James Madison University

JMU Scholarly Commons

Libraries Libraries

2-13-2020

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

French, Rebecca B., "Small Steps Make Giant Leaps: An Incremental Approach to Technical Services Workflow Automation" (2020). Libraries. 184. 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



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

```
<unittitle>Financial Files</unittitle>
  <unitid>Series 5</unitid>
 <unitdate normal="1941/1971" type="inclusive">1941-1971/unitdate>
 <lanqmaterial><lanquage langcode="eng">English</lanquage>.</lanqmaterial>
</did>
<scopecontent>
 <head>Scope and Content</head>
 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 & amp; 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.
</scopecontent>
<c02 level="file">
  <did>
   <unittitle>Bank Statements and Receipts</unittitle>
   <unitdate type="inclusive">1928-1969</unitdate>
   <container label="Mixed Materials" type="box">23</container>
   <container type="folder">5</container>
 </did>
</c02>
<c02 level="file">
 < did>
   <unittitle>Tax Documents</unittitle>
   <unitdate type="inclusive">1942-1961</unitdate>
   <container label="Mixed Materials" type="box">23</container>
   <container type="folder">6</container>
 </did>
</c02>
<c02 level="file">
 <did>
    <unittitle>Business Financial Documents</unittitle>
   <unitdate type="inclusive">1946-1970</unitdate>
   <container label="Mixed Materials" type="box">23</container>
   <container type="folder">7</container>
 </did>
</c02>
```

<c01 level="series">

<did>

FINDING AID

showing series description and container list

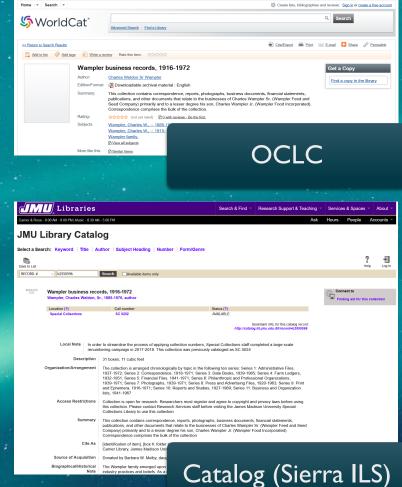
Series 5: Financial Files, 1941-1971

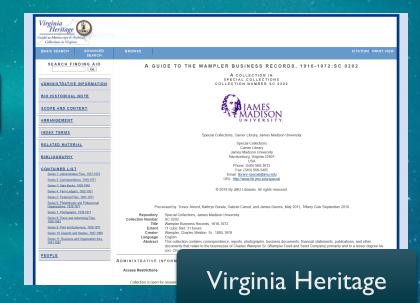
Series 5: Financial Files, 1941-1971, houses all loose financial documents relating to the Wampler businesses. The majority of the series consists of receipts from shipping, Wampler Feed & Seed, Sunny Slope Farm, travel expenses, and general purchases while the second part of the series contains miscellaneous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

Title/Description	Instances		
Bank Statements and Receipts, 1928-1969	box 23	folder 5	
Tax Documents, 1942-1961	box 23	folder 6	
Business Financial Documents, 1946-1970	box 23	folder 7	
General Sales Receipts, 1940-1953	box 23	folder 8	
General Sales Receipts, January-July 1954	box 24	folder 1	
General Sales Receipts, August-December 1954	box 24	folder 2	
General Sales Receipts, 1955-1971	box 24	folder 3	
Commonwealth of Virginia: Agency Purchase Order.			
Travel Receipts, 1942-1964	box 24	folder 4	

DISCOVERY PLATFORMS







People Accounts Hours

Wampler Business Records, 1916-1972

Browse manuscript collections

Request for use in **Special Collections**

Summary Information Bio/Historical Note

Scope and Content

Arrangement

Administrative Information

Related Materials

Index Terms

Bibliography

Collection Inventory

Series 1: Administrative

Series 2: Correspondence

Series 3: Date Books

Series 4: Farm Ledgers

Series 5: Financial Files

Summary Information

Repository: Special Collections, Carrier Library, James Madison University

Wampler, Charles Weldon, Sr., 1885-1976 Creator:

Title: Wampler Business Records

Collection No.: SC 0202

> Date: 1916-1972

Extent: 31 boxes, 11 Cubic Feet

Language of the Material: English

> Abstract: 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.

Preferred Citation

[identification of item], [box #, folder #], Wampler Business Records, 1916-1972, SC 0202, Special Collections, Carrier Library, James Madison University, Harrisonburg, Va.

