



IMPLEMENTATION OF A MODERN DIGITAL LIBRARY AT THE OHIO STATE UNIVERSITY LIBRARIES

White Paper

Submitted by the Strategic Digital Initiatives Working Group

April 15, 2014

Table of Contents

Executive Summary.....	2
Introduction	3
OSUL Digital Initiatives Program Guiding Principles	3
Implementation of the OSUL Digital Initiatives Program	4
Identification of Key Gaps.....	6
OSUL Digital Initiatives Program Guiding Principles	7
Implementation of the OSUL digital initiatives program	8
Functions.....	8
Digital Library Core Components Matrix	8
Core Component – Access:	9
Core Component – Content Selection/Management	10
Core Component – Preservation.....	13
Core Component – Interoperability	14
Architecture Development and Standards	15
Key Policies.....	17
Identification of Key Gaps.....	18
Recommendations	21
Appendix A: Ohio State University Libraries Digital Initiatives Program Guiding Principles	24
Principles:.....	24
Appendix B: DIGITAL PRESERVATION POLICY FRAMEWORK	28
References:	36

Executive Summary

The OSUL Executive Committee charged the Strategic Digital Initiatives Working Group with evaluating the Libraries' existing digital libraries environment and developing a set of guidelines to enable the Libraries to adopt a more modernized digital library environment. As part of this process, the SDIWG developed two documents: A set of *Guiding Principles* (Appendix A) that can be used to encourage a more uniform approach to digital project and infrastructure development and this formal *White Paper* outlining the core functional elements that need to be supported to enable OSUL to move its digital libraries program forward.

The *White Paper* is divided into multiple sections, providing the *Guiding Principles*, information about the Core Component Matrix, an evaluation of the Key Policies, an identification of gaps that need to be addressed, and finally, recommendations. **This *White Paper* does not make specific technical infrastructure recommendations,** but rather proposes a framework from which those decisions could be made to best support a flexible and innovative digital environment at OSUL.

Introduction

The Ohio State University Libraries (OSUL) has a long history of creating, digitizing, and supporting digital projects and collections. Projects like the Knowledge Bank and the Billy Ireland Image Database are just a couple examples of OSUL's long-standing commitment to making the Libraries' unique materials more accessible to the campus and global communities. As a result, much of the Libraries' current digital infrastructure, services, and policies support workflows and an environment that have traditionally been project driven. This approach has led to the implementation of a fragmented digital libraries environment that has become increasingly difficult to support and grow as the Libraries' digital portfolio has expanded.

In April 2013, the OSUL Executive Committee formed the Strategic Digital Initiatives Working Group (SDIWG). The group was created to help OSUL respond to the changes that have occurred in the digital libraries landscape since OSUL first started producing digital content. The SDIWG was charged with understanding how digital libraries have changed, and to develop a framework and recommendations to help OSUL's digital library program evolve to meet current and future needs.

In order to address this task, the SDIWG performed an environmental scan of OSUL's current digital environment, taking inventory of the Libraries' current policies and procedures. Using that information, the SDIWG identified two areas of significant need: 1) a set of principles around project development that shifts the emphasis from that of a project-centric model to more of a program-centric model, the *Guiding Principles*, and 2) this paper, a framework to redefine OSUL's digital libraries' infrastructure to provide key recommendations to enable the Libraries to evolve a more modern digital library program.

Finally, given the fast changing nature of technology, this *White Paper* has been developed as a tool to guide the development of the OSUL digital library program. The *White Paper* will be divided into three key areas: *OSUL Digital Initiatives Program Guiding Principles*, *Implementation of the OSUL digital initiatives program*, and *the Identification of key gaps*. It will conclude with a list of recommendations developed out of the SDIWG's evaluation of the current digital libraries program. A summary of the key areas follows.

OSUL Digital Initiatives Program Guiding Principles

The *Guiding Principles* represent a set of conceptual foundations undergirding the OSUL's digital library program. Evolving OSUL's digital initiatives program will require more than just a refresh of the technical infrastructure. In many ways, understanding and adapting to the technical challenges raised as part of the digital initiatives retooling effort will be the easy part of any transition. The more difficult part of this process will be facilitating a cultural change within OSUL. As noted above, OSUL's digital initiatives program has largely been driven through the development of siloed projects. This approach has enabled the Libraries early on to bring digital collections up quickly and create customized user experiences around a given

collection. At the same time, this process has had the undesirable effect of isolating the Libraries’ digital collections making it difficult to support discovery and access outside of and within OSUL. Additionally, to support this largely project centric approach, the Libraries has adopted a piecemeal approach to supporting digital projects. Decisions around what architectural components to build, borrow, or buy were largely driven by a specific project. The *Guiding Principles* seek to refocus the Libraries’ digital initiatives efforts by shifting the focus away from individual projects and refocusing them on the need to consider library infrastructure and project development as part of a singular unified program.

Implementation of the OSUL Digital Initiatives Program

Digital initiatives encompasses a wide array of topics and services, and it is important to provide some context around the specific set of topics that this *White Paper* will address. **When considering an implementation of a digital library program, the general inclination is to start by defining the technical infrastructure. The problem with this approach is that it defines a set of solutions before the organizational needs are well understood.** In some respects, this has been the process that the OSUL used to build its current digital infrastructure, selecting tools and software for specific projects, without considering the impacts within the larger context of the Libraries.

To better illustrate this concept, one needs to take a step back from the individual technological solutions, and look at the broader functional groups that make up the whole. Figure 1 is a simplified model that illustrates how these broad functional categories both frame and interact within the Libraries’ conceptualized infrastructure.

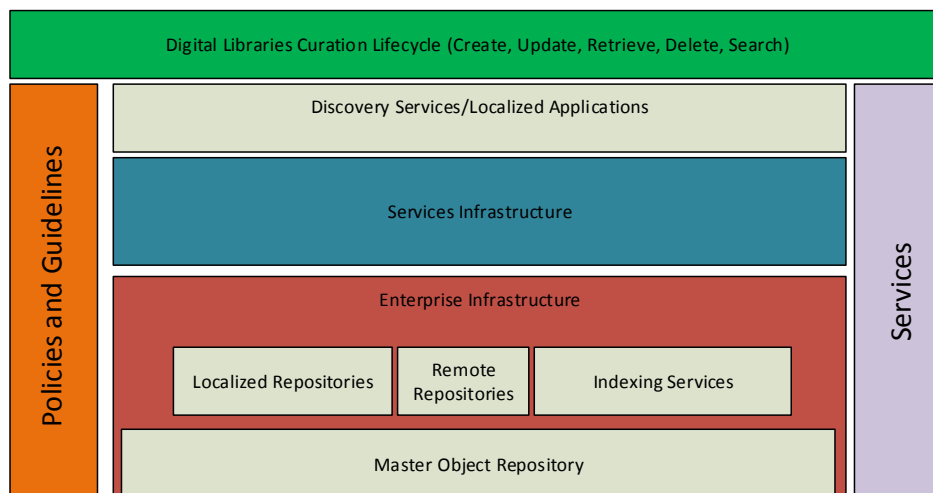


Figure 1

In the model, a number of broad categories have been defined. The inner most boxes contain Discovery/Localized Applications, Services Infrastructure, and Enterprise Infrastructure,

components that represent the Libraries' technical infrastructure and demonstrate information flows between the levels.

- **Discovery Services/Localized Application**

The Discovery Services/Applications level is where interface components are designed and developed. These tools mediate user access to the Libraries' digital collections. User-facing applications utilize the Services Infrastructure to interact indirectly with the Libraries' Enterprise Infrastructure. This abstraction isolates application development from the Enterprise Infrastructure, allowing the Libraries the flexibility to manage the Enterprise Infrastructure without fear of disrupting applications and services.

- **Services Infrastructure**

The Services Infrastructure is the glue between the Enterprise and Applications layers. The Services Infrastructure provides the application programming interfaces (API) and web services that enable both internal and external partners to access, discovery, and manage resources in the Enterprise Infrastructure. The development of a Services Infrastructure is one of the first steps towards the OSUL developing a digital library infrastructure that can be utilized as a platform for application development, rather than an amalgamation of individual components.

