The Impact of Electronic Publishing on the Academic Community

Session 5: Digital libraries and archiving of electronic information

Scholarly communication and the licensing of electronic publications

Ann Okerson

Yale University, Sterling Memorial Library, PO Box 208240, New Haven, CT 06520-8240, U.S.A., Ann.Okerson@yale.edu


Copyright Information

Introduction

My topic is the licensing of electronic information: what it means and, more to the point, how publishers of all kinds (societies, university presses and commercial publishers) are making a market for electronic journals and databases, at the same time as we here talk about how information ought to be free to users via the Web or Internet. While some important free spaces are indeed being carved out in cyberspace, in the main our society appears to be headed toward an information universe in which market forces and money are key and libraries or research institutions are funding a great deal of the new form of scientific and humanistic information.

Why do contracts or licenses (rather than copyright) govern electronic content? Can electronic information really be 'owned'?

Society now faces what seems to be a powerful competitor for copyright's influence over the marketplace of cultural products, one that carries its own assumptions about what intellectual property is, how it is to be used, how it can be controlled, and what economic order can emerge as a result.

For convenience's sake, the codification of intellectual property is assigned to the early 18th century. That is when the evolving notion of copyright was enacted into law, shaping a
marketplace for cultural products unlike any seen before. In that 18th century form, copyright legislation depended in four ways on the technologies of its time.

First, high-speed printing presses increased the printer's at-risk capital investment and greatly multiplied the number of copies of a given original that could be produced. Thus an author could begin to realize direct or indirect financial rewards through signing over copyright to a publisher. The visibility and the capital costs of establishing and operating a printing press meant that those who used such presses to violate copyright were liable to confiscatory punishment at least commensurate with the injury done by the crime itself. Last but certainly not least, a physical artifact (book) was produced and its subsequent replication was fairly easy to regulate.

While copyright (that is, the notion that creative works can be owned) is still and probably always will be recognized as a fundamental principle by most players in the information chain (whether they care to protect that ownership or not), many believe that currently articulated 'rules' about intellectual property ownership do not effectively address either the technical capabilities or reader's needs of a high-speed information distribution age. So, the gulf between copyright laws or treaties and the universe that those laws ought to address today feels to many vast and deep, and instead of relying on national copyright law, surrounding case law, international treaties and prevailing practice to govern information transactions for electronic information, copyright holders have turned to contracts (or licenses as they are more commonly called in the library world) as the mechanism for defining the owner, user and uses of any given piece of information.

When did this mode of doing business begin for libraries?

The concept of a license is old and fundamentally transparent. A license is essentially a means of providing use of a piece of property without giving up the ownership. For example, if one owns a piece of property and allows another to use it without transferring title, one may by law of contract stipulate the conditions one chooses; if the other party agrees to them, then a mutually agreeable deal has come into being. The owner of a movie theatre rarely owns the cans of film delivered weekly to the cinema, holding them instead under strict conditions of use: so many showings, so much payment for each ticket sold, etc. In the license mode of doing business (precisely defined by the legal contract that describes the license) the relationships are driven entirely by contract --- not copyright --- law: the owner of a piece of property is free to ask whatever price and set whatever conditions on use the market will bear. The ensuing deal is pure 'marketplace': a meeting of minds between a willing buyer and a willing seller.

Most academics began to be parties to license agreements when personal computer software (WordStar, WordPerfect) appeared in the 1980s in shrink-wrap packages for the first time. Purchasers of such software may have read the fine print on the wrapper detailing the terms and conditions of use, but for the most part they either did not or have ceased to do so. The thrust of such documents is simple: by opening the package, the purchaser has agreed to certain terms, terms that include limited rights of ownership and use of the item paid for \[1\]. In many ways, this mode of licensing raises problematic questions, but in others such as sheer efficiency, it suggests the kind of transaction that the scholarly information marketplace needs to achieve.
By the late 1980s, a number of important indexing and abstracting services (such as the well-known Biological Abstracts --- now BIOSIS --- Psychological Abstracts, Science Citation Index, Medline and many others) offered electronic versions directly to libraries via CD-ROM or through dial-up, and it was at this point, within the last ten years, that library licenses gradually became recognized as a means to a new and different sort of information acquisition or access. The license terms offered to libraries were accepted or not, the library customer regarding them mostly as non-negotiable. Non-acceptance was more often than not a matter of affordability, and there seemed to be little room for the library customer to affect the terms. Complaints about terms of licenses began to be (and persist in being) legion, for important reasons such as:

**Potential loss of knowledge**

By definition, licenses are arranged for specific periods of time. At the end of that time, librarians rapidly discovered, if the license is not renewed, prior investment can become worthless as the access ceases (for example, a CD-ROM must be returned or perhaps it stops being readable; connections to a remote server are severed).

