

Narrative

1. Project Justification

The Miami University Center for Digital Scholarship proposes a project to build a Prototype Digital Scholarship Applications dashboard that will allow Miami faculty and scholars to easily and quickly choose, configure and deploy any of a set of open source or commercially licensed web applications in a shared hosting environment. The service envisioned would be similar to those offered by commercial hosting providers for deployment of popular web applications, but would focus on tools more relevant to digital scholarship. This proposal supports the costs of system infrastructure and application portal development, creation of deployable web applications and associated documentation and video tutorials, and to support a one-day user group workshop.

Statement of Need

Centers for Digital Scholarship are an emerging service model in academic libraries that brings faculty and student scholars, technologists and librarians together to collaboratively develop digital projects supporting scholarship and research. A significant barrier faced by these collaborations is a tension between the needs of campus IT organizations to provide a homogeneous software and service portfolio to the broad campus community and the need for wild experimentation and exploration often required by faculty and students involved in digital scholarship projects.

In part, Digital Scholarship Centers have emerged as a possible solution to this tension. These centers within academic libraries often provide a forum for collaboration among librarians, technologists and both faculty and student scholars pursuing a digital project to support their research interests and may also help identify infrastructure resources using informal methods. Indeed, Digital Scholarship Centers excel in leveraging existing campus relationships to locate resources to accomplish client projects.

This is true at Miami University, in the Libraries' Center for Digital Scholarship. Our mission is to serve as a collaborative partner with faculty, students, and staff by providing digital library, data repository, multi-media, digitization, scholarly communication, geospatial and data management services so that members of the Miami community can accomplish their research, scholarly, and teaching goals. Staff in the Center leverage long term relationships with Library IT staff, Infrastructure providers in the College of Engineering, and others to locate and acquire hardware, network, and storage infrastructure with spare capacity. This may involve finding a mothballed server in one department, tasking a Virtual Machine from another provider and making heavy use of open source software to help accomplish a faculty or student scholar's project. By leveraging these existing relationships to "make a project just work" we are extending, as Lippincott describes in *Trends in Digital Scholarship Centers*, "the ways in which librarians and academic computing professionals relate to and work with faculty (and often advanced students) and their scholarly practices (Lippincott, 2014)¹.

¹ <http://www.educause.edu/ero/article/trends-digital-scholarship-centers>

Projects and activities include such scenarios as:

- Exploring the capabilities of MediaWiki to create a collaboratively edited publication.
- Creating an assignment where students digitize, organize, and create a digital exhibit of historical documents using Omeka, a popular web-based digital collection tool in the Humanities.
- Working with graduate students to create a website to supplement a print book project using WordPress.
- Creating an online, open access journal to expose students to the scholarly communication process using Open Journal System.
- Working with capstone students to test/implement changes to the source code of an interactive storytelling software such as ARIS.
- Working with student scholars to create a website to support their capstone experience.
- Creating an online visualization of spatial data in an interactive map.
- Accessing a traditional Unix command line shell to use command-line text analysis tools.

Successful collaborative projects like these with faculty and student scholars require exploration and experimentation with different technologies to identify the best match for any given project. That often means that some solutions are explored, found lacking and quickly discarded in favor of another approach. Herein lies the problem. Most academic IT environments, out of necessity, focus efforts on infrastructure and policy frameworks to serve the broadest audience with the greatest stability and reliability campus-wide. This often precludes the kind of experimentation which is a hallmark of the collaborations digital scholarship centers seek to foster.

