Developing the Read/Write Library

Lisa Goddard
University of Victoria

Scholarly and Research Communication
VOLUME 7 / ISSUE 2 / 2016

Abstract

Background: This article considers the use of Fedora-based library digital asset management systems (DAMS) as digital humanities (DH) research platforms.

Analysis: The features of DAMS are evaluated to identify the ways in which they can currently meet researcher needs, and to suggest areas where further development is necessary.

Conclusion and implications: Fedora-based DAMS hold great promise as the basis of digital humanities research platforms. Mature functionality is available for identity management, file and metadata management, versioning, publishing, social media sharing, discovery, interoperability, and long-term preservation. Further development is necessary in order to incorporate annotation, mark-up, and text analysis tools.

Keywords: Digital asset management systems; Digital humanities; Libraries; Digital preservation; Digital research environments; Hydra; Fedora

Lisa Goddard is Associate University Librarian for Digital Scholarship and Strategy, University of Victoria Libraries, PO Box 1800 STN CSC, Victoria, BC, Canada V8W 3H5. Email: lgoddard@uvic.ca .

CISP Press
Scholarly and Research Communication
Volume 7, Issue 2, Article ID 0201255, 13 pages
Journal URL: www.src-online.ca
Received June 1, 2016, Accepted August 10, 2016, Published November 8, 2016

Goddard, Lisa. (2016). Developing the Read/Write Library. *Scholarly and Research Communication, 7*(2): 0201255, 13 pp.

© 2016 Lisa Goddard. This Open Access article is distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc-nd/2.5/ca), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

We have undergone, and are still undergoing, a breathtakingly deep and vast transformation of the way in which information is produced, transmitted, exchanged, and stored. This sea change has enormous potential for reshaping the means of academic production, but also requires a completely different set of practices and tools that have to be accommodated within the scope of relatively static institutional and funding agency resource commitments. Digital humanists collect, describe, annotate, compare, and interpret large numbers of digital objects in the course of their research. Libraries have expertise in sustainable digital asset management and are making substantial investments in systems and infrastructure to facilitate the use, sharing, and preservation of electronic information. Scholars, librarians, and archivists could each reap substantial benefits by pooling their unique talents, but this is difficult to do across disciplinary boundaries, funding models, and organizational siloes. To facilitate this work we need to build new places to come together. In the context of the digital humanities it is consistent that these spaces would be, at least in part, virtual. These new environments are framed as the read-write library, in contrast to the more traditional read-only library. In the read-only library librarians and archivists make collection choices. Descriptive metadata is limited, neutral, and centred around the bibliographic properties of an information object. Very little contextual and interpretive information is provided to enhance collections. In the read-write library, collections are collaboratively created and curated with scholars, students, and the wider community. The read-write library allows anyone with content expertise to contribute contextual knowledge. It encourages patron involvement in tasks such as transcription, description, and mark-up, and it provides virtual spaces that remove barriers to that participation insofar as possible.

The new generation of Fedora-based digital asset management systems (DAMS) incorporate library best practices for interoperability and preservation; easy to use, web-based interfaces for managing collections of digital documents and artefacts across large, distributed teams of researchers; and engaging end-user interfaces for exploration and discovery. This article will consider how library DAMS can act as research environments that allow faculty and students to help build rich digital collections as part of the research process.

The challenge of digital stewardship

Technology has opened up huge new opportunities for research libraries, and for scholars. We can delve into artefacts at the micro-level of chisel marks (Levoy, 2015), and explore macro-patterns in huge text corpora (Moretti, 2005). There are, however, important pre-conditions for undertaking this type of computer-aided investigation. The information object (text, map, image, video, data set, etc.) must be available in digital form online. It must be in a standard format for which viewing and manipulation software is available. The information object must be described with sufficient metadata to allow a user to find that object and to identify it as being within the scope of investigation. Documents must be transcribed so that a plain-text version is available for automated search and processing. For most advanced applications the full text should be structured, marked-up, and indexed. Numerical data fields must be labelled and described so the meaning of the content is clear. Relationships between documents,

Goddard, Lisa. (2016). Developing the Read/Write Library. *Scholarly and Research Communication, 7*(2): 0201255, 13 pp.

people, places, and events should be encoded for computer readability. Unique global identifiers should be applied to works, expressions, and manifestations of objects. Citation information should be structured into discrete fields for ingest into citation managers. Machine-readable rights statements should be included, and access mechanisms must be in place to protect intellectual property. Versioning information should be accessible so that transformations over time can be clearly understood. Metadata standards must evolve, and crosswalks must be developed to allow the transformation of metadata from one platform to another for purposes of aggregation and comparison. Methods must be exposed to allow for ingest, export, and manipulation of objects and their associated data. Objects and their metadata must be preserved over time, identifiers must persist, and links must continue to resolve. While the research possibilities are endless, so are the challenges in preparing information for computer manipulation. With the digital age comes a host of new labour, some of which can be automated, but much of which requires the intervention of humans with deep subject knowledge, technological competence, and information management expertise.