**License restrictions on use and users**

Institutions are often asked to assure that only members of the institution can use electronic information, in order to reduce or curtail its leakage. Authorized institutional users are asked not to distribute the information to users in other universities, for example.

**Limitations on user's rights**

Initial license language not infrequently asks that institutional users themselves severely limit what and how much they may copy from the information resource and may prescribe the means by which such copying can be done.

**Cost**

In general, electronic licenses for indexing and abstracting services appeared, and still appear, to cost significantly more than print equivalents [2].

**What has happened to increase libraries' awareness of licenses?**

(i) Sheer numbers. Whatever their marketplace insecurities may be, thousands of information providers have jumped into the scholarly marketplace with electronic products of one sort or another: CDs, online databases, full-text resources, multimedia [3]. Many learned societies, scientific publishers, university presses, full-text publishers, vendor/aggregators, as well as new entrants to the publishing arena, now offer either beta ('early') or well-tested versions of either print-originating or completely electronic information. The numbers have ballooned in a short 2-3 years with no signs of abating. The Yale University Library licenses over 400 substantial electronic resources (generally these are not just single works but databases, collections of books or journals, etc.) of varying sizes, types, media and price, and reviews about two new electronic content licenses a week.
(ii) The attempts of various players in the information chain to create guidelines about electronic fair use or fair dealing have not so far proved fruitful [4]. Such stalemates should come as no surprise; in fact, they are healthy and proper. Any changes to national guidelines, let alone national law or international treaty, should happen only when the public debate has been extensive and consensus can be reached. In any case, instead of waiting on Congress or CONFU and allowing terms to be dictated to both parties by law, publishers and institutions are making their deals together, one step at a time.

(iii) Numerous formal partnerships and informal dialogues have been spawned by capabilities of new publication technologies. A number of libraries collaborate with the publishing and vendor communities as product developers or testers. Such relationships are fruitful in multiple ways. With regard to licensing, they encourage friction, pushback and conversation that leads to positive and productive outcomes. Libraries have been offered --- and have greatly appreciated -- the opportunity to discuss at length the library licenses of various producers, and many of us believe we have had the opportunity to shape and influence these with mutually satisfactory results [5].

(iv) In the United States, library consortia have aggressively entered the content negotiating arena. While library consortia have existed for decades, and one of their primary aims has been effective information sharing, it is only in the 1990s (and mostly in the last 2--3 years) that a combination of additional state funding (for state-wide consortia), library demands and, surprisingly, producers' willingness to negotiate with multiple institutions have come together to make the consortial license an efficient and perhaps cost-effective way to manage access to large bodies of electronic content. In Europe, we see national site licensing experiments happening in the U.K. and the Netherlands. In such cases, the publishers have crafted an arrangement that covers all the university and polytechnic libraries of an entire country, giving all students and faculties access to certain materials in electronic form. That said, the future of consortial licensing is no more certain than for individual library licenses, though for different reasons [6].

(v) The library community has organized itself to understand the licensing environment for its constituents. The Association of Research Libraries (ARL) has produced an introductory licensing brochure, The Council on Library Resources/Commission on Preservation and Access has supported Yale Library's creation of an important for-free World Wide Web site about library content licensing [7], and the Yale Library offers the library, publisher, vendor and lawyer world a high-quality moderated online list where the issues of libraries and producers are aired daily [8].

