University of Iowa

From the SelectedWorks of Mark F Anderson

January 2010

Scaling Back for an "Experimental" Collection

Contact Author

Start Your Own SelectedWorks

Notify Me of New Work

Available at: http://works.bepress.com/mark_anderson/4

Author Information

Mark F. Anderson is a Digital Initiatives Librarian at The University of Iowa Libraries. He has been on the staff of Digital Library Services since 2005, and is responsible for planning and coordinating the scanning and uploading of materials to the Iowa Digital Library and user accessibility for digital collections. He received his MLIS from The University of Iowa, and began his work in digital initiatives at St. Ambrose University.

Mark F. Anderson
Digital Initiatives Librarian
1015 Main Library
Iowa City, IA 52242
319-335-5685
mark-anderson@uiowa.edu

Abstract

Digital Library Services (DLS) at the University of Iowa Libraries has progressively worked toward coordinating more large-scale, "left-to-right" digitization projects both within the libraries and across campus, moving away from model of web exhibits that were often created before the department was formed in 2005. However, a variety of situations still call for small-scale projects. This chapter, describing the design and production of the "W9XK Experimental Television Digital Collection", attempts to show that small-scale digitization projects can bridge that gap, and yield collections that rise above the level of web exhibits in their usefulness to scholars and the general public by limiting exclusive selection and promoting comprehensiveness. While mirroring this approach of mass-digitization, digital librarians can also use curatorial decisions and software functionality to further assist users of these small-scale collections.

Keywords

Digital collections; Digital libraries; Digitization; Archives; Web exhibits

Scaling Back for an "Experimental" Collection

Introduction

Henry Wadsworth Longfellow wrote, "Most people would succeed in small things if they were not troubled with great ambitions." Even if the poet was not referring specifically to digital initiatives in libraries, it's an appropriate sentiment at a time when so much institutional effort is directed toward mass digitization projects such as the agreement between Google and the Committee on Institutional Cooperation (the consortium to which The University of Iowa belongs) to digitize no less than 10 million volumes from among members' collections (Committee on Institutional Cooperation, n.d.). At the same time, it has become common for institutions, or units within institutions, to work on building small, narrowly-focused collections as an initial foray into digitization, while securing funding for the necessary resources to ramp up to large-scale scanning projects and mass digitization, or to expose exceptional materials. This chapter will discuss one of these small digital collections, but will begin with the development of the digital library at the University of Iowa to which it belongs.

Digital Initiatives at the University of Iowa

The earliest digital collections hosted by The University of Iowa Libraries date back to the mid-1990s, and were a combination of small web exhibits and larger efforts to digitize materials from flagship collections. The web exhibits were considered "virtual versions" of exhibits displayed in the library's exhibit hall and which combined highlight images with essays and other supporting text provided by library staff and campus scholars (University Libraries. University of Iowa, 1999). In 1997, the Libraries were awarded Library of Congress/Ameritech National Digital Library Competition funds, which supported the digitization of nearly 8,000 talent brochures from the papers of the Redpath Lyceum Bureau, an agency that represented performers on the Midwest Chautauqua circuit. In many ways, this was the Libraries' first mass-digitization project (The Library of Congress, 2003). Other scholarly digital collections created at the time included the Center for Electronic Resources in African Studies and the International Dada Archive.

Figure 1. Keeping Our Word / North Lobby Exhibit. (© 1998, University of Iowa. Used with permission.).

In 2005, a new department, Digital Library Services (DLS), was formed to manage digital initiatives. Since that time, DLS has progressively worked toward coordinating more large-scale, "left-to-right" projects both inside the Libraries as well as with faculty and scholars on campus, moving away from the highly selective model of the early web exhibits, which tended to contextualize the digital materials for the user. Conversely, the model of comprehensive digitization aims to allow users the freedom to repurpose digitized materials for the creation of new knowledge and insight. This requires not only a thorough approach to digitization on the part of the library, but also the availability of tools to facilitate new uses for the items, especially tools that link, overlay, and share data. Unfortunately, many of these tools require extensive development and are not yet available as turn-key solutions.

