

---

# Case study: Streamlining digital archiving workflows at the Bentley Historical Library

Received (in revised form): 11th April, 2017



## Max Eckard

is the Archivist for Digital Curation at the University of Michigan Bentley Historical Library and serves on the ArchivesSpace Technical Advisory Committee, heading the Integrations sub-team. He is passionate about archives, digital curation and end users.

Bentley Historical Library, University of Michigan, Ann Arbor, MI 48109, USA  
Tel: +1 734 763 7518;  
E-mail: eckardm@umich.edu



## Dallas Pillen

is the Archivist for Metadata and Digital Projects at the University of Michigan Bentley Historical Library, where he is involved with the Bentley's implementation and use of ArchivesSpace and oversees its web archiving programme.

Bentley Historical Library, University of Michigan, Ann Arbor, MI 48109, USA  
Tel: +1 734 647 3559;  
E-mail: djpillen@umich.edu

**Abstract** In October 2016, the Bentley Historical Library, in partnership with the University of Michigan Library, wrapped up a two-year, US\$355,000 grant from the Andrew W. Mellon Foundation to enhance the interoperability of three open-source software systems: ArchivesSpace, a web application for managing archival information; Archivematica, a digital preservation system; and DSpace, a repository application that preserves and enables easy and open access to all types of digital content. The ArchivesSpace–Archivematica–DSpace Workflow Integration project integrates these systems into an end-to-end digital archives workflow, facilitating the creation and reuse of metadata among these platforms and streamlining the ingest and deposit of fully processed born-digital materials into a digital preservation and access repository. After outlining the institutional context and the Bentley's guiding principles for digital processing, this case study discusses the effort that went into developing this workflow, including the software development process, as well as the effort that went into migrating data from legacy systems and preparing archivists and other stakeholders to adopt new tools and workflows. The case study concludes by reflecting on project outcomes and lessons learned.

**KEYWORDS:** digital preservation, digital curation, systems integration, archival workflows, metadata

## INTRODUCTION

For more than 80 years, the Bentley Historical Library (the Bentley) at the University of Michigan has collected documents, films, photographs and other historical records related to the state of Michigan and the University of Michigan. The Bentley's more than 11,000 research collections range from the papers of governors of Michigan, to the records of student and faculty life at the university, to the entire historical record of intercollegiate athletics at Michigan, and much more. Holdings include over 70,000 linear feet of historical records, 55,000 books, 10,000 maps and 1,500,000 photographs and negatives. These days, holdings include nearly 80 TB of digitised and born-digital assets, as well as web and social media archives.

In October 2016, the Bentley, in partnership with the University of Michigan Library (MLibrary), wrapped up a two-year, US\$355,000 grant from the Andrew W. Mellon foundation to enhance the interoperability of three open-source software systems: ArchivesSpace (<http://archivesspace.org/>), a web application for managing archival information; Archivematica (<https://www.archivematica.org/en/>), a digital preservation system; and DSpace (<http://www.dspace.org>), a repository application that preserves and enables easy and open access to all types of digital content. This end-to-end digital archives workflow facilitates the creation and reuse of metadata among these platforms, streamlines the ingest and deposit of fully processed born-digital materials into a digital preservation and access repository, advances professional best practices, and addresses an essential finding of a 2012 Association of Research Libraries survey<sup>1</sup> that found that 43 per cent of respondents had not developed workflows for the ingest and processing of born-digital materials due to such obstacles as the selection of appropriate tools, staff training, and development of policies and procedures. This paper briefly outlines the

evolution of this workflow, including the software development process, as well as the substantial effort that went into migrating data from legacy systems and preparing archivists and other stakeholders to adopt new tools and workflows.

## Goals of digital processing at the Bentley

The Bentley has managed and preserved born-digital archives — a term used in libraries and archives to identify content produced in digital form, rather than having been digitised from physical or analogue content — since the accession of a former university president's personal computer, in 1997. Since 2006, the Bentley has provided online access to born-digital archives via DeepBlue (<https://deepblue.lib.umich.edu/>), the University of Michigan's permanent, safe, and accessible service for representing its rich intellectual community, based on DSpace Digital processing procedures<sup>2</sup> correspond to the Generate Archival Information Package (AIP)<sup>3</sup> function of the Reference Model for an Open Archival Information System's Ingest entity and prepare items for deposit to DeepBlue.