The value proposition for libraries

For centuries librarians and archivists have been arbiters of the cultural record. They have determined what information merits collection and preservation. Librarians have developed a host of measures to ensure that they are able to offer information that is factually accurate, rigorously researched and evidence-based, and produced by reputable authors and publishers. As purveyors of authoritative information, librarians favour sources that have been subject to formal processes of review and editing to ensure accuracy and authenticity. By enforcing high information standards and strong professional ethics, libraries have been successful in building a reputation as trusted, neutral information providers (De Rosa, 2010). This gatekeeper culture is evident in modern day library sites that offer little to no opportunity for faculty, students, and community to contribute information or to co-develop knowledge. In the new information economy, however, there are suddenly many more information artefacts than can possibly be vetted, organized, described, or managed using traditional library workflows and resources. Libraries need to broaden their approach without eroding their position as trusted providers in a sea of inexpert, inaccurate, and even deliberately misleading information sources. Libraries must find new models for undertaking collection building and curation in a more collaborative way, and academic libraries are fortunate to exist in an environment that is rich in expertise and highly qualified personnel.

Libraries are keen to have their collections digitally available for research, study, and creative endeavour, but the scope of the task is daunting. The requirements of digital scholarship are layered on top of traditional library services, such as research help, information literacy instruction, document delivery, and the acquisition, processing, and accessibility of core books and journals. Library budgets remain static, and journals costs continue to rise (Poynder, 2011). There is rarely additional money to support the new demands of digital scholarship, especially for tasks that are labour intensive and highly specialized. If libraries can engage scholars and students in the development and description of online collections, then they will greatly increase their capacity to offer rich aggregations of highly structured, vetted, machine-readable

information. In order to capture scholarly expertise, libraries must offer interfaces that encourage end-user contribution. Above all this means removing barriers to participation, and providing features that make it simple to add an object, enhance metadata, correct transcriptions, or create links. Ideally this would be built directly into the process by offering research platforms that funnel scholarly work into library discovery and preservation systems.

The value proposition for faculty

If libraries are to engage faculty as collaborators in digital collection building then it is critical to articulate the benefits that researchers will derive from the arrangement. Digital academic labour such as annotation, description, and exhibit building is not highly valued in the context of promotion and tenure, a process that is still deeply invested in traditional single-author book and article publications. There are ways, however, in which digital collection building can be incorporated directly into the research process. Much humanities research revolves around cultural artefacts, including books, essays, letters, journals, and a variety of texts, along with photographs, drawings, paintings, film, and related items. Scholars need a means to organize and manage digital objects that are collected or created in the course of their research. Traditional humanities departments offer little in the way of technological support. In some universities it is possible to pay for a virtual server that is administered by central computing services, but once grant funds disappear, so can the server. Canadian researchers can install software on Compute Canada's high-performance computing network (Simpson, 2015), but the command line environment demands substantial technical expertise. In many cases DH faculty find themselves trying to act as systems administrators, programmers, and Web designers, as well as researchers and subject experts. Library-supported research platforms can remove the burden of infrastructure and systems administration from researchers, by offering server, storage, and application-level management expertise.

Another well-documented problem is that of long-term sustainability and the preservation of digital humanities projects and their assets (Marcum, 2016; Rumsey, 2016). Canadian academic libraries are working with national organizations such as Compute Canada and Research Data Canada to devise a national strategy for the long-term preservation of digital research assets (Humphrey, 2016). Any robust strategy will require standard descriptions of objects and the standardization of formats, along with a workflow that ensures digital objects are submitted to preservation best practices. Libraries are campus leaders in the area of information preservation and long-term curation, and projects that are built on library digital asset management systems will automatically be subject to the same preservation expertise that the library brings to bear on its on local digital collections.

By providing a read-write research environment, the library can help to meet faculty research needs, and relieve scholars of the burden of systems administration, while also capturing project outputs for long-term access and preservation.

Features of the read-write library

The possible features of digital research environments are endless, but library resources are not. Although faculty may ideally prefer a technical workspace that is custom-built

for a given project, this is a hugely resource-intensive undertaking with significant sustainability issues. The read-write library can only be sustained if it is developed within the context of platforms that libraries are already building to manage their own mission critical services. Two of the most promising next-generation library platforms are Islandora CLAW (Ruest & Lamb, 2016) and Hydra Sufia (Sufia Project, 2016). Both are widely deployed open source systems that are in active development. The Fedora 4 (Duraspace Project, 2016) digital asset management system is the foundation of both of these products. Fedora is a highly scalable and durable digital repository that is used by cultural memory organizations around the world. Object metadata is stored natively in key value pairs that translate easily into Resource Description Framework (RDF) triples using standard ontologies such as Dublin Core and the Portland Common Data Model (Estlund, 2015). Fedora is quickly emerging as a gold standard for the management and dissemination of digital content.