- **Enterprise Infrastructure**

The Enterprise Infrastructure is the level at which the underlying repositories and asset management tools are managed. These repositories can take the form of remote repositories (like the HathiTrust) and localized repositories (like Fedora or DSpace), as well as enterprise indexing services like Elastic search or Solr. As OSUL reimagines its Enterprise Infrastructure, the Libraries will likely look to adopt an environment that makes use of specialized local repositories providing for the management of metadata, objects, and indexed content. These components work together to manage access/ingest into the Master Objects Repository.

A practical example of how these three components interact together could be seen through the Libraries' interaction with the Information Library System (ILS), Sierra. The Sierra database lives within the Libraries' Enterprise Infrastructure. OSUL can query the database directly, but that carries a number of risks, because the software vendor, Innovative Interfaces, could change the structure of the database at any time, breaking any local development. To insulate libraries against these types of changes, developers interact with the data through Innovative's Sierra API—a set of procedures that Innovative's developers built for its customers and that they agree will not change over time (Services Infrastructure). OSUL's developers can build web applications that call methods in the API (such as "Get Me Bib Number 123456") to present catalog data in a user-friendly format to the Libraries' user community (Localized Applications).

Architecting and building the Services and Enterprise Infrastructures is primarily the purview of the OSUL Information Technology Division. Designing the Discovery Services and Localized Applications, on the other hand, will require conversation and collaboration across all

departments of OSUL and with input from primary constituents. This *White Paper* seeks to provide the fundamental framework within which these important technology decisions will be made.

Here, rather than focusing on technical details, this document provides a high-level conceptual model of the Libraries' digital libraries environment, namely the concepts around the edges of the diagram: policies, services, and functions that make up the OSUL program.

Identification of Key Gaps

While the working group was discussing key functional areas necessary for implementation of a successful digital library program, a number of key gaps in the current OSUL digital libraries program began to emerge. These gaps represent areas of functional or policy need. In the case of policy gaps, many also lack a formal group responsible for the development, shepherding, and implementation of key policies.

OSUL Digital Initiatives Program Guiding Principles

A number of key principles have been identified that will guide the development of the Digital Initiatives Program and the Working Group as it evaluates current and future systems, policies, services, and best practices. These principles underscore the changing nature of digital environments and the need to embrace an environment that understands and anticipates the ephemeral nature of technology while ensuring the durability of managed content. Eleven key concepts were identified and defined (Appendix A):

1. We build services, not products
2. We carefully weigh when to Build versus Buy versus Borrow
3. We develop modular services, not monolithic systems
4. We develop for change
5. We don't keep everything forever
6. We build in assessment
7. We focus on the user
8. We work with partners
9. We embrace research as a core, fundamental value
10. We strive to stay grounded in the real world
11. We are driven by standards

The *Digital Initiatives Program's Guiding Principles* provide the OSUL with a framework for making and evaluating decisions related to the development and implementation of a long-term digital services architecture. These Principles reflect the changing nature of the OSUL's digital programs, shifting the focus from individual projects to a holistic program designed around an evolving services architecture. The Principles illustrate for the Libraries and its stakeholders the core values and long-term direction of the Digital Initiatives Program.

Implementation of the OSUL digital initiatives program

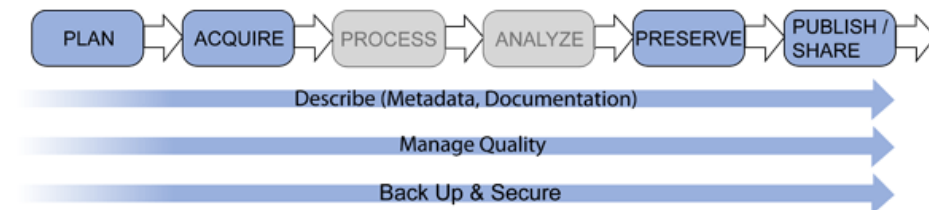
The *Guiding Principles* provide the foundation and strategic direction for ongoing digital library development at OSUL. Each guideline highlights a larger strategic goal in how the Libraries will address the long-term development of the Libraries' digital infrastructure, as well as how the Libraries select digital projects. While the guidelines provide the strategic framework for how OSUL's digital library program can move forward, the Principles do not address the functions, services, and policies that actually make up the digital library. This section of the White Paper will address key functional, service, and policy requirements necessary for implementing a modern digital initiatives program at OSUL.

Functions

The DNA of a digital library looks very much like that of a traditional library. At its core, a digital library's core functions can be summarized as the ability to:

- Collect, organize, and manage digital collections
- Discover and access information
- Support the long-term archiving and preservation of digital materials
- Provide a space for user engagement and innovation

Of course, digital libraries face a number of special challenges that make this type of simplification marginally useful. The categories can be expanded by placing the core functions within the larger context of object lifecycle management.

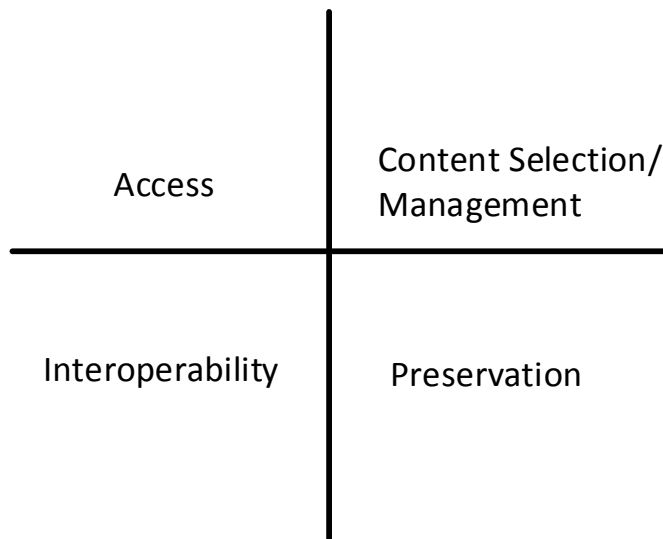


This diagram illustrates some of the complexities of processing and analyzing data, which will impact a digital library's ability to effectively search, share, and preserve its digital collections. However, this visualization doesn't provide a clear matrix for how new and existing projects and programs are evaluated as the OSUL evolves its digital library environment.

Digital Library Core Components Matrix

When examining the various curatorial and research lifecycles, core functions can be broken into the four basic components that they represent: *Access*, *Content Selection and Management*, *Interoperability*, and *Preservation*.

Digital Library Core Components Matrix



This core matrix provides the functional foundations around which roles, services, policies and technical infrastructure needed to support a more modernized digital library environment at OSUL can be defined. Additionally, the matrix provides a core vocabulary for discussing digital library development and projects at OSUL.

Core Component – **Access:**

Definition:

Drivers for making decisions to optimize discovery and use of digital content.

Within libraries, access is a widely understood concept. Libraries acquire or digitize materials and make them available. However, within the context of a digital library, especially in relation to necessary policies and services around this core component, this well understood concept encompasses a broad set of concerns.

- **Metadata** – data about data...metadata drives discovery. **Without exceptional, standards driven metadata, the ability to provide sophisticated, user-friendly discovery options becomes exceedingly difficult.** Additionally, good metadata drives present and future innovation. One present day example of this phenomenon has been the development of **faceted discovery** systems. This functionality exists today because catalogers include and utilize controlled terminology to support topical, geographic, and personal subject headings. Likewise, OSUL will seek to create and utilize the libraries' metadata to enhance the discovery experience and provide new tools and pathways to the Libraries' content.