While the Bentley's digital processing procedures have evolved dramatically since their early days of a highly manual workflow that involved more than 40 discrete steps, the operation of numerous standalone applications and saving output in various log files, the overarching goals and principles have largely stayed the same:

- *The appraisal and description applied to digital archives is not inherently different from the appraisal and description applied to physical or analogue archives:*<sup>4</sup> In other words, digital processing procedures reflect, above all, traditional archival functions such as appraisal, arrangement and description, as well as the Bentley's institutional policies. The Bentley employs ubiquitous archival content standards and processing strategies, such as Describing Archives: A Content

Standard and More Product Less Process,<sup>5</sup> in its arrangement and description of its archives. The amount of resources allocated to processing a particular collection and the level at which the collection is described is not dependent on that collection's format. By default, both digital and physical archives are appraised, arranged and described in aggregate in a way that delineates the essential functions and activities that the materials document. For example, formats such as MBOX are not arranged by default in a 'Digital Correspondence' or 'E-mail' series when a 'Correspondence' series exists and the files in a directory of office documents are not described at the individual file level.

- *Additional tools or procedures are necessary to ensure the authenticity, integrity and security of digital archives:* While these additional tools and procedures are necessary for digital preservation actions, the Bentley takes a common sense approach to appraisal: not all procedures will be required for all accessions and in cases where donors provide detailed descriptive information for fairly homogeneous content, appraisal may be fairly minimal.
- *The processing manual — and the workflows and digital processing procedures it records — is a living document:* Over time, systems, workflows and tools change, and users of the manual participate actively in the ongoing evolution of the manual by revising and improving it within the framework defined by professional consensus on contemporary archival theory. When possible — even when new systems are being introduced to workflows — changes are made iteratively and incrementally to ensure that this process of change and evolution is smooth and natural.

### **Previous workflow for digital processing at the Bentley**

Prior to integrating ArchivesSpace, Archivematica and DSpace into an

end-to-end digital archiving workflow and introducing new appraisal functionality into Archivematica, the Bentley's digital processing workflows involved a number of systems and tools that, while more efficient than their highly manual predecessors, were not completely satisfactory.

### ***Tracking accessions in BEAL***

In 2001, the Bentley began using an internally developed FileMaker Pro database called BEAL, an acronym that reflects the street on which the Bentley is located, Beal Avenue, to manage much of the Bentley's administrative metadata, including information about accessions, donors and locations of materials. Over time, the scope of this database grew to include descriptive and administrative metadata for audiovisual digitisation projects.

Accession records in BEAL included the information necessary to: document the provenance of an accession, including the date on which material was accessioned, the donor's name and contact information, the staff member responsible for receiving the accession and the status of the gift agreement associated with the accession; inform processing prioritisation by noting the relative importance of an accession; and provide processors with a brief description of the content and an indication of whether or not particular rights and privacy issues may be present in the accessioned material. Accessions containing digital content were also recorded in BEAL, along with the file size of the content and its location in the Bentley's digital processing backlog. The use of a structured and easy-to-use database application to manage accession records permitted accessioning archivists to record pertinent information about a collection, regardless of format, and for processors of physical and digital archives to later reference that information, providing a consistent, managed system that facilitated, rather than got in the way of, essential archival functions of appraisal and description.

### *Generating EADs for DLXS from Microsoft Word documents and macros*

The Bentley used Encoded Archival Description (EAD) to internally manage its finding aids and to provide users with online access to archival description for many years prior to adopting ArchivesSpace. The Bentley's nearly 3,000 EADs were authored as Microsoft Word documents with particular EAD elements corresponding to Microsoft Word paragraph styles. Those styles were referenced in macros that converted the Word Document into EAD XML (see: <http://web.archive.org/web/20150223135541/http://bentley.umich.edu/EAD/bhlfiles.php>), which was then transformed and indexed in the Bentley's finding aid delivery platform (<http://bentley.umich.edu/legacy-support/EAD/>) based on DLXS (<http://www.dlxs.org>).

The practice of using Microsoft Word documents to generate structured data significantly reduced the technical barriers to creating EAD at the library. Student processors, in addition to most of the library's professional staff, did not need to encode EAD by hand or know much about the schema's structure, elements or attributes, but were able to use ubiquitous word processing software to author structured archival description with ease.

### *Digital processing with AutoPro and archival storage with DeepBlue*

For many years after the initial accession of a former university president's personal computer in 1997, procedures for digital processing at the Bentley were largely manual. In 2009, the Bentley received a two-year grant, also from the Andrew W. Mellon Foundation, entitled 'E-mail Archiving at the University of Michigan'. An important outcome of this grant for the Bentley was that it added additional storage infrastructure (including the use of DeepBlue, the University of Michigan's DSpace repository) as well as two staff

members to enhance and standardise digital preservation policies and procedures.

