Received 20 February 2015
Revised 12 May 2015
6 July 2015
Accepted 10 August 2015

Investigating and implementing an extensible, adaptable game plan for digital initiatives at a large state university

Virginia A. Dressler
University Library, Kent State University, Kent, Ohio, USA

Abstract

Purpose – The purpose of this paper is to explore the process of assessing the needs of a digital repository. The study's institution – Kent State University, a large state university – is currently re-envisioning their digital library services to accommodate emerging technologies with an agile framework and platform. This paper includes discussion on the process of creating a sustainable digital initiatives program, a plan of action for migration, platform decision rationale and an overview of planning for future projects that is both flexible and extensible in nature.

Design/methodology/approach – The article's approach is to examine the process of review and selection of the digital repository at Kent State University through the needs and requirements checklist the team devised according to the institution's specifications. Literature published in the past five years is reviewed and applied to the selection process. The author maps out a plan that can be adaptable for growth and sustainability for a content management system.

Findings – Using models identified in the article, the team mapped a solution that can enhance the use and interaction by patrons to digital collections as well as provide a method to ensure the longevity of digital assets.

Originality/value – The article addresses issues that are inherent to digital content management systems of all shapes and sizes, and strives to provide a solution that would be relevant and applicable to many types of institutions in regard to digital repositories.

Keywords Digital libraries, Digital repositories, Open source software, Systems design, Digital collections management, Digital workflow

Paper type Conceptual paper

Introduction

As the inevitable consequence of progress and development, the day-to-day as well as long-term needs and requirements of a digital repository will change over time at any sized institution. It is important to address a periodic review process of these unavoidable changes and to reassess the platform and functionalities to ensure that the system is adapting to these changes. A system that was devised 10 years ago may still be operational and technically functional at a base level, but may not fit current scenarios in workflows and production of digital content. Formats, such as larger high definition digital video files, 3D files and born digital content, have advanced significantly in the past few years, have increased baseline file sizes and have impacted interface requirements. Digital platforms that have not adapted to changing formats and needs will not address current user expectations and may also be dismissing necessary transformations that need to take place. Updates to key points of the

The Electronic Library
Vol. 34 No. 4, 2016
pp. 588-596

0264-0473
DOI 10.1108/EL-02-2015-0034

workflow process, from ingestion to indexing to storage capabilities, will need to be addressed and include incorporating digital preservation tools at the root of these systems.

This past year at Kent State University (KSU), the new position of a Digital Projects Librarian was added to the professional staff at the university's library. One of the first tasks at hand for this new role was to address and assess the needs of diverse sets of digital media, digital collections and digital exhibits. Other staff and faculty members had previously done work of this nature, assigned in addition to their primary job duties. Consequently, no single person had been able to complete a thorough assessment and plan for digital initiatives. The library's current digital collections were created over the past few years on a part-time basis from a number of individuals as very much a group effort, but the digital collections as a whole did not have a framework of an overarching architecture for a solid digital repository. Small, boutique digital collections had been cobbled together using a home-grown approach by in-house developers using a number of solutions – further defined below – and a needs assessment of the whole system was overdue to assess its sustainability for future development.

As a result of the new position, a proposal was created for a digital preservation platform and repository with a recommendation for a solution to provide enhanced access to the content for a viable, long-term digital repository. Throughout the evaluation process, a big picture view was needed to address many of the varied and diverse needs of both the department and university at large to create a solution adapted to the existing needs with the potential to evolve to meet future needs. This paper addresses the evaluation process and shares some of the conclusions of the proposal for the scenario at the study's institution.

Background and literature review

During the initial process of assessing the scenario at Kent State's University Library digital collections, a research inquiry was made into the processes and workflows of other comparable digital initiatives at academic institutions. Articles and books written in the past five years were explored to help in advising and guiding the process for the institution, though curiously the literature has been found to be quite sparse on the topic. Digital management content solutions that addressed an agile platform design and increased workflow were also examined.

To begin, a 2009 case study from the University of Maryland was studied because they were in a similar conundrum as KSU when the staff looked to revamp the image content management system (Novara, 2010). The author wrote in-depth about the need for an improved workflow for digital collections to address effectiveness and efficiency. In particular, quality assurance was discussed in terms of staffing and the capability of the system to keep up with production. Additionally, digital initiatives were dispersed in terms of the physical location with multiple workstations and locations around campus, which added to complications in the already belaboured workflow.