- Discovery – the ability to find information within OSUL’s digital library from wherever a user is searching. Supporting a broad discovery experience requires the Libraries to consider policies around SEO (search engine optimization), metadata sharing, and permanent Uniform Resource Identifiers (URI) creation.
- Information Retrieval – the ability to retrieve discovered resources within OSUL’s digital library from wherever this information is found. On its face, one would assume that any digital initiatives project or system would provide adequate retrieval support. However, OSUL is being asked to provide access to more content, in more formats, from a greater diversity of discovery paths. Research faculty in the digital humanities are challenging the Libraries to provide access to not just the digital objects, but to the underlying metadata and in formats that can be machine readable to support large scale data mining efforts. The challenge for the Libraries moving forward is that as our research community increasingly seeks out access to the Libraries’ digital objects, the methods and diversity of how those materials are accessed and retrieved will continue to grow.
- Linking/Permanent Uniform Resource Identifiers (URI) – unlike most content on the web, OSUL’s digital library must strive to create durable, permanent URIs for all content made accessible through the digital library platforms, as well as support the ability to resolve to OSUL digital content through the use of well-established domain identifiers like Archival Resource Keys (ARK)[1] for archival materials or Digital Object Identifiers (DOI)[2] for publications. The need to support durable linking goes beyond the ability to support a wide range of discovery and information retrieval needs. In order for the OSUL digital library to become a trusted repository for digital content, the Libraries must be able to provide a durable linking solution to provide researchers with an assurance that data within the digital library can be cited and that users can always access the content via a system defined identifier.
- User Experience/Responsive Design – One of the most significant changes in technology over the past decade has been the rise of mobile computing. Devices such as tablets, cell phones, and phablets have redefined how users are interacting with the libraries’ resources. Likewise, this shift in user access habits underscores the fact that computing technology will continue to evolve and the libraries digital library services need to be capable of evolving with them. At the same time, issues such as ADA (American with Disabilities Act) compliance and general usability will remain constant and need to be balanced with the libraries need to support access to the fast changing consumer ecosystem.

Core Component – Content Selection/Management

Definition:

Drivers for making decisions on locally digitized or born digital content from Ohio State University or its partners.

A continual challenge faced by most digital library programs has been a lack of an organizational collection development policy around digital content. A collection development strategy for print or online materials identifies criteria to consider and articulates priorities in acquiring, processing, and storing materials. Additionally, as the Libraries expands from primarily curating digital documents and images and turns to larger media and data files, having a robust collection management and curatorial process will become even more important to successfully manage content throughout its lifecycle.

- **Selection** – Deciding what content should be digitized and what born digital materials should be collected, as well as determining the conditions of the commitment the Libraries is making to a digital object is fundamental to a well-managed digital library. As with their physical counterparts, born digital resources and digitized surrogates should be actively selected based on the Libraries' existing collection strengths, the University's research and teaching concentrations, user demand, and effectively calculated potential use. Decisions regarding when to create digital formats from analog collections, what kinds of digital objects to create, and what kinds of retention commitments to make to digital objects with analog counterparts are also important selection decisions. The collection manager's and curator's role in identifying content for acquisition or reformatting includes making judicious recommendations based on thoughtful consideration about how long a digital object should be managed, at what quality it should be maintained, what level of metadata is essential for basic retrieval and use by a user community, and other commitments that can be carried out systematically.
- **Rights Management** – Rights management remains one of the most challenging aspects around the development of a digital collection, and will be uniquely challenging for OSUL. Unlike many traditional special collections and archives found at other cultural heritage organizations, OSUL's Special Collections are largely made up of contemporary content. As one example, performance disciplines like dance and theatre include additional challenges, as rights extend beyond the principle performers, but include rights issues related to music, set design, etc. OSUL's Special Collections are principally composed of these types of challenging artifacts, complicating rights management activities. For the purpose of the digital library program, rights management is broken into two concerns:
 1. **The ability to secure preservation and access rights for specific items within a collection.** This relies on curators and collection managers to work with rights holders to ensure that OSUL has reproduction and digital access rights to a collection. Given the wide array of potential content that could be digitized by OSUL, strong collection policies specifying the importance of public access are

required to encourage curators to preference open collections for digitization.

2. **The ability to support granular access rights within the digital library's infrastructure.** While OSUL should pursue a digital collections development and selection policy that favors materials that can be made publically available, the reality is that some materials made available through the digital library will be sensitive or include limited access rights. The infrastructure will need to be able to address these varied needs by providing support for a roles based access and authentication system.
- **Distributed Access** – The ability to support distributed access rights is becoming more and more important in a world in which digital libraries are largely seen as small nodes within a much larger community. There was a time when libraries could largely concern themselves with managing access to a collection and content via a university library's website. However, this is no longer the case. OSUL is involved in a number of initiatives such as the HathiTrust and the Dance Heritage Coalition that require the Libraries to support distributed of access to content, supporting:
 1. **The ability to interact and trust credentials from partner organizations within a library's digital libraries authentication process.** An example of this type of negotiation occurs with projects like In-Commons – a shared community of Shibboleth users that allows researchers to access academic content from other campuses using their local identity and credentials.
 2. **The ability to provide distributed access to content from outside the Libraries' primary discovery mechanisms.** The Digital Public Library of America (DPLA) is a good example of how access can be distributed through a larger content aggregator. For example, the DPLA's exhibit functionality demonstrates how collections or items from collections at one institution might be utilized to create new exhibits or virtual collections within these larger aggregations.

Given the distributed and interconnected nature of many digital library efforts, the OSUL's content management strategy around digital assets needs to determine the level of involvement within these larger distributed efforts.

- **Publishing** – **The ability to support a wide range of publishing options, especially as researchers and authors become more comfortable with new and varying publication patterns and practices.** A quick survey of OSUL's current digital projects shows that the Libraries currently endeavors to support a wide range of publishing options ranging from traditional journal publishing options to digital exhibits building. Part of OSUL's collection management approach should be to determine what mechanisms for publishing digital content the Libraries wishes to provide and support long-term.

Definition:

Drivers for making decisions that optimize retention, integrity, and migration of digital content under the OSUL's custodial control.

The term preservation is often an overused and loaded term, unevenly applied within many different communities. Even in digital library communities, preservation all too often stands in for or is used interchangeably with byte level backup of content. Within OSUL's digital library program, the SDIWG, referencing the OSUL's Preservation Framework (Appendix B), wants to expand the concept of Preservation and address it more holistically as a key part of the larger curatorial lifecycle, and shift how preservation is thought about from that of a black box to something more dynamic that has significant implications around access and content management decisions. In fact, for the purposes of the digital library program, it would be instructive to think of preservation as continued access for the future. With this shift, a number of key functional needs emerge to support preservation functions within a digital library environment.

- **Data Audits** – Every preservation plan and preservation system is a good one until it is actually needed. It's when problems occur and content must be migrated or restored that an organization finds how good their preservation planning and systems in place actually are performing. **At a system level, one method for building confidence in the archival approach is to provide automated and interactive auditing support.** These resources protect against data rot and tampering, by creating a failsafe around the data managed by the system.
- **Distributed Storage** – One of the challenges related to the long-term preservation of digital content will be the capacity for organizations to maintain all their own physical storage. The need to evaluate and potentially utilize a wide range of storage options, from local to distributed, will likely become more important as the Libraries move to capture and preserve more large scale media content. Having a plan for how distributed storage is evaluated and its use integrated within OSUL's digital library infrastructure will be important in determining the scalability of the Libraries' preservation infrastructure.
- **Rights management** – Rights management overlaps a number of the core components. When thinking about preservation, rights management is directly related to the ability to preserve, migrate, and store the content and metadata utilizing local or distributed storage.
- **Migration** – The ability to identify and support the automated, wide scale migration of digital content from obsolete content formats.

- **Preservation** –The ability to provide indefinite support for the curation of OSUL digital content. The OSUL Preservation task force’s Preservation Framework defines the OSUL’s current thinking around preservation, as a topic.
- **Master Object Management** – Preservation must look at digital assets as more than just the image or document that has been digitized. Digital assets are really complex objects made up of a number of component assets including descriptive, structural, and administrative metadata, the asset and any derivative formats, and any ancillary items such as OCR’d content. Within OSUL’s preservation infrastructure, this complex nature of the relationships between metadata and assets needs to be retained. For many libraries, these relationships are retained using structural and administrative metadata like METS[3] and PRISM[4], while being stored within archival packages known as Bags[5] while within other environments, these relationships may be sustained through logical file mappings managed between an organization’s digital access management system (DAMS) and the underlying file storage system.