Expanding on the above anecdotal evidence, this project also builds upon the results of a previous collaboration with the Ohio Library and Information Network and Bowling Green State University on the NEH funded *Scholars Dashboard* project. That project resulted in three workshops that brought together librarians, technologists, and scholars to discuss current practice in digital scholarship and desired functionality in next generation digital library systems to support that scholarship. The resulting report described that desired functionality. However, the workshops themselves, and the frank sharing of actual digital scholarship practice, revealed that scholars engage in frequent experimentation with a wide variety of technology tools, adapting some and discarding others. Often, these tools were not originally designed for scholarly use and were not a complete fit for the scholar's needs. In most cases, they were good enough. One thread expressed during the workshops was a need for any system to be flexible enough to allow scholars to build tools that could not easily be defined or even imagined in advance. One of the original goals of that project was to conduct the workshops as "Design and Build" sessions where IT professionals would respond to requirements and feedback from the participants and begin the design of prototype systems and tools. Unfortunately, that goal was hampered by the same problems described earlier. Infrastructure systems could not be spun up, configured and experimented on within a reasonable timeframe. IT organizations, legitimately, need to evaluate any proposed software or hardware solution to determine the long term sustainability of the solution, potential security and support issues, and licensing costs to name just a few. IT organizations are optimized to handle these concerns with rigorous policy review, security audits, and change management procedures but these processes take time.

One solution is to allow faculty and student scholars to experiment with tools within a carefully controlled “sandbox” environment or within a desktop hosted virtualization container such as VMware, Parallels, or Oracle’s open source Virtualbox tool. In fact, this latter solution has been utilized in workshops offered by the Miami Center for Digital Scholarship in the past. Participants installed Virtualbox on a personal laptop and then activated a pre-configured virtual machine in a LAMP (Linux, Apache, MySQL, PHP) environment. From this, participants could install Omeka, WordPress and other web-based applications. The problem of course was that participants wanted to see how these tools worked in a real-life, Internet connected scenario. The solution was also only adequate for short workshops and not the development of longer term projects. What is needed instead is a system that leverages the stability and reliability provided by Campus IT organizations while providing a rich environment for rapid prototyping and experimentation.

Target Audience

While the initial project would benefit faculty and student scholars, IT professionals and digital scholarship librarians at Miami University, the techniques developed for curating software environments, deploying those environments on demand, and providing end-user documentation would broadly benefit information organizations seeking to provide a sustainable yet flexible environment to support digital scholarship. In addition to traditional white papers and reports, the project team intends to provide technical specifications, pre-configured machine images and a library of self-paced point of need documentation and tutorials. This audience is only expected to grow. As Lippincott observes, “Most U.S. universities and colleges do not yet have a digital scholarship center, and most existing centers were established only in the past few years. We expect to see an increasing number in the near future” (Lippincott, 2014).

Intended Results

The intended outcome of this proposal is to provide a framework for academic libraries to better support digital scholarship activities on their campus by removing barriers facing faculty and student scholars. Current research and best practice in digital scholarship supports the development of a much more sustainable and scalable environment, which may not conform to the requirements of traditional campus IT infrastructure deployment or to the lone researcher installing systems on their own hardware. For example, in the chapter *A Sustainable Repository Infrastructure for Digital Humanities: The DHO Experience* (Gourley & Battino, 2010)², the authors indicated that academic institutions planning to support digital projects must create an infrastructure that is easily sustainable on its own, using dynamic, scalable, and cost-effective hosting and development resources; they also recommended offering environments for creating rapid web site applications that can empower stakeholders to do as much customization as they are able using friendly and non-coding web administration interfaces. Further, as an emerging service model in academic libraries, Centers for Digital Scholarship are poised to provide leadership and direction on how to develop innovative models of engagement for libraries, information technology units, and academic departments. Dynamic and innovative digital scholarship projects have the potential to support and encourage new directions in research, teaching, and learning as well as provide an infrastructure that encourages experimentation in new areas of scholarship.

² http://link.springer.com/chapter/10.1007%2F978-3-642-16873-4_38

2. Project Work Plan

The Project will begin October 1, 2015 and end on September 30, 2016.

