## <EAD Export Service/>

Mark Custer ArchivesSpace Online Forum 2019-03-18 mark.custer@yale.edu

## Synchronization

ArchivesSpace itHub Itervesters (e.g. ArchiveGrid)

## Who harvests our finding aids?

American Institute of Physics

Connecticut Archives Online

National Library of Medicine

OCLC's ArchiveGrid

Yale

## Finding aid export options

Use the staff interface

Use the ArchivesSpace API

OAI-PMH (Open Archives Initiative – Protocol for Metadata Harvesting)

**ArchivesSpace Export Service** 

# How we used to provide access

Export and upload files manually

Files were moved around manually

Single directory of XML files

Each repository kept local backups

(what's the canonical version when you have a minimum of 3 versions floating around?)

### Index of /archivegrid

| [ICO]         | Name        | Last modified     | Size | <b>Description</b> |
|---------------|-------------|-------------------|------|--------------------|
| [DIR] Paren   | t Directory |                   | -    |                    |
| [TXT] arts.ac | ob.0001.xml | 11-Oct-2013 13:19 | 34K  |                    |
| [TXT] arts.ac | ob.0003.xml | 02-Feb-2012 13:57 | 63K  |                    |
| [TXT] arts.ac | ob.0005.xml | 24-Oct-2014 14:03 | 116K |                    |
| [TXT] arts.ac | ob.0008.xml | 31-Oct-2014 09:09 | 227K |                    |
| [TXT] arts.ac | ob.0009.xml | 02-Feb-2012 13:57 | 310K |                    |
| [TXT] arts.ac | ob.0013.xml | 02-Feb-2012 13:57 | 40K  |                    |
| [TXT] arts.ac | ob.0014.xml | 02-Feb-2012 13:57 | 12K  |                    |
| [TXT] arts.ac | ob.0018.xml | 13-Mar-2012 15:53 | 39K  |                    |
| [TXT] arts.ac | ob.0019.xml | 02-Feb-2012 13:57 | 16K  |                    |
| [TXT] arts.ac | ob.0028.xml | 02-Feb-2012 13:58 | 9.9K |                    |
| [TXT] arts.ac | ob.0031.xml | 01-Nov-2013 15:08 | 61K  |                    |
| [TXT] arts.ac | ob.0033.xml | 13-Nov-2014 13:22 | 304K |                    |
| [TXT] arts.ac | ob.0034.xml | 07-Nov-2013 14:21 | 173K |                    |
| [TXT] arts.ac | ob.0036.xml | 07-Oct-2014 16:37 | 255K |                    |
| [TXT] arts.ac | ob.0038.xml | 02-Feb-2012 13:58 | 15K  |                    |
| [TXT] arts.ac | ob.0039.xml | 07-Nov-2013 14:40 | 35K  |                    |
| [TXT] arts.ac | ob.0041.xml | 03-Dec-2014 15:36 | 24K  |                    |
| [TXT] arts.ac | ob.0042.xml | 02-Feb-2012 13:59 | 15K  |                    |
| [TXT] arts.ac | ob.0043.xml | 07-Nov-2013 14:53 | 31K  |                    |
| [TXT] arts.ac | ob.0044.xml | 07-Mar-2013 09:25 | 41K  |                    |
| [TXT] arts.ac | ob.0045.xml | 05-Mar-2013 14:42 | 19K  |                    |
| [TXT] arts.ac | ob.0047.xml | 23-Jun-2014 10:42 | 8.7K |                    |
| [TXT] arts.ac | ob.0050.xml | 02-Feb-2012 13:59 | 24K  |                    |
| [TXT] arts.ac | ob.0055.xml | 12-Nov-2013 11:52 | 348K |                    |
| [TXT] arts.ac | ob.0059.xml | 14-Jul-2014 14:02 | 161K |                    |
| [TXT] arts.ac | ob.0061.xml | 12-Jul-2013 09:15 | 76K  |                    |
| [TXT] arts.ac | ob.0062.xml | 02-Feb-2012 13:59 | 14K  |                    |
| [TXT] arts.ac | ob.0067.xml | 17-Apr-2014 16:05 | 13K  |                    |
| [TXT] arts.ac | ob.0069.xml | 02-Feb-2012 13:59 | 10K  |                    |
| [TXT] arts.ac | ob.0070.xml | 02-Feb-2012 13:59 | 9.0K |                    |
| [TXT] arts.ac | ob.0073.xml | 02-Feb-2012 14:00 | 17K  |                    |
| [mrrm]        | 1 0000      |                   |      |                    |