In 2011, Bentley staff developed and implemented a home-grown digital preservation tool written in Windows CMD. EXE shell scripts, the AutomatedProcessor (AutoPro), to make digital processing more efficient by automating essential workflow steps and reducing technical barriers, permitting archivists to focus their energies on the traditional archival functions of appraisal, arrangement and description. AutoPro (<https://deepblue.lib.umich.edu/handle/2027.42/97298>) automated essential portions of the digital preservation workflow and provided a single interface for more than 20 applications and guided users through a standardised workflow that included:

- a virus scan with MS Forefront Endpoint;
- creating a working backup with a batch file;
- opening archive files with extensions like ZIP and TAR with 7-ZIP;
- identifying and correcting missing file extensions with TrID and DROID;
- creating preservation copies for at-risk file formats with ffmpeg, ImageMagick, Ghostscript, Inkscape, Aid4Mail, and Office File Converter;
- scanning for personally identifiable information (PII) like credit card and Social Security numbers as well as other sensitive information with bulk\_extractor;
- appraising with QuickView Plus and TreeSize Professional, arranging and separating with a batch file, and packaging materials with 7-Zip, as needed;
- creating descriptive and administrative metadata with a batch file;
- extracting technical metadata and generating checksums with DROID;
- transferring material and associated metadata to long term-storage (if 'open', depositing an additional copy to DSpace) with BagIt and a batch file; and
- cleaning the processing directory and deleting temporary files with a batch file.

For many of these processes, AutoPro also standardised the creation of metadata and log files; this important information on provenance was ultimately included in the AIP.

### Shortcomings

While the use of BEAL to track accessions, Microsoft Word to author archival description and AutoPro to process digital content lowered technical barriers and introduced efficiencies into the Bentley's digital processing initiatives, there were numerous shortcomings to the tools.

The use of a custom FileMaker Pro database for tracking accessions and managing a variety of administrative and descriptive metadata meant that the Bentley's metadata were not in a widely used, open system. As archival institutions began to adopt collections management systems, first Archon and Archivists' Toolkit and later ArchivesSpace, the Bentley continued to manage its collections in BEAL, limiting the library's ability to take advantage of the affordances of more widely used systems, such as the ability to integrate with other tools.

Similar shortcomings were evident with the Bentley's tools and workflows for creating finding aids. The use of Microsoft Word made the process of generating EAD easier than hand-encoding XML, but training processors in Microsoft Word styles and macros made the process more localised and complicated than using industry standard tools for entering descriptive information. A reliance on Microsoft Word documents and EAD files, rather than a centralised collections management system, as the mechanism for storing authoritative descriptive metadata was problematic in that it ultimately led to a preponderance of finding aids in multiple locations, including printed copies and multiple versions of a finding aid's Microsoft Word document and encoded counterparts on shared drives and

individual desktop computers, complicating the processes of updating description and managing version control. An additional shortcoming of using EAD as the system for managing archival description was that the Bentley's description and encoding practices were limited to the style of EAD expected by DLXS and the kind of description that would be accessed by end users. This hindered the ability to structure metadata in a machine-parseable, rather than human-readable, way and complicated the management of metadata that continued to serve important internal functions after said data were no longer presented to end users, such as the management of containers for original materials after a collection had been microfilmed or digitised, in which case the finding aid was often updated to direct users only to access copies.

Finally, AutoPro was never intended to be a long-term solution. It had limited error handling, a poor user interface and various support issues. As acquisitions of digital archives continued to grow, it was clear that a more robust and user-friendly solution was required.

### Transition to ArchivesSpace and Archivematica

In recent years, ArchivesSpace and Archivematica have emerged as two of the most exciting open source platforms for working with digital archives. Upon review, the Bentley realised that ArchivesSpace and Archivematica addressed important aspects of the Bentley's approach to preserving and providing access to born-digital archives and had the potential to be integrated with the local DSpace repository in an end-to-end digital archives workflow. The Mellon Foundation-funded ArchivesSpace-Archivematica-DSpace Workflow Integration project, which concluded in October 2016, integrates these systems into an end-to-end digital archives workflow, facilitating the creation and

reuse of metadata among these platforms and streamlining the ingest and deposit of fully processed materials into a digital preservation repository. Before going live with these new systems and a new workflow, however, new features would need to be developed in Archivemata, archival description and legacy accession data would need to be migrated to ArchivesSpace, legacy born-digital archives currently in a backlog would need to be migrated to Archivemata's backlog, and accessioning and processing archivists at the Bentley would need to be prepared to adopt new tools and workflows.

#### *Working with Artefactual to Develop New Features in Archivemata*

The project had three major goals:

- Introduce functionality into Archivemata to permit users to review, appraise, deaccession and arrange content in a new 'Appraisal and Arrangement' tab in the system dashboard.
- Load (and create) ArchivesSpace archival object records in the Archivemata 'Appraisal and Arrangement' tab and then drag and drop content onto the appropriate archival objects to define submission information packages (SIPs) that will in turn be described as 'digital objects' in ArchivesSpace and deposited as discrete 'items' in DSpace.
- Create new archival object and digital object records in ArchivesSpace and associate the latter with DSpace handles to provide values for <dao> elements in exported EADs.

