



Statement on Digital Access Partnerships

CoSA Task Force on Online Content Providers (OCPs)

April 2007

Overview

Public/private partnerships are a growing option for making information available online. Such partnerships have the potential to broaden access to important documents, ensure long-term access to electronic information, create revenue-sharing opportunities, maximize existing resources for digital content, and serve as an excellent public relations piece for both partners.

To achieve these benefits, both the public and private partners will need to be flexible in order to meet the diverse needs of the private nonprofit or for-profit organization providing the technology and the institutions that hold public records. The Council of State Archivists (CoSA) has authorized a task force to consider the opportunities involved in such ventures and serve as a resource for institutions investigating such partnerships with the Genealogical Society of Utah/Church of Jesus Christ of Latter-day Saints (GSU/LDS), Ancestry.com, iArchives, and others.

Statement of Intent

The members of the OCP Task Force believe that private/public partnerships with online content providers present tremendous opportunities for state archives and for the content providers. Such partnerships also present challenges, which include: ensuring long-term access to and preservation of the records for state archives users, keeping public records public, protecting the rights of citizens, ensuring the soundness of tax dollar investment and providing a transparent view of the government process. Each state and local government will have different options to consider when discussing partnerships with private vendors. This Task Force intends to outline these contractual considerations and provide possible solutions or examples as developed by individual states and the vendor.

Contractual Considerations

Before signing any contractual agreement, a vendor should provide a proof of concept that their processes, product, and market research findings are feasible and support the viability of the project. In lieu of this, or in addition, a repository should explore the possibility of participating in a test bed or trial of the project before making a long-term agreement.

Digitization partnerships raise legal as well as technological concerns. It is not entirely safe to rely solely on another state's model, as laws about privacy and confidentiality vary greatly from state to state. In addition, laws change so any agreement with an OCP should take into account both current laws and emerging concerns regarding privacy. The Task Force strongly recommends close review of any agreement by counsel.

Each institution entering into a digitization partnership should consider the legal implications of a formal contract (a lengthy process with more assurances) versus a memorandum of understanding (shorter process with fewer assurances) with an OCP and also be aware of overarching laws, both state and federal, that govern such relationships. In addition, the following specific points should be considered in any contract with an online content provider:

Imaging guidelines

The contract should include clear guidelines or standards that address image quality, delivery, etc. Among the items to be addressed are:

- Equipment – the type of equipment used should conform to best practices for handling specific types of materials. For example, ensure that all equipment is large enough to handle oversized documents, and that the appropriate scanners are used for books and fragile materials. Document feeders are not recommended for imaging archival materials regardless of the age and condition of the originals. Many best practices documents are available that detail this topic.
- Resolution – Most archival imaging is performed at a minimum resolution of 300 ppi.¹ Digitization contracts often include a second image at 96 dpi for service on the web.
- Format – Most archives request images in TIFF, JPEG 2000 or JPEG format. Many request two copies, the TIFF for storage and the JPEG for delivery over the web.
- Image type – Will the images be color, black and white, or grayscale? Some images, such as maps, may be useless if they are not in color; others may be acceptable as black and white or grayscale.
- Alterations – Many archives prohibit alterations to the image (cropping, image enhancement, etc.), especially to the TIFF.
- Establish policies and procedures for dealing with the originals appropriately. For example, come to agreement about the scanning of blank pages or microfilm targets, reordering documents, rescanning errors and ensuring proper handling techniques by scanners. Generally, original archival documents should never be destroyed following imaging; make sure this is clearly articulated in any agreement.
- Multi-page documents – How will multi-page documents be handled? These can be scanned as separate images or as multi-page images. If separate images are produced, will they be linked in some way by metadata?

¹ ppi = pixels per inch, dpi=dots per inch. Dpi is a printing resolution

Indexing Guidelines

- Metadata and indexing – What data will be added by the vendor? Will the archives be involved in determining what information is available online? Will the archives receive some or all of this data along with their copies of the images? What use will the archives be permitted to make of this data? The value that most OCPs add to the archive’s images lies in this area, so this part of the contract should be considered carefully.
- Search ability – Will the archives have a say in how the data can be searched or reported online?
- Indexing – Outsourcing the indexing of scanned records is often a cost-effective way to complete this work. The scans themselves can be indexed remotely without access to the original records. But indexing historical records poses a variety of issues: Does the archives want the images indexed from the scanned copy rather than the original? Will the people who are doing the indexing understand the language in which the documents are written? Will they understand local usage and family names adequately to interpret questionable handwriting? Who will control the indexing and to what extent? Will the archives have the opportunity to suggest the components to be indexed? How will corrections be made, and for how long after the project is completed? These considerations are important because “inexpensive” indexing, if done poorly, can create long-term costs.
- File naming – create a clear, concise and consistent method for naming files that will assist staff in identifying content without having to open each file and will help make connections between separate images belonging to specific collections.

Copyright/exclusive use

Although most governments do not copyright their materials, many online content providers request exclusivity in the use of their images. What this means in practice is that the OCP vendor produces images which may not be given or sold by the archives to other vendors. Some OCPs may request exclusivity even in the delivery of images so that the public must visit their site (often for a fee) to view the images.

