



Principles to Guide Vendor/Publisher Relations in Large-Scale Digitization Projects of Special Collections Materials

June 2010

Principle 1: Distinct collections demand extra vigilance in digitization.

When digitizing distinct collections, special attention should be paid to the nature of the material being digitized. No blanket digitization standard should be applied to all materials. Instead, the inherent characteristics of the item should determine whether and how distinct collections are digitized. At a minimum, careful consideration should be given to copyright, privacy (including the presence of personal identifying information), and moral and cultural heritage concerns. Rare, unique, or fragile items should be digitized according to the highest professional standards in terms of handling, security, and scan quality. This may often require on-site conversion with specialized equipment or scanning by contractors selected by the library. All of these concerns should be communicated to the vendor.

Digitization should be conducted in a way such that it is not necessary to revisit the process in the future as repeated digitization may lessen the artifactual value of originals.¹ Material that is more common or does not contain significant artifactual integrity can be digitized in a fashion that fosters widespread access to the most amount of material.

Principle 2: Libraries must respect any donor-imposed restrictions on the digitization and use of materials.

Special collections material is often acquired from donors with express limitations on its use, even when that material is ostensibly in the public domain. In negotiating with commercial vendors, libraries must ensure that the terms of any applicable donor agreement are respected. In negotiating with donors, librarians and archivists should educate them about the desirability of making materials accessible online.

Principle 3: Libraries should seek the broadest possible user access to digitized content. This includes patrons of other libraries and unaffiliated researchers.

For over a century, libraries have participated in successful resource sharing cooperatives that have made content widely accessible. The same spirit should govern commercial digitization activities. Libraries should resist arrangements that result in onerous subscription charges for access to resources digitized from their collections. In the best of all possible worlds, there would be some level of free access to all content, with only special value-added services restricted to a subscription model. Libraries should not be restricted from adding or creating a cataloging record in any union catalog, including WorldCat, to facilitate user access to the digitized content and thus avoid duplicative effort and expense for other institutions.

¹ Task Force on the Artifact in Library Collections. *The Evidence in Hand: Report of the Task Force on the Artifact in Library Collections*. Washington DC: Council on Library and Information Resources, 2001.

Principle 4: Libraries should receive copies of all digital files generated from their collections, with the option for complete local access to the files (to the extent that copyright law allows).

Libraries should insist on the right to provide free local access to digitized materials from their holdings. They should determine on their own what constitutes a fair use of those digital files and make them available accordingly. Nothing in the contract with the commercial entity should limit the library's right to make a fair use determination. Material that is of uncertain copyright status should be excluded from commercial products. Additionally, libraries should ensure that the files they receive from the vendor meet standards necessary for deposit in an institutional or other appropriate repository.

Principle 5: Any enhancements or improvements to the digitized content including all metadata should be shared on a regular basis with the supplying library.

In addition to making material available to the public, research libraries should seek to provide context to aid in the understanding of that material. This is especially true with special collections materials, which often must be interpreted or analyzed. In order to preserve and provide context for digitized distinct collections, it is important that the contributing library receive on a regular basis copies of enhanced content and metadata about that content. This could include upgraded or replaced image files as well as corrected or improved OCR text.

Principle 6: Restrictions on external access to copies of works digitized from a library's holding should be of limited duration.

In order to allow a commercial partner time to recover its investment in digitization, it may be necessary to grant to that entity exclusivity over the specified commercial uses of the digital files for a period of time. The ultimate goal, however, is to "ensure the results are widely available for scholarship."² The period of exclusivity, with the exception of local copies and access at the library, therefore, should be limited, ideally to as short a time as possible. After that time period the library should be able to distribute freely any file digitized from its holdings. It should also be able to aggregate the content with other resources from its own collections and those of other institutions and to expose the content to data mining and other new ways of exploiting it.

Principle 7: Libraries should refrain from signing nondisclosure agreements (NDAs) as part of digitization negotiations.

The ARL Board has supported a resolution from the Scholarly Communications Steering Committee to "strongly encourage ARL member libraries to refrain from signing agreements with publishers or vendors, either individually or through consortia, that included non-disclosure or confidentiality clauses." The values of transparency and public disclosure that underlie state open records laws should guide library transactions whether their home institutions are public or private. Libraries should respect that commercial partners may need to protect certain business and technological secrets, but not agree to keep licenses or core financial arrangements confidential. Libraries must "insist on their own right to discuss aspects pertaining to their broader community."³

² Merrilee Proffitt and Jennifer Schaffner. *The Impact of Digitizing Special Collections on Teaching and Scholarship: Reflections on a Symposium about Digitization and the Humanities*. Dublin, Ohio: OCLC Online Computer Library Center, 2008.

<http://www.oclc.org/programs/2008-04.pdf>

Digitization in Special Collections Bibliography. Chicago, IL, RBMS, American Library Association.

http://www.rbms.info/committees/task_force/digitization/discbiblio.shtml

³ Peter B. Kaufman and Jeff Ubois. "Good Terms - Improving Commercial-Noncommercial Partnerships for Mass Digitization: A Report Prepared by Intelligent Television for RLG Programs, OCLC Programs and Research." *D-Lib Magazine* 13, no. 11/12 (November 2007). <http://www.dlib.org/dlib/november07/kaufman/11kaufman.html>

Principle 8: Libraries should ensure that the confidentiality of users is protected in the vendor's products.

The confidentiality of usage data is one of the guiding principles of the library profession. In almost every state, library usage data is also protected by law. If a library digitized and made accessible to its patrons resources from its holdings, it should hold in confidence any personally identifiable information associated with the use of that material. Certain technologies offer opportunities for personalization of digital products. Confidentially held user data maybe used for such limited purposes, but only as the user may opt in to such use. User data should not be released to third parties or used for anything other than the functioning of the content as suggested above. The same principle should apply to material digitized by a commercial entity working in partnership with the library. Commercial systems must offer library patrons the option of reading and working anonymously.

Principle 9: Libraries are encouraged to refrain from charging fees or royalties for access to or non-commercial use of public domain materials held in their collections.

The combination of digitization technologies and Internet distribution can radically transform how researchers make use of special collections materials. As the Budapest Open Access Initiative has noted, "removing access barriers... will accelerate research, enrich education, share the learning of the rich with the poor and the poor with the rich, make this literature as useful as it can be, and lay the foundation for uniting humanity in a common intellectual conversation and quest for knowledge."⁴ The most creative uses of our shared cultural heritage can only occur, however, if the public has the ability to access and use public domain source materials without onerous permissions processes or the imposition of fees. Therefore, in the spirit of the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, libraries are encouraged to grant all non-commercial users "a free, irrevocable, worldwide, right of access to, and a license to copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship."⁵ If fees are to be assessed for the use of digitized public domain works, those fees should only apply to commercial uses.⁶

Endorsed by the Association of Research Libraries Board of Directors on July 26, 2010.

An earlier version of these principles can be found online at: <http://www.arl.org/bm~doc/rli-267-kenney.pdf>.

⁴ Budapest Open Access Initiative, February 2002. <http://www.soros.org/openaccess/read.shtml>

⁵ Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, October, 2003. <http://oa.mpg.de/openaccess-berlin/berlindeclaration.html>

⁶ Peter Hirtle with Tricia Donovan, "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>