In the initial grant proposal, the Bentley intended to hire a software developer to handle the technical aspects introducing appraisal functionality into Archivemata and integrating it with ArchivesSpace and DSpace in a digital archives workflow. Due to a number of circumstances, however, the

Bentley was unable to secure any candidates after three months of intensive recruiting. At the time, the Bentley explored and ultimately decided to contract directly with Artefactual Systems (Artefactual), the lead developers of Archivemata, for the necessary development work due to the company's expert knowledge of Archivemata, extensive experience with the agile development of open source software for libraries and archives, and large following in the archives and digital preservation communities.

The partnership with Artefactual began in earnest with a site visit in early 2015 by Artefactual's President and its Director of Archivemata Technical Services. During the visit, a number of topics were discussed. The Bentley gave an overview of its existing digital backlog and digital processing workflows as well as its collections in DeepBlue. The existing features and functionality of ArchivesSpace and Archivemata were also discussed, as well as areas for future development and integration. The Bentley and Artefactual also met with MLibrary Library Information Technology (LIT) systems administrators to discuss the MLibrary's custom DSpace implementation and plans for an upcoming move to a Hydra environment. Artefactual demonstrated Archivemata for archivists, librarians, administrators and IT staff from the Bentley Historical Library, University of Michigan Library, Clements Library and Gerald R. Ford Presidential Library; they also reviewed the process of Archivemata installation and maintenance for LIT staff assisting with the installation of a local Archivemata instance.

Shortly thereafter, Artefactual prepared a basic Unified Modeling Language diagram of the proposed workflow. After a number of revisions by the Bentley and MLibrary, partners arrived at a version that served as a foundation for development work. Based on the site visit and follow-up calls, Artefactual prepared a

final consultation report that identified five essential development tasks:

- developing a new Appraisal and Arrangement tab in Archivematica;
- Archivematica–ArchivesSpace integration, including the ability for Archivematica to create archival object records in ArchivesSpace and associate digital objects, defining Archivematica SIPs and creating ArchivesSpace digital objects;
- repackaging of the AIPs — as part of the Bentley’s approach to providing access to AIPs and DSpace’s data model, Archivematica’s Storage Service needed to be able to package multiple files and/or folders into ‘objects’ and ‘metadata’ zip files to simplify patron access to and archival management of content;
- Archivematica–DSpace/DeepBlue integration, including the ability to upload data and metadata as well as the ability to update file URIs in ArchivesSpace digital object records with DSpace handles; and
- integration of external tools, including the ability to review transfer content using bulk\_extractor and the potential integration of other external tools for transfer analysis and file viewing.

After prioritising and refining the proposed development tasks, Artefactual began agile development cycles. Bentley archivists created user stories to inform feature development, and conducted weekly meetings with Artefactual to review current work, identify new priorities and plan next steps.

#### *Migrating archival description and accession data to ArchivesSpace*

The Bentley’s initial migration to ArchivesSpace, completed in August 2016, consisted of migrating approximately 3,000 legacy EAD finding aids into the collections management system. While ArchivesSpace does have a built-in EAD importer (<http://archivesspace.org/using-archivesspace>

/migration-tools-and-data-mapping), initial import testing revealed that 35 per cent of the Bentley’s EADs would fail on import and, even for those that imported successfully, some data elements were not mapped to the ideal ArchivesSpace field. As a result, Bentley staff decided to take the opportunity presented by migration to do a large amount of retrospective metadata cleanup across all of the Bentley’s EADs.

The large-scale metadata cleanup effort involved:

- reconciling EAD finding aids with their Microsoft Word and printed counterparts to create a single authoritative version;
- normalising extent types and splitting multipart extent statements into individual statements (eg the single statement ‘2 lin. ft. and 3 oversize folders’ became two statements: ‘2 linear feet’ and ‘3 oversize folders’);
- normalising dates into ISO 8601 format;
- normalising subject terms and the names of people, corporate entities, and families and attempting to reconcile all controlled access points with the Library of Congress Name Authority File and Library of Congress Subject Headings; and
- reformatting or moving elements in the EADs so that data would not be lost or inappropriately mapped during migration.

This large-scale cleanup effort involved modifying hundreds of thousands of EAD XML elements and necessitated a programmatic approach. Project staff employed the Python programming language (<https://www.python.org>) and tools such as OpenRefine (<http://openrefine.org>) to aid in the process. The Bentley also developed an ArchivesSpace plugin (<https://github.com/bentley-historical-library/bhl-ead-importer>) to modify portions of the application’s default EAD importer to meet local needs, with several bug fixes and generalisable improvements pushed upstream to the ArchivesSpace source code.

