Digital Library Collaboration

The Why and the How

By

Traci L. Keith

INFO 653-900, Prof. Xia Lin

Drexel University

I certify that:

• This paper/project/exam is entirely my own work.
• I have not quoted the words of any other person from a printed source or a website without indicating what has been quoted and providing an appropriate citation.
• I have not submitted this paper/project to satisfy the requirements of any other course.

Signature: Traci L. Keith
Date: 4 September 2012
Abstract

Collaboration between libraries has been a part of what libraries have been throughout history, digital libraries will be the next logical step if libraries are to remain relevant. The reasons behind why libraries choose to make the move into collaboration include cost benefits, reduction in duplication of online resources between facilities and increased user access. Libraries have to deal with creation of metadata standards, digitization standards, and staffing issues as they go about creating digital library systems that will work for their collaborative. These collaborations have the advantages that mirror the reasoning behind the move, those are cost reduction as well as added access by putting what patrons want in the place they want to find it. The disadvantages are led by issues with copyright, followed by logistical issues involving ensuring the library remains reliable through the changes that can and will occur in digital formats, as well as the time and expense that will be required to make the library work. All of it will be worth it in the end when the libraries which are part of the collaborative group remain reliable and relevant sources for their patrons.

Introduction

In recent years, as use of the Internet as increased, it has become commonplace for a search engine or other online resource to be the first thought when it comes to research of any kind. In order for libraries to remain relevant they have to follow this trend and create environments that will encourage searchers to make use of libraries instead of search engines. Whether the library is public or academic users expect to be able to access collections online, they expect the collections to be easily searchable and they expect the range of information to be extensive enough to satisfy their needs. If library access is not available online users are more likely to choose a less reliable Internet source.
Libraries have been collaborating in a variety of ways for a long time; in fact collaboration of one kind or another is a part of the foundation of what libraries are ("Conference proceedings documentation," 2001, p. 13). This traditional form of collaboration can be seen through interlibrary loan of full text items, as well as through the sharing of items through document sharing supported by such things as photocopying. The logical next step, therefore, would be collaboration in the online environment.

Collaborating can occur at a variety of levels; between university libraries, between departmental libraries within universities, on a state public library level, at national level with governmental or scientific libraries or even between libraries, museums and archival institutions. Regardless of the level at which the collaboration occurs at the idea behind it is the same; to reduce costs, to create access and to ensure the relevancy of libraries. This paper will cover the why, or the reasoning behind collaboration, the how, or what member libraries must do to create collaborative digital libraries, as well as the advantages and disadvantages collaborative digital libraries face.

The Why

The first and perhaps most important question that must be answered with regard to creating a collaborative digital library is, quite simply: Why? One answer to this is provided by Brancolini and Kowalczyk in their discussion of IN Harmony, Sheet Music From Indiana (2006), “Collaboration enables us to leverage the diversity of skill and knowledge that individuals have developed within the partner institutions” (p. 31). While this gives one important reason for the creation of collaborative digital libraries, there are other reasons beyond making use of the knowledge of all those individuals.
The first of these is to prevent duplication of resources. If each individual library were to create its own digital library there is a good chances there would be overlap between the collections. First of all this is a waste of money, most usually from budgets that are increasingly tight. It is also a waste of manpower and of digital storage resources. Wu, in Building a Collaborative Digital Collection: A Necessary Evolution in Libraries (2011) explained that “If each library group committed to preserving a portion of the world’s existing, printed knowledge in cooperation with one another, they could reduce duplication of effort and ensure an unbiased preservation of materials” (p. 545). While he is speaking specifically of printed materials, the concept is true for digitized and purely digital materials as well.

Following from, and connected to, this prevention of duplication the next reasoning for creation of digital libraries is cost savings. When only one copy of an item is being stored digitally there will be a significant reduction in costs. The collaboration ensures that the item in question is available to the library users who will need it, regardless of whether it was originally owned by their own or another library. This savings will initially be seen through a reduction in the amount of time actually spent digitizing, creating metadata and cataloging individual items. It will also be seen through the reduction in cost of storing items digitally. When more than one item is stored in two different places it is literally being paid for twice and taking up twice as much space.

The fact is, even in the best of times, budgets are limited, and libraries have to make the most of the funds available. Creating a full service digital library on just one library’s budget, according to Emerging Visions for Access in the Twenty-first Century Library, is prohibitive ("Conference proceedings documentation," 2001, p. 51). In addition to the costs of the actual creation and storage of digital items, there are other costs associated with digital libraries. These
include the creation and maintenance of the digital library system, staffing (which includes training of new or cross-training of current staff members) as well as the cost of subscription services. Subscription services are something most academic and many public libraries already invest in, however as Wu (2011) points out, dependence on these services as a main portion of a digital library leaves libraries at the whim of the subscription service (p. 533).