However, situations such as the anniversary of an important institutional event or a specific user-driven request have still led libraries to undertake small- and medium-sized digitization projects. The challenge for libraries is to create smaller digital collections that are more valuable for scholarship than boutique, highly-curated exhibits (see Figure 1 above). Unlike hand-picked images and carefully-crafted text, users should be confident that small-scale digital collections provide the same research potential as the entire physical collections from which

they are drawn. Digital librarians need to get out of the way of the user, and merely expose high-resolution images, standardized metadata and full transcriptions of textual materials to as wide an audience as possible.

Within the University of Iowa Libraries, DLS has cultivated a particularly close relationship with three content providers: Special Collections, Iowa Women's Archives, and University Archives. With the first two units, digitization projects are becoming more large-scale and left-to-right in their approach, but less so with the latter. This is mostly due to that fact that University Archives subscribes to the "principle of provenance", the archival practice of organizing collections by creator rather than subject, and collections are often selected for digitization based on a topic of interest rather than contributing entity, making it difficult to consider left-to-right scanning. For physical exhibits of University Archives materials, this has long led to "artificial collections" that pull from many different record groups, according to University of Iowa Archivist David McCartney. More recent *digital* collections from the archives are no exception. (D. McCartney, personal communication, October 27, 2009). This kind of intermediation by the selector is a major difference between the small-scale "boutique" digitization model and mass-digitization.

In "Shifting Gears: Gearing Up to Get Into the Flow," the authors direct libraries to stop obsessing about items and making curatorial decisions about what to digitize since the selection has already been done (Erway, R., and Schaffner J., 2007). Arguing that if items were important enough to acquire originally, they are important enough to digitize, the authors advocate a more programmatic approach of digitizing items as they are accessioned, rather than thinking of digitization in terms of discrete projects. Is it possible to work within these provocative guidelines and still develop small-scale digital collections that are "in the flow"? Yes. For example, libraries can push the decision making process from the item level to the collection level. After all, with so much catching up to do, there's no way all existing collections can all be scanned at once. Small collections can be selected for digitization, and worked through left-to-right, even if that means scanning hand-written grocery lists on backs of dry-cleaning receipts (a real-world example from an Iowa Digital Library collection). In this way, DLS has operated with the understanding that a digital collection, even a small one, should be a comprehensive surrogate of a physical collection, regardless of size or scope, which attracts not only general interest, but promotes new scholarly activities as well. Also, for text items, the combination of brief metadata and optical character recognition (OCR) can provide access without the need of extensive, handcrafted records.

Selecting the Collection:

At a time when video is so easily accessible through websites Hulu, Netflix, Boxee.tv, and YouTube as well as thousands of cable and satellite television stations, it's hard to imagine the technical and organizational challenges facing engineers and technologists involved in the infancy of television. Were these pioneers of a century ago similarly considering the scale of their new experimental medium? If so, then the small digital collection described more in detail in this chapter is particularly appropriate: "W9XK Experimental Television at Iowa" (http://digital.lib.uiowa.edu/w9xk).

