

## Special Collections and the Law in the U.S.<sup>1</sup>

Peter B. Hirtle

Why, you may be asking, should I care about how special collections in the U.S. are addressing legal issues? Europeans, after all, work in a vastly different arena than Americans. What possible relevance could US practice have for you?

In reality, and especially when it comes to unpublished works, the differences between the Europe and the US are surprisingly few. How US archivists are facing the issue may be of some interest. I can also share some of the concerns that do keep me awake at night, and which may be of concern to you as well.

So in this brief period of time, I want to discuss four things:

- The legal uncertainty that surrounds much digitization of special collection materials;
- The growing willingness of archivists and special collection librarians to act anyway based on their assessment that the risk to their institutions and to them is small;
- The development of “best practices” that support them in their efforts;
- Potential problems down the road. I can think of 3:
  - Possible new privacy legislation
  - TCEs
  - International jurisdiction issues

In the US, archivists have by and large failed to use new digital technologies to make the riches in their holdings more widely accessible. Money, of course, has a lot to do with it. The funding for massive archival digitization projects has just not been available. But there has also been a reluctance on the part of archivists to pursue the little funding that is there because of our professional practices.

The problem is that archivists and librarians may be the last people in the U.S. who blindly and passionately attempt to respect copyright. Our professional codes tell us we must. For example, the Code of Ethics for Archivists ([http://www.archivists.org/governance/handbook/app\\_ethics.asp](http://www.archivists.org/governance/handbook/app_ethics.asp)) asserts that archivists must obey all federal, state, and local laws – which would include copyright law. Furthermore, the Code of Ethics makes it plain that we must also respect personal privacy, especially when it involves third parties whose works may appear in archival collections but who have not actually donated the material to the repository themselves. Similarly, the ALA/SAA Joint Statement on Access to Research Materials in Archives and Special Collections Libraries (<http://www.archivists.org/statements/ALA-SAA-Access09.asp>), while encouraging archivists to make their holdings as widely available as possible and on an

---

<sup>1</sup> Copyright 2010 Peter B. Hirtle. Licensed under a Creative Commons Attribution-Noncommercial license. This talk is based in part on Peter B. Hirtle, “Undue Diligence?,” forthcoming in vol. 34 of the *Columbia Journal of Law and Arts*.

equitable basis, also states that we must respect copyright laws when making reproductions for users.

So archivists are imbued with the desire to respect the law. Yet in spite of our educational efforts, there is also great uncertainty in the profession as to what the law says, what is legal, and what is illegal. It is a riddle that most archivists do not know how to solve. Coupled with this is an understandable desire on the part of archivists to stay out of court. Because archivists are risk adverse, it is easiest to assume that the law stops us from doing many socially desirable things.

It doesn't help much that the law itself offers little in the way of assistance to archivists. There are explicit provisions in US law that allow archivists to digitize collections of unpublished material. It is often difficult, however, to tell the difference between what is published and what is unpublished, and the penalties for reproducing and distributing registered material can theoretically be great. In addition, the law only allows the use of the unpublished items on the premises of the library or archives, but not on the Web. US practice regarding fair use is perhaps more generous than "fair dealing" in the UK, and archivists did secure an explicit statement that it is possible to make fair use of unpublished works. Nevertheless, the actual determination of whether a particular use is fair is complicated, and great uncertainty surrounds its use.

And of course securing the permission of the copyright owner is not an option. Archival collections are perhaps the purest representation of the "orphan works" problem – those works still protected by copyright whose current rights owners cannot be found. (An aside: at a recent conference I attended, it was suggested that these items should more properly be called "zombie works" – works that will not die and yet still threaten us.) I do not need to explicate to this group the difficulties of living every day with a "life + 70" term. I would, however, draw your attention to a forthcoming article in the *American Archivist* that documents, using the example of the Thomas Watson papers at the University of North Carolina, the extreme difficulty in identifying and locating the copyright holders of the correspondence in a modern archival collection.<sup>2</sup> So given the professional requirement to follow the law, the absence of clear and useful exceptions for archives in the law, and the difficulties in locating copyright owners, it is not so surprising that few archivists are willing to risk engaging in large scale digitization of collections still technically protected by copyright.