The process for preparing and migrating legacy accession data did not involve the same level of retrospective metadata cleanup, but did require a similar degree of testing, data mapping and advance planning. There is not an encoding standard for archival accession records similar to the encoding standard for archival description, but ArchivesSpace does offer a CSV template for importing accessions metadata into the system. Initial testing of the accessions CSV importer revealed that it would not meet the Bentley's needs for migrating data into ArchivesSpace, so the project team used Python to develop a local solution that uses the ArchivesSpace application programming interface (API). The use of the API to migrate accessions metadata gave Bentley staff significantly more options and a greater degree of control over its migration of legacy accessions metadata from FileMaker Pro into ArchivesSpace.

#### *Migrating backlog material to Archivemata's backlog*

At the time the Bentley was ready to migrate to Archivemata, the library was in possession of over 3.5 TB of unprocessed born-digital archives in its backlog. On average, transfers were 17.2 GB and contained about 6,000 files apiece, although single transfers were as large as 806.6 GB and over 200,000 files. Aware that Archivemata had some scalability issues when indexing transfers with large numbers of files and when creating exceptionally large database tables with metadata for every single item, Bentley staff both reduced the size and volume of the largest transfers and also allocated more resources to the Archivemata environment. In the end, the Bentley settled on transfers of no more than 50–60 GB or 50,000 files and a virtual machine with 16 GB of RAM, four virtual CPUs, and about 350 GB of disk for processing. MLibrary staff also made a number of changes to optimise

Archivemata, including a pull request that drastically cuts down on the number of files that must be indexed by removing any empty bulk\_extractor logs (<https://github.com/artefactual/archivemata/pull/523>).

The Bentley also employed Archivemata's automation-tools (<https://github.com/artefactual/automation-tools>) to automate the processing of transfers in its Archivemata pipeline. In addition to base functionality, Bentley staff added a number of customisations to the automation-tools scripts. One customisation specified a processing configuration that would automate decisions to generate a transfer structure report (the archival principle of original order applies to digital materials as well as physical and analogue materials), use Siegfried to identify file formats, extract packages and delete original packages, check for PII, and send SIPs to Archivemata's backlog. Another customisation automatically associated accession numbers with transfers to facilitate the search and retrieval of all of transfers for a given accession. A shell script ran the automation-tools and various cron jobs ran the shell scripts at regular intervals, ensuring that archivists did not need to begin all 249 transfers manually or make individual processing decisions for them.

#### *Preparing archivists to adopt new tools and workflows*

The successful implementation of new tools and workflows involves more than just provisioning servers with the applications and migrating legacy data. More importantly, the adoption of new tools and workflows requires the development of clearly articulated guidelines and training procedures in order to communicate the general functionality of the new tools and their particular uses in the Bentley's workflows to internal users. The Bentley's processing guidelines for physical and digital materials underwent a nearly complete



overhaul to account for the changes brought to authoring archival description (<https://sites.google.com/a/umich.edu/bhl-archival-curation/processing-archival-collections/09-description>) and processing digital content (<https://sites.google.com/a/umich.edu/bhl-archival-curation/processing-archival-collections/08-digital-processing>) brought about by ArchivesSpace and Archivematica.

In addition to updating existing guidelines, members of the ArchivesSpace and Archivematica implementation teams regularly updated Bentley staff members on the progress of each system's implementation, presented on new tools, policies and procedures after the systems were in place, and developed several tiers of training for ArchivesSpace that ranged from high-level overviews of the system and its functions to more advanced training for those staff members who would be using the system in their day-to-day work.

Finally, preparing archivists to adopt new tools and workflows was not an entirely prescriptive process that took place after the systems were implemented; rather, it required involving numerous internal stakeholders throughout the implementation process and taking their feedback into account. Archivists who were not directly involved with implementing the systems, but whose work would be impacted by the systems post implementation, were consulted on relevant matters. User stories were gathered, for example, to ensure that the policies and procedures for accessioning materials in ArchivesSpace would meet the needs of field archivists and records managers for accessioning and processing materials and generating regular reports. The development of clear guidelines, frequent communication with staff about the progress of implementation, and the involvement of a variety of internal stakeholders were invaluable to ensuring that all archivists, from student processors to library administrators, were prepared for

the Bentley's adoption of new tools and workflows.