It can be difficult to quantifying the exact amount of money saved through collaboration. The main reason for this is that costs can be unpredictable from year to year. The most reliable way to calculate how much an individual library is spending each year is to average costs over several years. Wu (2011) described that member libraries in the TALLO (Taking Academic Law Libraries Online) consortium could calculate the cost of membership based on the concept that “...if each school provided $50,000 annually to the project, the operation would have a $10,000,000 annual operating budget. For that investment, each library would gain access to many times more titles than it could afford to purchase itself.” This annual contribution would provide each member with “...co-ownership of materials purchased that year” (p. 545), as well as providing those who access a much wider variety of materials than the individual library could provide. When compared with spending $50,000 as an individual library budget, each in the consortium can see how much more they will get for their money.

From the fact that a library can get more for its money when in collaboration with other follows directly the next reason for the creation of collaborative digital libraries: an increase of access to information. The fundamental core mission of every library throughout history has been to provide access to knowledge and information. Digital libraries take this core mission to the next level. Users want access to information from everywhere, digital libraries must provide what they are looking for, and in the environment they want to be a part of. The fact is that for
the average user today whether the information is reliable or not isn’t as important as how easily accessible it is. Even if a library has created its own digital collection, the fact of the matter is that it is not possible for any one library to have a collection that will cover every research need that will be asked of it. On the other hand, collectively a group of libraries has a better chance of being able to provide what it is users are looking for ("Conference proceedings documentation," 2001, p. 532).

This desire for access stems from several fronts. The first is basically the explosion of information that is available online. The second, and likely most important to university libraries that are considering entering into collaborative agreements, is the increase in the number of online and distance education students. These types of students want to be able to access the university library from wherever they are, be it at a great distance or near the actual institution. For this reason university libraries have to create digital libraries that include full text, freely available digital books (Hulton, 2008, p. 499), and doing this becomes more realistic when done in a collaborative fashion.

The last reason for libraries to take part in collaborations, that of relevance, could be considered one of the most important. As stated previously, users today gravitate towards the Internet to find answers. They will choose an online search over the use of a traditional library nearly every time. If libraries do not recognize this trend they will find themselves completely irrelevant. If libraries embrace the digital environment but fail to have comprehensive coverage of subject matter, they will find themselves irrelevant. Above all libraries must consider that “what is important is to be able to find the books, search across them, and integrate them with other information” (Levy & Turner, 2010, p. 32).
The How

When organizations are considering creating a digital library collaborative they need to take into account staffing issues and logistical issues. To begin with they have to ensure that the staff who will be responsible for the set-up and maintenance of the library itself have all of the required skills. These need to include specialists in development, in digitization, in indexing and organization, in cataloguing and in metadata creation and input. Additionally they would need staff members whose entire job would be to check whether or not items that are being considered for inclusion in the collection are already available at other digital repositories. When they find items that are already available their next job would be to work out an agreement with these repositories so that the item in question can be included in the library (Wu, 2011, p. 534). In addition to these considerations for what type of technical specialties are required organizations also have to take into account the personalities that are to be involved. An environment has to be created which will encourage those involved to exhibit a willingness for “trust, commitment, patience, letting go of egos, and open communication” if the collaborative effort is to succeed (Double, 2009).

Logistically the first step organizations need to take is to perform a SWOT analysis. Once a full understanding of all the strengths, weaknesses, opportunities and threats that exist with respect to the collaborative effort has been reached the group can move forward. This analysis will ensure that every member of the group goes into the collaboration comprehending the anticipated affect it will have on their individual library (Double, 2009, p. 129). After this initial stage is completed the actual library system which will be used in the collaboration must be created.
One of the most important aspects to an effective digital library is a consistently used metadata standard. If one standard is not specified from the initiation of the library it will become increasingly difficult to create a controlled vocabulary or to properly catalogue the collection. Without a controlled vocabulary and proper cataloguing users will not be able to make full use of the collection. When migrating items that already have metadata assigned to them the system is going to need to make use of metadata harvesting programs in order to transfer previously used metadata standards to the new standards being implemented (Brancolini & Kowalczyk, 2006, p. 29). Additionally, there must be standards set for the dimensions and pixel level of any images as well as standardization for the format documents should be created under (for example .pdf format rather than .doc). Standards should also be put into place for other types of items the library might contain, such as video or audio files. The creation of these standards will make the system more operator friendly from both the library side and the user side. This will, in turn, ensure that the system is actually used. People will not make use of a system that is overly complicated or difficult to understand. A last requirement should be that individual libraries are able to add to the collaborative from their own location, rather than needing to send items to a centralized location for digitization. Having to remove the original items from the collection of the library that owns it could be prohibitive to adding the item to the collection.