I would argue, however, that even in the absence of explicit and clear legal authorizations, repositories face little real risk in digitizing many collections that are still technically protected by copyright. I know this because many institutions have unknowingly and mistakenly assumed that they had the right to make available archival material in digital form, often to great acclaim. Two examples:

The Judaica Sound Archives at Florida Atlantic University makes available pre-1923 sound recordings on its web site; later recordings are accessible through research stations that

---

<sup>2</sup> <http://www.lib.unc.edu/dc/watson/>

are distributed to other universities.<sup>3</sup> Its efforts to make the voices of early cantors and rabbis better known has let to praise rather than lawsuits. They have done this in the mistaken belief that the material is in the public domain, when in reality it is all still protected in the U.S.

Second example: many archives have elected to add images to the Flickr Commons.<sup>4</sup> Repositories are supposed to assert that there are “no known copyright restrictions” on images in the Commons. There are four reasons for this:

1. The copyright is in the public domain because it has expired;
2. The copyright was injected into the public domain for other reasons, such as failure to adhere to required formalities or conditions;
3. The institution owns the copyright but is not interested in exercising control; or
4. The institution has legal rights sufficient to authorize others to use the work without restrictions.<sup>5</sup>

It would seem, though, that many repositories assume that if they do not know of copyright restrictions, then there are none. Here is an example of a photograph from the George Eastman House in Rochester, N.Y., entitled “Woman and boy sitting in a chair.” No author is given, and there is no evidence or indication that the photograph was ever published. Copyright law says that copyright in an unpublished anonymous work expires 120 years after creation. The Eastman House, however, says that since no copyright owner can be identified, one can treat the photograph as if it were in the public domain. To date, no one has objected. You find this sort of behavior everywhere: treating zombie copyrighted works as if they were in the public domain.

And there are elements in U.S. law that work to the advantage of repositories that make unpublished copyrighted works available. Chief among these is the absence of statutory damages or attorney’s fees for unregistered works. In addition, repositories are absolved of statutory damages if they can make a good-faith assertion that their use is a fair use. Because of these exemptions, it is much more likely that the repository will receive a take-down request than a lawsuit.

So what have we established so far? First, many archivists are reluctant to digitize their collections unless they are clearly out of copyright. Some archivists, though, have mistakenly or willfully mounted digitized collections of copyrighted material – and have suffered no consequences to date. It would appear that repositories, in their desire to obey the law and avoid litigation, may have been overly-cautious. To address this issue, OCLC sponsored an important workshop that led to the development of a document entitled “Well-intentioned practice for putting digitized collections of unpublished materials online.” You should have received a link to the document. It is intended to describe a community of practice to which all archivists can adhere. It is based as much in good archival practice as in law. By and large, archivists have good sense in recognizing material that can be problematic; the statement encourages archivists to follow the same procedures when putting material online. The document encourages

---

<sup>3</sup> <http://faujsa.fau.edu/jsa/home.php>

<sup>4</sup> <http://www.flickr.com/commons>

<sup>5</sup> <http://www.flickr.com/commons/usage/>

archivists to demonstrate sensitivity to the materials, use appropriate disclaimers, and have clear take-down provisions in the unlikely event that an individual might complain.

Tools to help archivists digitize material are also emerging. Most notable is the Society of American Archivists' "Orphan Works: Statement of Best Practices," a link to which was also sent to you. The statement, whose creation was funded by OCLC, was developed in anticipation of the passage of legislation permitting the use of orphan works. The proposed legislation would have required that one first conduct a "reasonable investigation" to identify and locate a copyright owner before using a copyrighted work. The statement was intended to establish the standard for good professional practice with regard to unpublished works. The legislation is still pending, but the statement is nevertheless valuable in helping archivists decide if they are being reasonable in their exploitation of copyrighted materials.

