

Binghamton University

## The Open Repository @ Binghamton (The ORB)

---

Library Scholarship

University Libraries

---

4-2019

### Rosetta Initiatives at Binghamton University Libraries

Erin Rushton

Jesse Russell

David Schuster

Follow this and additional works at: [https://orb.binghamton.edu/librarian\\_fac](https://orb.binghamton.edu/librarian_fac)



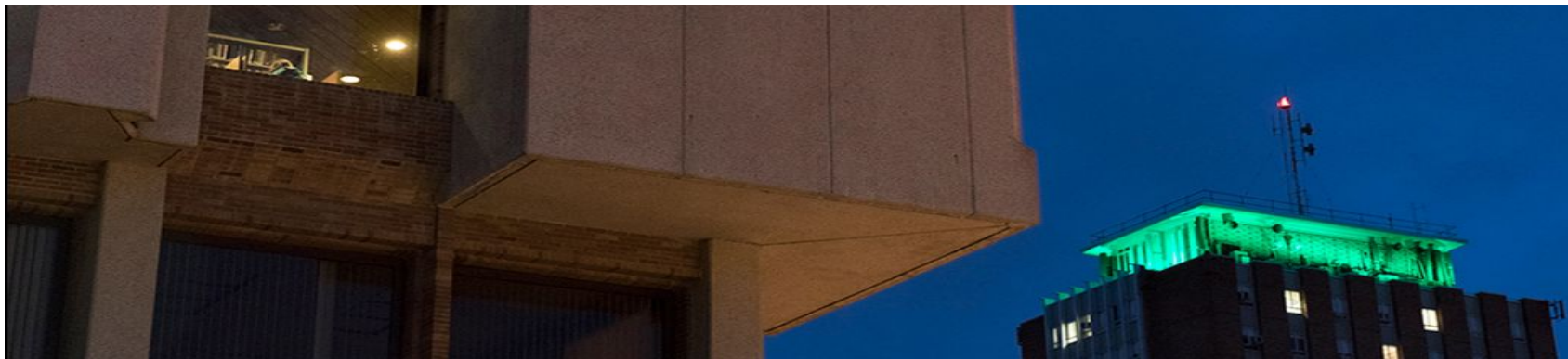
Part of the [Library and Information Science Commons](#)

---



# Rosetta Initiatives at Binghamton University Libraries

- **Erin Rushton**, Head of Digital Initiatives and Resource Discovery, [erushton@binghamton.edu](mailto:erushton@binghamton.edu)
- **Jesse Russell**, Library Systems Specialist, [russellj@binghamton.edu](mailto:russellj@binghamton.edu)
- **David Schuster**, Director of Library Technology and Special Collections, [dschuste@binghamton.edu](mailto:dschuste@binghamton.edu)





# Binghamton University Libraries and Rosetta

## Where it started

**2011:** Installed Rosetta - first university in US to do so  
**2011-2015:** Creating workflows and scripts, manually ingesting  
**2015:** Staffing changes, current staff less familiar with Rosetta technology  
**2016:** Servers getting old, concerns about space  
**2017:** Migrated to Virtual machines hosted by IT department

## What we currently do

-Ingest via manual and NFS (deposits of ~ 50 - 200 items)

-Migrating to Alma: attempting to manage metadata management efficiently

-Exploring large scale ingests (for us)

## What we plan to do

-Scale up production

-Continue working on integrations



# Digitized Special Collections Content in Rosetta



May 16, 1941

Mr. Michael P. Fellows,  
724 Officer's Mess,  
Box 10001,  
Ottawa, Ontario, Canada.

Dear Michael,

Thanks for your letter of May 15. I am very pleased to hear that everything is going along nicely and also that the personnel of the station is cooperating with you so nicely, and giving you information as to practical problems so that you will be able to simulate them on the trainer.

From your letter it sounds like Fort Albert is a very nice place and will be an enjoyable place to live for the summer.

I had had a discussion with MacGill in reference to the plan of assembly on the trainer and we have arrived at the following conclusion. You are to continue assembly and do all rigging and adjustment to the best of your ability, proceeding to read the gear lock and the cover. The four British boys will be available but need to assist you in this. We will not send Colwell unless you get a break. Bill Lonsdale will come in within the next two weeks to assist you in adjusting the cockpit part of the trainer so that you and he can get all the experience possible, and if you get stuck, Harold and Colwell will come up to assist.