Continuing public access is vital to public documents, so this question should be addressed carefully in the contract. There are several ways to address the issue:

- Insist on the archives’ right to distribute the images in accordance with its own policies.
- Place a limit on the time the vendor can hold exclusive rights to the images. The National Archives’ recent contract with an OCP included a five-year provision of exclusivity.
- Provide for free access to the images for a subset of users.
- Provide access to a version of the images; e.g. low resolution images online and redirect to the vendor or archives for high quality images depending on revenue sharing agreement.

Additional items to be addressed include:

- Ownership of the images – The images may be owned by the archives, by the OCP, or jointly.
- Use of the images – What can the archives do with the images it receives from the OCP? Among the things to consider: Can the archives provide access to the images on its own web site? Can the archives give or sell copies of the images to users? To other vendors?
- Prorated rights, e.g. exclusive right for a certain number of years and then non-exclusive.
- Assurance of non-proprietary image or metadata format.

The partners also need to specify what happens to these rights if the vendor goes out of business or sells their business. Even if a vendor secures exclusive rights to the images, that right should revert to the archives if the vendor goes out of business or sells the business. The rights to the physical property of the master image files, metadata, or metadata databases, and original materials created or housed by the vendor, should revert to the archives/institution.

Fees

While public access is of overwhelming importance to public institutions, a recent CoSA survey suggests that many state archives rely on research revenue to maintain operating budgets. Partnering with fee-based services, or those who reserve the right to charge fees, presents an interesting challenge to provide revenue for both parties.

- The first question to ask is whether fee sharing (or charging a fee at all) is compatible with state law.

Some options to consider include:

- Revenue sharing plan – In which the archives and the OCP share the revenues from fee-based services. The contract should address what percentage will be retained by the archives; who will deliver the images and the copies?
- Upfront royalty – In which the archives receives a one-time payment at the beginning of the project for use of the images. This option could be combined with revenue sharing.
- Access to data only – Redirect for copy of images.

Work site/work flow

Where the scanning work is performed – and by whom – is a very important consideration in any contract with an OCP. The archives should consider issues such as the following:

- Is use of offshore vendor indexing services in accordance with state law?
- Offsite vendor work – Who will cover insurance costs? Materials transportation costs?

- What type of equipment will be used?
- Does the equipment meet the state's imaging standards?
- Security – Can the security of records be ensured if they are moved off-site for scanning? Unless the archives has the means to assign staff to the scanning location for the duration of the project, removing the records from the archives reduces the archive's control over the materials. How will the materials be safeguarded against mishandling and even alteration while they are out of the archive's control? Records that are duplicated (such as duplicate microfilm, for instance) pose far fewer problems in this regard.
- On-site considerations – If the scanning work is to be done on-site at the archives, where will space be provided? Who will be responsible for providing scanning equipment? For hiring, training, and supervising the scanners? Who will oversee proper handling, and what recourse will the archives have if they feel the materials are being jeopardized?
- Quality control – How will the images and metadata be quality controlled? Will the archives have the ability to correct errors after the project is complete? Will the vendor ensure that such corrections are made within a reasonable amount of time, even long after the project is complete?

Delivery and storage of digital materials

Any archives or institution contemplating a digitization agreement should look at provisions for the long-term storage of digital materials. While it is appealing to get free scanning, institutions need to have migration plans in place to handle large volumes of images. Some things to consider:

- Volume – It is a good idea for the archives to get some estimate from the OCP of the volume of data they can expect to receive back, even though this may not be included in the contract itself. Will the archives have the capacity to store and service the volume of images?
- Delivery method – The contract should include how the images will be delivered by the OCP to the archives. Some options include magnetic tape, CDs, DVDs, or FTP. How will the OCP track shipments to ensure that they arrive at the archives securely? How long will the archives have to determine that the media can be read?
- Long-term retention – How will the archives maintain the images over a long period of time? What is the archives' migration plan? Although these probably will not be included in the contract, the archives should consider them in advance. The contract should include the OCPs written migration plan for maintaining their online content over long periods of time.
- Schedule of deliverables – The contract should include an itemized schedule of what the vendor will deliver including the images, metadata and original records.
- Digital materials should be delivered in non-proprietary format and on media that is accessible for the institution.

CoSA Task Force on Online Content Providers

This statement was developed by the Task Force on Online Content Development (OCP) of the Council of State Archivists. The Task Force is chaired by Mark Harvey (MI). Members of the Task Force who participated in the development of this document include Prudence Backman (NY), Ed Bridges (AL), David Carmicheal (GA), Conley Edwards (VA), Peter Gottlieb (WI), Gerald Handfield (WA), Robert Horton (MN), Joseph Klett (NJ), Dick Lankford (NC), Russell McCabe (DE), Patricia Michaelis (KS), Karl Niederer (NJ), Roy Tryon (SC), and Christine Ward (NY).

Council of State Archivists

308 East Burlington Street #189

Iowa City IA 52240

Telephone: 319-338-0248

<http://www.statearchivists.org>