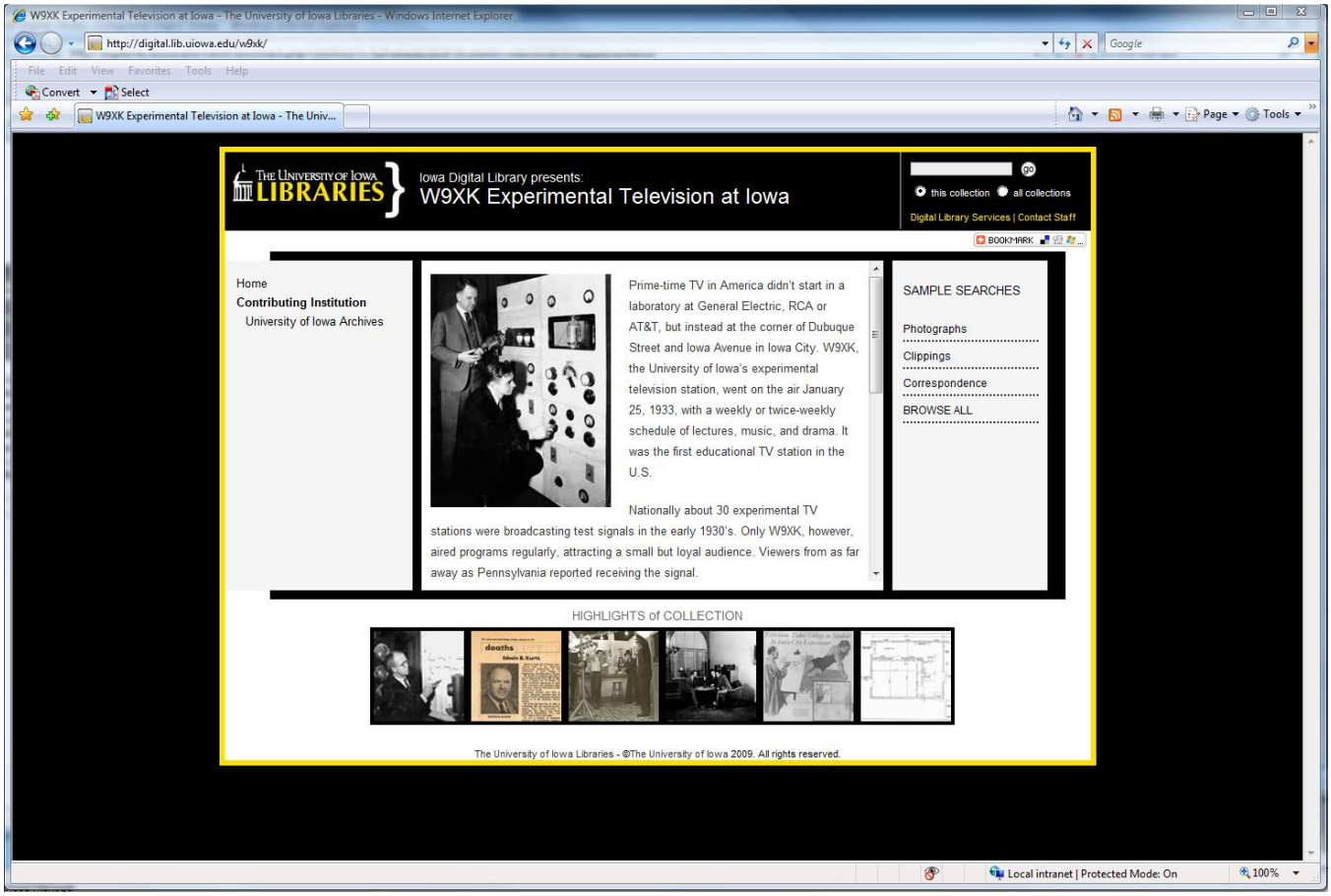

Figure 2. W9XK Digital Collection Homepage. (© 2008, University of Iowa. Used with permission.).

February 2009 was seen as an excellent opportunity to promote the history of the W9XK experimental television station on the University of Iowa campus by means of digitizing related archival materials, both because of the congressionally-mandated conversion to digital broadcast television, which was set to take place that month, and the 75th anniversary of the station the year before. W9XK, the first educational TV station in the U.S. went on the air January 25, 1933, with a weekly or twice-weekly schedule of lectures, music, and drama. Reception was reported back to the university from as far away as Oklahoma, Kentucky and New York (University Libraries. University of Iowa, 2009). Appropriate materials were gathered from throughout the archives including correspondence and newspaper clippings from faculty and electrical engineering department subject vertical files, still photographs from a prominent campus photographer, articles from a campus engineering journal, WPA federal writers project papers, and a student thesis.