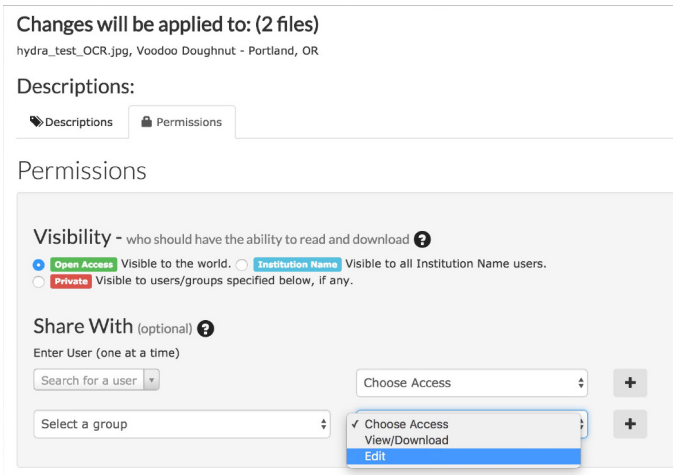
Islandora CLAW, the next major release of Islandora, couples more tightly with the Drupal environment to permit easier integration of external modules that can extend software functionality without the need for locally written code (Ruest & Lamb, 2015). Hydra is a Ruby on Rails environment that leverages Blacklight (Project Blacklight, 2015) discovery. Hydra is an extremely flexible development environment that offers limited out-of-the-box functionality, but allows libraries to design custom interfaces to suit their own needs. Sufia is the most common interface for Hydra repositories.

To a large extent these next-generation systems are concerned with aspects of digital asset management, description, publishing, discovery, and preservation. Although these activities do not encompass every aspect of the research process, there is significant overlap between the interests and needs of libraries and those of researchers, particularly researchers in humanities disciplines. It is worthwhile to consider how the main features of emerging digital asset management systems (DAMS) might support scholarly research projects, and to identify the limits of library DAMS as research environments.

Identity management and access

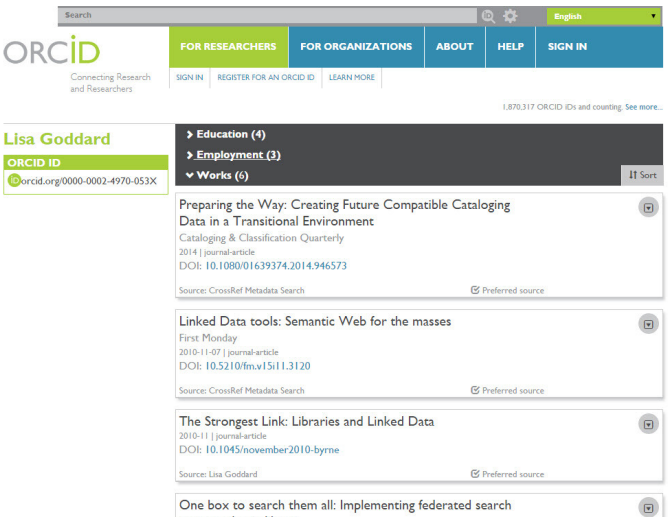
Because of the collaborative nature of digital scholarship, most large research projects require access for a team of researchers who will work together on a set of texts or objects. Library DAMs are capable of supporting multiple sites and collections on a single platform. This means the library can offer dedicated research environments to multiple groups without installing additional software. Create, read, update, and delete (CRUD) permissions can be managed in granular ways for individuals and groups, and DAMS platforms can be integrated with the university's central identity management systems using Lightweight Directory Access Protocol (LDAP), Central Authentication Service (CAS), or Shibboleth (see Figure 1).

Figure 1: A screenshot from Hydra's Sufia DAMS interface showing the granularity and scope of file permission options in digital asset management systems.



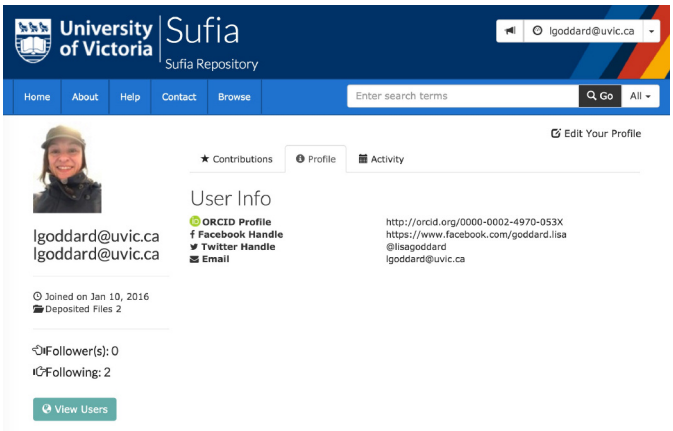
Ideally all of our research systems would use a unique, university-verified, global identifier to represent each researcher. Readers could thus be assured of the expertise and authority of contributors. It would also be a means of clearly distinguishing the contributions of individual faculty members and graduate students to allow for proper citation and attribution. Global identifiers could follow researchers across the Web as they engage in a variety of research endeavours in different platforms and with different groups. This identity would not only promote trust within scholarly applications but could also be used to link together the various contributions of a given researcher, and to reveal networks of collaboration. One of the most promising systems for global researcher identification is ORCID (Brown, Wilmers, & Haak, 2015), a system that links unique researcher Uniform Resource Identifiers (URIs) with details about education, employment, and publication (see Figure 2). Library DAMS are capable of storing and exposing these kinds of identifiers, but this functionality will be of limited use until such identifiers are more universally adopted by researchers.

Figure 2: The ORCID site allows researchers to obtain a unique URI that acts as an unambiguous identifier and links an author's research across multiple platforms and aggregators.