Another point of reference that was used during the initial investigation was the digital preservation outreach and education curriculum outline on the Library of Congress's site (www.digitalpreservation.gov/education/curriculum.html), which proved to be an important framework to the discussion at the institution. This outline provided a good basis for a framework and a point of reference while

conversations with collection managers, content and research specialists and other library staff were carried out to ascertain scenarios and define the needs of the library and the institution at large. The outline proved to be very helpful in providing a straightforward tactic that would be useful for any size digital initiative. The outline provided the following steps:

- identify the types of digital content in the library;
- select the portion of the digital content to be preserved;
- store selected content long term;
- protect the content from everyday threats and emergency contingencies;
- manage and implement requirements for long-term management; and
- provide access to the digital content over time.

Another tremendously useful resource that was examined during the investigation process was to look at another large state university in Ohio, which had also recently undergone a similar transition and review of the digital project ideology and underlying framework. Under the guidance of the Digital Initiatives Librarian, Ohio State set out 11 guiding principles to help frame the process through discussions of a special working group (Reese, 2013). The resulting guidelines provided a realistic and practical set of directions when applied to a large university in a modern setting. The principles consider many facets and aspects present in digital content management, such as addressing the digital life cycle and creating modular services for users. The ideas outlined in these principles can more adeptly serve the needs of a variety of departments and individuals over time through its design, while providing a realistic service model with numerous solutions for media and also setting expectations (and limitations) of service, particularly within the model of academic libraries.

The most interesting aspect of the approach from Ohio State is the focus on flexibility and change, fully addressing the likelihood of migrations, upgrades and issues of obsolescence that plague digital repositories of any shape or size. There is an underlying theme throughout the guiding principles that address user needs through modular services that work with and complement existing frameworks. This approach blends a practical angle with a more intriguing, adaptable modernized view of systems and applications.

In a similar vein to Ohio State's model, Awre (2012) wrote about the emerging open-source Hydra Project in 2012, a framework based on the Fedora repository structure. Awre addressed the needs of the University of Hull and how an open-source product is able to tackle institutional requirements and also provide some distinct opportunities through the nature of shared development with like-minded partners at other academic institutions. While the Hydra Project has made large strides in development since the article was written (and has also added many partners to the open-source project), the article is a great example of how the topic of adapting to changing technology needs and requirements was addressed at an early stage.

Finally, a 2013 title on the changing research landscape by Simons and Richardson examined current needs of the modern researcher and how the digital repository may address these needs. Issues of data sharing, digital life cycles and sustainability were

explored, as well as formats, such as data sets, streaming media and web archiving. The other interesting issue that is broached is not only the changing media landscape, but how metadata needs may vary in these formats, as well as the appropriate platform to host varied content. Further, the authors contend that training and development should be an on-going process, not a stagnant one.