With no single collection to process in a left-to-right manner, this project is a departure from DLS's more recent production model described earlier, and runs somewhat contrary to the challenges in "Shifting Up," but it is unlikely that these materials would be scanned together in any other way short of a heroic effort to mass-digitize the entire University Archives. The collection is truly small, containing just over 50 digital objects

(although some of the objects are composed of several items grouped together, their organization to be discussed below), but it tells an important story of the station's development and operation in the broader context of early advances in television technology, using the variety of formats listed above. Unfortunately, no video or film footage of W9XK's broadcasts could be included in the digital collection since none is known to exist today.

Building the Collection:

From 2005 to 2007, digital initiatives at the University of Iowa Libraries were completed mostly by staff in DLS in what is referred to as "in-house" production. Scaling up to support larger digitization efforts has required production tasks to occur in other library departments with both the staffing and expertise to process not only more materials, but also address unique handling and cataloging issues that invariably arise from digitizing entire collections. Metadata is now mostly applied by cataloging staff, except in cases where digital collections come to the library already packaged with existing records, which requires more in the way of reorganization and standardization rather than creating new records from scratch, which as a brand new collection, is the way metadata for the W9XK collection was handled.

In a similar way, the scanning that had been done by staff and students working directly under DLS has mostly moved to the Preservation and Reformatting department. However, scanning for the W9XK collection broke from this workflow model. A staff member in University Archives was interested in learning about scanning and image editing, especially the operation of a new top-down book scanner the Libraries had just purchased. The potentially small size of the digital collection presented a good opportunity for that. In this way, small-scale digital collections afford opportunities to deviate from established workflows for staff to develop new skills. But with other parts of production spread out, it may not make a considerable difference in time to do just part of the production "in-house". The time for an item in the W9XK collection to be completely processed and added to the digital collection was neither considerably faster nor slower than other collections.

The University Archivist selected and pulled appropriate materials from the physical collections listed above and prepared them for the staff member who scanned them on either an 11x17 flatbed scanner or the newly acquired top-down scanner, which accommodates items up to 24 x 36 with a book cradle for bound materials. Scans were then edited and display files derived by DLS staff who then uploaded the images to CONTENTdm, the digital library software that the Libraries use for many of its digital collections. Lastly, cataloging staff added descriptive, administrative and technical metadata to the objects. Metadata is almost always applied to the digital items following upload to the digital content management system in order to take advantage of software functionality that allows for easier batch cataloging processes, increasing efficiency.

Collection highlights and sample searches were selected for a collection homepage, and the university archivist provided an introductory paragraph, which takes the place of the more contextual and interpretative writing that used to accompany web exhibits directly alongside images. A lingering concern about small-scale digital collections is packaging them in beautiful, but deep, web portals that users must navigate to find the information they seek, rather than letting digital objects live in a large repository with thin logical collections built around them. A compromise is to maximize collection exposure to search engines and aggregators. The

Libraries use Google Sitemaps to help assure collections are crawled and ranked highly by Google. OAI harvesting of collections is enabled for outside service providers, and the Libraries also implemented a next-generation catalog that brings together records from the traditional catalog as well as online books and journals, local web pages and digital collections such as W9XK.

One recent change in how DLS delivers large-scale, left-to-right digital collections is to use folder-level organization of the digital objects. The digital library software used for most collections in the Iowa Digital Library, CONTENTdm, uses "compound objects" to deliver multi-page items, from the two sides of a postcard, pages of a document, a monograph, or an entire folder's worth of materials arranged hierarchically. This decreases the number of items for users to browse through up front, or consider in results sets. While the W9XK collection did not fit the folder-structure organization model since its items came from many different archival collections, one of the digital collection-delivery decisions made was to display correspondence items (letters between engineering faculty, university administrators, government authorities, and the public) so that letters and their replies were in the same compound objects, rather than as separate objects (see Figure 3 below). One could argue that this mediation is too much along the lines of a web exhibit, but it was thought to be helpful in navigating the correspondence. Feedback from users would help to support whether this is a good decision. In any case, digital librarians often have to make organizational decisions based on software functionality and with the intention of assisting users and researchers.

