.[19]

CAMBIA has developed two open licences relevant to data – the BiOS Plant Enabling Technology Licence and the BiOS Genetic Resource Technology Licence. Paragraph 2.1 of each licence gives licensees

A worldwide, non-exclusive, royalty-free right and licence to make and use the IP & Technology for the purpose of developing, making, using, and commercializing BiOS Licensed Products without obligation to CAMBIA, including a sub-licence…[20]

This gives licensees the right to sub-licence the material, as long as it is sub-licensed under the same terms as contained in the original licence agreement.

CAMBIA's model allows researchers to obtain patents over inventions that build upon CAMBIA's research data. However, instead of using patent licences to "extract a financial return from a user of a technology," CAMBIA advocates using a patent licence to "impose a covenant of behaviour."[19]

According to CAMBIA, the purpose of the BiOS licences is that:

Instead of royalties, BiOS licensees must agree to legally binding conditions in order to obtain a licence and access to the protected commons. These conditions are that improvements are shared and that licensees cannot appropriate the fundamental "kernel" of the technology and improvements exclusively for themselves. Licensees obtain access to improvements and other information, such as regulatory and biosafety data, shared by other licensees. To maintain legal access to the technology, licensees must agree not to prevent other licensees from using the technology in the development of different products.[19]

By making the licence cost-free, CAMBIA hopes to encourage what founder Richard Jefferson terms:

The most valuable contribution to the license community: "freedom to innovate."[19]

CAMBIA is currently developing a new version of the BiOS licence, which to our understanding will remove any positive obligation to share improvements in return for some type of covenant not to enforce rights in relation to patented improvements against members of the CAMBIA community.

Conclusion

Any research project should adopt a "mission-driven approach." The question to be asked is, "What do we want to achieve?" The goal may be commercial gain, may simply be the advancement of research for the public good, or both. Open access to research data and publications should always be considered, especially in the case of publicly funded research.[21] The level of access to and reuse of research data and

publications that is to be allowed should ideally be determined at the outset of a research project.

From the commencement of a research project, it is imperative to have appropriate policies and frameworks in place. Policies must cover copyright management and data management. Copyright management policies should deal with copyright ownership rights and how copyright protected material is to be shared. Researchers should consider the various open content licensing models that can be applied to their copyright material. Data management plans should deal with how data is to be generated, managed and stored; data ownership rights and legal controls that may apply to data (including patents); and how access will be provided to the data and how the data will be disseminated.

Interestingly, some argue that, while open access in terms of copyright material will allow us to read that material and potentially to reproduce and electronically communicate it to colleagues, it most likely will not provide permission to use or exploit related patented material. One of the challenges for the near future will be to consider to what extent open access to publicly funded knowledge (e.g., that makes up tools or platform technologies in biotechnology) requires an accompanying commitment to allow a certain level of use of patented material. In this regard, the CAMBIA project provides an interesting approach that deserves close attention in coming years.

As lawyers, we hope that the law can adapt to facilitate the very great potential cyberinfrastructure promises us. To this end, we need to think of legal tools as being part of the infrastructure and work towards providing innovative models for the future.

[1]For more information, see Professor Brian Fitzgerald, Dr Anne Fitzgerald, Professor Mark Perry, Scott Kiel-Chisholm, Erin Driscoll, Dilan Thampapillai and Jessica Coates, 'OAK Law Project Report No. 1: Creating a Legal Framework for Copyright Management of Open Access Within the Australian Academic and Research Sector' (2006), http://eprints.qut.edu.au/archive/00007306/01/Printed_Oak_Law_Project_Report_No_1.pdf and Dr Anne Fitzgerald and Kylie Pappalardo, 'Building the Infrastructure for Data Access and Reuse in Collaborative Research: An Analysis of the Legal Context' (2007), http://www.oaklaw.qut.edu.au/files/Data_Report_final_web.pdf
[2]Lambert Working Group on Intellectual Property - http://www.innovation.gov.uk/lambertagreements/index.asp?lvl1=1&lvl2=0&lvl3=0&lvl4=0
[3]University Industry Demonstration Partnership - http://uidp.org/
[4]Seventh Research Framework Programme (FP7) - http://cordis.europa.eu/fp7/home_en.html See further, European Commission, "Intellectual Property and Technology Transfer" http://ec.europa.eu/invest-in-research/policy/ipr_en.htm
[5]Fitzgerald, B., Fitzgerald, A., Perry, M., Kiel-Chisholm, S., Driscoll, E., Thampapillai, D., Coates, J. "OAK Law Project Report No. 1: Creating a Legal Framework for Copyright Management of Open Access Within the Australian Academic and Research Sector,"(2006), at 1.22, http://eprints.qut.edu.au/archive/00007306/01/Printed_Oak_Law_Project_Report_No

_1.pdf
[6]Creative Commons - http://creativecommons.org/
[7]Public Library of Science (PLoS) - http://www.plos.org
[8]PLoS One - http://www.plosone.org
[9]Nature Precedings - http://precedings.nature.com/
[10]Nature Precedings – Copyright - http://precedings.nature.com/about
[11]For more information, see Kylie Pappalardo and Dr Anne Fitzgerald, 'A Guide to Developing Open Access Through Your Digital Repository' (2007), available at http://www.oaklaw.qut.edu.au/node/32
[12]Science Commons - http://sciencecommons.org/
[13]See http://sciencecommons.org/resources/faq/databases
[14]UniProtKB - http://www.uniprot.org/terms
[15]Driscoll, C. T. "NIH data and resource sharing, data release and intellectual property policies for genomics community resource projects," *Expert Opin. Ther. Patents* (2005), Vol. 15, no.1, pp. 4.
[16]Eisenberg, R., Rai, A. "Harnessing and Sharing the Benefits of State-Sponsored Research: IP Rights and Data Sharing in California's Stem Cell Initiative," (2006), 21 *Berkley Law Journal* 1187 at 1202.
[17]HapMap, Project Public Access Licence, previously at http://www.hapmap.org/cgi-perl/registration. Users are no longer required to enter into this licence to use the HapMap database. See also [16].
[18]CAMBIA - http://www.cambia.org/daisy/cambia/home.html
[19]Jefferson, R. "Science as Social Enterprise: The CAMBIA BiOS Initiative," (2006) *Innovations*, Vol. 13 at 22 available at http://www.bios.net/daisy/bios/3067/version/default/part/AttachmentData/data/INNOV0104_pp13-44_innovations-in-practice_jefferson.pdf
[20]CAMBIA, Biological Innovation for Open Science, About BiOS (Biological Open Source) Licences, http://www.bios.net/daisy/bios/398
[21]OECD, Declaration on Access to Research Data from Public Funding http://www.oecd.org/document/0,2340,en_2649_34487_25998799_1_1_1_1,00.html

URL to article: http://www.ctwatch.org/quarterly/articles/2007/08/the-law-as-cyberinfrastructure/