With reference to your expenses for entertaining, I don't know just how we can set a stipulated amount, for in all probability it will vary. Of course we want you to do the necessary amount of entertaining for convincing good will and interest in the trainer. The amount of that will have to be up to your discretion as I go and see how we are going to draw up any iron fist rules for amounts you must abide by. The only thing I would suggest is that you submit such expenses week by week so if we do not agree with you on such expenses, we can advise you before it continues too far, as I have mentioned, entertaining is something you must draw on from time to time. I also think that it can be carried to excess as well as not doing enough.



RUSSIA, BITUMINOUS COALS  
ONE OF WESTERN AND COALS  
PREPARED SMITHING COAL  
& SPECIALTY COALS

**GEO. E. GREEN**  
BITUMINOUS COALS AND COKE  
BINGHANTON, N. Y.  
NEW YORK OFFICE: 170 BROADWAY

NEW YORK  
Nov 28<sup>th</sup> 1905.

My dear friend Diebig,  
Please regard Enclosures as  
Confidential. They speak for themselves,  
and give the best answer I can at  
present of Laque's situation. Really my  
dear fellow this is going to break  
out and finally give us more than our  
money back. Best wishes.  
Sincerely Yours,  
Geo. E. Green.





# Born Digital University Content in Rosetta

HOME ARTS RESEARCH SPORTS NEWS BRIEFS KUODS SUBSCRIBE ABOUT BINGHAMTON UNIVERSITY



English professor Liz Rosenberg reads from her new children's book *Tyrannosaurus Dad* at Barnes & Noble in Vestal on June 12. Photo by Jonathan Cohen

## English professor introduces young readers to a larger-than-life father

By Eric Colter | JUNE 20, 2011

Tweet 1

Recommend 8 recommendations. Sign Up to see what your friends recommend.

### Recent News

- SOM honors faculty excellence
- Scientific writer/author discusses 'risk' allies
- Baseball rallies past Cornell on the road
- Softball team completes three-game sweep of UMBC
- Union renovation to begin in summer

### Most Popular Stories

- Union renovation to begin in summer
- Accounting professor advocates for additional corporate oversight
- Plank leads men's lacrosse team past Canisius
- Stenger discusses growth with SOM alumni
- Jennifer Jensen named American Council on Education Fellow

### News Sources

- Events Calendar
- Dateline
- B-Line
- Ask A Scientist

### Follow Binghamton

- Facebook
- YouTube

Online journal of rural nursing and health care : Fall 2012 - Volume 12, Number ...

PRESERVATION\_MASTER:VIEW

SEARCH

Abstract: The Provost Award for Excellence in Undergraduate Research honors undergraduate students graduating during the 2014-2015 academic year who have excelled in research, scholarship and creative activities that extend beyond traditional coursework. Nominations must be made by the university's undergraduate director or chair. All nomination materials should be submitted electronically by March 13, 2015, to Ashley Sorbonich, via e-mail. For more information, go online. For questions, contact Janice McDonald at 607-777-4324.

**Provost's Award for Excellence in Undergraduate Research**

The Provost's Award for Excellence in Undergraduate Research honors undergraduate students graduating during the 2014-2015 academic year who have excelled in research, scholarship and creative activities that extend beyond traditional coursework. Nominations must be made by the university's undergraduate director or chair. All nomination materials should be submitted electronically by March 13, 2015, to Ashley Sorbonich, via e-mail. For more information, go online. For questions, contact Janice McDonald at 607-777-4324.

**Bearcats win #OurHouseRules contest!**

We did it, Bearcats! We took down Stony Brook and Albany (just to name a few) to win the #OurHouseRules contest! Watch the video and thanks for voting.

**Defiant Obedience: The Religious Acoustics of Boston**

Wills Young, associate professor of humanities and religion at Endicott College, Beverly, Mass., will present "Defiant Obedience: The Religious Acoustics of Boston" from 10-11:30 a.m. Monday, March 2, in LH-12.

The talk is part of the larger Spring 2015 Speaker Series in Religious Studies, "The Study of Religion in the Modern University."