^ Top

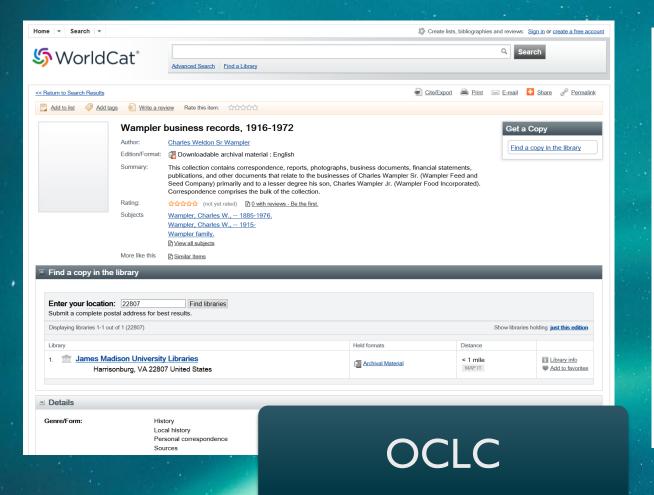
Bio/Historical Note

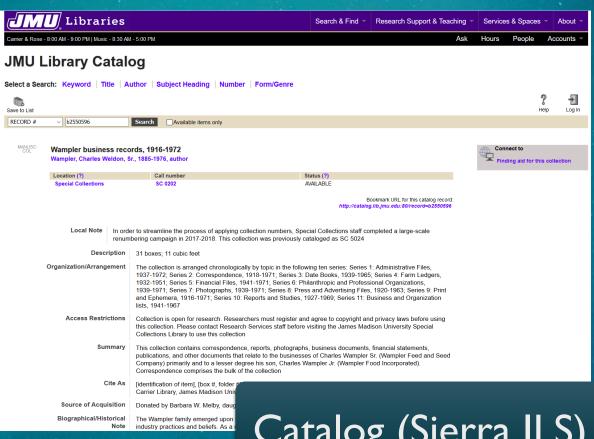
The Wampler family emerged upon the American poultry industry in the early twentieth century changing the industry practices

and beliefs. As a result of and helped the Shenando

In 1811, John Wampler at county during a period w

Libraries Website





Catalog (Sierra ILS)



BASIC SEARCH

ADVANCED SEARCH BROWSI

CITATION | PRINT VIEW

SEARCH FINDING AID

Go

ADMINISTRATIVE INFORMATION

BIO HISTORICAL NOTE

SCOPE AND CONTENT

ARRANGEMENT

INDEX TERMS

RELATED MATERIAL

BIBLIOGRAPHY

CONTAINER LIST

Series 1: Administrative Files, 1937-1972

Series 2: Correspondence, 1918-1971

Series 3: Date Books, 1939-1965

Series 4: Farm Ledgers, 1932-1951

Series 5: Financial Files, 1941-1971

Series 6: Philanthropic and Professional Organizations, 1939-1971

Series 7: Photographs, 1939-1971

Series 8: Press and Advertising Files,

920-1963

Series 9: Print and Ephemera, 1920-1971

Series 10: Reports and Studies, 1927-1969

Series 11: Business and Organization lists,

1941-1967

PEOPLE

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

A COLLECTION IN SPECIAL COLLECTIONS COLLECTION NUMBER SC 0202



Special Collections, Carrier Library, James Madison University

Special Collections Carrier Library James Madison University Harrisonburg, Virginia 22807 USA Phone: (540) 568-3612 Fax: (540) 568-3405 Email: library-special@jmu.edu

URL: http://www.lib.jmu.edu/special

© 2018 By JMU Libraries. All rights reserved.

Processed by: Trevor Alvord, Kathryn Barela, Gabriel Camut, and James Dennis, May 2011; Tiffany Cole September 2018

Repository Special Collections, James Madison University

SC 0202

ber SC 020

Title Wampler Business Records, 1916-1972

Extent 11 cubic feet; 31 boxes

Wampler, Charles Weldon, Sr., 1885-1976

Language English

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

con Char

ADMINISTRATIVE INFORM

Access Restrictions

Collection is open for research

Virginia Heritage

WHO IS INVOLVED?