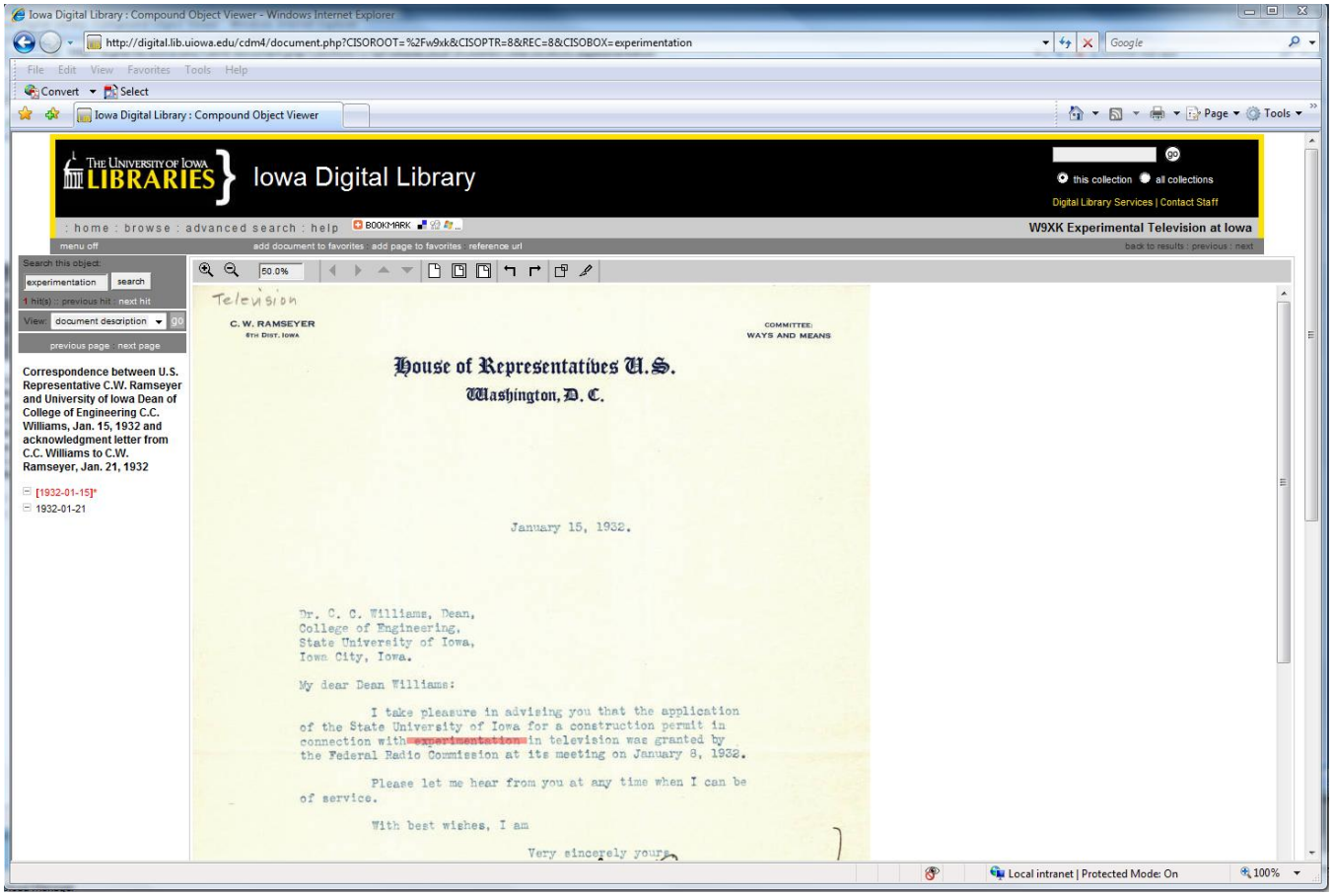

Figure 3. W9XK Correspondence. (© 2008, University of Iowa. Used with permission.).