One example of a digital library system is CoreDev, a digital library system which is in use in Malaysia in the national collaborative library which was created for in use in their secondary school system. This system was created to allow group members to contribute by creating original digital items, digitizing existing monographs and providing links to sites on the Internet which will prove relevant to other group members. As suggested above, CoreDev
allows group members to input objects themselves, a practice which “enhances the sharing &
discovery of information between various user groups” (Edzan & Abdullah, 2004, p. 17).

A second is within the California Digital Library (CDL). This system, again, allows
member libraries to contribute directly from their own library while keeping the original on file
in the library where it is owned. They have created standards for metadata and specifications for
how images are to be digitized. Specifically they have chosen to make use of the standards on
the digitization of images as created by the National Archives and Records Administration
(NARA) in order to ensure the quality of images is maintained across all items (Levy & Turner,
2010, p. 13-14).

The Advantages

Many of the reasons libraries choose to collaborate can also be listed as the advantages of
collaboration. A reduction in the cost of the development, creation and maintenance of a digital
library is the advantage that will mean the most to the accounting department as well as
managers in charge of budgeting and is likely to assist in ensure their backing when a
collaboration proposal is initially presented. A loss of duplication will be a benefit that will
become increasingly valuable as more and more digital libraries come online. At this point many
institutions are creating digital collections as part of their library and the creation of full digital
libraries is still a relatively new concept. This is in fact resulting in a duplication of resources.
As institutions join forces they will do away with the duplication and make better use of the
digital storage space that is available.

The advantages provided by digital collaborative libraries with regard to access are
manifold. As stated above individuals doing research want to be able to access information from
anywhere at any time. On time access is of particular importance in today’s library. In the past
patrons were willing to wait days, if not weeks, to receive materials that were being borrowed from another library through interlibrary loan. This is no longer the case, users expect to instantly be able to find and access information sources which will support their research.

One additional aspects of collaboration which was mentioned by Edzan and Abdullah (2004) is that it allows local content to become available to a larger audience than they would if only made available through a localized digital library. When this local content is placed on a collaborative digital library it is pushed out to a larger audience (p. 19). What this really means for users is that they are going to be offered a richer variety of items, they will receive the expected, popular items and then will also be able to mine deeper into lesser known titles and information.

When standards are created for one collaborative program the effects are seen throughout the library community. This is a long term advantage that will continue to be felt for many years. When other individual or collaborative libraries are looking at creating their own digital library they will, as a rule, look at existing metadata and digitization standards. They will investigate to find what aspects of these standards works and what does not work and will then use this to aid in the creation of their own standards. The result over the long term will be the creation of increasingly effective standards that will spread throughout the digital library environment.

A last advantage, which actually connects with one of the disadvantages, is the fact that with digital items there is the potential for more than one person to be able to view an item at any given moment. With traditional monographic items each item can only be checked out or made use of by one person at any given moment. The very nature of digitization allows items to be
viewed by more than one person at a time. This however leads to questions of copyright which must be dealt with by every digital library today, collaborative or not.

**The Disadvantages**

The disadvantages for collaborative digital libraries fall under two main headings: copyright and logistics. An entire book could be written on the subject of copyright and digital libraries, this paper will touch only on the main points of concern with acknowledgement that the issue is much deeper. Copyright is, without a doubt, one of the most pressing issues facing the creation of digital libraries today, especially the collaborative type of library. As Wu (2011) states so well “Copyright has been and continues to be the greatest hurdle to transforming library collections” (p. 536).

As indicated above, the issue of copyright starts with the issue of access. With traditional monographs the library purchases an item and lends it out for set periods of time, staying well within current copyright laws. Any photocopying of books is kept to a minimum, due in part to the prohibitive cost involved in copying an entire book, and so there is no worry from publishers or authors that their items are being used outside of the laws of copyright. When an item is digitized it creates a completely different issue when it comes to copying since each time the book is opened on a computer a copy is created. This has led to questions regarding whether it is legal, under copyright law, to loan a digitized book that was purchased in monograph form. Answers to questions like this are going to have to be made over time, and in order to get to them “...libraries must be willing to test the assumptions behind copyright protection” (Wu, 2011, p. 536).