Beyond file access permissions and basic identity management, the new systems are designed to support social scholarship. The Hydra/Sufia interface allows contributors to expose their user profiles publicly, including their ORCID URI and other social media accounts. The user profile includes an automatically generated list of contributed objects and permits users to follow one another's activity feeds (see Figure 3).

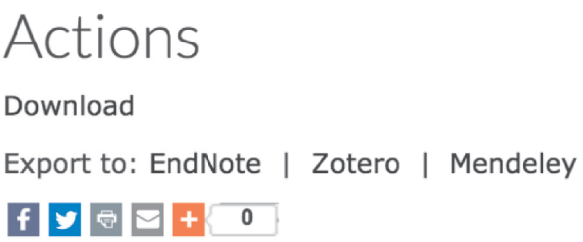
Figure 3: Hydra Sufia offers "social scholarship" features, including user profiles where researchers can share their object streams and follow other users. ORCID identifiers and information about other social media accounts are also featured.



Each object record includes the ability to share that document, photograph, or other asset in a variety of social media platforms or to direct export to bibliographic metadata and object URIs to citation managers so that articles, reports, or research objects can easily be shared and linked back to the object URI. Citation managers are increasingly becoming hubs of social scholarship. One of the new "Altmetric" measures is the extent to which a particular article has been saved in scholarly citation managers (Priem, Taraborelli, Groth, & Neylon, 2010). Software such as Mendeley goes one step further in providing recommended readings based on articles that are popular with other readers whose citation collection suggests similar research interests. By facilitating the easy sharing of research outputs in social scholarship applications, and by allowing scholars to easily find and

follow others working in related areas, DAMS help to integrate faculty research into scholarly networking platforms (see Figure 4).

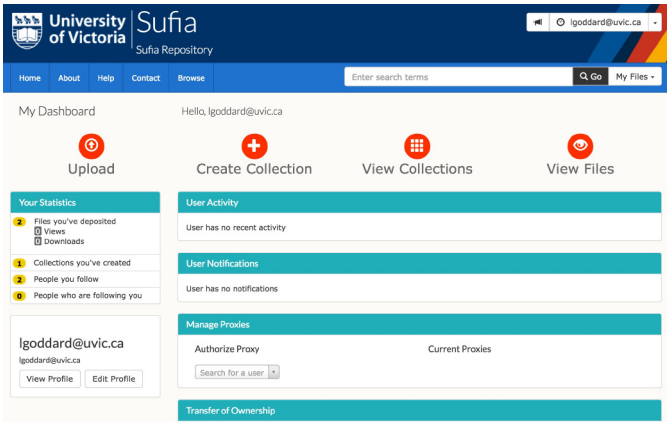
Figure 4: Each record-level object can be easily shared in popular social media platforms. Structured citation information can be exported directly to citation managers.



Object ingest and transformation

DAMS software is fundamentally designed to support the ingest and organization of digital objects. The term "digital object" comprises a huge spectrum of information carriers including text, still images, video, software packages, and computer files of all kinds. Digital asset management software can benefit any project that relies on collecting, aggregating, or otherwise amassing objects of study for annotation and manipulation. This is also true for projects that produce large numbers of digital files that need to be stored, shared, and preserved over the long term. DAMS typically offer options for the upload of individual objects, or the batch upload of many objects simultaneously. There are options to capture object metadata during upload, and options to transform objects during upload. Transformation might include methods to improve the Web accessibility of digital objects, including video compression and streaming, the creation of thumbnails and lower-resolution images for Web browsing, or the transformation of large-scale map and poster files for use with pan and zoom software. Optical character recognition is offered to convert digitized images of text documents into plain text files that can be indexed for full text search or marked up in a variety of ways (see Figure 5).

Figure 5: Hydra Sufia offers an intuitive Web-based dashboard that allows for easy management of research objects and collections.



While library digital asset management systems can ingest a wide variety of files for storage and retrieval, there is a limited amount of file manipulation that can happen within the platform. It is possible, for example, to upload a spreadsheet of data, but queries and visualizations cannot then be run automatically over that data. The spreadsheet would have to be downloaded for manipulation in statistics software, such as SPSS, R, or Excel. Data-intensive projects may be better served by a repository such as Dataverse (Crosas, 2011), which is specifically designed to share and search numerical datasets. Another limitation is that although software packages could theoretically be uploaded to the system for access and preservation, it would not be possible to execute that code within the DAMS. The DAMS would simply allow a user to find and download the software package for execution in a local environment.

Object description

In order to be able to identify, search, and sort a collection of digital objects, the research team has to provide descriptions of the objects. Basic descriptions are usually entered using a Web form with separate fields for common elements such as title,

creator, date, subject, and provenance. This information is especially important for the search and discovery of any file that cannot be keyword searched, including images, audio, video, and hand-written manuscripts. DAMS systems can usually accommodate custom metadata forms, so each project can capture the information that is of interest in the context of a particular research question (see Figure 6).

Figure 6: Library DAMS support the easy creation of structured metadata using a standard such as Dublin Core that will maximize the machine readability and interoperability of records.

