In January 2008, riding on the coattails of a six-month intensive strategic planning process, the New York Public Library deconstructed its Digital Library Program (DLP) and established an entirely new organizational unit, NYPL Labs. Drawing talent from across the organization, the new unit was formed to take a broader, more user-focused approach to the ever-evolving digital environment and the challenges it represents for service-centric cultural organizations.

NYPL Labs represents a shift in the approach to library digitization in that “digital” is understood as an integral part of core activities—an organic feature woven into the life of the library. As David Ferriero, the Andrew W. Mellon Director of the New York Public Libraries has often said, “Digital is the way we work today. There is no separation between what we do and what they do. It is all of a piece.”

Digitization production, which had been part of the digital library program, is now concentrated in a dedicated reformatting unit. This unit functions as a coordinating body as well as a production center, helping to shepherd digital content produced elsewhere in the organization through the ingest process and into the repository. Metadata creation is governed in a centralized set of systems and policies, but its creation is distributed across the library.

Decoupling production from the service elements of the digital economy frees the library to imagine a more satisfying and useful interaction with deep content. NYPL Labs is a place for experimentation, a place to try things out. And these “things” are not just software but “environments” where software and content collide. In this way, NYPL Labs is expected to focus on the “use” of library materials and the “use” of library space, both physical and virtual.

Failures are acceptable and expected, because failure provides invaluable intelligence for both developers and managers. And failures provide a “line in the sand”—a benchmark from which developers and managers can establish an appropriate means of measuring success.
NYPL Labs is not an engineering farm. It is dedicated to understanding and working with tools that are already available to the community, and, where no tools exist, to leveraging relationships to get those tools made. NYPL Labs is not attempting to build everything from scratch but to make systems and content work well together. NYPL Labs is not a traditional digital library unit, if there is such a thing.

Background

The impetus for the reorganization was simple: the existing structure and the funding were untenable. Over the last decade, the digital library program at NYPL was marked by grant-funded, project-driven development. Strategically, this was sufficient for making near-term goals, such as getting the DLP up and running, but it did little to address the long-term viability of both the program and the projects it produced. Furthermore, the program and projects were understood to be “fixed-cost” items that had clear beginning and ending dates that fit nicely into fundraising scenarios.

However, as early as 2000 it was recognized that funding models were changing and that a “special projects” staff was neither affordable nor desirable. That said, it was inconceivable that the library should abandon its digital library program.

Taking the library outside the four walls and into the community has been as much a part of the organization’s history as providing resources for scholars and researchers within the bounds of the various reading rooms. The New York Public Library has been involved in outreach services as long as it has been in existence.1 So it was important for the library to continue its outreach activities as a piece of its long-term strategic plan and to integrate digital activities into the core of library practice to guarantee that the good work already done would be preserved and accessible to future generations.

Furthermore, since 1998 the library had generated a significant amount of digital content, and that content required curatorship. But this was not an easy sell to the broader NYPL organization, since the lion’s share of the work in the DLP projects had been conceived, executed, and managed by resources dedicated only to these activities. Moving the rest of the NYPL community toward a more holistic, distributed approach was a challenge.

NYPL was well versed with large-scale projects. In the past ten years the library had produced over 1 million archival images; had created a large body of digital video and audio concentrated in the area of performing arts; had converted several

---

1. NYPL History, available online at www.nypl.org/pr/history.cfm (Accessed February 27, 2009).
hundreds of thousands of microfilm frames to digital format; and had participated in the Google and MSN Books projects. But the vision and energy behind these projects was not married to other activities in the library. The goals of the digital projects were siloed, with no relationship to the direction and planning of other library initiatives, blessedly oblivious to work carried on under the same roof.

Consequently, selection of materials for digital projects was also siloed. Digital projects had their own selection committees, oversight committees, and experts. The idea of including digital content in Collection Development or vice versa was not considered. But the question kept arising: “How do you make choices about what gets digitized?” And the answers were as numerous as the curators who had participated in the digital projects.

The most compelling answer came from Sharon Farb, Associate University Librarian at UCLA, who said, “The choice is simple; you have already made it.” In other words, the hard work has already been done by Collection Development (CD). In the process of selecting materials for the library, the CD librarians have evaluated the content in terms of its complementarity with the existing collection and the existing collection policies. Any choice is the right choice.

While this may not be practical advice, it is strategically excellent advice. Perhaps retrospectively, choices may still be difficult, but prospectively, including digital selection (e-resources, born-digital content, digital archives, and the like) into the collection development workflow is the best model to follow. By shifting the selection process to CD, the library is able to take advantage of existing best practices, bring digital resources into line with the rest of the collections, exploit existing cataloging and cataloging expertise, leverage the strengths of the organization, and provide a coherent body of materials to its users.

This last point is the most critical to the overall strategy of NYPL. By reorienting the approach to digital content by focusing on the user experience, the hard questions become somewhat easier, or at least the path to the answers becomes somewhat clearer.

What makes sense from a user’s perspective is very different from what makes sense from a curator’s perspective. While the curator is critical to the conversation, that conversation has to include all the stakeholders. Selection is bigger than any given curator or selector, since the user sees a slice of all the past selection processes. Selection is part of the overall profile and reputation of the library and the corpus.

---

2. NYPL Annual Reports, available online at www.nypl.org/pr/annualrpt.cfm (Accessed February 27, 2009).
represented by that library and, as such, provides ever-increasing opportunities to engage the audience and draw relationships among the community and the content.