In conjunction with these initiatives, we are seeing more archives deciding to digitize and make publicly available copyrighted unpublished material. Most notable is the Archives of American Art. As part of their Collections Online initiative, they are in the process of digitizing parts or all of 105 collections in their holdings. While some of the material is in the public domain, much is still protected by copyright. Nevertheless, the Archives has yet to receive any complaint as to its digitization.

One important element that protects such digitization efforts is that the material is made freely accessible on the web. It is possible under American law to make a fair use defense for the digitization and distribution of copyrighted material. Doing so for commercial purposes, however, greatly decreases the likelihood that one's use will be found to be fair. In part for this reason, there is a growing movement among some US libraries and museums to do away with publication permission fees for the use of special collection materials. I am proud to say that the Cornell University Library joined this trend in 2009, when it decided to charge only service fees when making reproductions, but otherwise allow the unfettered use of any online or reproduced items from its holdings.

So there is a positive movement in which archivists and special collections librarians are overcoming both their fear of litigation and their desire to monetize their holdings in order to make unique material widely available on the Internet. It sounds all good, doesn't it? Except that as I said at the beginning of my talk, there are things that may derail these efforts. There are three of particular concern to me.

The first is proposed Internet privacy legislation. I assume that we are all interested in protecting privacy, both our own and that of others. The US does not have anything like the European Privacy directive or the UK's data protection act, and so legislation may be appropriate. Yet some of the proposed bills do not recognize that historical data may be of a different nature than one's web searching habits. One bill, for example, would require archivists to scan every page of every digitized document to determine if there is private information such as Social Security numbers included in the document. Mandated item-level review would bring the mass digitization of archival collections to a grinding halt.

The second issue of concern is the initiative underway in the World Intellectual Property organization to create a new intellectual property right in what is labeled “traditional cultural expression.” Unlike copyright, this protection would be perpetual. It would allow representatives of “traditional cultures” to control the reproduction and use of the cultural heritage of those cultures. In some cases, it might require the return of documents to the control of those cultures, even if they didn’t create them (for example, photographs taken by Westerners of traditional sacred ceremonies, even if the photographers had the groups permission at the time the photos were taken). In spite of the fact that there seems to be poor definition of what constitutes a traditional culture or who can speak for that culture, as well as disagreement on the scope of the rights to be granted to indigenous cultures, WIPO is promising quick action on a new international treaty. I am scared.

The third issue that keeps me awake at night are international jurisdiction issues. By and large, I know the copyright rules in the US. I am quite willing to thumb my nose at Europeans who know nothing about US copyright law, including the Cambridge University Press, contact me and tell me that Cornell is violating their copyrights by selling works on Amazon. I respond by politely telling them that we do not sell those books on Amazon.co.uk., but only on Amazon.com, and it is not our problem if a UK customer orders from the American site rather than the British site. Yet some recent rulings in France suggest that a French author might still be able to sue us there for copyright infringement. At a recent conference at which I spoke at Columbia University Law School, a prominent American law professor suggested that we will have to impose geographic filters on all of our web sites, restricting access only to those countries in which we know that our material is “safe.” This seems inimical to the existence of the “Republic of Letters” of which we are all part, and would be an immense impediment to scholarship.

Let me conclude. There is talk of fundamental copyright reform in the wind, but I see little chance any time soon that the law will change in meaningful ways to aid archival and special collection repositories. It is more likely that new laws that threaten access to our holdings will be implemented. We therefore need a different, non-legal, approach. Just as the great documentary editing projects of last century ignored the copyright laws in order to make access to the papers of the U.S. Founding Fathers more widely available, so too must archivists and special collection librarians give up their hope of black-letter rules on copyright and instead embrace responsible risk management as the appropriate way of managing our social goals.