## How we provide access now

We're using the ArchivesSpace Export Service

Which provides a configurable publication pipeline

Intended to be a "set it and forget it" service

But wait, there's more...

### How does the service work?

- ArchivesSpace Plugin Endpoint (backend)
- Export Service (exporter\_app)

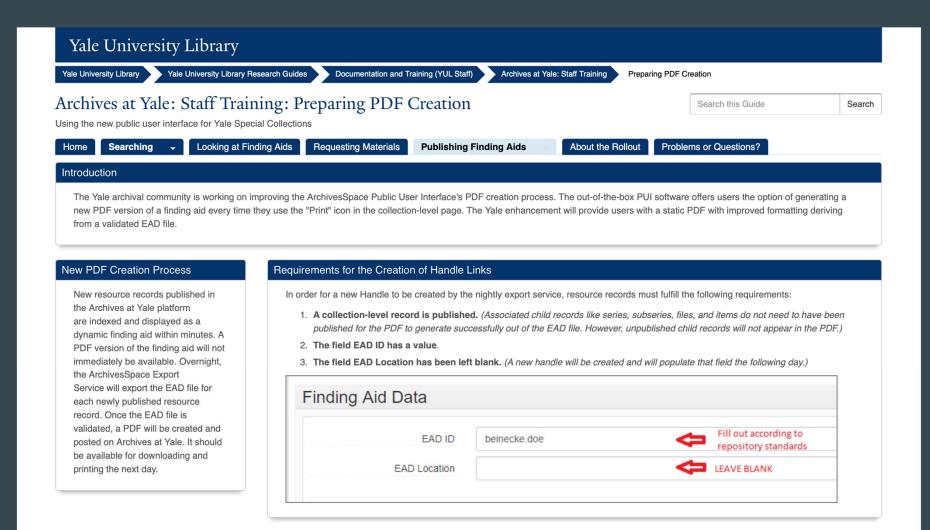
```
"Adds" list: newly published or recently updated Resource records
```

"Removes" list: deleted, suppressed, or unpublished Resource records

## What happens to files in the "adds" list?

- 1. Create a "Handle" link optional
- 2. Download the file
- 3. XML cleaning step
- 4. XSLT Transformations optional
- 5. XSD Validation
- 6. Schematron Tests optional
- 7. Generate PDF optional
- 8. Git Push

