

Removing All Restrictions: Cornell's New Policy on Use of Public Domain Reproductions

Peter Hirtle, Senior Policy Advisor, Cornell University Library,
with Tricia Donovan, Administrative Assistant, ARL

Introductory Note

Restrictions on the use of public domain work, sometimes labeled “copyfraud,” are generating increasing criticism from the scholarly community. With significant collections of public domain materials in their collections, research libraries are faced with the question of what restrictions, if any, to place on those who seek to scan or otherwise reproduce these resources with the intention of publication.

Cornell University Library has responded by adopting new permissions guidelines that open access by no longer requiring users to seek permission to publish public domain items duplicated from its collections. Users planning to scan and publish public domain material are still expected to determine that works are in the public domain where they live (since public domain determinations can vary internationally). Users must also respect non-copyright rights, such as the rights of privacy, publicity, and trademark. The Library will continue to charge service fees associated with the reproduction of analog material or the provision of versions of files different than what is freely available on the Web. The new guidelines are found at <http://cdl.library.cornell.edu/guidelines.html>.

Below is an interview with Peter Hirtle, Cornell University Library's Senior Policy Advisor, who gives some insight into Cornell's decision to change their “text and image use” permissions guidelines.

What prompted you to think about alternatives to Cornell's text and image use guidelines?

Two recent events—concern over the commercial use of images from the Core Historical Literature of Agriculture collection and our decision to add thousands of public domain scans to the Internet Archive—raised the issue of what limitations, if any, Cornell wished to place on scans of its public domain books.

What were Cornell's previous practices regarding the use of digital scans of public domain materials?

Our previous practice was to try to limit some uses. Cornell has made a

substantial investment by acquiring and preserving printed material and then converting that material to digital form. By requiring payment for certain uses of that material, we hoped to recoup some of that investment. So the previous guidelines allowed only personal or research use. Commercial or scholarly use of the material—such as republishing via print on demand or including a page in a university press book or scholarly article—required the permission of the library and a possible payment. Different library units had differing policies: Catherwood Library in the School of Industrial and Labor Relations and the Rare and Manuscript Collections had procedures similar to the general Cornell guidelines, whereas Mann Library allowed free use of its scans.

What were the options Cornell considered when developing the new guidelines?

We identified three options: attempt to restrict in a legally enforceable manner certain uses of digitized material, and thus preserve a possible revenue stream; post an aspirational but unenforceable restriction; or make it free to use. We went with the third option. We get into the specifics later, but our library felt strongly that this was the right choice both for philosophical and logistical reasons.

How would a legally enforceable restriction be implemented?

Copyright is normally used to control subsequent use of written material, but this material was in the public domain and therefore not under copyright restrictions. Scanning alone is not creative enough to warrant its own copyright, and so we have no copyright in scans of material in the public domain. We could, however, have used a contract with potential users that would legally restrict the downstream use of public domain scans. Some institutions, for example, have click-through licenses on their Web sites that require users to agree not to use or redistribute the scans for commercial purposes. Alternatively, some institutions have created a “Terms and Conditions” statement, also called a “browse-wrap” license, which governs subsequent use of the material.

What sorts of problems did you see with this approach?

We identified a number of problems:

- Browse-wrap licenses are of uncertain enforceability, and with either a click-through or browse-wrap license, we would have had to be willing to bring legal action if we found a violation.

- The contract would be only with the individual who used the images from our site. If he/she gave them to a third party who used them commercially, Cornell would have no legal recourse against that third party. Cornell could only bring legal action against the person who actually downloaded the images.
- It would be very difficult to identify which individual actually downloaded the images that were subsequently distributed in violation of the license. As part of its settlement agreement with authors and publishers, Google will do this by embedding identifying information into prints and downloads. Given our library's commitment to confidentiality, I am not sure that Cornell would want to emulate what Google is going to do, even if technically we could.
- The Internet Archive indicated that it has no mechanism to place a click-through license in front of content or a browse-wrap license on the site. This option would therefore only work with Cornell's Web site. It would not have made sense to have different terms for the same material depending on whether it is coming from Cornell or the Internet Archive.

Was there an option that would be less restrictive than a contract?