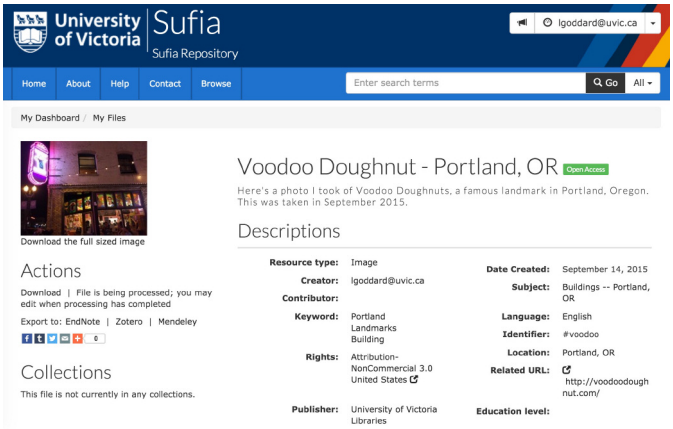


Figure 7: DAMS automatically generate an audit history for each digital object, including checksums and other means of ensuring file integrity over time.

File Details

Depositor: lgoddard@uvic.ca
Date Uploaded: 2016-01-10T23:39:16+00:00
Date Modified: 2016-01-10T15:52:52-08:00
Audit Status: Audits have not yet been run on this file.
Characterization: File format: jpeg (JPEG File Interchange Format)
Mime type: image/jpeg
File size: 1742910

Last modified: 2016:01:10 15:39:21-08:00
Filename: IMG_0332.jpg
Original checksum: 1400a4e9f032bbb08b6d56810e846daf
Orientation: normal*
Scanning software: Photos 1.3
Exif version: 0221
Gps timestamp: 04:40:07
Pages:

User Activity	Date
User lgoddard@uvic.ca has updated Voodoo Doughnut - Portland, OR	less than a minute ago
User lgoddard@uvic.ca has updated Voodoo Doughnut - Portland, OR	3 minutes ago
User lgoddard@uvic.ca has updated Voodoo Doughnut - Portland, OR	4 minutes ago
User lgoddard@uvic.ca has updated Voodoo Doughnut - Portland, OR	6 minutes ago
User lgoddard@uvic.ca has updated Voodoo Doughnut - Portland, OR	7 minutes ago
User lgoddard@uvic.ca has updated Voodoo Doughnut - Portland, OR	11 minutes ago
User lgoddard@uvic.ca has deposited IMG_0332.jpg	14 minutes ago

It is also possible to apply controlled vocabularies to certain fields, so that the project lead can ensure that values such as subject headings or geographic place names are entered consistently. Controlled fields are particularly important in the creation of faceted search interfaces, and for interoperability between projects. Administrative information about system activity is logged by DAMS, so the record creator, date of creation, file format, and edit history are captured automatically (see Figure 7).

Fedora 4-based DAMS systems, such as Hydra and Islandora, store digital object information as RDF triples in a key-value store (Cramer, 2013). This means that unique, resolvable, global identifiers will automatically be applied to every object within the library DAMS, and descriptive metadata will be modeled as RDF triples. Middleware, such as Apache Camel, can export that metadata to an RDF triple store (Woods, 2015). Linked data and RDF are becoming increasingly popular in DH applications of all kinds, and are the basis of projects such as the Canadian Writing Research Collaboratory (Brown & Simpson, 2015), Linked Modernisms (Ross, Christie, & Sayers, 2014), Linked Jazz (Patuelli, Miller, Lange, and & Thorsen, 2013) and the InPho Project (Sztyler, Huber, Noessner, Murdock, Allen, & Niepert, 2014). DAMS offer the ability for projects to create and expose RDF metadata without the need for any expert knowledge in this technology. Simply by entering descriptive information into a Web form, researchers will be creating linked data representations of their records.

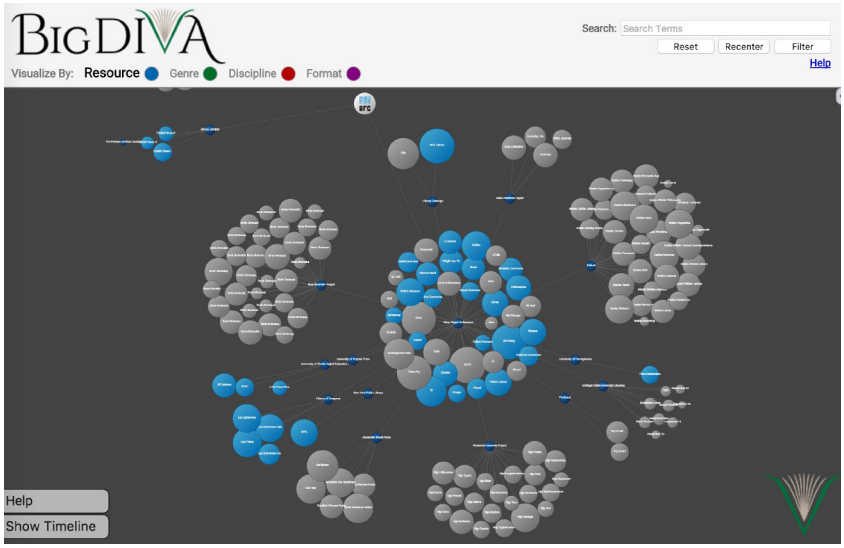
Metadata management is a core library function, and DAMS systems include a variety of ways to expose metadata for export, querying, and reuse. This includes XML-based export formats, such as MODS/METS and Dublin Core/OAI-PMH. While DAMS can accommodate a great diversity of metadata, a certain amount of standardization is necessary for interoperability and file interchange. It is likely that libraries will insist on the inclusion of certain basic fields that are common across a variety of library standards, such as MARC, Dublin Core, and MODS. Extensive data

interoperability is a major benefit of using a library DAMS to store digital research objects. Project data can easily be exposed for aggregation in a larger project such as Networked Infrastructure for Nineteenth-Century Electronic Scholarship (NINES), 18thConnect, or Renaissance English Knowledgebase (ReKN) (Mandell, 2012).