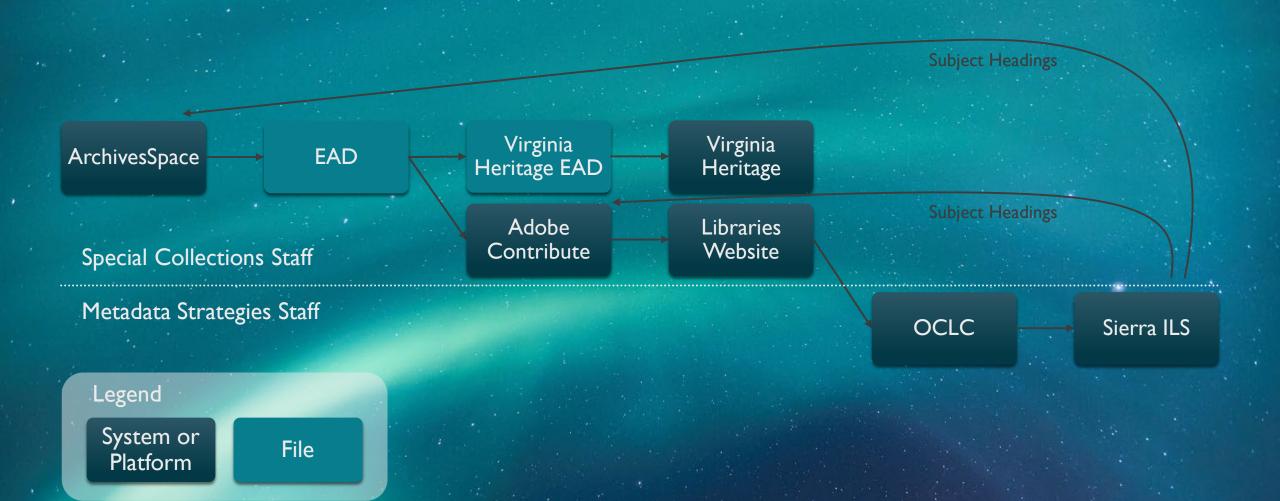
Special Collections

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

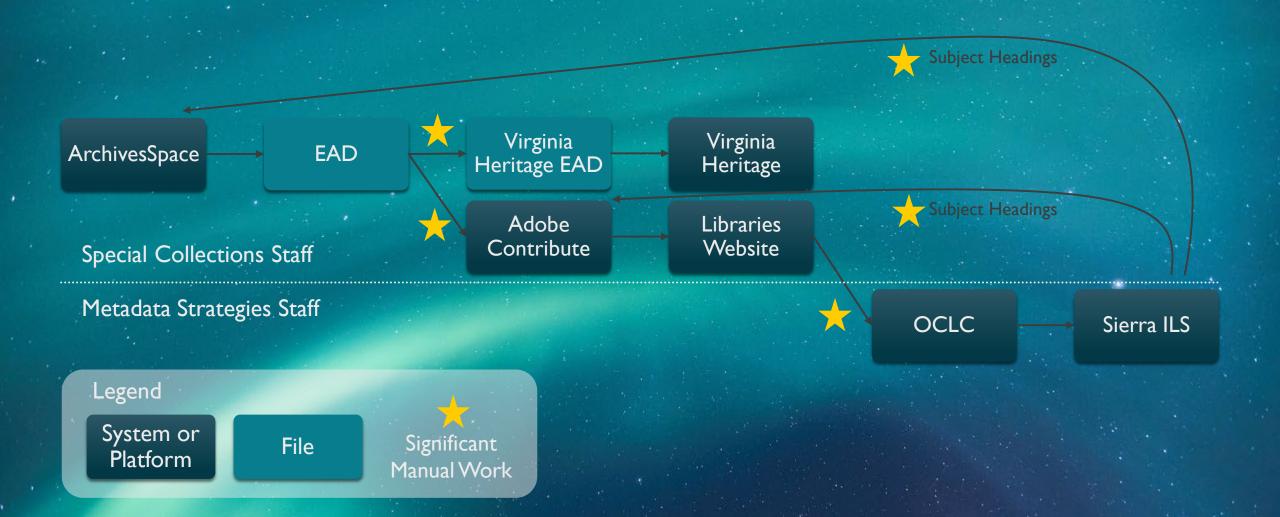
Metadata Strategies

- 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

ORIGINAL WORKFLOW



ORIGINAL WORKFLOW





Time consuming



Time consuming



Opportunities for human error



Time consuming



Opportunities for human error



Hard to keep different platforms in sync