### **ArchivesSpace–Archivematica–DSpace workflow integration**

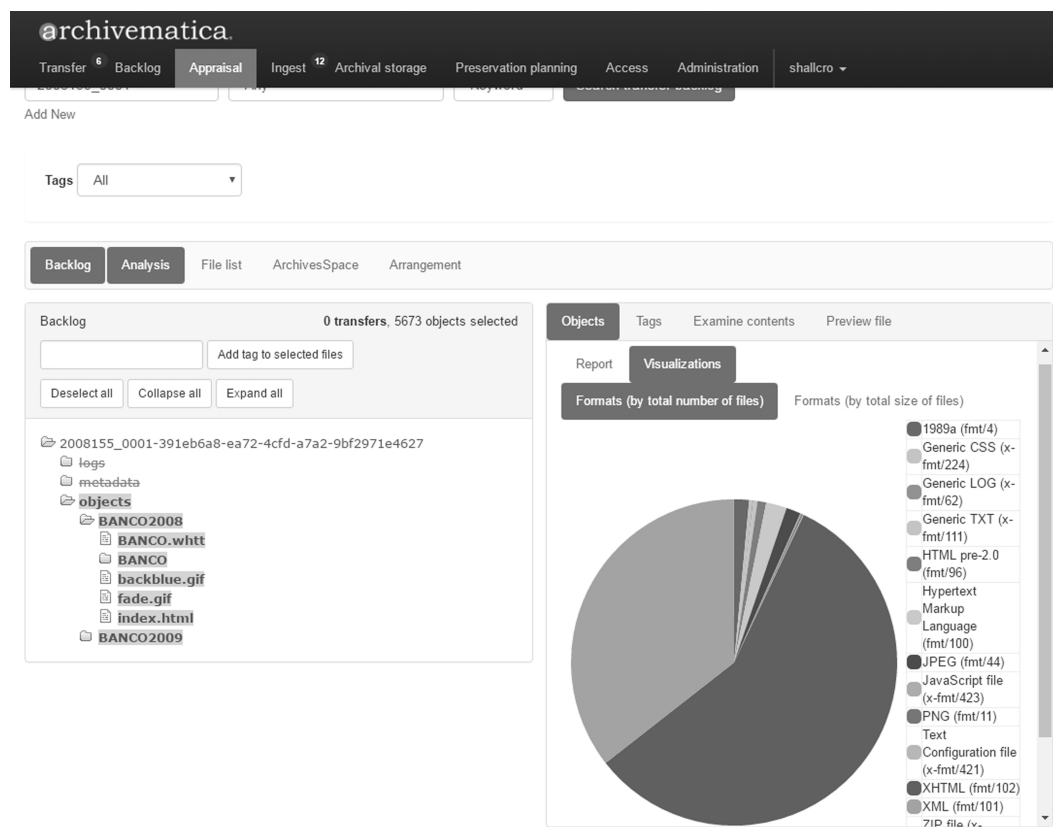
When the ArchivesSpace–Archivematica–DSpace Workflow Integration project was completed in October 2016, the Bentley and its partners at MLibrary had achieved all three of the major development objectives previously outlined. An end-to-end digital archives workflow was in place.

Now, after archivists record an accession in ArchivesSpace and transfer digital content to Archivematica's secure backlog — running SIPs through several micro-services including UUID assignment, checksum verification when checksums are present, package extraction, virus checking, indexing, format identification and validation, and metadata extraction — they navigate to a new Appraisal tab in Archivematica.

The creation of this tab in Archivematica permits archivists to characterise, review, arrange and describe digital archives. Archivists may browse folder hierarchies in a 'Backlog' pane, for example, and perform a number of analyses in the 'Analysis' pane, such as identifying file format distributions in a tabular view or a visualisation (Figure 1) and reviewing PII.

Individual digital objects may be displayed in a 'File List' pane, with contents updated based on facets selected in the Backlog and Analysis panes; they can also be previewed (using available web browser plugins) within the Analysis pane or downloaded so that a user may use the utility of his or her choice to render it. Finally, ephemeral tags may be applied in a number of these panes to aid in the appraisal of content, archival description, identify sensitive data, record deaccession decisions, etc.

The integration of Archivematica and ArchivesSpace permits archivists to create, retrieve, update and delete archival description in ArchivesSpace within



**Figure 1** Identifying file format distributions in a visualisation.

Archivematica. Metadata are written immediately to ArchivesSpace, the system of record for archival description, via its API. Archivists associate digital objects from Archivematica's backlog with archival description in ArchivesSpace by dragging-and-dropping digital objects to the appropriate archival object (Figure 2), thereby associating data with metadata and defining a SIP for Archivematica's ingest process.

It is important to note that other institutions making use of this new functionality in Archivematica may also opt out of ArchivesSpace integration, simply arranging content into SIPs using Archivematica's existing arrangement functionality without employing ArchivesSpace.

