Florida International University
FIU Digital Commons

Works of the FIU Libraries

FIU Libraries

7-10-2017

dPanther: Benefits and Challenges of Building and Managing Your Own Digital Repository

Kelley Rowan
Florida International University, krowan@fiu.edu

Rebecca Bakker
Florida International University, rbakker@fiu.edu

Follow this and additional works at: https://digitalcommons.fiu.edu/glworks

Part of the Archival Science Commons, Cataloging and Metadata Commons, and the Collection Development and Management Commons

Recommended Citation

dPanther: Benefits and Challenges of Building and Managing Your Own Digital Repository

Rebecca Bakker, Digital Collections Librarian, Digital Collections Center, Florida International University,
Miami, Florida, rbakker@fiu.edu

Kelley Rowan, Digital Archives Librarian Digital Collections Center, Florida International University,
Miami, Florida, krowan@fiu.edu

Abstract

This article addresses the benefits and challenges associated with building your own digital repository
platform and managing content, technical hurdles, and partners. The authors discuss the advantages of
building a repository that can evolve as standards in the field do and how to handle the problems that
arise when you are both the metadata manager and technical support. The authors also discuss
difficulties their partners have experienced with the platform and lessons learned. It is the hope that
this information will provide those who may be either managing a digital repository or considering a
new digital platform, relevant information that will prove to be helpful in the decision-making process.

dPanther: Benefits and Challenges of Building and Managing Your Own Digital Repository

There are a multitude of vendors available for digital repositories, so the question becomes, "Why host
your own?" Vendor solutions may indeed be appropriate for many digital library initiatives. In fact, a
vendor solution was available to our university through a statewide consortial membership. In the state
of Florida, academic libraries are part of the Florida Academic Library Services Cooperative,[1] which
provides services for both the statewide digital repository and the digital preservation archive. However,
Florida International University (FIU) participates in a number of community-based funded projects that
require a high degree of development and design flexibility. Open-source software, while still requiring
funding for development and maintenance, was more flexible and financially realistic than choosing
another vendor. So in 2014, FIU, along with two other universities, opted to host its own digital
repositories based on available open-source software rather than adopting the newly chosen state
repository.

Although building and managing a homegrown digital repository is not for the fainthearted, we believe
it has been a learning experience that has benefited FIU as well as both our internal and external
partners. dPanther,[2] the digital repository at FIU, was created by the developers in FIU's Geographic

---

[1] About FALSC," Florida Academic Library Services Cooperative: Services and Information for Florida's College and
University Libraries, https://libraries.flvc.org/about-falsc (accessed September 18, 2017).
[2] Boyuan (Keven) Guan, Julian Gottlieb, and Meng Ma, "DPanther Repository," Home—Florida International
University (FIU) Library Digital Collection Repository System (dPanther), http://dpanther.fiu.edu/dPanther/home
(accessed September 18, 2017).

Information System (GIS) center[3] and is maintained and managed by the digital librarians and assistant director in the Digital Collections Center (DCC).[4]

dPanther became a live site in 2014 and currently has 54,816 records in 63 collections and 17 subcollections, which include oral histories, audio and video recordings, photographs, government reports, newsletters, yearbooks, and maps of various cultural and historical partners to FIU. The repository is an amalgam of content from both internal and external sources that range from FIU-created items to global publications to the digital holdings of local archives and museums. A significant portion of the content has a regional focus, including Florida, the Caribbean, and Latin America.

Although it has not been without its obstacles, building our own digital repository allowed us to include interface features that enhance the user experience but were not available in Digitool, the digital repository in use by the state of Florida prior to 2014, or in the new repository the state was considering. By building dPanther, we have had the luxury to both customize the platform to our own specifications and enable specific features that funding partners want. Some of these interface attributes include embedded audio and video resources, a zoomable JPEG viewer for images, geospatial and temporal search functionality, a metadata-based machine learning resource recommendation function, and customizable interface design/branding for partners.

Conversely, the homegrown repository comes with challenges. For instance, when the inevitable problems arise, it requires diligent sleuthing and trial and error to resolve the issues internally. We occasionally need the help of the GIS programmers, but maintaining and updating dPanther is one of their many tasked projects, and we must prioritize our requests strategically. And, as with most project management, time considerations are often the most prevalent obstacles that need to be resolved. As digital librarians, our job generally involves copious amounts of quality control, troubleshooting, metadata creation, and dissemination of digital content through dPanther. This includes various text documents, such as rare books, manuscripts, maps, audio files, video, and images (Figure 1).

