

James Madison University

JMU Scholarly Commons

Libraries

Libraries

5-6-2020

To Discovery and Beyond: Using Workflow Automation as an Opportunity for Collaboration and Education

Rebecca B. French

James Madison University, frenchrb@jmu.edu

Follow this and additional works at: <https://commons.lib.jmu.edu/letfspubs>



Part of the [Adult and Continuing Education Commons](#), [Archival Science Commons](#), [Cataloging and Metadata Commons](#), [Training and Development Commons](#), and the [Vocational Education Commons](#)

Recommended Citation

French, Rebecca B., "To Discovery and Beyond: Using Workflow Automation as an Opportunity for Collaboration and Education" (2020). *Libraries*. 187.

<https://commons.lib.jmu.edu/letfspubs/187>

This Poster is brought to you for free and open access by the Libraries at JMU Scholarly Commons. It has been accepted for inclusion in Libraries by an authorized administrator of JMU Scholarly Commons. For more information, please contact dc_admin@jmu.edu.

To Discovery and Beyond

Using Workflow Automation as an Opportunity for Collaboration and Education

Rebecca B. French, Metadata Analyst Librarian, James Madison University

Situation

JMU Libraries Special Collections makes finding aids for manuscript collections accessible through four discovery platforms.

- [Libraries website](#)
- [Virginia Heritage](#), a statewide finding aid database
- OCLC's [WorldCat](#)
- [JMU Library catalog](#) (Sierra ILS)

The old workflow was very manual, requiring a lot of copying and pasting data from one system to another.

- Inefficient and time consuming
- Many opportunities for human error
- Difficult to keep different platforms in sync because all updates had to be made manually in multiple places
- Boring
- Does not use the unique skills and strengths of staff members

Participation

Special Collections department

- Two archivists arrange and describe manuscript collections, put finding aids on Libraries website, and submit finding aids to Virginia Heritage

Metadata Strategies department

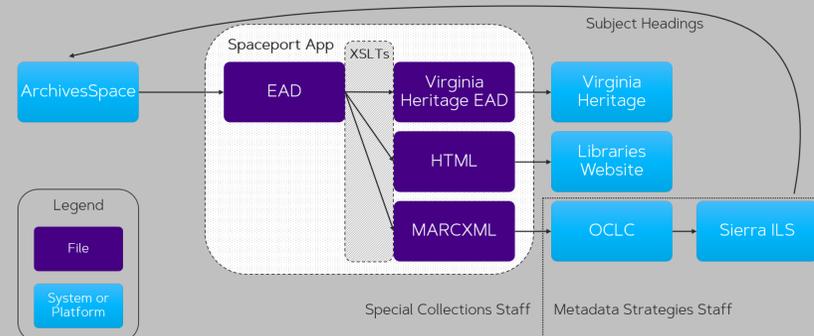
- Cataloger creates bibliographic records for manuscript collections, performs subject analysis, and adds records to OCLC and local catalog
- Metadata librarian develops scripts and tools to automate parts of the workflow

Automation

Project stages

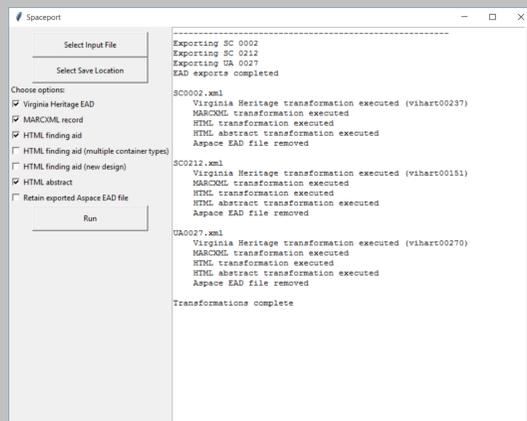
1. XSLT transformation scripts to create web page HTML from EAD, create MARCXML catalog record from EAD, and modify EAD for Virginia Heritage
2. Spaceport Python app
 - Uses ArchivesSpace API to get EAD from ArchivesSpace (replacing manual export)
 - Runs XSLTs using Saxon XSLT processor
 - Makes additional non-XSLT modifications to Virginia Heritage EADs
 - Graphical user interface allows user to select files they want to create and shows progress as the app runs; also shows error messages
3. Additional modifications to Spaceport for usability and stability
4. Python script to sync subjects between catalog and ArchivesSpace (in development)

Code available at https://github.com/frenchrb/JMU_ArchivesSpace



Propagation

- App is packaged into an executable using [Pynstaller](#)
- Hosted on a shared network drive
 - Single location to update
 - Updates are automatically available to all users without any action needed on their part
- Users run the app from a desktop shortcut



*Space icon by Mello from thenounproject.com

Collaboration

Incremental approach

- Initial meeting between Special Collections and metadata librarian to understand the current workflow, its challenges, and goals for improvement
- Met frequently to check in throughout project
- Developer shared updates and stakeholders provided feedback via email and in person
- Implemented automation in stages
- Similar to agile project management methodologies

