Small Steps Make Giant Leaps: An Incremental Approach to Technical Services Workflow Automation

Rebecca B. French
*James Madison University, frenchrb@jmu.edu*

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

Part of the Archival Science Commons, and the Cataloging and Metadata Commons

**Recommended Citation**

https://commons.lib.jmu.edu/letfspubs/184

This Presentation 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.
SMALL STEPS
MAKE GIANT LEAPS
AN INCREMENTAL APPROACH TO TECHNICAL SERVICES WORKFLOW AUTOMATION

Rebecca B. French, James Madison University
A SERIES OF STEPS

1. XSLTs
2. Spaceport Python app
3. Refining Spaceport
4. Python scripts for syncing subjects
5. Record uploads via API
SOME DEFINITIONS

- **Collection**
  - “A group of materials with some unifying characteristic”
  - E.g., items from a person or organization.

- **Finding aid**
  - “Guide or inventory of the contents of an archival collection”
  - Contains historical or biographical information, description of collection arrangement, and container list (inventory)
  - Encoded Archival Description (EAD) is the standard machine-readable format
Series 5: Financial Files, 1941-1971

Series 5: Financial Files, 1941-1971, houses all loose financial documents relating to the Wampler businesses. The majority of the series consists of receipts from shipping, Wampler Feed & Seed, Sunny Slope Farm, travel expenses, and general purchases while the second part of the series contains miscellaneous bank statements, account information, tax documents for state and federal income, business taxes, and some Wampler family business documents.

<table>
<thead>
<tr>
<th>Title/Description</th>
<th>Instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bank Statements and Receipts, 1928-1969</td>
<td>box 23</td>
</tr>
<tr>
<td>Tax Documents, 1942-1961</td>
<td>folder 5</td>
</tr>
<tr>
<td>General Sales Receipts, 1940-1953</td>
<td>box 23</td>
</tr>
<tr>
<td>General Sales Receipts, January-July 1954</td>
<td>box 24</td>
</tr>
<tr>
<td>General Sales Receipts, August-December 1954</td>
<td>box 24</td>
</tr>
<tr>
<td>General Sales Receipts, 1955-1971</td>
<td>box 24</td>
</tr>
<tr>
<td>Commonwealth of Virginia: Agency Purchase Order</td>
<td>box 24</td>
</tr>
<tr>
<td>Travel Receipts, 1942-1964</td>
<td>folder 4</td>
</tr>
</tbody>
</table>
DISCOVERY PLATFORMS

Libraries Website

Catalog (Sierra ILS)

OCLC

Virginia Heritage
Wampler Business Records, 1916-1972

Summary Information

<table>
<thead>
<tr>
<th>Repository</th>
<th>Special Collections, Carrier Library, James Madison University</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creator</td>
<td>Wampler, Charles Weldon, Sr., 1885-1976</td>
</tr>
<tr>
<td>Title</td>
<td>Wampler Business Records</td>
</tr>
<tr>
<td>Collection No.</td>
<td>SC 0202</td>
</tr>
<tr>
<td>Date</td>
<td>1916-1972</td>
</tr>
<tr>
<td>Extent</td>
<td>31 boxes, 11 Cubic Feet</td>
</tr>
<tr>
<td>Language of the Material</td>
<td>English</td>
</tr>
<tr>
<td>Abstract</td>
<td>This collection contains correspondence, reports, photographs, business documents, financial statements, publications, and other documents that relate to the businesses of Charles Wampler Sr. (Wampler Feed and Seed Company) primarily and to a lesser degree his son, Charles Wampler Jr. (Wampler Food Incorporated). Correspondence comprises the bulk of the collection.</td>
</tr>
</tbody>
</table>

Preferred Citation


Bio/Historical Note

The Wampler family emerged upon the American poultry industry in the early twentieth century changing the industry practices and beliefs. As a result of his efforts, John Wampler Sr. prospered and helped the Shenandoah Valley become a prizewinning poultry region.

In 1881, John Wampler and his family moved to Hume, a small town in the Shenandoah Valley. He was a part owner of the Valley Poultry Company and flourished on the business. However, the company went bankrupt in 1896 and John Wampler Sr. made several attempts to restart the company. By 1900, John Wampler Sr. was able to start a new company called the Shenandoah River Poultry Company which became successful.