Project Personnel and Project Management. This project is a joint effort of the Miami University Center for Digital Scholarship and the Miami Information Technology Services. The Libraries have a history of working well together on joint technology partnership. One of the latest being efforts to migrate Library systems and technology to virtualized computing and storage infrastructure, including the Library Catalog, Digital Collections and Institutional Repository hardware, to the IT Services data center. Key personnel include:

John Millard is the Head of the Miami University Libraries Center for Digital Scholarship. His research interests and activities focus on developing digital library systems to serve distinct user communities. He is a former NASA and USGS funded investigator working for the OhioView project, a congressionally sponsored project to further the use of U.S. civilian satellite data by the public. He serves as an adjunct faculty member in the interactive media studies program where he teaches IMS201: *Information Studies in the Digital Age* and the forthcoming IMS203: *Scholarship in the Digital Age*.

Elias Tzoc is the Digital Initiatives Librarian at Miami University Libraries. His recent work includes: developing and prototyping web interfaces for digital projects using CONTENTdm, DSpace, WordPress, OJS, and Omeka; researching new access points and mobile apps for digital library programs using the jQuery mobile framework and Responsive Web Design standards; researching and publishing on technical issues and open source apps for libraries; developing web plug-ins using PHP, HTML5, XSLT, CSS, and jQuery. He collaborates extensively with faculty and student scholars on Digital Scholarship projects in the Center for Digital Scholarship.

Brian Henebrey is the Associate Director of Enterprise Systems and Operations at Miami University. He provides leadership and accountability in the development, implementation, maintenance and retirement of tier one infrastructure supporting the University's academic and administrative mission and objectives. He manages strategic operations involving cross platform systems, servers, services, storage, middleware and databases that require high attention to standardized practices, automation and monitoring.

The project will be split into 4 phases. The first phase will be focused on planning, evaluating a starting list of software environments to provide and conducting a one day workshop with project end-user participants. Phase 2 will consist of software development and deployment of hardware to support the project. Phase 3 will include the development and deployment of the on-demand dashboard and deployment of initial virtual machine images. Phase 4 will focus on dissemination of project results to the broader user community.

Phase 1: Planning (October 2015-December 2015)

The goal of phase one is to validate the initial list of software environments by gaining feedback from librarians and technologists and by conducting a one-day workshop with Miami faculty scholars. Participants will be recruited from the cohort of faculty engaged in digital scholarship projects already in the center, interested faculty nominated by subject liaison librarians, and library and project staff. The workshop will involve discussions of individual research interests and scholarly practice as well as the features of existing tools

that may apply to individual projects. Specific use cases and requirements will be identified and documented and will inform the decisions of what additional software environments to create and how those environments should be configured. Tzoc and Millard will lead this effort.

Phase 2: Software Development (January 2015 - March 2015)

During phase 2 of the project, IT Services will purchase and integrate the required virtualization server hardware into the existing data center. In order to meet the needs of the dashboard project, the infrastructure will need to leverage existing virtualization technologies present in the data center; in particular, the VMWare virtualization framework already present. Leveraging this existing investment also allows for storage and backup services to be integrated easily into the project. Once the virtualization environment is installed and configured, IT Services will create a general purpose operating system blank image. Libraries' team members will use this blank image to create customized environments for each of the validated software tools. These customized images will in turn be made available to IT Services for deployment in the portal. This activity will be led by Henebrey. Phase 2 activities also include creation/curation of end-user documentation and tutorials of each selected software environment. Short video tutorials will be created explaining how each tool can be used and possible applications beyond those for which it was originally intended as well as in basic use and configuration of each tool. Tutorials will also be created on the use of the Dashboard itself including how to search for, select and deploy a virtual machine as well as how to gain access to that VM once deployed. Henebrey will lead efforts on the IT Services side while Millard and Tzoc will provide configuration of customized virtual machines and create documentation and video tutorials.