Core Component – Interoperability

Definition:

Drivers for making decisions and supporting the integration of OSUL content and services both within the Libraries as well as with on and off campus partners.

One of the most seismic changes that has occurred in digital libraries over the past decade has been the shift in perspective of the digital library as a destination to that of a node on the wider information network. This shift in thinking has permeated much of the library community, but has been especially important in the digital libraries community as it has led to realignment in how digital library infrastructure is viewed and envisioned. While issues related to user experience and interface design are still important, there has been widespread recognition that libraries need to make their digital content available outside of their own digital library systems, by providing multiple methods for users to discover and link to their content. It could be argued that the concept of an individual digital library is quickly giving way to a world view in which organizational digital library systems are seen as minor nodes within a larger digital libraries ecosystem. The challenge for these nodes, and for digital library programs like OSUL, is to develop an infrastructure that encourages the type of promiscuous data sharing between systems and allows the Libraries to be an important member within this larger digital libraries ecosystem. In considering the role that interoperability should play in the OSUL’s digital library program, the following high level topics need to be addressed.

- **Mediation** – In this context, mediation refers specifically to support for machine-to-machine interactions. These interactions could occur within our own digital infrastructure or represent interactions between key partners and our digital infrastructure. Within a digital library environment, libraries are asked to mediate

requests for a wide range of content, including metadata and complex digital objects, while still enforcing specific rights management rules. These types of automated users represent a key user community supporting the wide spread distribution and access to both the Libraries' digital assets and underlying metadata. While the Libraries' digital library infrastructure will support a wide range of communication protocols, many of which can be used to support various levels of mediation between outside entities and the Libraries, sophisticated mediation support will require the development of a comprehensive set of APIs and documentation to support both our internal and external development partners.

- Communication Protocols – One of the most challenging aspects around the support for widespread integration with one's digital library is determining the types of communication protocols to support. Libraries and partners have a wide range of protocols and standards that are being utilized and supported. Within OSUL's digital infrastructure, it would make sense to identify and support key communication protocols within the library and archives community, as well as key standards supported by stakeholders and data aggregators, to support the widest community of users.

Today, this list of protocols is relatively short:

1. Sitemaps[6] (search engines)
2. OAI-PMH[7] – to support aggregations such as DPLA.
3. ResourceSync[8] – to support distributed preservation efforts like DPN[6].
4. HTTP[9]
5. JSON[10]

The above protocols represent some of the most widely utilized within the library and digital libraries community. One protocol notably absent from this list is Z39.50, an older library community protocol supported largely by integrated library systems (ILS). While this protocol remains important within the library community, JSON and to a lesser degree, SRU[11], have replaced Z39.50 as the communication protocol of choice.

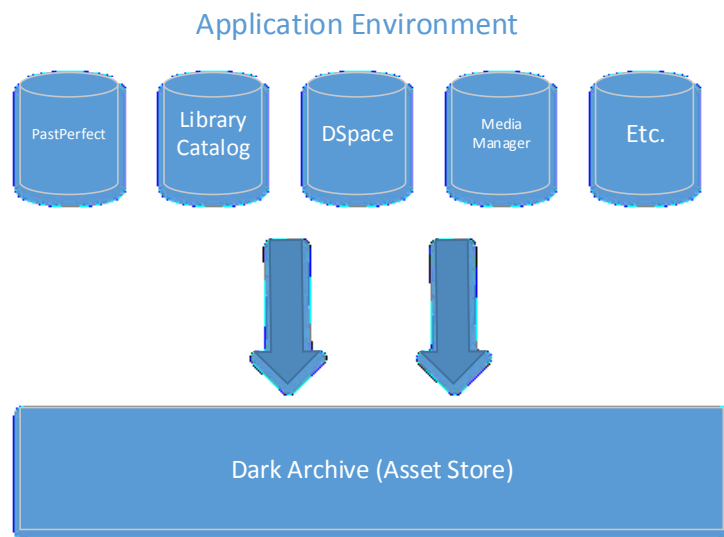
- Commerce – Presently, OSUL supports options for users to purchase high resolution digital objects. Since this will likely continue to be a need into the future, the ability to communicate with OSU's commerce systems will remain a core dependency.

Architecture Development and Standards

One of the challenges associated with reimagining OSUL's digital library architecture will be dealing with the long legacy of systems currently in use. Over the past decade, OSUL's digital infrastructure has primarily been project centered. This means that infrastructure decisions were primarily determined by the specific project being processed, resulting in silos of the Libraries' digital assets, and an inability to integrate various pieces into a single coherent

system. Additionally, some of the core components that the Libraries rely on are not managed by the Libraries and today, are largely unsupported.

After conducting an environmental scan of the Libraries' current digital infrastructure, a diagram slowly began to emerge.



In the current digital library environment, the application environment was largely driving system development. While the Libraries has attempted to make in-roads in normalizing some of the architectural decisions by utilizing the Libraries' Institutional Repository software, DSpace, as an asset delivery mechanism, the process of digitizing and reformatting materials still had the overall result of fragmenting the existing digital library infrastructure.

As the Libraries retools its digital initiatives infrastructure, it is now time to examine how decisions regarding infrastructure development are made and adopt a philosophy that places a premium on services and modular development. Within this reimagined environment, the application environment would develop out of a common services environment. Utilizing the Guiding Principles, the Libraries should evaluate tools and infrastructure components based on:

- the ability to meet a defined user need
- the ability to do something that isn't already provided by the current components
- the ability to utilize standards and integrate within existing digital infrastructure
- the ability to be utilized for more than a single task – i.e., not just project specific
- the ability to be sustained long-term

Whenever possible, infrastructure decisions should not be made simply to support a single program or project, but rather, will be evaluated by the SDIWG based on their capacity to support a wide range of needs within the Libraries in line with the principles outlined in this *White Paper*.

In considering the Libraries' digital library infrastructure, the SDIWG considered the necessary library wide policies that will need to be in place to support and enable a more unified digital library program. One of the past challenges has been the absence of established policies and organizational standards, leading departments or project teams to develop and adopt their own. If the Libraries is to succeed in creating a unified digital library platform, a shared set of policy documents need to be at the heart of that process. The SDIWG identified the following necessary policy documents:

- **Identifier Management Policy** – The Libraries currently follows a number of different naming conventions when creating digital assets. These names impact how items are stored within the master object archive as well how items are made available to the public. Sometimes identifiers are tied to permanent identifiers, as in the instance of handles, while other times, they represent internal naming conventions that become a way of capturing descriptive metadata about an item. **The SDIWG and key Libraries stakeholders need to have the following discussions:**
 1. **A uniform procedure around the naming of digital assets**
 2. **A discussion around the type of unique identifiers that will be supported within the Libraries' infrastructure. Currently, the Libraries supports handles, but questions arose around support for alternative identifiers like ARKs and DOIs.**
- **Content Format Policy** – Presently, the Libraries' Reformatting Working Group works with curators and digital project submitters to determine appropriate format standards for digital materials. However, there appears to be no organizational wide content format standard describing acceptable best practices for supported archival and derivative content types.
- **Metadata Standards Policy** – **Like Content Formats, the Libraries has no organizational metadata standards policy detailing metadata standards currently supported by the libraries or the context in which these standards are used.** Rather, metadata support is handled by individual content silos. For example, within the Libraries' integrated library system (ILS), MARC metadata is supported, while in the Libraries' institutional repository, Dublin Core is favored. Ideally, the Libraries will support a wide range of metadata standards, providing curators and content managers the ability to utilize the descriptive metadata schema that best fits their content. But within this approach, it is vitally important for the Libraries to have clear guidelines of what metadata standards are in use, as these standards have significant impact on how digital content is indexed to support Access and Discovery, as well as modelled, to support long-term preservation

of the resource.