Yes. Although we may not have been able to restrict legally what people could do with the scans, we could have included a statement that indicated Cornell's preferences. We simply could have asked individuals not to use the scans in ways that concerned us and explained why we were asking for their cooperation. We could also have implemented certain technical measures that would limit problematic activities—for example, the bulk downloading of books from Cornell's Web sites. But we would not have had a legally enforceable contract with users. If users could get copies, they could offer our scans for sale through print-on-demand sites or other commercial ventures.

Does any other major service or database use a voluntary license?

Google uses an aspirational statement with its books. There's a page that they place in the front of their scanned books in which they ask that people only use the books for personal, non-commercial purposes.

But that's not without its issues, correct?

Google has been harshly criticized for including this statement in its books.

See, for example, <http://www.earlham.edu/~peters/fos/2007/01/testing-googles-restrictions-on-google.html>.

Jessamyn West, a noted library blogger and critic, recently wrote:

The thing I've found so weird about Google Books lately is that if you download a public domain book, you still get a little Google cover page asking you to "play fair" or some nonsense and not use the book for commercial purposes and a whole bunch of other codicil type stuff that's not legally binding [it's PUBLIC domain] but just asked in the "do us a favor" sort of way. I think they'd prefer to have a lot more restrictions on the use of public domain materials than they do.

In addition, an aspirational license would do nothing to stop bad actors who could take the books, cut out an aspirational statement, and sell them if they wanted. The people who would have been really harmed by this option are the ones who would have responsibly respected our wishes.

But there's an even less restrictive option than the aspirational statement, correct?

That's right. Cornell's final option was to free our scans of public-domain items from any restrictions. Our books would then appear in the Internet Archive, for example, with metadata crediting Cornell for their creation, but with no implicit or explicit limitations on further use.

How do other research organizations and libraries use the "free use" model?

In the Hathi Trust, Michigan books have watermarks indicating who digitized the book and from which library, but there is no page inserted in the scans limiting use, nor is there a "Terms and Conditions" page that describes restrictions (though some individual volumes may have licenses or other contractual terms that may restrict further distribution). Hathi Trust also limits printing and downloading to 10 pages at one time, however, and only as a PDF. The Internet Archive makes a broader set of files available, including PDF downloads of entire books. The PDFs I have examined on the Internet Archive site from the University of Chicago and Johns Hopkins University do not contain any restrictions on subsequent use.

What are the concerns associated with the “free use” model?

The downside to this approach is that nothing can stop someone from taking the PDFs or other versions of Cornell books from the Internet Archive and producing their own print-on-demand or electronic versions for sale.

Of the three options identified, why did Cornell adopt Option 3?

The benefit to scholarship, research, and learning from having free access to public domain books is considerable. Making this material free is consistent with

Making this material free is consistent with the University's mission to promote a culture of broad inquiry, as well as our library's commitment to making information free and disseminating knowledge as widely as possible.

the University's mission to promote a culture of broad inquiry, as well as our library's commitment to making information free and disseminating knowledge as widely as possible.

In addition, revenues from the permission process may not even have outweighed the costs associated with managing permissions. And as we

have already discussed, there are legal, technical, and ethical issues with the other options. Lastly, it would be ironic if, at the same time the library was advocating open and free access to the copyrighted works of faculty members, it also was restricting use of its digital files of public domain books.

Do the new guidelines only apply to digital files?

No. It would not make sense to charge permission fees for the use of reproductions made from physical public domain items found in the library if we are not charging permission for items that have already been digitized. We do, however, continue to charge for reprographic services.

Do you see other libraries adopting this sort of policy in the future?

I would assume so, for the same reasons that led Cornell to this position. There are reasonable forces working against others following our lead, however. First, there is the desire to control the use of “my materials” in order to ensure that they are treated with respect. In addition, there is the allure of what appears to be an easy revenue source. We are having to license access to public domain works found in other libraries that have partnered with commercial ventures;

it seems odd to give away our own public domain works to those same institutions that are profiting from us. In the end, however, I assume that most libraries will focus on their mission to support scholarly communication, and will adopt policies similar to ours.

© 2009 Peter Hirtle



This work is licensed under the Creative Commons Attribution 3.0 United States License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/3.0/us/>.

To cite this article: Peter Hirtle, "Removing All Restrictions: Cornell's New Policy on Use of Public Domain Reproductions." *Research Library Issues: A Bimonthly Report from ARL, CNI, and SPARC*, no. 266 (October 2009): 1–6. <http://www.arl.org/resources/pubs/rli/archive/rli266.shtml>.