## What triggers a Handle to be created



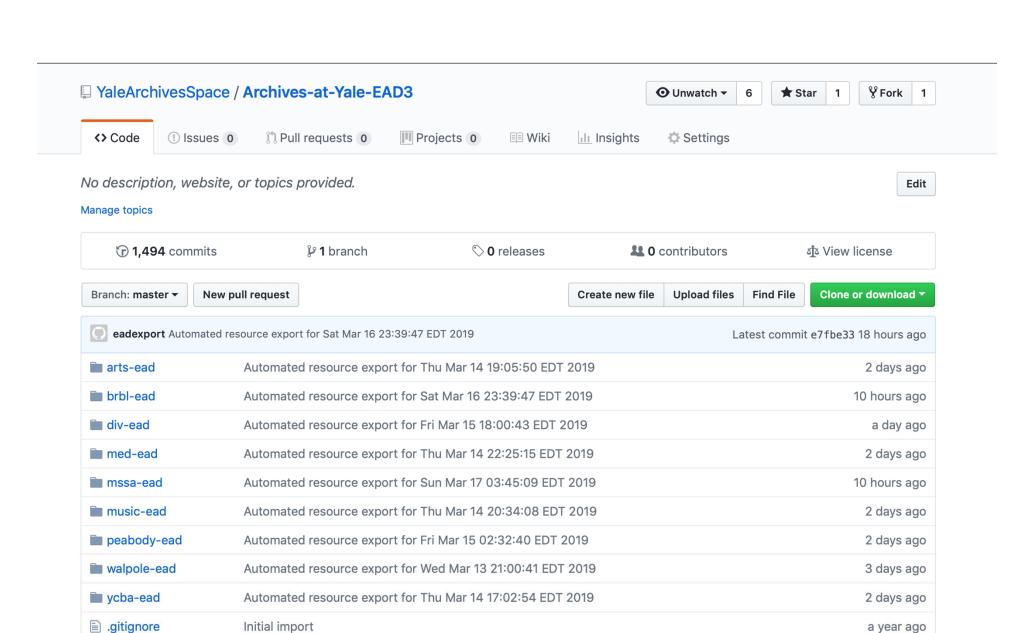
## A note on post-export cleanups

EAD3 export add-on (to handle dates):

https://github.com/YaleArchivesSpace/ead export addon/blob/master/backend/model/ead3 exporter addon.rb#L55-L61

XSLT changes: <a href="https://github.com/YaleArchivesSpace/EAD3-to-PDF-UA/blob/master/xslt-to-update-the-ASpace-export/yale.aspace-v2-to-

- Add CCO data license (now that we're using EAD 3 1.1)
- Collapse containers (e.g. 1, 2, 3, 4, 5 → 1-5)
- Reunite notes that were split due to character limits in ArchivesSpace
- ...plus lots more little changes to ensure valid EAD files.