Following Archivematica's ingest process, during which digital objects are packaged

into SIPs and run through several micro-services, including normalisation, packaging into an AIP and, optionally, generation of a dissemination information package, the integration of Archivematica and DSpace permits users to select a DSpace collection to which Archivematica will automatically deposit the AIP from available storage service locations. The Archivematica Storage Service splits the AIP into two packages: one for digital content, which will be publicly accessible by default; the other for administrative metadata and log files, which will not be publicly accessible by default. Deposits are made to DSpace via the SWORD v2 protocol and are stored as an item with two bitstreams (one each for the objects and metadata packages) in the DSpace collection (Figure 3). The DSpace API is employed to update the permissions on the metadata package.

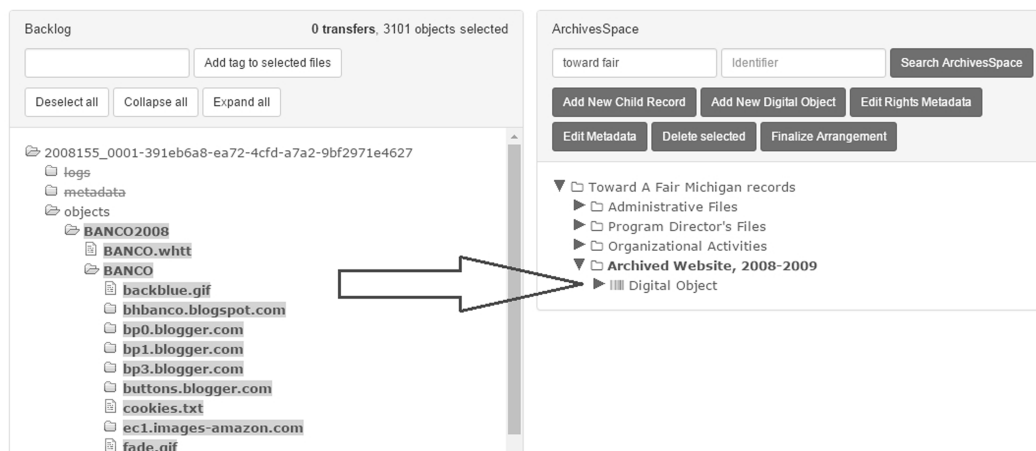


Figure 2 Associated digital objects from Archivematica's backlog with archival description in ArchivesSpace.

### "Meditations" Lest We Forget, 2003

Blanco, Jose  
2016

**Abstract:** Fiscal governance in the EU is an exemplary case of the regulatory state; the EU governs member states' fiscal and public policies through rules rather than expenditure. The weaknesses of the EU fiscal regulatory state were apparent to observers from before the i... [more]

**Handle:** <http://hdl.handle.net/TEMP-BOGUS/304776>

**Series/Report no.:** Publicity - Promotional Audio Commercials and Interviews

[Show full item record](#)

This item appears in the following Collection(s)

- [DSpace API Test](#)

[View/Open](#)  
Name: objects.7z  
Size: 21KB  
Format: Unknown  
Description: Archival mate ...

[View/Open](#)  
Name: metadata.7z  
Size: 507.3KB  
Format: Unknown  
Description: Administrative ...

Figure 3 A DSpace item with two bitstreams — one each for the objects and metadata packages.

It is important to note that, like ArchivesSpace integration, DSpace integration is optional, and the use of the SWORD v2 protocol for initial deposit of the AIP ensures that this integration is repository agnostic, as the SWORD v2 protocol is compatible with a number of repository applications.

Finally, a new digital object record, containing the unique DSpace handle URL for the item, is created in ArchivesSpace and associated with the appropriate archival object, so that links to the items appear within

their appropriate intellectual context in ArchivesSpace as well as the Bentley's online archival collection discovery interface.

This workflow is described in two screencasts created by Bentley Staff: 'Configuring ArchivesSpace and DSpace Integration within Archivematica'<sup>6</sup> and 'Appraisal, Arrangement to ArchivesSpace and Deposit to DSpace'.<sup>7</sup>

To supplement the work done by Artefactual, and to further blend the default ArchivesSpace-Archivematica-DSpace workflow (and default affordances of the

respective software applications) with the Bentley's digital processing workflow, a number of additional tools and scripts were developed by Bentley staff and are available in the Bentley's GitHub organization (<https://github.com/bentley-historical-library>), including:

- *Donor Details Plugin*: an ArchivesSpace plugin that adds additional donor details to agents;
- *Add BHL Acts Add-on*: a Greasemonkey/Tampermonkey script that ensures archivists add machine-actionable, Bentley-specific PREMIS Rights Statements to AIPs;
- *DSpace [REST] API Python Programming [Language] resource (DAPPr)*: a client to communicate with a remote DSpace installation using its backend API;
- *Resource Collection*: a command-line tool to create DSpace collections and Archivemata locations from ArchivesSpace resources;
- *Archivemata AIP to DSpace Item*: a script to interpret PREMIS Rights Statements on Archivemata AIPs and deposit items and bitstreams (with appropriate GroupIDs) to DSpace collections;
- *BHL ArchivesSpace Translations*: Bentley-specific tooltips and user-defined fields for ArchivesSpace.