- **Preservation Policy** – The Libraries presently has a Preservation Framework that outlines the issues and aspirational goals related to how the Libraries will implement and support long-term preservation issues related to digital assets. This framework outlines broad guidelines and best practices. Moving forward, the Libraries needs to determine how this framework will be developed further and implemented in a manner that address some of the pressing concerns around long-term preservation, including management of the masters object repository (formally dark archive), workflows around submission of content for preservation, including the development and management of metadata utilized in creating the archival digital objects. And finally, discussions around long-term processes for audits, remote storage, and potential partnerships and consortia that support the preservation of digital content.
- **Selection Policy** – As OSUL seeks to formalize content management processes for digital assets, the Libraries needs to provide clear guidance around the selection policy of items for digitization and acquisition. This is especially important for OSUL, given the relative contemporary nature of the Libraries’ special collections holdings. The primary goal for digital projects and digitization efforts should be to provide access to distinctive content that is uniquely available at OSUL.
- **Master Objects Retention policy** – Many curators and digital project managers remain confused around the process for archiving preservation, master objects. Ongoing work needs to be completed to help data managers and content creators identify and define the master object, and put workflows in place for ingesting these objects into the Master Objects Repository.
- **Discovery policy** – Presently, the Libraries has a working group that examines discovery, as it relates to the Libraries’ access to vended content. At present, no policy or group is responsible for supporting discovery and access of the Libraries’ growing local digital content. Given the significant resources being spent to create, preserve, and serve these digital assets, a group and policy dedicated to ensuring their discovery should be a priority.

Identification of Key Gaps

A number of key gaps were identified when looking at the present digital library environment. Utilizing the core components matrix, the SDIWG evaluated a number of projects and programs to determine some of the greatest areas of need. While an evaluation of specific projects or programs is outside the scope of this report, some general trends emerged.

1. OSUL's current digital projects do not support integration as defined in this report. In many cases, systems utilized by the Libraries to serve digital assets only provide access to materials via a single web interface. The ability to improve Search Engine Optimization (SEO), integrate the content within our own discovery services, or expose our digital collections to larger aggregators such as the DPLA are simply not supported by many of our current architectural components.
2. If OSUL is to meet strategic goal 6.1, *Increase the amount of OSU-generated digital content management by the Libraries* or 6.3, *Enhance the impact of our campus partners' digital content by broadening access to it*, the Libraries needs to make addressing the lack of integration support between our legacy platforms a priority.
3. OSUL's current digital projects do not support preservation as defined in this report. At present, how a project or collection supports preservation is largely defined by the platform being used to serve and manage the content. Within the Knowledge Bank, for example, digital assets are managed through the Master Objects Repository. However, for resources stored in the Libraries' Open Journal System, Media Manager, PastPerfect, etc. – no such preservation of digital assets is taking place outside of general server backups. This uneven approach to preservation has made a uniform policy difficult to create and complicates the Libraries' ability to fully understand what digital assets are under its custodial control.
4. OSUL's current digital library environment suffers from a lack of clear roles and responsibilities. When considering policies and standards around the digital library, one thing that became clear to the SDIWG members was that many of the identified policies have unclear ownership or fall under workgroups that are unsure of their roles. Examples of this lack of clarity can be seen when discussing current metadata and content format standards supported by the Libraries. The Libraries currently has no group responsible for organizationally managing or answer questions around these standards. As the Libraries seeks to move away from the project centric approach to building a digital library to a more uniform programmatic approach, providing clear guidelines and standards will be one of the primary glues that will enable a unified digital program.
 - a) OSUL's current approach to metadata and content standards must change. In preparing this *White Paper*, it became apparent that no organizational policies exist, and that different units are utilizing different metadata and content standards as part of their digitization processes. While the Libraries will want to strive to support a diverse array of content and metadata standards, it will be imperative for the Libraries to develop and organize set of guidelines to provide consistency between different departments and units.
 - b) OSUL's current approach to URI minting is inadequate for its current needs. At present, permanent URIs are only supported as part of the Knowledge Bank's handle minting service. These handles are tied to the software platform, which

could potentially be problematic into the future. Additionally, the Libraries is currently discussing the support for alternative object identifiers, like DOIs. To support the growing requirements for permanent URI minting, the Libraries needs to investigate the decoupling of the existing handle service from the existing application platform and develop a system-wide component that could be utilized to support a wide array of identify management systems.

5. OSUL's current efforts around discovery have focused primarily around purchased journal content and the ILS (Integrated Library System). Very little practical attention has been paid to how the Libraries surfaces its digital content, creating a significant barrier to discovery for the Libraries' users. OSUL should make a deliberate effort to simplify the discovery process of its digital assets, by seeking to promote greater access through wider metadata distribution and indexing, but also create a more intuitive approach to discovery from within the OSUL web presence.
6. The work done by the OSUL Preservation Task Force around the development of the OSUL Preservation Framework brings to light many of the challenges around doing sophisticated resource preservation and curation. Currently, OSUL provides byte level preservation for assets stored within its "dark archive". Additionally, management of this archive requires a good deal of human mediation from both Information Technology and the curators responsible for the collections. While having byte backup helps to provide some level of protection of the digital content, the Libraries currently utilizes no significant auditing mechanism or utilizes any mechanism for preserving digital resources as complex digital objects; linking the item, descriptive, structural, preservation, and administrative metadata together.
7. One of the challenges within the current OSUL digital environment is that different units and departments utilize locally developed and specialized workflows to manage accession, description, and curation of similar content. In many cases, these variant workflows developed due to the lack of a cohesive environment from which to manage shared content. Today, they represent ingrained processes that will need to be evaluated and brought into alignment as the Libraries' shared digital infrastructure takes shape.
8. The reimagining of the Libraries' digital environment will impact a wide range of Libraries' users and services. The process will require learning new procedures, policies, systems, and technologies. The process will also introduce the Libraries to new models of service development, specifically around the need for incremental service development. It will be vital that the Libraries work to empower staff to understand and be active participants in this change through sustained education and communication. Providing this education should be a Libraries-wide initiative, but the SDIWG will need to take an assertive role in providing these sustained activities.

Recommendations

As the Libraries works to evolve its digital library infrastructure, it is worth noting that the Libraries' greatest asset in this process will be its people. The OSUL has a long history creating, curating, and managing digital resources. The OSUL's Special Collections and the Libraries' Digital Content Services have created exemplar resources. Resources like the Billy Ireland's Image Database and the Knowledge Bank represent significant achievements that provide incredible value to the OSU community, as well as serve as beacons within their respective research communities.

Based on the background and information provided in this *White Paper*, the SDIWG proposes the following recommendations:

- 1) Continue to provide education around the *Strategic Digital Initiative Guiding Principles*. These *Guiding Principles*, where were formally adopted by the Libraries' Executive Committee, provide the basic underpinnings for the development of a program-centric approach to digital library development at OSUL, as well as highlight key concepts around the need for a less monolithic approach around infrastructure development, and the adoption of an iterative approach when it comes to tools and services development.
- 2) Continue to provide education around the OSUL Preservation Framework. The Preservation Framework outlines the necessary challenges, issues, roles, and preservation activities that OSUL could potentially support. However, the Preservation Framework is just a first step. The Libraries needs to continue to build off this work to provide an organization statement around the Libraries' digital preservation program and the various issues that surround the preservation of digital content. This includes the current investigation around curatorial workflows for submitting materials in the Master Objects Repository, as well as early discussions between Dan Noonan and Libraries' Information Technology related to the development of complex archival packages for the Libraries' master digital objects.
- 3) Adopt the concept of broad functional categories to both frame and focus the development of the Libraries' infrastructure.
 - a. Adopt the concept of a Services Infrastructure model as a mechanism for normalizing interactions with the Libraries' digital resources and Enterprise Infrastructure.
 - b. Adopt the concept of a localized digital asset management infrastructure to support the management of our local digital resources. The selection and implementation of specific infrastructure components will primarily be evaluated and selected by Library Technology (pp. 22, 3c).