Time consuming



Opportunities for human error



Hard to keep different platforms in sync

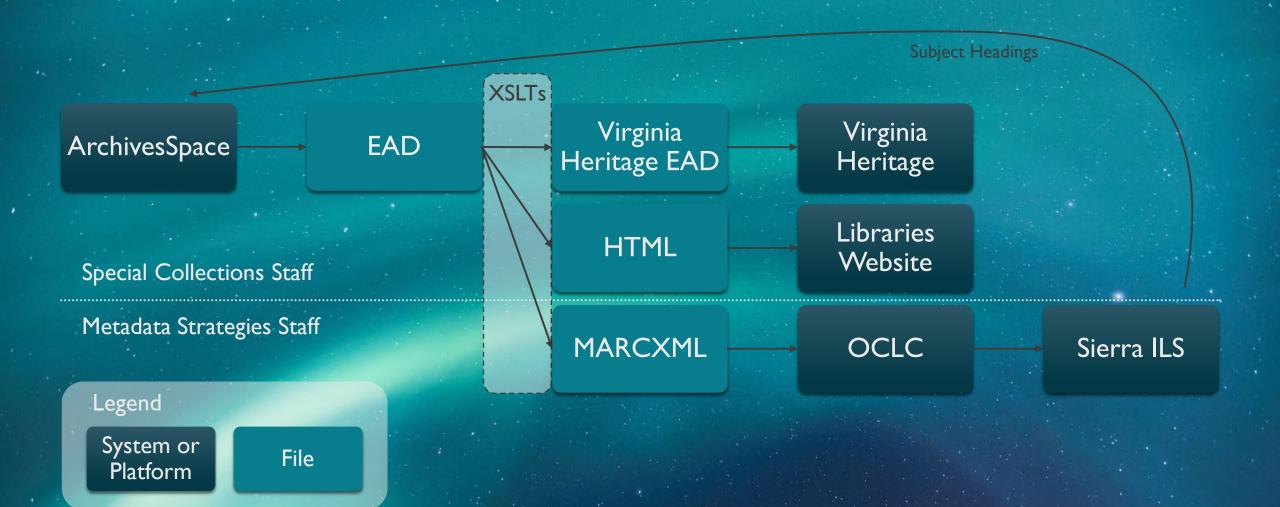


Boring

STEP I: 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 I: XSLTS

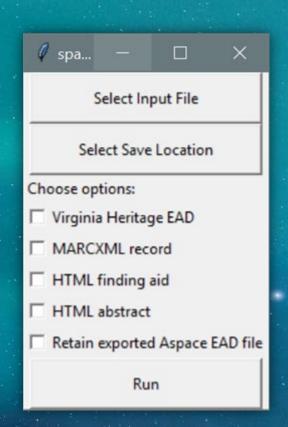


"...this is a DRASTIC improvement over my previous workflow."

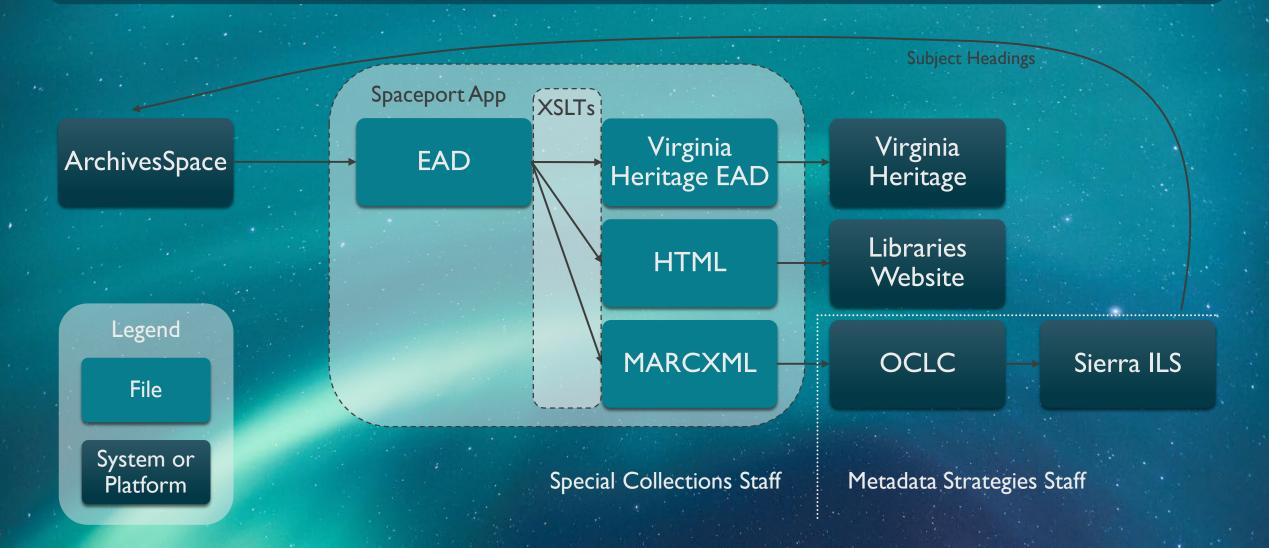
Tiffany Cole, Special Collections Archivist