Some notable challenges for institutional licensing

Terms of use

This area needs to be mentioned at the outset, as it has caused some of the most anguished discussions between publishers and libraries. Initially, many publishers' contact language for electronic information was highly restrictive about both permitted users and permitted uses. Assumptions and requirements about how use ought to be contained were at times ludicrous, for example, in phrases such as "no copies may be made by any means electronic or mechanical".
Through dialogue between librarians and producers, who are usually genuinely eager to market their work to happy customers, much of this language has disappeared from the first draft contracts presented to library customers. Where libraries are energetic and aggressive on behalf of their users, the terms of use can indeed be arranged to facilitate educational and research goals.

**Scalability**

Institutional electronic content licenses are now generally regarded as negotiable, mostly because the library/customer side of the marketplace is now treating them as such (which publishers seem to welcome) and successes of different sorts have ensued (success being defined as a mutually agreeable contract), making the parties feel that they can work together effectively in this new mode. However, negotiations are labour intensive. Negotiation requires time (to develop the expertise and to negotiate), and time is a major cost here. The current mode of one-on-one negotiations between libraries and their publishers seems at the moment necessary, for many reasons, and at the same time it places new demands on institutional staff. Scalability is the biggest challenge for the licensing environment.

Clearly, it is too early to shift the burden onto intermediaries such as subscription agencies or other vendors who have vested interests of their own. So far their intervention has been absent or not particularly successful. In fact, in some of the situations where intermediaries purvey electronic databases, library customers secure less advantageous use terms than those libraries could obtain by licensing directly from the publishers.

**Price**

Pricing models for electronic information are in their infancy; they tend to be creative, complicated and often hard to understand. Consortial pricing can be particularly complex [9]. Each new model solves some of the equity or revenue problems associated with earlier models but introduces confusion of its own. While pricing of electronic resources is not, strictly speaking, a problem with the license itself, price has been a major obstacle in making electronic agreements. The seemingly high price tags for certain electronic resources leave the 'serials crisis' in the dust [10]. In addition, pricing models have been very diverse and often quite complicated for customers to understand very well [11].

**The liability/trust conundrum**

One of the most vexing issues for producers and their licensees has been the producer's affirmation that institutions can and ought to vouch for the behaviour of individual users and that individual users' abuses of the terms of a license can, in fact, kill the deal for a library or a whole group of libraries. Working through this matter with provider after provider in a partnership/co-operative approach poses many challenges. In fact, this matter may be a microcosm of a larger issue: the development of the kind of trust that must underlie any electronic content license. Generally the marketplace for goods is not thought of in terms of trust; it regarded as a cold-cash (or virtual cash) transaction environment. Yet the kinds of scaled-up scholarly information
licenses that libraries are engaging with now depend on mutual understanding and trust in a way not needed for the standard trade --- or even the print --- market to work.

The aggregator aggravation (and opportunity)

The technological investments that producers need to make to move their publications on to an electronic base, the publishing processes that are being massively re-conceived and re-organized, and not least, the compelling vision of digital libraries that proffer information to the end user through a single or small number of interfaces, with a single or modest number of search engines, gives rise to information aggregators of many sorts and lead publishers and vendors to package and then license materials as groups or collections: all the journals of a single publisher, all English poetry of the 18th century, and so on. Single titles (large scientific journals of record may be the exception) are rarely cost-effective in this environment. From the viewpoint of academic research libraries, it appears that the electronic environment has the effect of shifting transaction emphasis from single titles to collections or aggregations of electronic materials as marketplace products. In turn, licensing collections from aggregators makes libraries dependent on publishers and vendors for services in a brand new way [12].

The electronic collections offered to the academic library marketplace are frequently not in configurations librarians would have chosen for their institutions, had these resources been unbundled. This has been an issue in several of Yale Library's negotiations. Say that the publisher of a large number of quality print journals makes only the full collection available in electronic form. By this means, the Yale Library recently 'added' 50 electronic journal titles to its cohort, titles it had not chosen to purchase in print. The pricing model did not include a cost for those additional 50 titles; it was simply easier for the publisher to include all titles than to exclude the less desirable ones.