This broader view has led to a new set of best practices, or tenets, for digital activities at NYPL. First, digitization is not a department; it is something that everyone should participate in. Second, lower the barrier for participation by building systems that allow librarians at all levels to create and contribute—don’t make the digitization arm of an organization the privileged unit. Third, proselytize through social networking tools. If you develop radical transparency in your digitization efforts, and don’t privilege the process, you will open the process up to everyone in your organization.

Tools and Standards

One of the quickest ways to surface good work is to make the work tools available to the entire organization.

NYPL Labs developed a digitization nomination system that allows anyone with an organization ID to put a request into the queue. This system, called “Q,” not only accepts requests but has a built-in dashboard interface. The purposes of the dashboard are to expose the list of requests to the entire community and to track the progress of requests in the system. The community may not establish priorities, but the simple display of information provides staff with an overall view of activities in a certain area. By not overformalizing the process, the barriers to entry are lowered.
"Q" was developed in part to meet our commitment to radical transparency; in part to allow all staff to have access to tools; and in part to see if we could build a useful application in a very short time. By building "Q" in a lightweight software platform (Ruby on Rails), NYPL Labs was able to "open the door" and experiment in a more timely fashion.

More important, "Q" underscores NYPL’s commitment to open up digitization practices to the entire library. There are no organizational roadblocks to nominating collections or reviewing the dashboard. Furthermore, there is a synergy that is developed by viewing the choices and requests of one’s colleagues. Knowledge increases creativity and discussion, furthering staff engagement with the digital agenda at the library.

Hand in glove with this model is the idea that a variety of formats may be suitable for ingest into the repository. It has long been the practice of digital libraries to control the flow of content into their repositories by raising the barrier to entry so high that very few materials qualify. While this is laudable for standard preservation purposes, perhaps not everything in the repository can or should be preserved.

It is entirely feasible that an institution may want to have an open repository that functions more in line with electronic records management than with traditional digital archives. The principles and “rules” should reside in the policies that govern the repository rather than in the controls over the ingest specifications. In other words, a broader digitization model may also include the ingest of nonpermanent content into the repository. NYPL may make and use nonarchival standard digital content for a number of years and migrate it through a number of formats, but it may also develop retention schedules that address these data rather than barring them in the first place. Perhaps by rethinking the institutional obligations incurred by digital preservation, the obligation to keep everything indefinitely may be obviated.

**Guiding Principles**

NYPL is building its collections in the following ways:

First and foremost, all digitization activities are distributed across the library. There is a commitment on the part of upper management to embrace these processes and incorporate them into the everyday life of the library. Internal resources are being devoted to building workflow systems to facilitate communication among these distributed pods and to shuttle data and content among systems.

Selection of new digital content to purchase, reformat, or acquire is being folded into the overall collection development strategy. Digital content will be accessioned
according to the library’s mission and should be complementary with the library’s collection and staff strengths. Digitization should not be viewed as an opportunity to develop a new collection area.

When at all possible, digitize according to existing best practices and procedures. If not possible, modify practices and procedures to accommodate digital resources into ordinary workflow.

Use digital collections to supplement, or fill out, subject areas. Expand user-generated areas of interest to build relationships between subject areas.

Be as authoritative and comprehensive as possible. Cherry-picked digital collections do not represent the strength and breadth of a library and, from a user’s standpoint, are less than satisfactory. Users want to know what is available, how much is available, and whether it is everything that the library holds on a given topic.

Leverage every possible venue for digital content. If patrons request copies, make digital copies and send them to the repository. If staff produce in-house vehicles for teaching or instruction, capture the content and move it into the repository. Use cataloging and interlibrary systems to gather metadata to be repurposed for digital content.

Enlarge the pool of options so that a variety of selectors and a variety of collections can be put forward for digitization. Allow staff at all levels to make suggestions so that those who work most closely with the processing and paging of materials can contribute to the overall digital strategy.

Past-use patterns are a guide, but not a definitive measure. Simply because something does not get a lot of use, it does not necessarily follow that it is not useful or would not be used if presented in a more compelling manner. Build from the user’s perspective, not from raw numbers. User behavior must be analyzed, not just measured.

Before archival processing begins, inventory large collections with an eye to “handling once.” By making the decision ahead of time to do item-level descriptions on some part or parts of the collections, the double processing of these collections is avoided. Broaden the view of the hierarchical organization of archival collections. Perhaps the hierarchy is suitable for certain purposes, but not all purposes. Consider making entry points into the archival structure at many different levels.

Consider metadata a multipass, multcreator system. Instead of a single cataloger, consider using paraprofessionals to do brief record creation that will be enhanced by subject specialists.
In the end, digitization is an enterprisewide activity. It should involve all areas of the library that are concerned with the processing and delivery of library materials. It must also involve those professionals who are engaged in user services, since their input is fundamental to understanding patron needs.

With this level of engagement, NYPL Labs is able to work diligently to bring better services to its patrons, both public and professional. By distributing the creation and management of digital content across a broad range of players, NYPL Labs is freed to focus on the user experience and to work to enhance the library digital space.

The continuing goal of NYPL Labs is to think broader and deeper and, by doing so, to slightly redirect the conversation around digital libraries. Digitization is necessary, access is critical, but “doing” is the direction of today. It is no longer good enough to make materials available—we must now consider how our readers want to use our materials and whether we are providing the most that we can to make this happen.