---

[3] GIS Center at FIU, January 1, 1970, http://maps.fiu.edu/gis (accessed September 18, 2017).
[4] "Digital Collection Center: The World's Best Collections," http://digitalcollections.fiu.edu/dcc (accessed September 18, 2017).

Figure 1. The Front end of Florida International University's dPanther digital repository showing thumbnails of selected collections. www.dpanther.fiu.edu

dPanther: An Origin Story

Generally, the two most serious encumbrances to building your own digital repository are funding and expertise, with a close third being the sobering thought of having to migrate all the content, a task that would have been done for us had we gone with the state consortium. Despite these obstacles, inevitable and unforeseen changes occurring both at the state level and at FIU led us to seriously consider this undertaking.

In 2014, the Florida Academic Library Services Cooperative was preparing to adopt a new statewide digital repository system for the state of Florida, making the impending data migration of digital files inevitable for our DCC staff. As the state explored options, various ideas circulated throughout the FIU Library, including opting into the state consortium and using whichever platform they would eventually choose, partnering with another university, or creating our own repository.

At the same time, FIU Libraries was going through its own changes with a new dean and a reorganization that placed the DCC under the GIS department, a department that, through necessity, employed a staff of developers. With this particular reorganization, the option of creating and maintaining our own repository began to look more feasible. What was still lacking, of course, was the funding.

Rather serendipitously, the nearby city of Coral Gables was already partnering with FIU's DCC to create a geo-enabled virtual history project[5] and had avail-able funding for the endeavor. The state, meanwhile, opted to go with Islandora, an open-source digital repository based on Fedora Commons and Drupal, which, while highly functional, did not include the geo-functions required for the Coral Gables project. Thus, dPanther, with willing developers from GIS and funding from the city, was born. (The name dPanther was created to indicate the digital aspect of the repository [the lowercase "d"] and to recognize FIU's mascot, the Florida panther.)

Under the Hood of dPanther

dPanther was built on the framework of SobekCM,[6] an open-source digital con-tent management system that uses the programming language C# as well as the open-source Web framework known as ASP.net.[7] SobekCM had been created by the George A. Smathers Library team and was already in use by

---

[5] "Coral Gables Virtual History—The City Beautiful," http://maps.fiu.edu/cgm (accessed September 18, 2017).
[6] "Sobek Repository," SobekCM: Digital Content Management System, http://sobekrepository.org (accessed September 18, 2017).
[7] "Learn ASP.NET," The Official Microsoft ASP.NET Site, https://www.asp.net/learn (accessed September 18, 2017).

the University of South Florida and the Digital Library of the Caribbean before FIU adopted it, allowing us to view instances of its application in other repositories. The plat-form is relatively adaptable and allows for full-text searches, multiple display, and viewer options and accommodates records in English, Spanish, and French, which was an important feature considering the regional diversity of our collection (Figure 2).

Figure 2. Image showing dPanther web form and associated metadata with diacritics.

SobekCM records allow descriptive metadata to be expressed with various metadata schemas, such as MODS (Metadata Object Description Schema),[8] as well as qualified Dublin Core,[9] a metadata schema that provides information about other data, giving high levels of flexibility for the metadata needs encountered in digital collections. Like most digital library software, it also allows for structured data integration into the Open Archives Initiative, Harvesting Protocol (OAI-PMH),[10] allowing for repository interoperability.

What Users See

dPanther consists of a customized front-end user interface created by FIU's GIS developers that is simple and user friendly. Thumbnails represent selected collections, while the full list is found in the advanced search. At the top of the page, there is a prominent search bar for quick access to content. When users search for an item, they get a two-column display with multiple refinement options on the left and re-cord information and thumbnail image on the right. Users can refine their search by collection, topic, geographic area, author, publisher, type, and language (Figure 3).

---

[8] "Official Web Site," Metadata Object Description Schema: MODS (Library of Congress), http://www.loc.gov/standards/mods (accessed September 18, 2017).

[9] "Dublin Core Metadata Initiative," DCMI, http://dublincore.org (accessed September 18, 2017).

[10] Carl Lagoze, Herbert Van De Sompel, Michael Nelson, and Simeon Warner, "The Open Archives Initiative Protocol for Metadata Harvesting," Open Archives Initiative—Protocol for Metadata Harvesting—v.2.0. June 14, 2002, http://www.openarchives.org/OAI/openarchives-protocol.html (accessed September 18, 2017).

Figure 3. User view in dPanther of a search for fishing which shows the aggregators on the left side.