Use of the Collection:

The W9XK Digital Collection was officially launched on February 17, 2009 through press releases and library blogs. While we must expect that our users will tend to find our digital collections through the back doors of search engines and aggregators rather than our finely-crafted web portals, only 10% of visits have come directly through Google, while more than half of all users are referred to the collection from promotion on university and library pages as well as the library catalog, although the search engine could have sent users to these local pages first. Over 75% of page views to date occurred during the first two months, owing much to the initial publicity (see Figure 4 below). While somewhat discouraging, most early use may have just been casual interest, and it may take longer for research and scholarly use to increase. Owing to Lally and Dunfords assessment of the impact of Wikipedia in directing users to library digital collections, links to the W9XK collection have recently been added to Wikipedia articles on experimental television and may have an effect on traffic over time. In "Using Wikipedia to Extend Digital Collections" (Lally, A. and Dunford, C., 2007), the authors note that Wikipedia is "a prime candidate for … efforts at pushing information about the Libraries out to where users conduct their research," as their server statistics indicate it is indeed driving a significant amount of traffic to their collections.

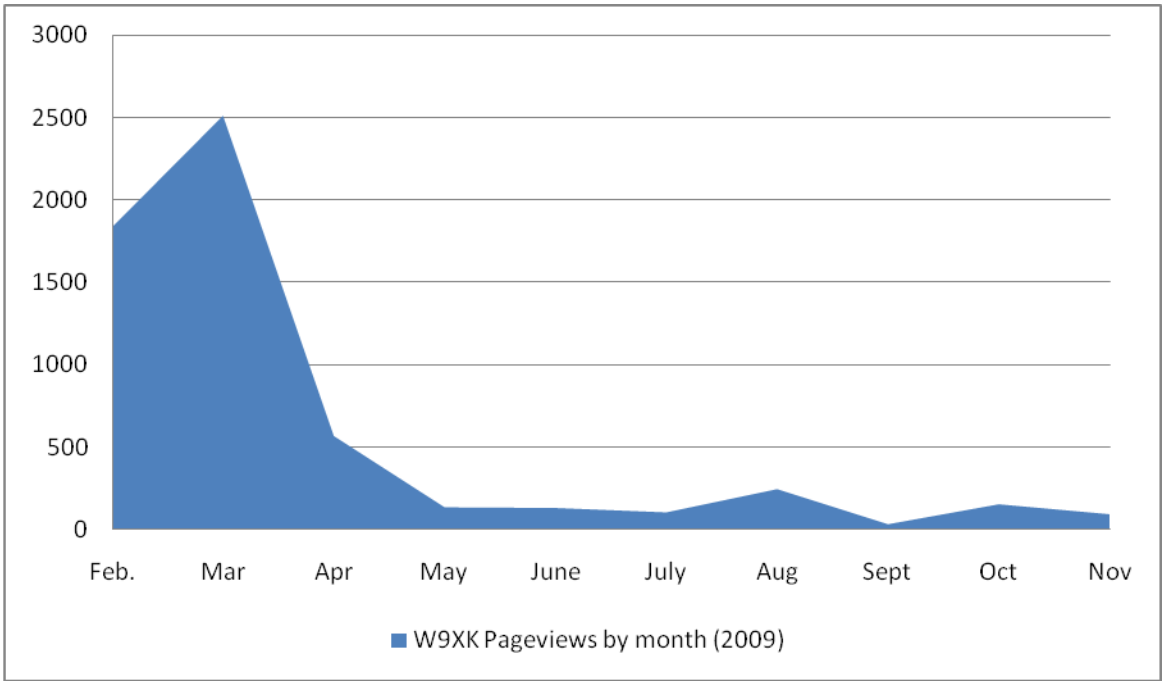

Figure 4. W9XK Pageviews by month (2009)