**SUNY BEST: "Inspiring Energy Initiatives: The Private Sector"**

The SUNY Business and Education Cooperative of the Southern Tier (SUNY BEST) will host "Inspiring Energy Initiatives: The Private Sector" at 10 a.m. Thursday, March 5, at the University Downtown Center, 67 Washington St., Binghamton.







# Rosetta content is harvested by Primo

FindIt! Provided by BINGHAMTON UNIVERSITY LIBRARIES

NEW SEARCH DATABASES SUBJECT GUIDES COLLECTION DISCOVERY STARTS WITH/BROWSE NEWSPAPER SEARCH

anna wilcox Local Digital Collections ADVANCED SEARCH STARTS WITH/BROWSE

All Items anywhere in the record

Sign in to get complete results and to request items [Sign in](#) [DISMISS](#)

Refine my results

Sort by Relevance




Material Type   
 Images (3)

Creation Date   
 From To   
 1800 1900 [Refine](#)

Topic   
 Wilcox, Anna E.--Portraits (3)   
 United States--History--Civil War, 1861-1865--Portraits (2)   
 Chenango County (N.Y.) (1)   
 Pike, John--Portraits (1)   
 Smyrna (N.Y.) (1)

Author/Creator

3 Results

-  **Photograph of Anna Wilcox (Standing)**  
Unknown  
Binghamton University Libraries 1864?  
10026838  
[Online access](#)
-  **Photograph of Anna Wilcox (Close-up)**  
Unknown  
Binghamton University Libraries 1868?  
10026838  
[Online access](#)
-  **Group photograph of Anna Wilcox, John Pike, Mrs. J. [Emma Rose] Pike, and Mrs. W. S. [Sarah E.] Pike**  
Unknown  
Binghamton University Libraries 1868?  
10026838  
[Online access](#)



# Challenges

- **Expertise:**
  - Understanding Rosetta
  - understanding digital preservation
- **Resources:**
  - Not one person “devoted” to Rosetta
  - Competing priorities/projects
- **Time:**
  - projects often move slower than anticipated
- We are “getting stuff” into Rosetta but we know there are better ways/more “right” ways

# On a positive note....

- We are regularly ingesting content into Rosetta which means we are engaging in digital preservation
- We are constantly seeking new ways to make Rosetta more scalable and efficient for our institution
  - Alma - Rosetta Migration
  - Kitodo
  - Web-based ingests
  - Ingest of large loads (university photographs)







# Alma-Rosetta Integration

- Along with 64 other SUNY institutions - moving to Alma/Primo June 2019
- Our initial understanding of Alma-Rosetta Integration:
  - Continue to deposit into Rosetta
  - Content harvested into Alma via OAI
  - In Alma, update digital representation bib records & manage metadata
  - Updated metadata harvested by Rosetta
- Expectations:
  - Easier/more efficient to edit metadata
  - Synchronized metadata between Rosetta and Alma
  - For digitized print collections (catalogued in MARC), Alma would support basic crosswalking during both publishing and importing, allowing Rosetta to store a basic Dublin Core version of a MARC record.



# Reality of Alma-Rosetta Integration

- Found out recently that *“As you don't have Alma-D we're going to configure and load Rosetta as an external data source”*
- This first point wasn't clear from the documentation or from ExL
- Next steps
  - Focus until mid-2019 is implementation of Alma and Primo VE
  - As long as Alma harvests Rosetta (similar to how our current Primo-Rosetta integration works) we will be satisfied
  - Still interested in integration - need to determine if we really need Alma-D

# **B** Kitodo - What is it

- An open source software product for cultural institutions, libraries, museums, etc. designed to help control and monitor all stages of a digitisation project.
- We are interested in Kitodo.Production
- Features of Kitodo
  - Workflow management
  - Manage access rights and roles for scan operators, librarians, controllers, etc.
  - Digitisation projects are governed by a set of rules determining the appropriate structural and metadata sets
  - Individual process steps are precisely defined: including metadata import from the catalogue system, scanning the object by means of defined scan parameters (e.g. OCR)
  - Metadata files (METS/MODS) generated by the system are subsequently transferred to the presentation system (in our case Rosetta) or a search engine.