- c. Appoint a group within Information Technology and including identified stakeholders to flesh out the Application, Services, and Enterprise Infrastructure models to provide more detailed recommendations around specific infrastructure components that will support the integration of how information flows within the OSUL digital environment.
- 4) One of the key gaps to be identified as part of this *White Paper* is the lack of responsibility around key organizational policies. Sometimes these gaps exist because no formal group within OSUL has ownership of a specific topic; sometimes these gaps exist because the ownership of a specific policy or topic has been defused across many different units and departments. We recommend the following:
 - a. Expand and empower the Digital Reformatting Working Group to take organizational leadership around the development and management of an organizational wide content formats and standards policy. In addition to developing and managing the organization's content format standards policy, the group would provide a centralized location where questions regarding Libraries' supported content types could be handled.
 - b. Appointment of a group responsible for the development and management of an organizational wide metadata standards policy. In addition to developing and managing the organization's metadata standards policy, the group would provide a centralized location where questions regarding Libraries' supported metadata types, as well as questions general metadata questions related to digitization, curation, or discovery could be handled.
 - c. A shift in the Discovery Systems Management Working Group's charge to move beyond discussions related to discovery issues focused on the Libraries' implementation of WorldCat Local, to a more holistic approach to discovery, with special focus being given to the Libraries' digital resources.
 - d. Develop clear selection criteria to enable the selection of materials for digitization that would have the greatest impact for access across the OSUL's various user communities. The fact is that OSUL has many more materials and collections that would benefit from digitization than the Libraries has resources. And while a formal selection policy for digital assets may not be necessary, a clear statement of values around the materials that will be prioritized for digitization would provide clear guidance to both curators and the Reformatting Working Group. Criteria such as:
 - Are the materials damaged or in significant threat and in need of preservation?
 - Are the materials to be made publicly available?
 - Do these materials serve more than a single user community?

- Can these materials be accessed online?
 - Can these materials or their metadata be shared with other partner communities or projects?
 - Does the digitization of the selected materials offer opportunities for collaboration with other groups on (or off) campus?
- e. Develop clear selection criteria to enable to selection of materials acquired from outside the OSUL. While OSUL's institutional repository and publishing programs have documented procedures around the types of materials selected for curation, the Libraries' is often approached by various individuals and organizations looking for places to host resources and content. Determining a selection process for evaluating and acquiring this type of born digital content will likely become more important as OSU and OSUL address the need to collect, manage, and disseminate research data.
- 5) Utilizing the OSUL Preservation Framework, determine the OSUL's preservation/archiving scope and long-term commitments – and then make that information available as OSUL's trusted repository program.
 - 6) Recognizing the many different submission workflows and processes currently in use by departments and units throughout OSUL, it is apparent that some level of education and workflow management will need to take place as the Libraries moves to integrate its infrastructure.
 - 7) Long-term education and institutional buy-in will be significant issues around any restructuring of the Libraries' digital library infrastructure. The Working Group recognizes that while the technological challenges will be significant, they are generally easier to manage than individual expectations or biases. In order for this effort to be successful, the Libraries will need to make a concerted effort to provide numerous avenues for interested faculty and staff to be involved in this process. Likewise, the Libraries will need to make a deliberate effort to provide training and educational opportunities to ensure that OSUL faculty and staff remain fully engaged.
 - 8) The OSUL can no longer treat our digital collections and infrastructure as islands. As the Libraries implements new infrastructure components and looks to engage in targeted tools and services development, the Libraries should make an active effort to seek like-minded partners, both internal and external to the university, and look for avenues of collaboration or research.

Appendix A: Ohio State University Libraries Digital Initiatives Program Guiding Principles

Permanent URL: http://library.osu.edu/documents/SDIWG/di_principles_v2.pdf

Principles:

A number of key principles have been identified that will guide the development of the Digital Initiatives Program and the working group as it evaluates current and future systems, policies, services, and best practices. These principles underscore the changing nature of digital environments and the need to embrace an environment that understands and anticipates the ephemeral nature of technology while ensuring the durability of managed content.

- ***We build services, not products***

The OSU Libraries' digital infrastructure will be conceived around the philosophy that the Libraries should be investing in the development of services, not one-off development projects. This represents the natural maturation of the OSUL's Digital Initiatives Program; moving from an environment that was tied to individual digitization projects and products, to a programmatic architectural approach that considers the issues of long-term sustainability, interoperability, preservation, and accessibility of resources at OSUL and beyond. This approach should allow for the Libraries' sharing and leveraging of data to provide better integration, access, and discovery of content.

- ***We carefully weigh when to Build versus Buy versus Borrow***

The development of the Libraries digital infrastructure must recognize the needs of the institution and weigh them against the Libraries' desire to promote a culture of open access and open source development, and the Libraries' current capacity to develop and support new services and systems. The development of the Libraries' Digital Initiatives Program will occur along a continuum of options, where tools and systems are selected based on the Libraries' needs, support capacity, and a system or tool's "fit" within the present environment.

- ***We develop modular services, not monolithic systems***

The Libraries' Digital Initiatives Program will be developed around the philosophy that systems designed around modular services and components ultimately will provide a greater level of flexibility and encourage long-term innovation over traditional monolithic systems. The Libraries' will utilize a "plug and play" philosophy, where systems and tools are simply building blocks of a larger whole. This approach will provide the Libraries with the greatest flexibility by allowing the Libraries to continuously evaluate the blocks that make up the larger architecture, and replace components and tools as new, better options become available.

- ***We develop for change***

In order to be successful over the long-term, the Libraries' Digital Initiatives Program will adopt a philosophy of continual change. The simple truth is, nothing is forever. Whether it is the digital object and the need to refresh or migrate formats as standards change, or the iterative and incremental development of systems, services, and tools, the technology and processes that make up the Libraries' Digital Initiatives Program will always remain in a state of flux. By recognizing this fact, the Libraries' will adopt a vision of iterative development. This vision will be marked by the following tenets:

- All things evolve – Regardless of the format, the system, or the people managing and developing a system, things will always be in a state of flux. We acknowledge that and accept it as a part of the architectural design.
- Plan for continuous improvements – The Libraries will use an agile approach to systems development, adopting a philosophy of iterative development: release soon and release often, allowing the Libraries to move quickly and anticipate new needs and services.
- It's OK to fail – In systems development, a certain amount of risk exists. It's easy to react to trends, but difficult to anticipate future needs without experimentation. Failure in this context is acceptable and periodically expected, so long as the Libraries is able to learn through the process, fail fast, and fail forward, and not lose managed content.
- Anticipating End-Of-Life – Whether it is a data format, a tool, a system component, services, or a best practice – anticipating the migration of data, systems, and workflows is a universal constant when working within a digital environment. The Libraries will minimize the impact of these disruptions by proactively planning and preparing for these inevitable technology shifts.

- ***We don't keep everything forever***

While libraries understand the need to curate traditional print collections, few apply those principles of curation to the digital environment. The simple fact is, the Libraries will not maintain all digital data forever, and the Libraries' digital programs will need to provide a clear process for managing the life cycle of the content that the Libraries curates as part of the OSUL's Digital Preservation Framework. However, this goes beyond curation of content. Technology moves at a rapid pace, and the projects and services created, used, and supported by the OSUL will need to change as well. The Libraries' Digital Initiatives Program must be able to thrive in this ever changing environment, and include clear assessment metrics and documented exit procedures.

- ***We will build in assessment***

The Digital Initiatives Program will embed assessment into its processes. Utilizing available data, stories, and user feedback, the Program will seek to continually assess Program strengths, weaknesses, and needs as it moves forward.

- ***We focus on the user***
 The OSUL will strive to create a user friendly environment, developing intuitive services, and systems that support the Libraries' strategic mission in advancing discovery and learning. Utilizing techniques such as focus groups, user feedback, and testing, the OSUL will strive to provide a superior, user-focused environment.
- ***We work with partners***
 A digital initiatives program is a very complex undertaking, and its success will depend upon leveraging partnerships within the Libraries, across the University, and with stakeholders external to OSU.