One example of how copyright issue is being dealt with is discussed by Brancolini and Kowalczyk in their article “IN Harmony: Sheet Music From Indiana” (2006). They discuss the
fact that many sheet music repositories simple do not make available any music that was drafted after 1923 because items after that were still in copyright (at the time of the writing of the article). IN harmony however is intent on testing a theory, one which states that 90% of the music created between 1924 and 1964 is not in copyright because the original license at that time had to be renewed and the creator of the music did not do so. This is an excellent example of libraries pushing the boundaries of the assumptions behind current copyright laws. It is requiring the library to keep employees on staff whose job it is to check the copyright status of any piece of music which falls within that possibility window, but it is also allowing them to make available a much wider spectrum of music (2006, p. 30).

The first logistical disadvantage to digital libraries is the issue of reliability. It has become apparent over the last 15 to 20 years that digital formats have the propensity to change rapidly. Each time such a change takes place there is the potential of losing information and resources that are currently stored in the last digital format. When going into collaboration libraries must take this into account and upgrades of this type have to be allowed for in budgets and planning or the collaborative library could find itself seen as irrelevant. It will be both costly and time consuming to make the transition from one format to another and for this reason some will just want to avoid it. This head in the sand approach will result in the end of the library at the next evolution in digitization. Instead collaborative members must write into the original proposal how new formats will be dealt with, including funding and manpower estimates should a large scale changeover of formats be necessary.

A second logistical disadvantage that must be dealt with from the beginning is the expense that will be required in the original creation and startup of the library. Even if all the libraries involved already have an extensive repository of digitized and digital items it will take
time, it will take effort and it will cost money. The system must be created, or imported if the
group is fortunate enough to find a system that fits their needs. Personnel must be hired and
trained on how to use the system as well as current personnel being brought up to speed on this
new system. Metadata and digitization standards must be created, as discussed earlier. Lastly
the actual digitization and metadata transitions must take place. All of this before the library
actually becomes available to users.

The metadata standard creation is the third logistical disadvantage of the digital library
collaboration. If the group libraries are not already using the same standard it creates the
possibility of two problems. The first is that the standard has to be created which is time and
budget intensive. It is very likely that there will be a “...necessity for building new indexing
architecture both at the engineering and the discovery levels. [As well as] the necessity for
mining and mapping data to build linkages that are interactive...” (Brantley, 2008, p. 38). The
second is that each library has to be convinced to let go of the standard they have been using and
embrace one that works for all the libraries in the collaborative. Once everyone buys in each
library has to be trained on how to properly use the new standard.

This last area with regard to metadata standards touches on the fourth logistical
disadvantage in digital library collaboration: the people and cultures within each individual
library. As a rule no two cultures are going to be exactly the same. Practices are also likely to
differ from library to library. Each library is going to have to be able to set aside their own
culture and practices in order to be effective contributors to the collaboration. They will need to
embrace the idea that all of the libraries in the group are working toward a common goal
(Brancolini & Kowalczyk, 2006, p. 29). If even one library in the group does not accept these
goals, if they place their own individual goals over those of the collaborative, it could undermine
the entire project.

While it may appear there are many disadvantage to overcome, the advantages and the
reasons libraries will become a part of a digital library collaboration largely outweigh them. The
fact of the matter is “...the solutions to these problems will come about through the ability of
libraries to collaborate among themselves and also with IT communities, publishers, search
ingines, for-profit content providers and others” (Brantley, 2008, p. 38). If libraries go into
collaborations knowing what to expect, understanding what is to be faced and with solutions in
mind, the disadvantages become merely speed bumps.

Conclusion

In the end the reality libraries have to face is that the Internet is where their users are
going for information today. In order to remain relevant libraries have to stay on top of this and
create gateways that put the information needed by users where they want to look for it. The
days of patrons being willing to come into a library and sit for hours sifting through microfiche
or monographs have passed. Creating individual digital libraries is an expensive time consuming
task that will result in a great deal of overlap of items and information between libraries.
Collaboration will reduce the risk of duplication, it will reduce the costs involved for each
individual library and it will build collections which are deeper and exists out where the users
are. In the end “Libraries are successful to the extent that they can bridge communities and can
leverage the diversity of the quest, the research, and the discovery. Libraries are successful when
they offer new services and when they help others discover services provided by others”
(Brantley, 2008, p. 38).
References