Indexing, search, and discovery

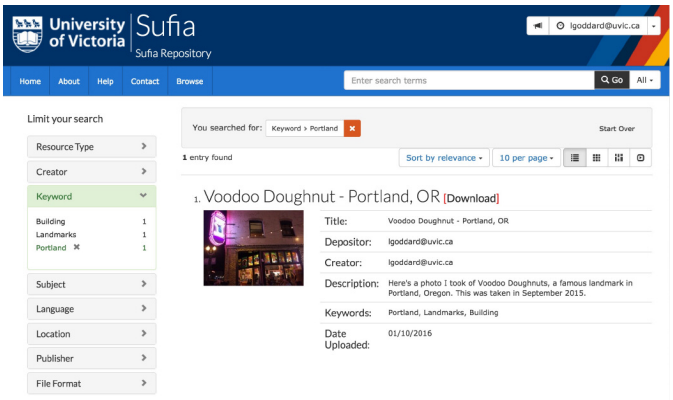
As descriptive metadata is entered in an object record it is immediately indexed for search and retrieval. DAMS metadata is entered in forms that allow for the advanced searching of specific field combinations, as well as the general keyword searching of the metadata and full-text content of each object. Both Hydra and Islandora rely on Apache Solr as the core indexing and search server. Solr is world-class enterprise search platform that enables phrases, wildcards, facets, and much more across any data type. Object metadata is exposed for harvesting in Web search engines to facilitate broad discovery.

Figure 8: Fedora-based DAMS expose options for exporting objects, links, and records in a host of standard formats that can plug in to data aggregations, such as the Big Diva catalogue, administered by the Advanced Research Consortium (ARC).



Search interfaces within digital asset management systems are still fairly traditional keyword, browse, and faceted interfaces. Graph-based discovery interfaces, such as Big Diva (Grumbach & Mandell, 2014) [see Figure 8] or Linked Modernisms (Ross et al, 2014), are not yet common. Data visualization tools are becoming more common, but are still not particularly sophisticated within DAMS. Text analysis and mining tools, such as Voyant (Sinclair & Rockwell, 2016) or the Software Environment for the Advancement of Scholarly Research (SEASR) (Ashton, 2011), offer much more sophisticated analysis options than a DAMS interface (Figure 9) would provide. In the long term it is likely that data visualization and text-mining tools will be integrated into digital library interfaces. In the short term researchers will have to export their data in order to apply advanced analysis and visualization techniques. In the context of analysis, the function of the DAMS is to ensure that data is entered in a consistent, structured way, and output in standard XML or JSON formats that will enable more sophisticated analysis in other tools.

Figure 9: DAMS discovery interfaces offer a variety of keyword and browse search options. Objects are also



The Fedora 4 object store includes a RESTful application programming interface (API) that exposes methods for reading, exporting, adding, and updating objects and descriptive metadata (Armintor, 2014).

This provides a great deal of interface flexibility. The DAMS interface could be used for data entry and storage, but the use of APIs would allow a research team to develop Web interfaces and presentation layers that are customized for a project, or to plug DAMS objects and functionality into an existing project website.

Sustainability and long-term preservation

Limited-term funding cycles have lead to massive sustainability problems for early digital humanities projects (Dobreva, 2013; Johnston, 2013; Kretzschmar & Potter, 2010; Munoz & Flanders, 2015; Reed, 2014). Server access, network storage, link maintenance, backups, and software upgrades have annual costs that will persist long after the initial funding is gone. Projects and their outputs become orphaned as graduate students and programmers move on from the university, and faculty members become involved in other projects. Most projects are abandoned to decompose benignly online, but obsolete software platforms and unpatched operating systems can become security risks for the whole organization.

Research that is developed and stored within the library DAMS will automatically be subject to the library's best practices for digital object preservation. The DAMS enforces standard file formats, and standard metadata that can easily merge into existing digital preservation strategies. The library has full access to the DAMS system, and so has the ability to intervene to preserve materials. The DAMS is a critical library application, so there is a strong commitment to maintaining and upgrading the platform over time. When the current DAMS becomes obsolete, faculty data can be migrated to the new system using the same method that the library will develop to port its own digital library content. Use of the DAMS will build research projects directly into the library's preservation workflow. As funders increasingly require formal action on long-term sustainability (NEH, 2015; NSERC, 2015), the availability of a library DAMS can help to address the requirements of data management plans within grant application.