Phase 3: Deployment and Evaluation (April 2015 - September 2015)

In order to meet the needs of the Scholars Dashboard the infrastructure will leverage existing virtualization technologies in place at Miami University, currently VMware and supporting suite of tools. These tools provide a foundation to provide the following:

- 1) Access portal (browser based) to select and deploy pre-configured applications
- 2) Catalog of available virtual machines
- 3) End-user management for access and authorization
- 4) Administrative management to allow for adding and removing available resources
- 5) Self service by end-users without administrative intervention
- 6) Reporting and trending

During this phase, the Dashboard service will be developed and deployed to the Miami community and other stakeholders and workshop participants. The Dashboard will be used by students in a credit class IMS 203: Digital Scholarship in the Digital Age taught by Millard. Students will deploy web applications from the Dashboard catalog to 1) implement specific class assignments and 2) compare and contrast the features and capabilities of multiple software tools. The dashboard will also be used and tested with existing clients in the Center for Digital Scholarship. Throughout Phase 3, effectiveness and use of the dashboard will be measured through automatically collected usage and resource allocation reporting and qualitatively measured through end-user surveys.

Phase 4: Dissemination (June 2015 - September 2015)

Activities in Phase 4 support dissemination of project results. Activities include developing an end-user project website housing project documentation, findings, whitepapers, and catalogs of provided virtual applications. A github repository will be established for dissemination of project code and virtual machine images. Investigators will seek to present at appropriate Library conferences and venues as well as submit peer reviewed publications. Tzoc and Millard will lead this effort.

Budget Resources and Rationale

Project budget is \$25,000 which covers direct expenses of server hardware, video production, student assistants and a modest travel budget to support project dissemination. Personnel costs of lead investigators and key personnel to support the project are being supported by existing institutional budgets.

3. Project Results

As mentioned in the Justification section of this proposal, recent studies such as the 2014 EDUCAUSE report indicates that there is a generalized expectation that the number of Digital Scholarship Centers will not only continue to grow but these centers will also need to adjust their priorities and services as the nature and requirements for digital scholarship projects continue to evolve. We believe our proposal has the potential to cost-effectively change the way faculty, librarians, and technologists discover, browse, and deploy curated web applications for their digital scholarship projects. In addition, the project white paper and documentation will provide instruction and guidance required to replicate the model at other institutions.

The dissemination of results will be carried out on two phases. For the short term, the activities in our proposal have a measurable outcome. At the end of the grant period, we intend to work and develop at least 5 projects with existing and future faculty clients. The IMS 203 class will allow us to test the system for a class setting with 20+ students. Additionally, workshops with faculty and researchers will allow us to fully test and provide a list of pre-selected software applications, which will include short descriptions and examples of possible uses in digital scholarship projects. For the mid-long term, the project success can be measured by the number of successful conference presentations (e.g. Code4Lib, Brick and Click, ALA) and peer-review publications (e.g. The Journal of Academic Librarianship, D-Lib Magazine) about the project, the number of views/downloads/inquires on our project website and, perhaps most importantly, the number of faculty requests and new projects and collaborations initiated.

Lastly, while the initial project would only benefit faculty and student scholars, IT professionals and digital scholarship librarians at Miami University, the techniques to be developed for curating software environments, deploying those environments on-demand, and providing end-user documentation would broadly benefit information organizations seeking to provide a sustainable yet flexible environment to support digital scholarship.

DIGITAL STEWARDSHIP SUPPLEMENTARY INFORMATION FORM

Introduction:

IMLS is committed to expanding public access to IMLS-funded research, data and other digital products: the assets you create with IMLS funding require careful stewardship to protect and enhance their value. They should be freely and readily available for use and re-use by libraries, archives, museums and the public. Applying these principles to the development of digital products is not straightforward; because technology is dynamic and because we do not want to inhibit innovation, IMLS does not want to prescribe set standards and best practices that would certainly become quickly outdated. Instead, IMLS defines the outcomes your projects should achieve in a series of questions; your answers are used by IMLS staff and by expert peer reviewers to evaluate your proposal; and they will play a critical role in determining whether your grant will be funded. Together, your answers will comprise the basis for a work plan for your project, as they will address all the major components of the development process.