John Wampler Sr. was also a farmer and owned a farm in the Shenandoah Valley. He had a small farm location for growing food for his family and producing surplus for sale. He also raised chickens and sold their eggs and chickens to local businesses.

In 1916, John Wampler Sr. and his son, Charles Wampler Jr., started the Wampler Feed and Seed Company. They produced manure and sold feed and seed to local farmers. The company was successful and continued to grow under the leadership of John Wampler Sr.

In 1920, John Wampler Sr. and his son, Charles Wampler Jr., started the Wampler Food Incorporated. They produced manure and sold food products to local businesses. The company was successful and continued to grow under the leadership of John Wampler Sr.

John Wampler Sr. passed away in 1930, but his son, Charles Wampler Jr., continued to run the company. Charles Wampler Jr. passed away in 1976, but the company continued to grow under the guidance of his son, Charles Wampler Jr., and his daughter, Jane Wampler.

Charles Wampler Jr. and Jane Wampler continued to run the company until 1990 when they sold it to another company. The company is now called Wampler Poultry and is located in Harrisonburg, Virginia.

^ Top
A GUIDE TO THE WAMPLER BUSINESS RECORDS, 1916-1972: SC 0202

A COLLECTION IN SPECIAL COLLECTIONS
COLLECTION NUMBER SC 0202

JAMES MADISON UNIVERSITY

Special Collections, Carrier Library, James Madison University
Special Collections
Carrier Library
James Madison University
Harrisonburg, Virginia 22807
USA
Phone: (540) 568-3612
Fax: (540) 568-3450
Email: library-special@jmu.edu
URL: http://www.lib.jmu.edu/special/

© 2010 By JMU Libraries. All rights reserved.

Processed by Trevor Award, Kathryn Barlet, Gabriel Caram, and James Dennis, May 2011. Tiffany Cole September 2018

Repository
Collection Number
Extent
Creator
Language
Abstract

Special Collections, James Madison University
SC 0202
Wampler Business Records, 1916-1972
11 cubic feet, 31 boxes
Wampler, Charles Wampler, Sr., 1865-1976
English
This collection contains correspondence, reports, photographs, business documents, financial statements, publications, and other documents that relate to the businesses of Charles Wampler Sr. (Wampler Feed and Seed Company) primarily and to a lesser degree his son, Charles Wampler Jr.

Virginia Heritage
WHO IS INVOLVED?

Special Collections
- Two archivists
  - Arrange and describe manuscript collections
  - Put finding aids on Libraries website
  - Submit finding aids to state-wide Virginia Heritage database

Metadata Strategies
- Cataloger
  - Creates bibliographic records for manuscript collections, based off of finding aids
  - Performs subject analysis
  - Adds records to OCLC and our local catalog
- Metadata librarian
  - Workflow automation
PROBLEMS WITH MANUAL WORKFLOW

Time consuming
PROBLEMS WITH MANUAL WORKFLOW

Time consuming

Opportunities for human error
PROBLEMS WITH MANUAL WORKFLOW

- Time consuming
- Opportunities for human error
- Hard to keep different platforms in sync
PROBLEMS WITH MANUAL WORKFLOW

Time consuming
Opportunities for human error
Hard to keep different platforms in sync
Boring
**STEP 1: XSLTS**

- Deal with the biggest pain point (hand-encoding for Virginia Heritage, copy and paste to create web page)
- XSLT: language for transforming XML documents into other XML documents
- Wrote XSLT transformation scripts
  - Make (most of) the changes needed to create a Virginia Heritage EAD
  - Create web page HTML from EAD
  - Create MARCXML catalog record from EAD
- Special Collections staff used Oxygen XML Editor to run the XSLT transformation scripts
STEP 1: XSLTS

Legend:
- System or Platform
- File

Special Collections Staff
Metadata Strategies Staff

ArchivesSpace → EAD → XSLT's → Virginia Heritage EAD → HTML → MARCXML → OCLC → Sierra ILS

Virginia Heritage
Libraries Website

Subject Headings
“…this is a DRASTIC improvement over my previous workflow.”