65% of visits to the W9XK collection home page have come from within the state of Iowa, a higher percentage of local visits than for two other small-scale digital collections in the Iowa Digital Library, which were prompted by high use and requests for reproduction from unique physical collections, especially where appeal went beyond the local area (see Figure 5 below). Two examples of these are the Victor Animatograph pamphlets (http://digital.lib.uiowa.edu/victor) and the Rand and Leopold Desk catalogs (http://digital.lib.uiowa.edu/desks), both with fewer than 30 items each. The Victor Animatograph Corporation of Davenport, Iowa made the world's first 16mm motion picture projector, and was a leader in the area of film technology throughout the 20th century. People in possession of one of these projectors would often contact the Libraries requesting copies of the product pamphlets, which contain detailed schematic diagrams and operating instructions. This collection was built before resources allowed for left-to-right digitization, so rather than digitize the entire collection of papers from the Victor Corporation, just the 21 highly-requested pamphlets were scanned. Since launching the collection, 20% of visits to the collection have come from outside the United States, and over 70% of visits to the collection have come from outside the state of Iowa according to site statistics.

The Rand & Leopold Desk Company was an office furniture manufacturing firm which operated out of Burlington, Iowa for 102 years before closing in 1990. Furniture collectors had always been interested in the Libraries' collection of 27 original product catalogs, and there were many requests for reproduction. After scanning them and making them available through the Iowa Digital Library, only 7% of site visits to the digital collection have come from outside the US, but over 90% have come from outside the state of Iowa. Perhaps more so than the W9XK collection, these two small collections meet the goal of facilitating information needs of users who cannot easily travel to the UI Libraries.