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 Pylnstaller
- Hosted on shared network drive for ease of updating
- Code available at https://github.com/frenchrb/JMU_ArchivesSpace

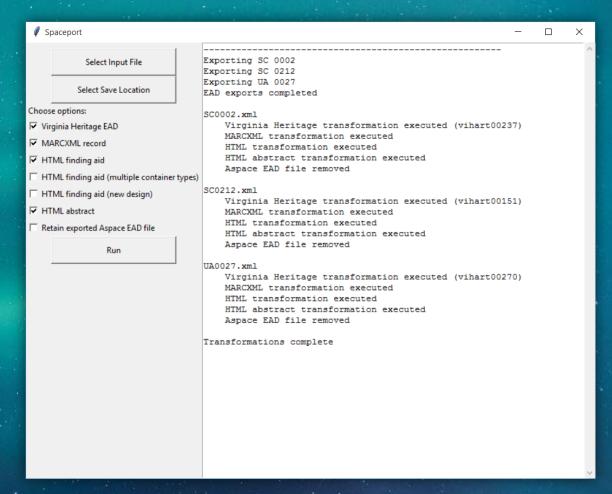


STEP 2: SPACEPORT PYTHON APP



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

ArchivesSpace

Spaceport App Virginia Heritage

Libraries Website

OCLC

Subject
Sync Script

Sierra ILS

Special Collections Staff

Metadata Strategies Staff

Legend

System or Platform

App or 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

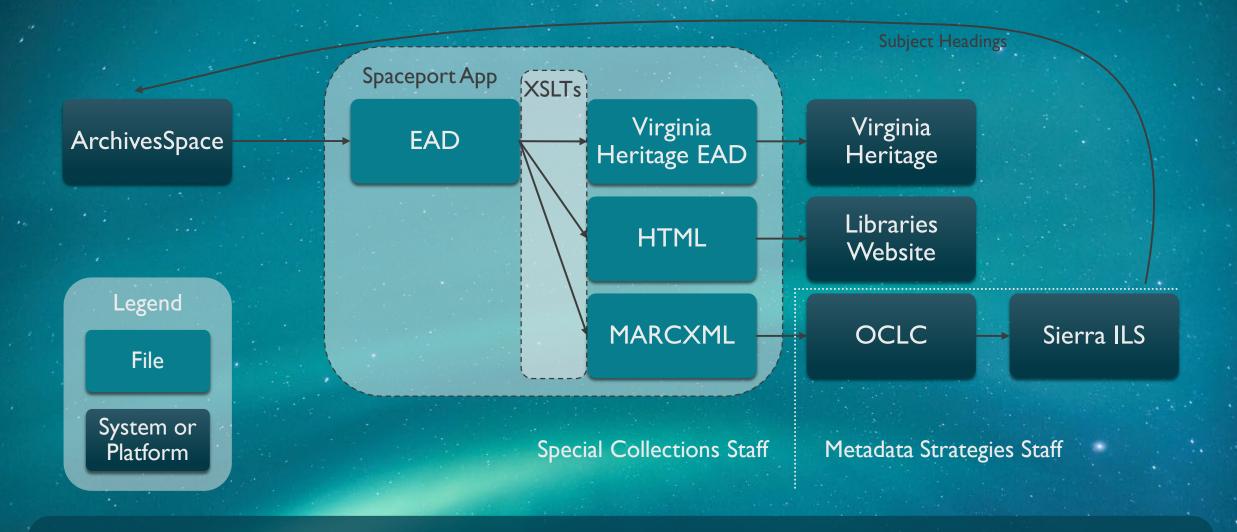
- Gradually build skills and develop new ideas that couldn't have been foreseen from the beginning
- Generate excitement and buy-in by making most impactful change first
- Easier to adjust to smaller changes
- Easier to test changes in small chunks
- Fit project work in with other priorities rather than 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
- <u>"Time" icon</u> by Richard de Vos from <u>the Noun Project</u>
- "Error" icon by Vaibhav Radhakrishnan from the Noun Project
- "Sync error" icon by Xinh Studio from the Noun Project
- "boring" icon by Pham Thanh 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 – <u>frenchrb@jmu.edu</u>

Code for Spaceport: https://github.com/frenchrb/JMU_ArchivesSpace