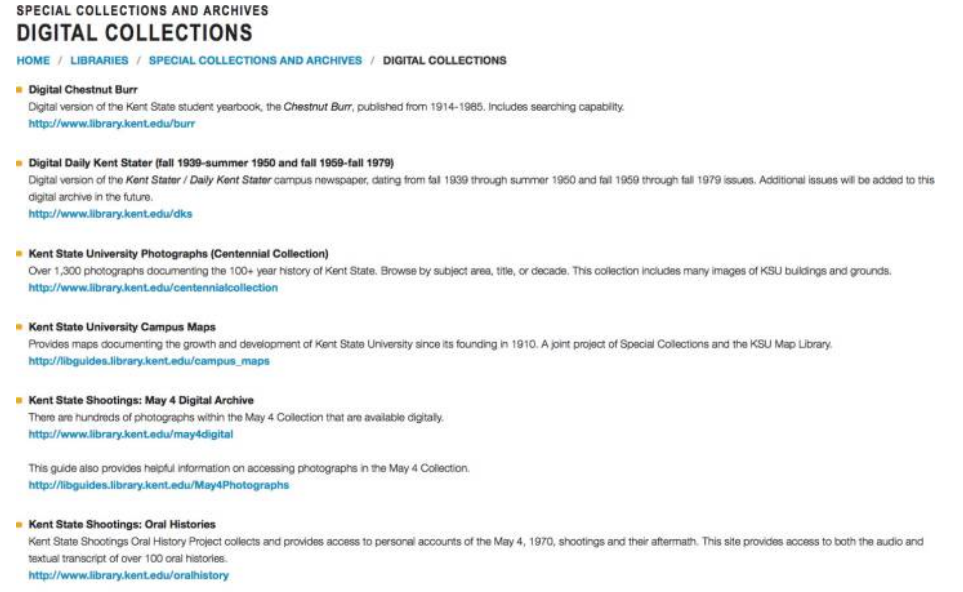
Methodology

The internal team at KSU looked to the Library of Congress outline mentioned above in the literature review to begin the process of review and planning. The first step in the outline is a process of identification, and also one of definition, as to the scope of what digital collections encompass at the university library (or, possibly an easier question to answer at this point, is what it does not entail). Many of the existing digital collections were centred mainly on special collections and archives materials as the starting point (which is likely a common scenario for the genesis of many similar digital collections), though KSU's team wanted to be sure at this point in the review to step back and consider digital projects at a wider scale during this first stage. This process may not have previously been considered in regard to non-typical digital content, when collections first begin to develop and this digital content may have been overlooked. This consideration included other library collections within and outside of the main library collection, a process that engaged with outside departments for input as well.

One clear goal for the institution at the onset of the first step in the review was to recommend a working group to define the mission and scope of the digital repository that had not been present in previous work at the library. Digital collections at KSU are currently encompassed within a home-grown application that has some peculiar features. Collections are put in silos within separate MySQL databases and cannot be cross-searched, as well as other disadvantages further outlined below. From the start of the review process, the working group needed to address these (and other) crucial issues involving both long-term and short-term goals. The group also provided feedback from other viewpoints of multiple department and service points that had not previously been considered. Staff members assembled in the working group has, thus far, provided many opinions and thoughts, providing a broader perspective more representative of the university at large.

After the creation of the working group, current digital collections were inventoried into a spreadsheet with scope details to clearly define each collection. During this inventory process, conversations arose that proved to solicit ideas for future projects, as one unexpected result. Potential candidates from the list were marked for transfer into the more permanent content management solution proposed below. In other scenarios, some projects were selected to remain on their current platforms for various reasons relating mainly to vendor and contract commitments already in place. The exclusions here were mainly the student newspaper and yearbook digital collections that were hosted on Veridian and internet archive, respectively). Specifically, this step also provided a place to start brainstorming with staff as informal tours and conversations took place to start taking note of potential future projects and collaborations. This inadvertent side note of the identification process proved to be valuable in identifying future needs and trends of which the institution was currently unaware.

Digital collections at the library were either organized on external hosted platforms or an internally hosted library web page (Figure 1) and, in one case, on a LibGuide