Experimental mindset

- Approached this project as an experiment, with the option to go back to the old workflow if automated solutions turned out to be unsatisfactory
- Process of trying different solutions was valuable professional development, increasing developer's skills and knowledge of strengths and weaknesses of different technologies
- Goals evolved organically as we found successful solutions and gained new skills

Observations

Benefits of improved workflow

- More efficient and sustainable workflow
- More consistent discovery experience for patrons
- Facilitated completion of additional projects, both by freeing up time for them and providing the tools to make the work possible
 - Updates to legacy description
 - Migration of Libraries' website content management system
 - Update look of finding aid web pages
 - Aeon implementation
- Automation of routine tasks freed up time for staff to do work they are uniquely skilled at (e.g., subject analysis, description)

Benefits of incremental and collaborative approach

- Provided space for professional development and for gradually building skills needed to complete project work
- Supported evolving project goals
- Facilitated adapting to the new workflow
 - Easier to test changes in small chunks
 - Easier for staff to adjust to smaller changes
- Generated excitement and buy-in by making the most impactful change first
- Strengthened relationships between departments
- Enabled project work to fit in with other priorities rather than being put off until it could all be completed at once

Thankfully, hand encoding finding aids is a distant memory at this point!
I can't even imagine my life without Spaceport.

-Tiffany Cole, Special Collections Archivist

Education

The metadata librarian used this project as an opportunity to develop coding and software development skills, including

- Working with ArchivesSpace and Sierra APIs
- Building a simple GUI with Tkinter
- Packaging Python scripts into an executable file
- Further developing existing XSLT knowledge

Our incremental approach to the project provided space for professional development that responded to evolving project needs.

- Able to build on existing knowledge piece by piece
- Did not have to learn everything all at once
- Use of skills in the context of a real-world project led to more effective learning and retention

Recommendations

Courses, workshops, websites, and other resources and documentation for learning more about the tools and technologies used in this project

XML and XSLTs

- Chavez, R. (n.d.). *Introduction to XML* [Online course]. Library Juice Academy. <https://libraryjuiceacademy.com/shop/course/019-introduction-to-xml/>
- Chavez, R. (n.d.). *Transforming and querying XML with XSLT and XQuery* [Online course]. Library Juice Academy. <https://libraryjuiceacademy.com/shop/course/052-transforming-querying-xml-xslt-xquery/>
- Chavez, R. (n.d.). *XSLT fundamentals* [Online course]. Library Juice Academy. <https://libraryjuiceacademy.com/shop/course/189-xslt-fundamentals/>
- *Saxon documentation*. (n.d.). Saxonica. <http://www.saxonica.com/documentation/>
 - In particular, see the section [Running XSLT from the Command Line](#)

Programming

- Library Carpentry. (2020). *The UNIX shell* [Lesson]. Library Carpentry. <https://librarycarpentry.org/lc-shell/>
- Library Carpentry. (2019). *Introduction to Git* [Lesson]. <https://librarycarpentry.org/lc-git/>
- Chavez, R. (n.d.). *Introduction to JSON and structured data* [Online course]. Library Juice Academy. <https://libraryjuiceacademy.com/shop/course/161-introduction-json-structured-data/>

Python

- Library Carpentry. (2020). *Python intro for libraries* [Lesson]. <https://librarycarpentry.org/lc-python-intro/>
- Sweigart, A. (2019). *Automate the boring stuff with Python: Practical programming for total beginners* (2nd edition). <https://automatetheboringstuff.com/>
- Severance, G. R. (2016). *Python for everybody: Exploring data with Python 3*. <https://www.py4e.com/book>
 - The related website <https://www.py4e.com/> also links to online lessons and courses based on the book.
- Lungh, F. (2005). *An introduction to Tkinter*. <http://effbot.org/tkinterbook/>
- Klein, B. (2020). *Tkinter tutorial*. https://www.python-course.eu/python_tkinter.php
- *Pynstaller manual*. (n.d.). <https://pynstaller.readthedocs.io/en/stable/>
- *Pynstaller wiki*. (n.d.). <https://github.com/pynstaller/pynstaller/wiki>

APIs

- ArchivesSpace. (n.d.). *API reference*. <http://archivesspace.github.io/archivesspace/api/>
- Innovative Interfaces, Inc. (2020). *Sierra API documentation*. <https://techdocs.iii.com/sierraapi/Content/titlePage.htm>
- Addonizio, V., & Custer, M. (2018, August 10). *ArchivesSpace API workshop* [Slides and Python scripts]. <https://github.com/archivesspace/api-training>
- Wiedeman, G. (2019, August 2). *Introduction to Python and ArchivesSnake for Archivists Workshop* [Slides]. <https://gregwiedeman.com/presentations/introPythonASnake.html>
- Tillman, R. K. (2019, March 18). *Series of tubes: Moving subjects from MARC to ASpace records*. Presentation at the ArchivesSpace Online Forum. <https://youtu.be/t9ll73TOXos>