Limits of the read-write library

Libraries are actively developing a new generation of digital asset management tools that can have significant application in faculty research projects. These new virtual collaboration spaces form the basis of the read-write library – a cultural knowledge base that will allow libraries to combine traditional collection development and description work with the subject expertise of scholars, students, and the wider community. Most projects require core environments within which to collect, share, describe, and organize digital files of all kinds. The management of these collections requires identity management, indexing, Web publishing, batch import and export utilities, and long-term preservation services. Library DAMS can offer all of these services via existing software infrastructure that is administered by information technology (IT) professionals and librarians within the current confines of library budgets. Faculty who use the library DAMS as a research environment reap significant project benefits, while helping to enrich library digital collections.

Library digital asset management systems such as Hydra and Islandora are a significant first step toward a more inclusive read-write library model, but they cannot satisfy all of the technological needs of digital scholars. It is unlikely that libraries will allow

Goddard, Lisa. (2016). Developing the Read/Write Library. *Scholarly and Research Communication, 7*(2): 0201255, 13 pp.

scholars to modify the code in production DAMS environments, nor will they necessarily allow scholars to execute local code on library-managed servers. It is also unlikely that faculty members will be given high-level administrative access to the system, so customization of the research environment will be limited. DAMS are not currently optimized for the collaborative annotation and mark-up of full text objects, and they lack sophisticated visualization and analysis tools. Hydra and Islandora are both highly extensible open source environments, however, so it is possible to develop them in these directions if the academic community chooses to direct resources to that end. Library DAMS can provide an inexpensive, fully managed solution to many of the challenges of digital research project management; they can fill an important niche in research infrastructure for digital humanists; and they can act as the basis of collaborative read-write environments for knowledge mobilization. These systems have reached a level of maturity that enables a suite of critical user and data management tools, while remaining extremely flexible in terms of interface development and tool integration. There is a huge opportunity for faculty and librarians to come together to fund, direct, and participate in the further development of projects such as Hydra and Islandora in order to shape them into full-featured research environments for digital scholarship.

Websites

18thConnect, http://www.18thconnect.org/
Advanced Research Consortium, http://idhmcmain.tamu.edu/arcgrant/
Apache Solr, http://lucene.apache.org/solr/
Big Diva, http://bigdiva.org/
Canadian Writing Research Collaboratory, http://www.cwrc.ca/en/
Hydra Project, https://projecthydra.org/
Indiana Philosophy Ontology Project, https://inpho.cogs.indiana.edu/
Islandora, http://islandora.ca/
Linked Jazz, https://linkedjazz.org/
Linked Modernisms, http://linkedmods.uvic.ca/
NINES, http://www.nines.org/
Renaissance Knowledge Network, http://rekn.itercommunity.org/
SEASR, http://www.seasr.org/
Voyant, https://voyant-tools.org/

References

Ashton, Andrew Thomas. (2011). *Semantically rich tools for text exploration: TEI and SEASR.* Digital Humanities 2011 Conference Abstracts (pp. 270–271).

Armintor, Benjamin. (2014). Fedora project documentation: RESTful HTTP API. *Duraspace Wiki.* URL: https://wiki.duraspace.org/display/FEDORA40/RESTful+HTTP+API [December 13, 2015].

Brown, Josh, Wilmers, Catalina, & Haak, Laurel. (2015). Final report: Sloan ORCID adoption and integration program 2013–2014. *Figshare.* URL: https://dx.doi.org/10.6084/m9.figshare.1290632.v1 [December 5, 2015].

Brown, Susan, & John Simpson. (2015). An entity by any other name: Linked open data as a basis for a decentered, dynamic scholarly publishing ecology. *Scholarly and Research Communication, 6*(2).

Cramer, Tom. (2013) Fedora 4 as the Hydra RDF store. *Duraspace Wiki.* URL: https://wiki.duraspace.org/display/hydra/Fedora+4+as+the+Hydra+RDF+store [December 13, 2015].

Crosas, M. (2011) The dataverse network: An open-source application for sharing, discovering and preserving data. *D-Lib Magazine, 17.* URL: http://dx.doi.org/10.1045/january2011-crosas [December 13, 2015].

de Rosa, Cathy. (2010). *Perceptions of libraries: Context and community* (p. 40). Dublin, OH: OCLC. URL: https://www.oclc.org/reports/2010perceptions.en.html [December 13, 2015].

Dobreva, Milena. (2013). *Preservation scenarios for projects producing digital resources in the area of digital humanities.* Bath, UK: Digital Curation Centre. URL: http://www.dcc.ac.uk/sites/default/files/documents/dcc_dh_final.pdf [December 13, 2015].

Duraspace Project. (2016). *About Fedora.* URL: http://fedorarepository.org/about [March 15, 2016].

Estlund, Karen. (2015). *Portland common data model.* URL: https://github.com/duraspace/pcdm/wiki [December 13, 2015].

Grumbach, Elizabeth, & Laura Mandell. (2014). Meeting scholars where they are: The Advanced Research Consortium (ARC) and a social humanities infrastructure. *Scholarly and Research Communication, 5*(4).

Humphrey, Chuck. (2016, January 27). *Compute Canada and CARL join forces to build a national research data platform.* Ottawa, ON: Canadian Association of Research Libraries. URL: http://www.carl-abrc.ca/news/compute-canada-carl-join-forces-build-national-research-data-platform/? [May 11, 2016].