In addition, the Bentley maintained a project blog, Archival Integration (<http://archival-integration.blogspot.com/>) and continues to use it to explore innovative archival practice and emerging technologies to curate its collections.

## LESSONS LEARNED AND CONCLUSION

As a result of the migration of legacy data to new systems and the integration of those systems in an end-to-end digital archives workflow, staff at the Bentley have achieved increased efficiency in processes related to

the management, arrangement, description and preservation of archival materials. The outcomes of the project, however, have impacted aspects of the Bentley's activities beyond the scope of this particular project, as project staff have learned invaluable lessons in project management, developed new technical skills that can be applied to a variety of projects, and have strengthened the Bentley's capabilities to centre the needs of users, including internal stakeholders and end users, in all that it does.

The agile development methodology employed by Artefactual Systems during the development of the Archivemata Appraisal tab has influenced approaches taken to other projects at the Bentley. The practice of making small, iterative and incremental changes, rather than substantial one-time changes has influenced projects to gradually, but significantly, improve the functionality of the Bentley's website and other access systems. Additionally, user stories, which seek to define specific roles, tasks and desired outcomes for a project, have not only been used for field archivists and records managers accessioning and processing materials, but also in the development of an internal digitisation project tracking database to ensure that the database meets the needs of Bentley staff members who oversee those projects.

The development work done by Bentley staff detailed in this paper, including Python scripting to aid in legacy metadata migration, the development of ArchivesSpace plugins, and work with the ArchivesSpace and DSpace APIs, has increased the Bentley's technical capabilities and the scale at which it can work. Project team members are now more adept at programming and have applied those skills to assisting university faculty with research projects using the Bentley's structured metadata, working with MLibrary to prepare metadata for digitisation projects more efficiently, visualising archived websites and social media data to demonstrate their research potential to undergraduate students.

Finally, while the Bentley seeks to foreground the needs of users in all that it does, the user-centeredness of the Appraisal tab development process has strengthened that commitment and the Bentley's ability to achieve that end. Drawing on lessons learned about agile development, strengthened technical skills, and a focus on end users and their needs has led to numerous improvements to the ways in which the Bentley approaches changes to user experience and how those changes are achieved including conducting interviews with both internal stakeholders and external users regarding future access systems, examining web analytics and server logs to understand better how users interact with the Bentley's current access platforms, and making regular, small improvements to finding aid navigation, search documentation and user-facing terminology.

## REFERENCES

1. Nelson, N. L., Shaw, S., Deromedi, N., Shallcross, M., Ghering, C., Schmidt, L., Belden, M., Esposito, J., Goldman, B. and Pyatt, T. (2012) 'SPEC kit 329: managing born-digital special collections and archival materials', available at: <http://publications.arl.org/Managing-Born-Digital-Special-Collections-and-Archival-Materials-SPEC-Kit-329/> (accessed 10th March, 2017).
2. Bentley Historical Library. (n.d.) 'Digital Processing', available at: <https://sites.google.com/a/umich.edu/bhl-archival-curation/processing-archival-collections/08-digital-processing> (accessed 10th March, 2017).
3. Consultative Committee for Space Data Systems. (2012) 'Reference Model for an Open Archival Information System', available at: <https://public.ccsds.org/Publications/Archive/650x0m2.pdf/default.aspx> (accessed 10th March, 2017).
4. Greene, M. A. and Meissner, D. (2005) 'More product less process: revamping traditional archival processing', *The American Archivist*, Vol. 68, No. 2, pp. 208–263.
5. Greene, M. A. (2010) 'MPLP: it's not just for processing anymore', *The American Archivist*, Vol. 73, No. 1, pp. 175–203.
6. Bentley Historical Library. (2017) 'Configuring ArchivesSpace and DSpace Integration within Archivematica', available at: [https://bentley.mivideo.it.umich.edu/media/t/1\\_jhjo10ru](https://bentley.mivideo.it.umich.edu/media/t/1_jhjo10ru) (accessed 10th March, 2017).
7. Bentley Historical Library. (2017) 'Appraisal, Arrangement to ArchivesSpace and Deposit to DSpace', available at: [https://bentley.mivideo.it.umich.edu/media/t/1\\_a7epibum](https://bentley.mivideo.it.umich.edu/media/t/1_a7epibum) (accessed 10th March, 2017).