Beyond licensing

This introductory discussion was meant to give participants in the Academia workshop a flavour of the emerging digital marketplace for scholarly information, a market that many practising scientists are probably not aware of. In itself, the licensing environment is neither bad nor good; it can be as beneficial as the buyers (often libraries) and the users insist that it be. Making an effective marketplace for the participants in the information chain is an important activity. Why, one might ask? Is libraries' licensing of electronic information not a capitulation to commercial interests that inhibit the progress of science and education? My answer to this is, not at all. If we accept that a marketplace is a useful thing, that it generates the necessary capital for publishers and vendors to continue work that the library and scholarly community deem valuable, then pursuing user-friendly, affordable licenses is surely one practical way to invest in a robust information future --- by assuring its economic stability. On the other hand, if the scholarly and scientific communities no longer value the work that other sectors or portions of other sectors contribute to the information chain, then making a statement within the marketplace is a fundamentally important way of withdrawing support for activities of those players the scholarly community no longer wishes to support. In fact, decisions about the value that users or patrons find important are decisions that pertain to all media, both traditional and electronic, even though
such decisions may be expressed in slightly different ways (by not renewing a print journal subscription versus not licensing an electronic publication).

Enough now of licensing and let us give the last word to a category one may term 'beyond licensing'. Given the current successes, from modest to spectacular, of certain electronic resources generated from within the research and academic communities and shared for free via the Internet or World Wide Web (the LANL preprint archives are extolled, quite rightly, here and in other venues as the most dramatic of these successes), one has to wonder what sorts of scholarly and scientific information will indeed be created and thrive under a different model: one in which characteristically institutions of higher education or government agencies bear the creation and distribution costs as part of their research mission. Numerous such resources already exist, though few are well known outside specialist communities. For such 'free' electronic resources, often ones of great substance and power, no care is taken about matters of licenses, copyrights or cost recovery. Ideally, this is a subset of the information universe all of us here should be eager to embrace and advance however we can. It is not that such a model implies disrespect for copyright ownership or denigrates the value of intellectual property; in fact, such a model celebrates the value of intellectual property by sharing it widely with those who will use and value it. While librarians already operate wisely and energetically in the commercial information world and its licensing requirements, they too look to alternative methods of information sharing and recovery. But that is a topic for a different paper.

In the meanwhile, I hope the thoughts here have offered an insight into yet another way that the work of libraries is changing and some of the efforts librarians engage in to effectively and smoothly manage such changes.

Acknowledgements

The thoughts for this talk were developed en route to a longer paper on this topic commissioned by the Andrew W. Mellon Foundation and given at its conference, Scholarly Communication and Technology, on 24 and 25 April 1997 at Emory University. We thank the Foundation for allowing this version to be produced in this volume.

Notes and references

1. See the recent decision ProCD versus Zeidenberg in the United States Court of Appeals for the Seventh Circuit, 30th June 1996. The question posed was: must buyers of computer software obey the terms of shrink-wrap licenses? The district court had said not. The Seventh Circuit reversed this decision. ProCD (the plaintiff) compiled information from 3'000 or more telephone directories into one database, with additional information such as zip code extensions, with their own searching software. They packaged it as a CD-ROM for personal sale in a shrink-wrap box. They also sold it in other ways to commercial companies as mailing lists and so on. On the basis that factual information cannot be copyrighted, Mr Zeidenberg bought a package of SelectPhone™ at a shop in Madison, WI. He formed a company to re-sell the information which
he made available over the World Wide Web, apparently quite cheaply. Zeidenberg argued that one cannot be bound by the shrink-wrap license because the terms are not known at the time of purchase. They are inside the package and the purchaser cannot be bound by terms that are secret at time of purchase. The judge’s decision was that the shrink-wrap license is legal and a buyer is bound by it. The full decision and rationale may be read at the following URL: http://www.sgpdlaw.com/cases/procd_op.html

2. M. Kellogg (1991) CD-ROM Products as Serials: Cost Considerations for Libraries. Serials Review 17(3), 49--60. Using the tables in this article as a basis for comparison between print reference or indexing-and-abstracting works and their CD equivalents shows a difference of about 30% where resources are comparable. Recent e-mail from the University of Michigan Library suggests differentials between print and electronic as high as 60%.