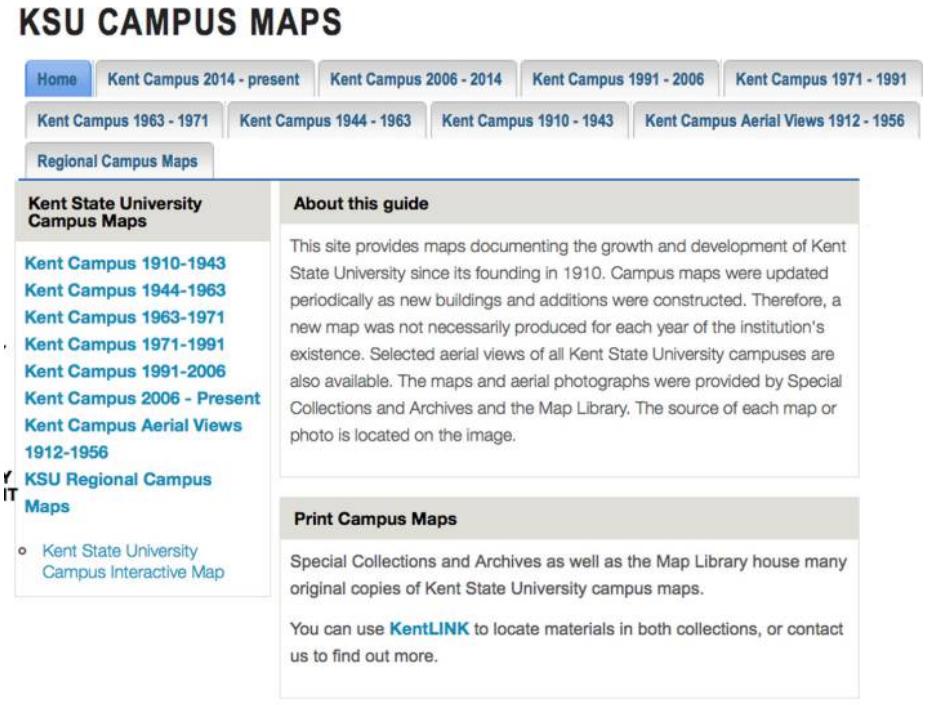


Source: www.library.kent.edu/page/16454 (accessed 10 February 2014)

Figure 1.
Current interface for digital collections at Kent State University

(Figure 2). The internally hosted collections have a tailored backend interface for single-object ingest by content specialists and were heavily tailored in each collection by the library development team. For the internally hosted collections, metadata was entered into customized web forms and then stored in a MySQL database. Larger image files were compressed or resized into smaller files before ingestion if the original file was too large (90 MB or higher), and ultimately pointed to a system that focused on access to lower-quality content over providing full resolution files when available. This issue is just one technical limitation that was quickly hindering the production and ingestion of content, but illustrated the need to review and assess current infrastructure. Further, visitors to the internally hosted digital collections could search each collection individually, but there was no feature available to search across the collections, and some of these identified digital collections had a similar content base that would be greatly improved with an overarching search mechanism (the May 4 Digital Archive and Oral History collections, in particular in KSU's scenario).

The second step in the Library of Congress outline was to isolate the collections needing preservation. In particular, this included an in-depth consideration and sorting out of the recently installed Digital Commons instance at KSU. One clear goal from the outset was to define what types of content would be present in each platform solution. One solution is in the arena of Digital Commons/Bepress, whose focus is on faculty and university publications that is primarily a presentation platform. Conversely, the other architecture solution proposed would be on a platform that can handle a wider variety of formats, which require a more long-term approach on a solid digital preservation platform, and also provide some options for user interface, display and access. For example, there may be a project that includes



Source: http://libguides.library.kent.edu/campus_maps (accessed 20 February 2014)

Figure 2.
LibGuide for the campus maps digital collection

not only large, high-quality video content, but also has major copyright restrictions on use that the latter option would better address in this scenario (note to reader: this example is taken from a current project with these specifications). A plan for migration will need to take place once a new platform solution is in place and the current metadata will need to be mapped into a selected metadata schema. This step will ideally include obtaining the full resolution image files from the existing collections to ingest into a new content management system (Figure 3).

The third step of the Library of Congress digital preservation outline involved a proposal for the implementation of a digital repository platform as a solution for a long-term content management solution, which is currently under review at KSU. The proposal recommends the open-source Fedora repository platform for the long-term storage solution for digital media (http://fedorarepository.org). By creating a central storage location, the digital collections could be joined into one backend solution. The recommendation further suggests using either the Omeka (http://omeka.org) or Hydra (http://projecthydra.org) option to provide multiple choices in a variety of flexible user interfaces that can be built upon the Fedora infrastructure. The newest release of Fedora 4 has ingrained key tenants of digital preservation into its core development, such as file fixity and auditing features. In addition, the new release is poised to adapt to linked data initiatives and advanced storage potential. While the adoption of Fedora 4 is recommended, the development of the necessary updates to either the Hydra or Omeka

Figure 3.
Example of collection page with search and browse features limited to an individual collection

components will likely be delayed while programmers contributing to these open-source options make the mandatory fixes to ensure the various components will work fluidly together. In addition to Fedora, our team also looked at DSpace, ContentDM and the VITAL digital asset management system. Ultimately, Fedora fit many of the key tenets of the desired extensible system and the ability to create custom interfaces.