Instructions:

If you propose to create any type of digital product as part of your proposal, you must complete this form. IMLS defines digital products very broadly. If you are developing anything through the use of information technology – e.g., digital collections, web resources, metadata, software, data– you should assume that you need to complete this form.

Please indicate which of the following digital products you will create or collect during your project.

Check all that apply:

	Every proposal creating a digital product should complete ...	Part I
	If your project will create or collect ...	Then you should complete ...
<input type="checkbox"/>	Digital content	Part II
<input checked="" type="checkbox"/>	New software tools or applications	Part III
<input type="checkbox"/>	A digital research dataset	Part IV

PART I.

A. Copyright and Intellectual Property Rights

We expect applicants to make federally funded work products widely available and usable through strategies such as publishing in open-access journals, depositing works in institutional or discipline-based repositories, and using non-restrictive licenses such as a Creative Commons license.

A.1 What will be the copyright or intellectual property status of the content you intend to create? Will you assign a Creative Commons license to the content? If so, which license will it be? <http://us.creativecommons.org/>

Materials created will be licensed under a Creative Commons Attribution 4.0 International license, unless prohibited by University Policy or State or Federal law.

A.2 What ownership rights will your organization assert over the new digital content, and what conditions will you impose on access and use? Explain any terms of access and conditions of use, why they are justifiable, and how you will notify potential users of the digital resources.

Software product, where possible, will be made available freely for download with no additional restrictions on use other than those imposed by the CC license.

A.3 Will you create any content or products which may involve privacy concerns, require obtaining permissions or rights, or raise any cultural sensitivities? If so, please describe the issues and how you plan to address them.

No

Part II: Projects Creating Digital Content

A. Creating New Digital Content

A.1 Describe the digital content you will create and the quantities of each type and format you will use.

A.2 List the equipment and software that you will use to create the content or the name of the service provider who will perform the work.

A.3 List all the digital file formats (e.g., XML, TIFF, MPEG) you plan to create, along with the relevant information on the appropriate quality standards (e.g., resolution, sampling rate, pixel dimensions).

B. Digital Workflow and Asset Maintenance/Preservation

B.1 Describe your quality control plan (i.e., how you will monitor and evaluate your workflow and products).

B.2 Describe your plan for preserving and maintaining digital assets during and after the grant period (e.g., storage systems, shared repositories, technical documentation, migration planning, commitment of organizational funding for these purposes). Please note: Storage and publication after the end of the grant period may be an allowable cost.

C. Metadata

C.1 Describe how you will produce metadata (e.g., technical, descriptive, administrative, preservation). Specify which standards you will use for the metadata structure (e.g., MARC, Dublin Core, Encoded Archival Description, PBCore, PREMIS) and metadata content (e.g., thesauri).

C.2 Explain your strategy for preserving and maintaining metadata created and/or collected during your project and after the grant period.

C.3 Explain what metadata sharing and/or other strategies you will use to facilitate widespread discovery and use of the digital content created during your project (e.g., an Advanced Programming Interface, contributions to the DPLA or other support to allow batch queries and retrieval of metadata).

D. Access and Use

D.1 Describe how you will make the digital content available to the public. Include details such as the delivery strategy (e.g., openly available online, available to specified audiences) and underlying hardware/software platforms and infrastructure (e.g., specific digital repository software or leased services, accessibility via standard web browsers, requirements for special software tools in order to use the content).

D.2 Provide URL(s) for any examples of previous digital collections or content your organization has created.

Part III. Projects Creating New Software Tools or Applications

A. General Information

A.1 Describe the software tool or electronic system you intend to create, including a summary of the major functions it will perform and the intended primary audience(s) the system or tool will serve.