Once users click on a record, they see a picture of the item at the top (zoom-able if it is a photograph or image), followed by record information below. Off to the right, a menu, including file viewer, record information, maps/geo-reference, and item usage, allows viewers to avoid scrolling. Below the menu is the relatively new "recommender" feature, which includes "related items," which will take you to similar items. There is also an "others liked" feature that uses an algorithm involving the user's IP address to ascertain recommendations based on similar viewers and, finally, a "newly added items" feature. Partner branding information is shown through the use of word marks, allowing for appropriate recognition, as are the options to share items via social media (Figure 4).

The maps/geo-reference section near the bottom of the record was originally developed for the city of Coral Gables and the Coral Gables Memory Collection. The city was interested in a virtual tour function and having each image geo-located on a map. Not all of our partners have been interested in taking advantage of this function, but the map is available whether location information is entered or not (Figure 5).

Figure 4. User view at the top of a dPanther record, showing additional files on left and the recommender system and wordmarks to the right of the image. Image: Photographs depicting alligators and crocodiles, approximately 1931. Identifier: CM00500006. dpanther.fiu.edu.

Figure 5. Bottom view of a record in dPanther showing geo-location map tool.

If You Build It, They Will Come

Once we had our own digital repository in place, we were ready to add more con-tent. Migrating over records from the previous repository took time and effort but allowed DCC staff members to hone our skills in creating batch records, trouble-shooting common issues, and ironing out the kinks of the fresh, new system. It was during this time that we learned that hidden files on Mac computers would sometimes show up as images within the record of dPanther. Items would not dis-play in the sequence we had structured (e.g., item 11 would show up after item 1, as the system read the numbers from left to right). Images would become distorted and discolored when zoomed in through the viewer. Our server would crash when too many records were uploaded at once. Each of these required work-arounds, system upgrades, or simply conscientiousness to overcome.

Initially, in the more traditional model of digital collections departments, the DCC staff took responsibility for uploading, curating, and describing all digital content. However, as departments such as FIU's Sound and Image were gifted enormous collections, we realized quickly that we could no longer be as hands-on with content. dPanther allowed us to set various levels of permissions, so we began to give our associated partners the ability to upload directly into the system with-out giving them the ability to delete items or entire collections.

This allowed the DCC to begin a new model: training our FIU internal partners and outside cultural institutions to add content directly to dPanther themselves. The DCC would no longer be the bottleneck in digitization projects. In-stead, we began to transition to the role of liaisons. We no longer determined what items were to be digitized, we usually did not do the actual digitization, and we did not, for the most part, create the metadata for their records.

We did, however, create multiple instructional materials, group workshops, and one-on-one trainings with our partners and their staff to enable them to add content directly to dPanther. We designed work flows and automation strategies. We provided services for text recognition, transcription of audio and video, file folder and unique identifiers creation, and image processing when needed. We created a dPanther working group, giving our partners a platform to share their experiences and advise us of any difficulties that needed to be addressed. We attempted to guide our partners every step of the way in the decision making and technical skills needed to successfully ingest content into the repository. By empowering and training out-side partners to add their own content and metadata, DCC librarians could

concentrate on quality control, standardization, long-term digital preservation, and solving technical problems. Using this model, our collections began to grow exponentially.

Managing Partners, Not Content

dPanther allows our partners to create records directly in the repository using a simplified Web form. The Web form does not include all the rich metadata options found in MODS or qualified Dublin Core but has the benefit of a being straight-forward enough for nonlibrarian/archivist partners to create manageable meta-data and upload content from their own computers. Larger batches of records are added through use of a METS Editor,[11] a freestanding application that creates an uploadable XML file and allows for more nuanced and granular metadata.

Uploading high-resolution TIFFs or PDFs is also easy for partners: the dPanther system automatically creates thumbnails. Multimedia files, such as mp3s and mp4s, are hosted on a separate video server and are embedded into dPanther with a simple URL link. However, this separate media server requires authentication at sign-on, requiring DCC staff to do the work of uploading these types of files.

Allowing non-library partners the ability to create and upload content came with an inherent risk: inconsistent appearance of collections, messy metadata, and so on. However, it also becomes essential as partners and content grow. All digital repositories require managers to make the classic "quantity versus quality" decision at some point, depending on the needs of the institution. Like many other institutions, the DCC is working with a small staff, with many projects to manage. The DCC is fortunate in that we generally have volunteers, part-time staff, and federal work study students who can be trained to perform quality control on our collections and in this way provide assistance in the managing of our digital repository.