a year ago

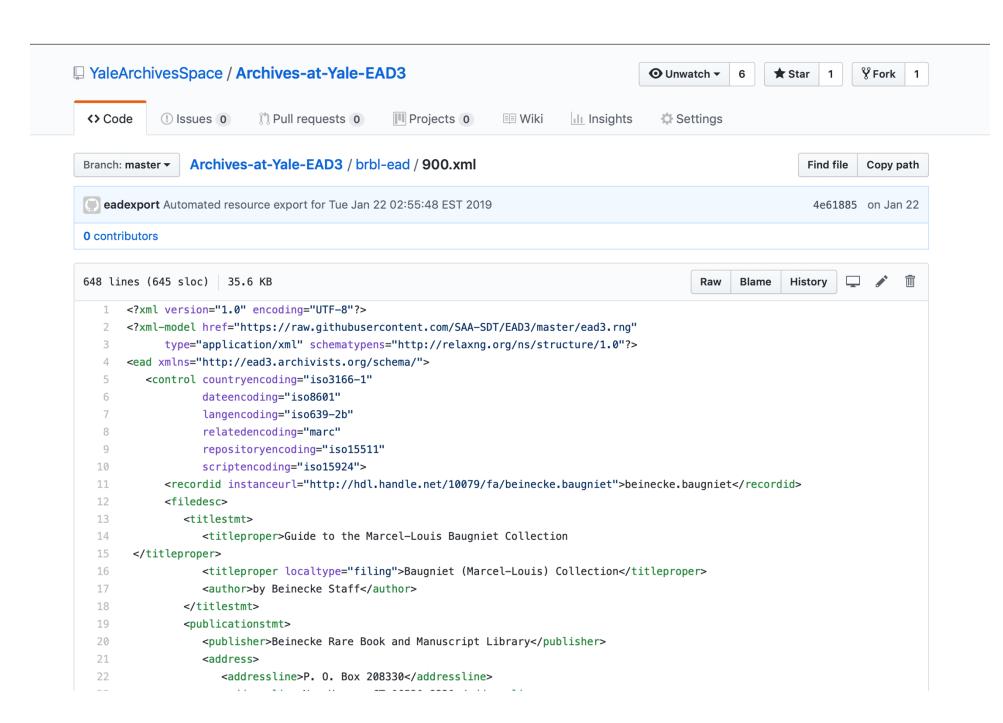
7 months ago

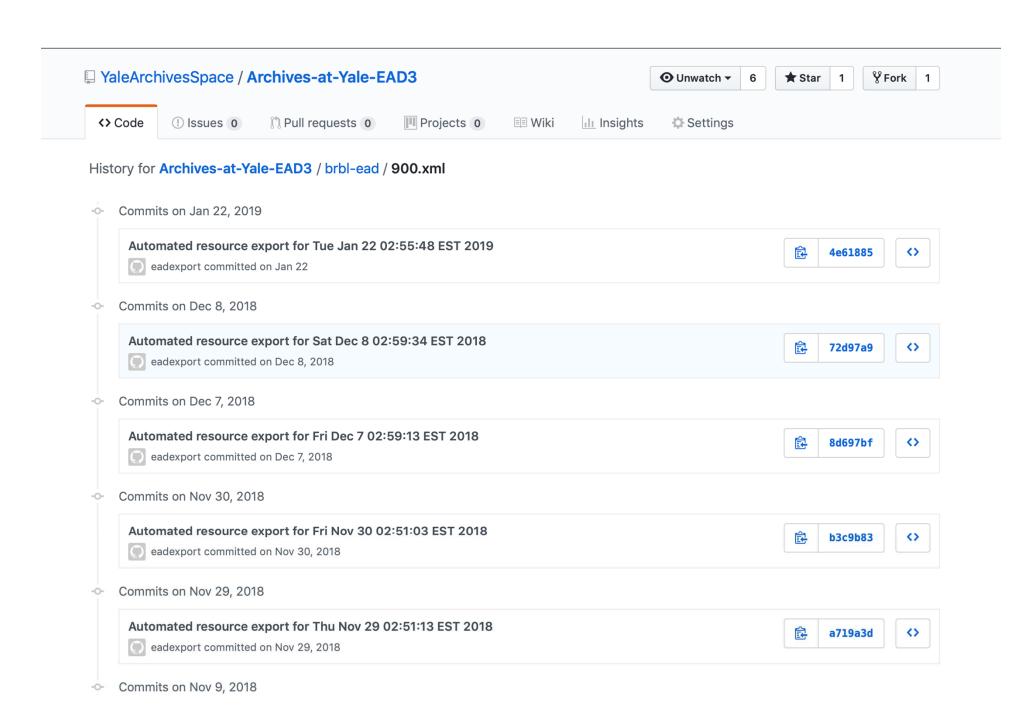
LICENSE.md

README.md

Update additional repository files

Update additional repository files





### **Marcel-Louis Baugniet collection**

### Collection

Call Number: GEN MSS 900

**Collection Overview** 

Finding Aid View

**Container List** 

## PDF Finding Aid

**99** Citation



### **Scope and Contents**

The collection includes material designed by Baugniet including original drawings for advertisements and other ephemera, printed versions of advertisements and sheet music, posters and ephemera related to Baugniet's exhibitions, and a periodical featuring an article on his wife Akarova.

### **Dates**

1922 - 1936

#### Creator

• Baugniet, Marcel-Louis, 1896-1995

### Language of Materials

In French and Dutch.

### **Conditions Governing Access**

The materials are open for research.

#### Extent

0.33 Linear Feet (2 boxes)

### **Catalog Record**

A record for this collection is available in Orbis, the Yale University Library catalog

### **Persistent URL**

http://hdl.handle.net/10079/fa/beinecke.baugniet

| Se | earch Collection     |                       |
|----|----------------------|-----------------------|
| Fı | rom year             | To year               |
| Se | earch                |                       |
| Na | vigate the collec    |                       |
|    |                      |                       |
| Ma | arcel-Louis Baugniet | collection            |
| Ma |                      | collection<br>22-1930 |

## So, every night...

Published resource records that have been recently updated are exported as EAD and added to our GitHub repository

Each EAD file is converted into a PDF/UA file

Deleted, unpublished, or suppressed records are removed

If a Handle is required, one is created and written back to ArchivesSpace

Errors are recorded for invalid files but those files are not updated

## Try it out

Yale's EAD3 files:

https://github.com/YaleArchivesSpace/Archives-at-Yale-EAD3

ArchivesSpace Export Service:

https://github.com/hudmol/archivesspace export service