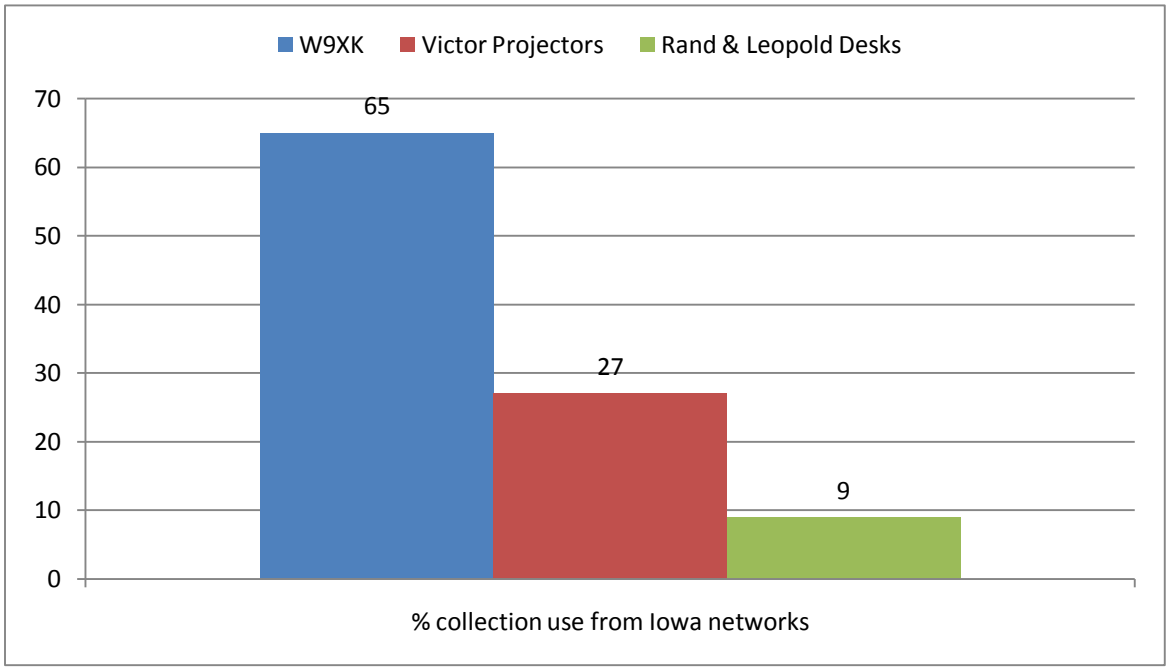

Figure 5. % collection use from Iowa networks

Difficulties in tying a bow

Even when original goals are met, small-scale digital collections don't often stay "done" for long. Scope-creep is a common danger, for even when items haven't been cherry-picked for a digital collection, the source collection almost always logically connects with more collections. In trying to connect users with a broader universe of resources, libraries are compelled to extend the scope of their digital collections. In an example of how digital collections never seem to stay finished, a collection of an artist's papers in the Iowa Women's Archives were digitized to complement her digitized artwork. When discussing publicity of the collection, IWA staff noted that they had found several audio oral histories and also a video made about the artist. Even when those multimedia materials are digitized and added to the collection, it's possible that the archive will acquire additional items that would need to be added to the digital collection in order for it to remain comprehensive.

Conclusions:

Even as mass-digitization projects begin and are sustained at libraries through funding and organizational commitments to scaled-up scanning workflows, requests for creating small-scale digital collections will persist. This is because many unique collections, especially those with preservation issues or spread out in archives will not be included in mass-digitization workflows. Likewise, the needs of libraries' clientele and donor relationships do not necessarily coincide with broad institutional objectives. Responding to these requests can help strengthen relationships inside and outside of the library and build support for digitization programs. Even one-off collections should fit within programmatic guidelines, although in the case of the W9XK collection, some liberties were taken in its creation. Small-scale digital collections can be useful to researchers and the general public, but they have to be accessible. Once found, the collections should be complete to the point where users

can be confident that they have the entirety of a collection's information at their disposal. Although the W9XK digital collection required the hand-selection of items, steps were taken to ensure the fullness that is necessary in today's digital libraries.

Reference List:

Committee on Institutional Cooperation (n.d.). CIC-Google Agreement. Retrieved from http://www.cic.net/Home/Projects/Library/BookSearch/CIC-Google.aspx

Erway, R., and Schaffner, J. (2007). Shifting Gears: Gearing Up to Get Into the Flow. Retrieved from http://www.oclc.org/programs/publications/reports/2007-02.pdf

Lally, A., and Dunford,C. (2007). Using Wikipedia to Extend Digital Collections. *D-Lib Magazine*, 13 (5/6).

Library of Congress (2003). Library of Congress and Ameritech Competition (1996-1999): Awards and Collections. Retrieved from http://lcweb4.loc.gov/ammem/award/collections.html

University Libraries. University of Iowa (2009). Iowa City Town and Campus Scenes. Retrieved from http://digital.lib.uiowa.edu/ictcs

University Libraries. University of Iowa (1999). Keeping Our Word / North Lobby Exhibit / University of Iowa. Retrieved from http://www.lib.uiowa.edu/exhibits/keeping/intro1.htm

University Libraries. University of Iowa (2009). Map/diagram of Midwest, 1930s. Retrieved from http://digital.lib.uiowa.edu/u?/w9xk,67

University Libraries. University of Iowa (2009). Rand and Leopold Desk Company Catalogs. Retrieved from http://digital.lib.uiowa.edu/desks

University Libraries. University of Iowa (2009). Victor Animatograph Company Pamphlets. Retrieved from http://digital.lib.uiowa.edu/victor