3. For example, NewJour, the online forum for announcing new e-journals, magazines and newsletters reports 3,634 titles in its archive as of 5th April 1997, this without the 1/100 science journal titles that Elsevier is now making available in electronic form. NewJour is a joint project of the Yale Library, the University of Pennsylvania, and the University of California San Diego Library. Its fully searchable archive is located at: http://gort.ucsd.edu/newjour/

4. A good example of a so far unsuccessful process is CONFU in the United States. CONFU was a legislatively encouraged gathering of about 70 stakeholders in the information chain, representing numerous producer and user groups, who met monthly for over two years in order to try to achieve consensus on electronic fair use. A good summary of the flavour, debates and progress of CONFU can be found at URL: http://www.utsystem.edu/OGC/IntellectualProperty/confu.htm The CONFU interim report is available at URL: http://www.uspto.gov/web/offices/dcom/olia/confu/

5. A. Okerson (1996) Buy or Lease? Two Models for Scholarly Information at the End (or the Beginning) of an Era, Daedalus 125(4), 55--76. This is the special issue on libraries called Books, Bricks, and Bytes. I suggest that one possible outcome of the new trend to scaled-up consortial licensing activities is that the library marketplace will gain significant power and that publishers of scholarly information could find themselves in quite a different position to the 'captive' marketplace of today. It is possible to argue that such an outcome is healthy; on the other hand, even librarians and scholars might find it undesirable in that it would put today's specialized scholarly publications, with their attendant high prices, out of business. It seems to me that such publications are already at risk as they offer a perfect opportunity for scholars, universities and libraries to devise a different mode of publication and distribution. The Daedalus piece can also be found at the URL: http://www.library.yale.edu/~okerson/daedalus.html


7. See "LIBLICENSE; Licensing Digital Information --- A Resource for Librarians" at URL: http://www.library.yale.edu/~llicense/index.shtml. This Web resource contains license
vocabulary, licensing terms and descriptions, sample publishers licenses, links to other licensing sites and a bibliography about the subject.

8. LIBLICENSE-L is a moderated list for the discussion of issues related to the licensing of digital information by academic and research libraries. To join the LIBLICENSE-L list, please do the following: send a message to listproc@lists.yale.edu and leave the subject line blank. In the body of the message type: subscribe LIBLICENSE-L, First name, Last name.

9. A LIBLICENSE-L message of 12th February 1997 enumerated a dozen different pricing models for electronic resources, and correspondents added several more in subsequent discussion.

10. Several reasons are advanced for the higher cost of electronic resources over comparable print resources: (i) the producers are making new research and development, and technology investments whose significant prices are passed on to the customer; (ii) producers of journals generally offer a package which includes print plus electronic versions, giving the customer two different forms of the same information, rather than only one; (iii) the functionality of electronic resource is arguably higher than of the print version; (iv) electronic resources are not marketed as single journals or books but as scaled-up collections, often of substantial heft (consider the corpora of humanities full texts marketed by Chadwyck-Healey, the large backfile collections of JSTOR, the full collection of Academic Press titles available under its IDEAL program: it seems that there is little incentive for producers to create and sell one electronic item at a time); and (v) in becoming the source, site or provider, the electronic information is taking on many of the library's roles and costs.

11. Imagine, for example, a large journal publisher who markets only to consortia, or groups of libraries. This publisher offers a pricing model based on the prices each of the subscribing libraries paid at a baseline set for 1995, for a license period of three years. No titles can be cancelled for credit within the license period. Furthermore, the prices are for electronic versions only, calculated at 90% of the current year's list price (the list of titles being the 1995 list taken at that time). If the library customers desire to buy print versions, they can do so at various deep discounts, i.e. as much as 75% off. The discounts, though very large, are not identical in percentage. One could go on with the details of this particular offering, but the point is to show a high degree of complexity which I have only begun to describe here, and to characterize it as quite different from the comparatively simple print subscription pricing models.

12. A LIBLICENSE-L message of 14 March 1997 defined aggregators in the following way: "'aggregation' as used on this list means the bundling together or gathering together of electronic information into electronic collections that are marketed as a package". For example, DIALOG@CARL aggregates 300 databases; Academic Press's IDEAL aggregates 170+ journals; Johns Hopkins University Project MUSE is an electronic collection of 40+ journals, and so on. But the term 'aggregator' is more usually used in describing the supplier who assembles the offerings of more than one publisher, so one is more likely to hear Dialog, OCLC (Online Computer Library Center), Information Access and UMI spoken of as aggregators than the Johns Hopkins University Press.