Johnston, Leslie. (2013, April 12). Digital humanities and digital preservation. *The Signal.* Washington, DC: Library of Congress. http://blogs.loc.gov/digitalpreservation/2013/04/digital-humanities-and-digital-preservation/ [November 27, 2015].

Kretzschmar, William, & Potter, William. (2010). Library collaboration with large digital humanities projects. *Literary and Linguist Computing, 25*(4), 439-445. http://dx.doi.org/10.1093/llc/fqq022 [December 5, 2015].

Levoy, Marc. (2015). *The digital Michelangelo project.* Stanford, CA: Stanford University. URL: http://graphics.stanford.edu/projects/mich/ [December 13, 2015].

Mandell, Laura. (2012). Promotion and tenure for digital scholarship. *Journal of Digital Humanities, 1*(4). URL: http://journalofdigitalhumanities.org/1-4/promotion-and-tenure-for-digital-scholarship-by-laura-mandell/ [December 13, 2015].

Marcum, Deanna. (2016, April 26). Due diligence and stewardship in a time of change and uncertainty. *Ithaka S+R Issue Brief.* URL: http://www.sr.ithaka.org/publications/due-diligence-and-stewardship-in-a-time-of-change-and-uncertainty/ [May 11, 2016].

Moretti, Franco. (2005). *Graphs, maps, trees: Abstract models for a literary history.* London, UK: Verso.

Munoz, Trevor, & Flanders, Julia. (2015). *Digital humanities data curation guide.* Graduate School of Library and Information Science (GSLIS), University of Illinois, Urbana-Champaign. URL: http://guide.dhcuration.org/contents/ [December 13, 2015].

National Endowment for the Humanities. (2015). *Data management plans for NEH Office of Digital Humanities proposals and awards.* URL: http://www.neh.gov/files/grants/data_management_plans_2015.pdf [December 5, 2015].

Natural Sciences and Engineering Research Council of Canada (NSERC). (2015). *Draft tri-agency statement of principles on digital data management.* URL: http://www.science.gc.ca/default.asp?lang=En&n=83F7624E-1 [December 13, 2015].

Pattuelli, Cristina, Miller, Matt, Lange, Lea, & Thorsen, Hilary. (2013). *Linked Jazz 52nd street: A LOD crowdsourcing tool to reveal connections among jazz artists.* Proceedings of Digital Humanities, 337–340. URL: http://dh2013.unl.edu/abstracts/ab-254.html [November 27, 2015].

Poynder, Richard. (2011). The big deal: Not price but cost. *Information Today, 28*(8). URL: http://www.infotoday.com/it/sep11/The-Big-Deal-Not-Price-But-Cost. [December 13, 2015].

Priem, Jason, Taraborelli, Dario, Groth, Paul, & Neylon, Cameron. (2010, October 26). *Altmetrics: A manifesto.* URL: http://altmetrics.org/manifesto [December 8, 2015].

Project Blacklight. (2015). *Blacklight: A multi-institutional open-source collaboration building a better discovery platform framework.* URL: http://projectblacklight.org/ [December 5, 2015].

Reed, Ashley. (2014). Managing an established digital humanities project: Principles and practices from the twentieth year of the William Blake archive. *Digital Humanities Quarterly, 8*(1). URL: http://www.digitalhumanities.org/dhq/vol/8/1/000174/000174.html [December 13, 2015].

Ross, Stephen, Alex Christie, & Jentery Sayers. (2014). Expert/crowd-courcing for the Linked Modernisms Project. *Scholarly and Research Communication, 5*(4).

Ruest, Nick, & Daniel Lamb. (2016). About Islandora CLAW. URL: http://islandora-claw.github.io /CLAW/ [May 11, 2016].

Ruest, Nick, & Daniel Lamb. (2015). *Islandora and Fedora 4; The Atonement v3: The Atonermenter.* iCamp CT. Hartford, CT. October 20-22, 2015. URL: http://hdl.handle.net/10315/29505 [December 13, 2015].

Rumsey, Abby Smith. (2016, May 4). The risk of digital oblivion. *The Chronicle of Higher Education.* URL: http://chronicle.com/article/The-Risk-of-Digital-Oblivion/236342 [May 11, 2016].

Simpson, John. (2015). *Digital humanities working group.* North York, ON: Compute Canada. https:// www.computecanada.ca/research-portal/digital-humanities-working-group/ [December 5, 2015].

Sinclair, Stéfan, & Rockwell, Geoffrey. (2016). *Text analysis and visualization* (pp. 274-290). A new companion to digital humanities. New York, NY: Wiley.

Sufia Project (2016). *About Sufia.* URL: http://sufia.io/ [December 13, 2015].

Sztyler, Timo, Jakob Huber, Jan Noessner, Jaimie Murdock, Colin Allen, & Mathias Niepert. (2014). *LODE: Linking digital humanities content to the web of data.* Proceedings of the 14th ACM/IEEE-CS Joint Conference on Digital Libraries, pp. 423-424.

Woods, Andrew (2015). Fedora project documentation: External triplestore. *Duraspace Wiki.* URL: https://wiki.duraspace.org/display/FEDORA4x/External+Triplestore [December 13, 2015].