 - Partners, not customers: One of the benefits of adopting an agile development environment is the necessity for close partnerships between IT departments (Digital Initiatives, Applications Development & Support, and Infrastructure Support) and the Libraries as a whole. Going forward, the expectation is that librarians and IT will work collaboratively throughout the life-cycle of a project.
 - Excellence to Eminence: As the University and the OSUL strive to embrace the move from excellence to eminence, one aspect of that transformation is the participation in large, national and international collaborations. Nationally and internationally, issues related to digital initiatives, preservation, semantic data, repositories, etc. are being actively researched. OSUL must strive to find a place at that table, and look for strategic opportunities to build new and exciting partnerships outside the University, and with the larger library community.
- ***We embrace research as a core, fundamental value***
 One of the fundamental purposes of a university is to cultivate research at all levels, including systems development. This goes beyond the development of new services to considering the fundamental shifts occurring in libraries and positioning the OSUL to be a significant contributor and valued partner as the library community wrestles with these issues. The need to research, experiment, and push boundaries must be a core part of the program.
- ***We strive to stay grounded in the real world***
 The pace of technological change can be dizzying as is the wide range of new services, systems, and best practices available. The Digital Initiatives Program will seek to cultivate an environment of experimentation and innovation, while staying grounded in the knowledge that local needs must be met. While these two values can sometimes clash, the Program will seek to provide a balance between innovation and the need to bring together collections and people through the engagement and integration of scholarship and instruction.
- ***We are driven by standards***
 Forecasting the future is never an exact science. Like a meteorologist, we look at available information and make the best decisions that we can. And while we may not always hit the mark, standards provide the touchstones that will keep the OSU Libraries

grounded. Standards apply to all areas of a digital initiatives program. They can be found in project management and the need to utilize systematic approaches to planning and development. Standards also inform our decisions around object formatting and reformatting, metadata schemas and vocabularies, and system interoperability and design. The Digital Initiatives Program will work with the OSUL and its partners to help the Libraries navigate these sometimes muddy waters to ensure that we are following a best set of standards and practices. Utilizing well-established international and national standards over locally created best practices, OSUL will strive to create an environment that will support long-term sustainability and promote interoperability between services, assets, and systems, both locally and with our partners.

The Digital Initiatives Program Principles provide the OSUL with a framework for making and evaluating decisions related to the development and implementation of a long-term digital services architecture. These Principles reflect the changing nature of the OSUL's digital programs, shifting the focus from individual projects to a holistic program designed around an evolving services architecture. The Principles illustrate for the Libraries and its stakeholders the core values and long-term direction of the Digital Initiatives Program.

Submitted by the Strategies for Digital Initiative Working Group, June 24, 2013.

Approved by the Libraries Executive Committee, June 25, 2013.

Appendix B: DIGITAL PRESERVATION POLICY FRAMEWORK

Permanent URL:

http://library.osu.edu/documents/SDIWG/Digital_Preservation_Policy_Framework.pdf

Purpose

This statement formalizes The Ohio State University Libraries (OSUL) continuing commitment to the long-term stewardship, preservation of and sustainable access to its diverse and extensive range of digital assets. In alignment with the OSUL mission to create, acquire, organize, disseminate, and preserve scholarship, this policy makes explicit OSUL's long-term commitment to The Ohio State University (Ohio State) community as its trusted digital repository. The OSUL's digital stewardship efforts contribute to Ohio State's mission to build a world-class faculty, develop academic programs that define Ohio State as the nation's leading public land-grant university, improve the quality of the teaching and learning environment, enhance and better serve the student body, create a more diverse university community, and help build Ohio's future by ensuring access to this corpus of information over time.

Objectives

The primary purpose of digital stewardship and preservation is to preserve the intellectual and cultural heritage important to The Ohio State University, while at the same time making sure that it is accessible and held in trust for future use. The objectives in this statement define a framework to:

- identify, through systematic selection, digital assets to be preserved across new generations of technologies
- maintain access to reliable data at bit-stream level, the digital assets encoded in the bit streams, and access to the intended contextual and intellectual meaning of the digital assets
- include in the scope of the program materials that originated in digital form and those that were converted to digital form
- protect OSUL's digital investments through a fully-implemented digital preservation program
- demonstrate organizational commitment through the identification of sustainable strategies
- develop a cost-effective program through means such as, system-wide integration, shared responsibilities, and automating human-intensive efforts, when possible
- comply with prevailing community standards for digital preservation and access
- seek, expand, and develop digital preservation methods that are appropriate for Ohio State and promote inter-institutional collaboration

Mandate

OSUL's mandate for digital preservation is at least five-fold:

- **Scholarship:** As an institution of higher education, The Ohio State University is obligated to support scholarship, teaching, and learning. As more resources and services associated with these functions become digital, OSUL's responsibilities must expand to include the identification, stewardship, and preservation of designated digital content.
- **Institutional records:** Ohio State has charged OSUL with maintaining the University Archives by collecting and preserving university records that best document the history of Ohio State, including those in electronic format.
- **Legal obligations:** Ohio State has mandated responsibilities to preserve and maintain access to certain digital objects, as well as responsibilities as a designated land grant institution. Some legal obligations derived from Federal and State laws require us to maintain files in an archival fashion.
- **Organizational commitment:** OSUL's commitment to digital preservation is explicitly cited in the initiatives in OSUL's current strategic plan:
 - 1) which calls for OSUL to develop and implement a cross-divisional plan for supporting curation, storage and dissemination of library-created or library-managed digital content
 - 2) which calls for OSUL to build a robust, reliable, secure technical infrastructure base including both human and technology resources.
- **Consortia and contractual commitments:** OSUL has commitments to consortia (e.g. OhioLINK and Hathi Trust) and contractual agreements to assume or share in the responsibility for preserving designated digital content.

Scope

This policy addresses preservation of digital collections and resources for which OSUL is the primary custodian. Although this policy only covers digital collections and resources for which OSUL is the primary custodian, OSUL has responsibility for informing, consulting, and as appropriate coordinating with other units of The Ohio State University to assure that Ohio State faculty, staff, and students will have adequate ongoing access to administrative, scholarly, and other digital resources created at Ohio State outside of the Libraries. Further, OSUL personnel will also work externally through consortia (e.g. the Committee of Institutional Cooperation (CIC) and OhioLINK), licensing agreements, etc., to assure that Ohio State faculty, staff, and students will have adequate ongoing access to all currently available digital resources. OSUL, however, cannot guarantee preservation for materials that we do not own and manage.

Challenges

There are recognized challenges in implementing an effective digital preservation program, including but not limited to:

- **Rapid growth:** Technology that enables the variety of formats and dissemination mechanisms changes rapidly. As different types of materials are submitted (data sets, complex digital objects), monitoring different needs (storage size, metadata, etc.) of the materials and maintaining procedures and policies based on these needs is necessary.
- **Sustainability:** Developing a sustainable digital preservation model that will respond to technological and staffing changes as needed, without under- or overestimating the needs imposed by these changes. The need for good cost models and affordable programs is widely acknowledged, yet still not fully addressed. OSUL requires sufficient funding for operations and major improvements for digital asset management, as well as designated library funding to sustain ongoing preservation efforts. Further, there are administrative complexities in ensuring cost-effective and timely action to implement preservation strategies. The scale of funding is based on the level of commitment, therefore the program should reflect reasonable expectations of requisite resources, i.e., OSUL should not promise more than can be delivered.
- **Management:** Moving from well-managed digital collections to preserved collections in the true sense of the term requires institutional effort, partnership development, and a financial commitment. OSUL should provide a thoughtful balance between access and preservation, while being mindful of preservation's core role in maintaining access.
- **Partnerships:** OSUL must work with creators and providers of crucial content to employ appropriate maintenance prior to deposit that will facilitate future preservation.
- **Flexibility:** The digital preservation plan must continually revise its abilities to respond to the evolving technological capabilities and changing user expectations without jeopardizing the ongoing care of the digital content.
- **Expertise:** OSUL must commit to continually updating staff expertise, where appropriate, as technologies change.
- **Rights:** There are a myriad of intellectual property and other rights-based constraints on providing access that impact digital preservation efforts.

Principles

Guiding principles

OSUL will use consistent criteria for selection and preservation as for other resources in the libraries. Materials selected for digital stewardship and preservation carry with them OSUL's commitment to maintain the materials for as long as needed or desired.