In order to meet the needs of the trial servers and dashboard the infrastructure will need to leverage existing virtualization technologies. Currently three organizations have products available that should be considered: 1) VMware and supporting suite of tools, 2) Microsoft Hyper-V, 3) Openstack suite of tools.

Each of these options would be used to provide a foundation to provide the trial for 35 Participants (including , Students and Staff) the following:

Access portal (browser based) to select and deploy virtual machines on demand from a defined catalog

Access portal (browser based) to select and deploy pre-configured application servers on demand from a defined catalog

Catalog of available virtual machines

End-user management for access and authorization

Administrative management to allow for adding and removing available resources and server templates

Self service by end-users without administrative intervention

Reporting and trending

A.2 List other existing digital tools that wholly or partially perform the same functions, and explain how the tool or system you will create is different.

Oracle VM and Redhat Enterprise Virtualization suites are similar in nature to the proposed solutions but lack Dashboard feature sets and wide acceptance in the industry.

B. Technical Information

B.1 List the programming languages, platforms, software, or other applications you will use to create your new digital content.

Infrastructure: VMware vSphere, VMware Automation Center, Microsoft Hyper-V, OpenStack, Oracle Enterprise Linux

Applications: MediaWiki, Omeka, WordPress, Open Journal System, ARIS

B.2 Describe how the intended software or system will extend or interoperate with other existing software applications or systems.

The dashboard infrastructure will provide a portal for participants in the trial to access content and self serve applications and systems on an as-needed basis.

B.3 Describe any underlying additional software or system dependencies necessary to run the new software or system you will create.

Existing infrastructure may be accessed to provide storage and back-up needs. Storage utilizes IBM SVC to provide disk space, back-up is provided through IBM Tivoli.

B.4 Describe the processes you will use for development documentation and for maintaining and updating technical documentation for users of the software or system.

ITIL v3

B.5 Provide URL(s) for examples of any previous software tools or systems your organization has created.

N/A

C. Access and Use

C.1 We expect applicants seeking federal funds for software or system development to develop and release these products as open source software. What ownership rights will your organization assert over the new software or system, and what conditions will you impose on the access and use of this product? Explain any terms of access and conditions of use, why these terms or conditions are justifiable, and how you will notify potential users of the software or system.

Unless prohibited by Federal/state law or University Policy, products created will be released under an open source/access license.

C.2 Describe how you will make the software or system available to the public and/or its intended users.

As a prototype, the system will be available to participants on campus and accessible using enterprise authentication and authorization methods

Part IV. Projects Creating Research Data

1. Summarize the intended purpose of the research, the type of data to be collected or generated, the method for collection or generation, the approximate dates or frequency when the data will be generated or collected, and the intended use of the data collected.

2. Does the proposed research activity require approval by any internal review panel or institutional review board (IRB)? If so, has the proposed research activity already been approved? If not, what is your plan for securing approval?

3. Will you collect any personally identifiable information (PII) about individuals or proprietary information about organizations? If so, detail the specific steps you will take to protect such information while you prepare the research data files for public release (e.g. data anonymization, suppression of personally identifiable information, synthetic data).

4. If you will collect additional documentation such as consent agreements along with the data, describe plans for preserving the documentation and ensuring that its relationship to the collected data is maintained.

5. What will you use to collect or generate the data? Provide details about any technical requirements or dependencies that would be necessary for understanding, retrieving, displaying, or processing the dataset(s).

6. What documentation will you capture or create along with the dataset(s)? What standards or schema will you use? Where will the documentation be stored, and in what format(s)? How will you permanently associate and manage the documentation with the dataset(s) it describes?

7. What is the plan for archiving, managing, and disseminating data after the completion of research activity?

8. Identify where you will be publicly depositing dataset(s):

Name of repository: _____

URL: _____

9. When and how frequently will you review this data management plan? How will the implementation be monitored?