The remaining steps will be worked into the overall plan as the team moves forward with design and implementation for a digital repository in the coming months here at our institution, and many of these steps are addressed with the configuration and operation of the Fedora repository at the heart of its structure and mantra. The open archival information system model is inherent to the Fedora underpinnings, which are crucial in the building of a sustainable digital repository that will be agile over time. The key to the project will be to have a system that is capable of building out different interfaces that can adapt to the changing technologies and subsequent future file formats that will undoubtedly be called for and needed down the road. Having a core structure that is efficient for evolving interfaces and adaptable features is at the heart of the plan for a digital repository. The main goal of the migration will be to enhance and improve the current scenario and provide a robust platform that can adapt with emerging file formats and provide enhancements, such as cross-collection searching and making improvements to the user experience through the interface. As such, Fedora by design and structure has proven to be a sustainable, adaptable platform solution that fits our particular institution's needs, though the scalability of the product also makes this a viable solution for many other scenarios.

Conclusion

During the initial investigation at our institution, some other sets of priorities began to both emerge and evolve to help define the individual steps and goals to enable the process of building a sustainable digital repository for the future. This process began by examining the granular level of operations throughout the university library, which quickly took shape into a broader, more encompassing scope and plan for the university

digital repository. A need to map the architecture for digital projects began to arise and move forward at the university library, one that is both flexible in nature and adaptable to a growing repository was essential to lay the groundwork and foundation of our digital projects. Our institution will require a system that is both agile and accommodating for a growing digital repository that is capable of substantial growth and development over time.

Over the coming year, the KSU development team will begin to build and migrate existing digital collections to a more dynamic platform, while the newly built workflows and staff members will continue to develop and evolve as new content and formats are created and submitted. This process proved to be a valuable one that should be done routinely, in part as a way to check if the existing infrastructure is both sustainable and meeting expectations, but also to generate ideas for enhancing the repository at both a platform level and at the surface. A strong recommendation is made to make this review a continual one to avoid creating a stagnant system. The process also involved many individuals across departments and succeeded in making connections for new collections and projects. The actions involved in the process described above could easily be used by digital initiatives at any stage of production, regardless of the shape or size of the institution or staffing.

It is often a very difficult task to look forward with any degree of certainty and clarity when it comes to anything related to technology planning. The frameworks and workflows that are put in place today can be quickly outdated and uprooted by new advancements and tools. One fear that is common to those involved with planning and development of digital repositories is to make a decision and select a choice that is both sustainable and adaptable to the best of your ability. Migration and upgrades over time are quite inevitable in the world of digital media for a number of reasons. We can strive to make well-informed decisions with the information we have available today, but it becomes problematic when the future may hold the solutions for many of the current issues we face today. The best one can do is to create a system that is adaptable to new challenges and obstacles, but also create a method of checks and balances to re-visit and review solutions periodically to ensure that needs are being addressed.

References

Awre, C. (2012), "Hydra UK: flexible repository solutions to meet varied needs", *Ariadne*, No. 70.

Novara, E. (2010), "Digitization and researcher demand: digital imaging workflows at the University of Maryland Libraries", *OCLC Systems & Services: International Digital Library Perspectives*, Vol. 26 No. 3, pp. 166-176.

Reese, T. (2013), *OSU Digital Initiatives Program Guiding Principles and You*, available at: http://library.osu.edu/blogs/digitalscholarship/2013/09/05/osu-digital-initiatives-program-guiding-principles-and-you/ (accessed 22 November 2014).

Further reading

Calhoun, K. (2014), *Exploring Digital Libraries: Foundations, Practice, Prospects*, Neal-Schuman, Chicago, IL.

Cervone, H.F. (2012), "Understanding the elements of a digital library project plan", *OCLC Systems & Services: International Digital Library Perspectives*, Vol. 28 No. 4, pp. 176-179.

Joubert, M. and Wishart, J. (2012), "Participatory practices: lessons learnt from two initiatives using online digital technologies to build knowledge", *Computers & Education*, Vol. 59 No. 1, pp. 110-119.

Simons, N. and Richardson, J. (2013), *New Content in Digital Repositories: The Changing Research Landscape*, Chandos Publishing, Oxford.

About the author

Virginia A. Dressler is currently the Digital Projects Librarian at Kent State University. Virginia has over 10 years of experience working with digital projects in both academic libraries and museums. She holds a Master's in Library and Information Science from Kent State University and a Master's in Art Gallery and Museum Studies from the University of Leeds. Virginia A. Dressler can be contacted at: vdressle@kent.edu