Tiffany Cole, Special Collections Archivist
STEP 2: SPACEPORT PYTHON APP

- Spaceport, a custom Python app
  - Uses ArchivesSpace API to get EAD from ASpace (replacing manual export)
  - Runs XSLTs using Saxon XSLT processor (replacing Oxygen)
  - Does additional Virginia Heritage non-XSLT work
  - Graphical user interface includes selection of options
- Packaged into an executable with PyInstaller
- Hosted on shared network drive for ease of updating
- Code available at https://github.com/frenchrb/JMU_ArchivesSpace
STEP 2: SPACEPORT PYTHON APP

Legend:
- File
- System or Platform

Special Collections Staff
- EAD
- Virginia Heritage EAD
- HTML
- MARCXML

Metadata Strategies Staff
- ArchivesSpace
- Virginia Heritage
- Libraries Website
- OCLC
- Sierra ILS

Subject Headings

Virginia Heritage EAD

MARCXML

HTML
STEP 3: REFINING SPACEPORT

- Additional modifications to Spaceport
  - Moved status reporting to the GUI to show progress as the app runs (was previously in a separate window)
  - Added error reporting to the GUI
- Waited to make these changes until after basic functionality was thoroughly tested
- No change to workflow
BENEFITS OF IMPROVED WORKFLOW

- Able to use ArchivesSpace as system of record
- Improved metadata accuracy
- Time savings
  - Only takes seconds to format a finding aid for upload to discovery platforms
  - 99% decrease in time to produce Virginia Heritage EAD, compared to hand-encoding
  - Generate web pages for all 340+ collections in only 20 minutes
- More staff time to dedicate to standardizing description and updating legacy finding aids
- Facilitated completion of other projects
PROJECTS FACILITATED BY IMPROVED WORKFLOW

CMS Migration to WordPress
PROJECTS FACILITATED BY IMPROVED WORKFLOW

CMS Migration to WordPress

Update Finding Aid Web Pages
PROJECTS FACILITATED BY IMPROVED WORKFLOW

- CMS Migration to WordPress
- Update Finding Aid Web Pages
- Aeon Implementation
“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
STEP 4: PYTHON SCRIPT FOR SYNCING SUBJECTS

- Currently in development
- Uses APIs for ILS and ArchivesSpace
- Subject sync script
  - Finds manuscript collection bib records that have been updated in the catalog
  - Parses LC subject headings
  - Adds any new subjects to ArchivesSpace and links them to the corresponding resource record
- Would schedule to run regularly, not requiring human intervention other than monitoring the results
- More accurate subject records in ArchivesSpace; could enhance with URIs

STEP 4: PYTHON SCRIPT FOR SYNCING SUBJECTS

Legend
- System or Platform
- App or Script

Special Collections Staff

Metadata Strategies Staff

ArchivesSpace

Spaceport App

Virginia Heritage

Libraries Website

OCLC

Sierra ILS

Subject Sync Script
STEP 5: RECORD UPLOADS VIA API

- Future step we could consider
- Uploading records to various discovery platforms doesn’t involve a significant amount of work (time or difficulty) in original or current workflow, so is lowest priority for automation
- Potential to use Python scripting and APIs to automate uploads to OCLC, catalog, and/or website
BENEFITS OF INCREMENTAL AUTOMATION

- Gradually build skills and develop new ideas that couldn't have been foreseen from the beginning
- Generate excitement and buy-in by making most impactful change first
- Easier to adjust to smaller changes
- Easier to test changes in small chunks
- Fit project work in with other priorities rather than putting it off until it could all be completed at once
“… Because it was so difficult and unwieldy to make changes across platforms, much of our description has languished for years … Spaceport has changed all of that. We are now much more able to uphold our ethical obligations to the historical record.”

Kate Morris, Head of Special Collections
IMAGE CREDITS

- Slide background photo by Johny Goerend on Unsplash
- "Time" icon by Richard de Vos from the Noun Project
- "Error" icon by Vaibhav Radhakrishnan from the Noun Project
- "Sync error" icon by Xinh Studio from the Noun Project
- "boring" icon by Phạm Thạnh Lộc from the Noun Project
- WordPress logo from https://wordpress.org/about/logos/
- "Html Fle" icon by BomSymbols from the Noun Project
- Aeon logo from Atlas Systems
THANK YOU!

Rebecca B. French – frenchrb@jmu.edu
Code for Spaceport: https://github.com/frenchrb/JMU_ArchivesSpace