- The Libraries are committed to the long term preservation of selected content.
- Digital preservation is an integral part of OSUL's processes
- Processes, policies, and the institutional commitment are transparently documented.
- Levels of preservation and time commitments determined by selectors, curators, in consultation with technical experts
- OSUL will participate in the development of digital preservation community standards, practice, and solutions.

Operating principles

The Library will strive to:

- Develop a scalable, reliable, sustainable, and auditable digital preservation infrastructure
- Manage the hardware, software, and storage media components of the digital preservation function in accordance with environmental standards, quality control specifications, and security requirements
- Comply with the Open Archival Information System (OAIS) and other appropriate digital preservation standards and practices
- Ensure that the digital archive is as interoperable as possible by utilizing open source options whenever feasible
- Ensure the integrity of the data
- Secure metadata (e.g. administrative, descriptive, preservation, provenance, rights and technical) necessary for the use of the digital assets
- Comply with copyright, intellectual property rights and/or other legal rights related to copying, storage, modification and use of digital resources.

Standards

- Ohio State is best served when distributed and disparate systems conform to standards and best practices that make communication between these storage systems possible.
- To utilize the OAIS Reference Model for as the basis for developing and implementing strategies and tools for long term digital information preservation and access.

Categories of commitment

OSUL's levels of commitment as outlined below recognize that developing solutions for "born digital" materials informs solutions for the other categories; it does not imply that these assets are inherently more valuable or important than any of the other categories and/or our traditional, analog materials.

- **Born digital materials:** Rigorous effort will be made to ensure preservation in perpetuity of material selected for preservation, both library resources and institutional records.
- **Digitized materials (no available analog):** Every reasonable step will be taken to preserve materials without a print analog, when re-digitizing is not possible or no analog versions are located elsewhere. Also included are digitized materials that have annotations or other value-added features making them difficult or impossible to recreate.
- **Digitized materials (available analog):** Reasonable measures will be taken to extend the life of the digital objects with a readily available print analog. However, the cost of re-digitizing as needed will be weighed against the cost of preserving the existing digital objects

- **Commercially available digital resources:** OSUL has responsibility for working externally through consortia, licensing agreements, etc. to assure that one party or parties provides the necessary infrastructure to provide for preservation activities, so that Ohio State faculty, staff, and students will have adequate ongoing access to commercially available digital resources. If the resources are external to OSUL, there needs to be an articulated exit strategy in the event of the cessation of the consortia or licensing agreements. Particular emphasis should be given to resources which exist in digital form only.
- **Other items and materials:** No preservation steps will be taken for materials requested for short term use such as materials scanned for E-reserve and Document Delivery, or for content that is deemed unessential.

Levels of Preservation

Digital Archiving Maturity Model

- Level 1 - safe storage –simple bit-level storage on magnetic or optical storage with some level of reassurance that the bits are protected against simple storage failure.
- Level 2 - storage management – moves the bits to the most appropriate location. The decision on which bits are located where maybe done on the basis of storage durability, cost reduction, or performance.
- Level 3 - storage validation –multiple object storage plus fixity checking to validate storage durability. Object fixity is checked on storage, access and at regular intervals to confirm objects have not been tampered with. If bit failure is identified, self healing from an alternative copy will occur.
- Level 4 - information organization – incorporates information hierarchy organization, descriptive data management, and simple processes for uploading, locating and downloading information. Basic information security is also included.
- Level 5 - information processes – efficient and flexible business processes to automate the activities associated with information management. These include interfaces to the information sources and dissemination to information consumers using flexible workflows and programmer interfaces. They also include high-throughput capabilities and integration with a third party identity management system. Non-archiving processes such as object versioning, should be excluded
- Level 6 - information preservation – capabilities to ensure that the information stored is usable when it is needed by the audience that requests it. Includes variety of strategies to ensure the information is accessible for as long as it is needed.

This *Digital Archiving Maturity Model* was developed by Tessella Technology and Consulting (© Tessella plc 2012)

Roles and Responsibilities

OSUL has identified the following stakeholder categories for the digital preservation program. The terminology is adapted from the OAIS Reference Model (CCSDS 650.0-M-2 (2012))

- **Producer:** The role played by those persons or client systems that provide the information to be preserved. Producers include faculty, students, staff, alumni, collectors, creators of content, publishers, and others. Producers can also be other OAISes or internal OAIS persons or systems. Producers will be responsible for complying with established deposit requirements and working with the management of the digital archive to ensure a successful transfer. (expanded OAIS definition)
- **Management:** The role played by those who set overall OAIS policy as one component in a broader policy domain, for example as part of a larger organization. The OSUL Executive Committee will be responsible for setting digital preservation policies and integrating them into broader organizational contexts. (expanded OAIS definition)
- **Administrators:** Content stewards (designated staff responsible for selection and for ongoing curation of specific collections), digital preservation specialists and working teams. Administrators will be responsible for the establishment of the digital preservation program and for day-to-day management of the digital archive(s). *[Note: OAIS uses Administration Functional Entity: The OAIS functional entity that contains the services and functions needed to control the operation of the other OAIS functional entities on a day-to-day basis.]*
- **Co-operating Archives:** (OAIS definition) Those Archives that have Designated Communities with related interests. They may order and ingest data from each other. At a minimum, Co-operating Archives must agree to support at least one common Submission Information Package (SIP) and Dissemination Information Package (DIP) for inter-Archive requests. Examples include: the CIC, OhioLINK, DuraSpace. (expanded OAIS definition) At OSUL we think of this group as collaborators.
- **Consumer:** The role played by those persons, or client systems, who interact with OAIS services to find preserved information of interest and to access that information in detail. This can include other OAISes, as well as internal OAIS persons or systems.
- **User Groups /Client Groups:** The various types of clients who use OSUL's digital collections.

Collaboration/Cooperation

OSUL acknowledges digital preservation as a shared community responsibility, and as such has long-standing and emerging partnerships with similarly committed organizations (e.g. CIC and OhioLINK) and is committed to collaborating with other institutions, as well as with units within Ohio State to:

- advance the development of the digital preservation program
- share lessons learned with other digital preservation programs
- extend the breadth of our available expertise
- extend the digital content that is available within a broad information community to OSUL users through cooperative efforts

Generally, in working, cooperating and collaborating with others, OSUL desires to:

- understand the goals, objectives, and needs of the communities of creators and the communities of consumers of its digital resources
- identify appropriate partners and stakeholders to contribute to national and international efforts in digital preservation
- help develop national and international strategies and initiatives that enable the distribution of collecting, description, service delivery, digitization and preservation activity
- work actively with creators of digital materials to encourage and promote standards and practices

Access and Use Criteria

OSUL acquires, manages, and preserves digital resources so that they remain accessible to its constituents over the long term. Certain limitations may be placed on access due to legal, donor and/or other reasons, but, in general, in so far as possible, OSUL endeavors to make its digital resources accessible to all users.

Implementation

Implementation of this policy framework is contingent upon the infrastructure (technological and human resources) provided by Ohio State and OSUL, the availability of cost-effective solutions, the adoption of standards, and development of best practice and procedures.

Review Cycle

This policy will be reviewed at minimum annually to assure timely revisions as technology progresses and preservation strategies and experience mature.

References:

- [1] ARK Specification: <https://wiki.ucop.edu/display/Curation/ARK>
- [2] DOI Specification: <http://www.doi.org/factsheets/0607DOIIdentifierSpec1-1.pdf>
- [3] METS Specification: <http://www.loc.gov/standards/mets/>
- [4] PRISM Specification: <http://www.idealliance.org/specifications/prism-metadata-initiative/prism>
- [5] BagIt Specification: <http://en.wikipedia.org/wiki/BagIt>
- [6] Sitemap Specification: <http://en.wikipedia.org/wiki/Sitemaps>
- [7] OAI-PMH Specification: <http://www.openarchives.org/pmh/>
- [8] ResourceSync Specification: <http://www.openarchives.org/rs/toc>
- [9] Digital Preservation Network (DPN): <http://www.dpn.org/>
- [10] JSON Specification: <http://www.jsonrpc.org/specification>
- [11] SRU Specification: <http://www.loc.gov/standards/sru/>