Benefits/Challenges and Lessons Learned

Despite the great successes within this new model, challenges were inevitable. Our partners and colleagues had varying levels of expertise in digitization, meta-data, archives, and preservation. Many of our cultural institution partners, such as museums and archives, were skilled at traditional forms of metadata creation and archiving but lacked the technical skills to deal with a sometimes unforgiving computer platform. Other of our partners had high technical skills but were not librarians and thus had trouble with standardized language, authority control, consistency in date format fields, or difficulties in organizing the structure of their digitized files.

In an attempt to mitigate these problems, we have instituted a metadata training initiative for all of our member partners who now meet with us prior to any creation of records in dPanther. We also provide them with access to our on-line training materials, called Libguides,[12] which detail metadata guidelines. However, this is a continual process, as other organizations are often using part-time help and their turnover can be significant, requiring frequent retraining for the same project. We also face our own changing requirements, such as new demands from outside organizations needing to harvest our records, meaning that we must update tutorials, retrain dPanther users, and perform retrospective quality control to update and standardize records according to new guidelines. Despite these initiatives, mistakes still happen, and continuous quality control remains crucial.

Another of our challenges with dPanther arose when trying to ingest records with diacritical marks in them. Miami is a diverse city, and much of our content comes from South America and the Caribbean,

---

[11] Mark V. Sullivan, "METS Editor (SobekCM)," SourceForge, https://sourceforge.net/projects/ metseditor (accessed September 18, 2017).

[12] "FIU Libraries: Research: FIU Digital Project Guidelines and Help Materials: Preliminary

where Spanish, Portuguese, and Haitian Creole are dominant languages. It is common for our records and metadata to include diacritical marks, especially in personal names. Unfortunately, those accent marks cannot be recognized by our system in file names, requiring quality control on the front end before items are ingested. While partners are trained to avoid these, we often run a check for diacritics or other problematic characters in file names using Adobe Bridge (a digital asset management application from Adobe) before ingesting. This, of course, works only with entities or departments for whom we do the ingesting of content.

Having photographs and images scanned and digitized outside of our purview can also cause issues that are not immediately apparent. When items are scanned off-site, often by interns and students, consistency can be lacking. An-other unique but significant issue that affects the performance of dPanther is the inclusion of a bracket in front of a title when the metadata are created by the cataloger. We are fortunate to have partners who have expertise in creating archival metadata, but it can be a problem for a digital repository such as dPanther, which can no longer alphabetize the records and simply puts all bracketed items first. It can be difficult for librarians in other departments who are trained with a specific set of standards, such as machine-readable cataloging, to leave those schemas behind while creating metadata for digital records in dPanther. Fortunately, dPanther allows us to perform a large-scale find and replace to quickly clean up this issue in batches.

Our most recent challenge has been modifying records to meet the requirements for sharing dPanther content with the Digital Public Library of America (DPLA).[13] This is where records that do not have a standardized date format (YYYY-MM-DD) can become problematic when migrated to the larger DPLA platform. Additionally, DPLA has specific requirements for certain metadata, such as the "Rights Statements" field. Early on in this process, we discovered that dPanther would not allow for multiple rights statement fields despite the MODS recommendation for multiple fields. We have had to allow our developers many months to develop the ability to have multiple rights statements in dPanther. When we added a secondary field in the XML file, dPanther simply picked one of them. Although this has definitely been a challenge, the benefit is that we absolutely can make this change; it will just take some time. Overall, this is preferable to finding ourselves committed to a repository that also may not have this ability but that we have no control to change in order to meet newly evolving standards.

As mentioned previously, an unfortunate by-product of having a home-grown system is that when we face issues, we often face them alone. Many of the problems we face are not unique to digital repositories overall, requiring creativity and flexibility to determine the particular solutions within the dPanther platform. Ultimately, having the ability to customize our digital repository as needed has thus far outweighed the challenges we have navigated.

Last but not least, as digital librarians, one of the biggest dilemmas for those of us in the DCC was relinquishing control and coming to terms with the fact that not every record was being created exactly as we would have done it ourselves. Not all of the metadata in dPanther are perfectly up to our standards, and that is some-times a difficult thing to acknowledge. Our mission in the DCC is to build online collections of enduring value for the university and broader user community. We are making valuable content accessible and have determined that the perfect is too often the enemy of the good.

---

[13] Guidelines & Workflow," FIU Libraries: Research at Florida International University, http://lib guides.fiu.edu/digitalprojects (accessed September 18, 2017).