# B Kitodo - Our Vision

- Our vision for Kitodo
  - Physical storage, digitization, metadata creation exist in different areas of the Libraries we are interested in a tool that facilitates the transfer of files and metadata between departments
  - Ability to export from Kitodo via API to Rosetta
  - Take advantage of digitization templates for books, manuscripts, etc



# **B** Kitodo - Challenges

- Challenges
  - Not widely (not at all) installed in the United States
  - Documentation is in German
  - Open source *but* requires time/customization for implementation
  - We would need support for integration and training
  - Little documentation on how others have integrated with Rosetta
  - Can't tell where Kitodo stopped and the other services provided by Zeutschel exist
  - Kitodo is in flux from a stable, used v2 platform, to v3 that is currently in beta
  - Cost for training and setup?



# Kitodo - Current Status

- In early 2017 worked with ITS to install on server
  - Issues with file structure
  - Small successes
    - Integrated Alma BIBS API for search and retrieval
      - doesn't retrieve Network Zone records
      - needs more investigation

**Still interested but at what expense?**





# Web based Rosetta ingest

## The challenge

- DI Librarian not comfortable using back-end command line process to ingest content into Rosetta, a process which can be more efficient for uploading large batches of files and metadata

## The solution

- Systems specialist created a web interface that allows upload of CSV template and file content to be ingested in to Rosetta
- This tool creates the necessary directory structure needed to upload CSV template and ingest files (via sFTP), and sets all necessary permissions

## The process

- DI Librarian creates CSV and stages files for ingest
- Logs on to web interface to upload CSV file
- Upload ingest content
- Log in to Rosetta to complete ingestion



# How It Works

**BINGHAMTON UNIVERSITY** | LIBRARIES

### Rosetta CSV NFS Ingest

Directory:  \*

(letters and numbers only, special characters will be ignored, maximum length: 32)

Password:  \*

\*Required

The following ingest folders already exist:

1. `civilwars`
2. `link_diaries`

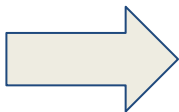


**BINGHAMTON UNIVERSITY** | LIBRARIES

### Rosetta CSV NFS Ingest

`/exlibris/incoming/reinhardtpromptbooks/` has been created.  
`/exlibris/incoming/reinhardtpromptbooks/content/` has been created.  
`/exlibris/incoming/reinhardtpromptbooks/content/streams/` has been created.

Select CSV to upload:  No file chosen



**BINGHAMTON UNIVERSITY** | LIBRARIES

### Rosetta CSV NFS Ingest

The file `1548274831_3343398.csv` has been uploaded to `/exlibris/incoming/reinhardtpromptbooks/content/`

Enter the password and click submit to continue with the existing file:

[Start Over](#)



**BINGHAMTON UNIVERSITY** | LIBRARIES

### Rosetta CSV NFS Ingest

Permissions have been set. You can now [FTP](#) your ingest content.

When the upload is complete, [login](#) to Rosetta and choose the `reinhardtpromptbooks` folder to ingest the uploaded content.

[Start Over](#)

# University Photographs

- Many years ago, the Libraries' agreed to add all of the university photographer's photographs to Rosetta
- Only a few hundred photographs were added before the project stalled because of staffing changes - also realized workflow was not scalable for the volume of photographs





# Photographs - Challenges

- 20000 + TIFF files stored on hard drives
- Collection not curated
- Collection will be discoverable - should each photograph be a single IE (could saturate search results) or group photographs into topics e.g. "Convocation 2018" but then lose unique metadata
- If we don't do something to preserve photographs, no one will



20180914\_bass03\_jwc.j...

JPEG Image



20180914\_bass04\_jwc.j...

JPEG Image



20180914\_bass05\_jwc.j...

JPEG Image



20180914\_bass01\_jwc.j...

JPEG Image



20180914\_bass02\_jwc.j...

JPEG Image



# **B** Photographs - Status

- Exploring photography batch ingesting
  - Reads the metadata out of the EXIF using ExifTool
  - Creates the CSV with metadata
  - Verify the subject headings
  - Use the Rosetta ingest web interface
- How to create single IE for multiple images - since the photos are done by activity





# Moving Forward

Enhance the existing CSV NFS ingest to allow for automatic CSV creation of campus photography using EXIF data

The screenshot shows a web interface for Binghamton University Libraries. At the top, there is a dark green header with the text "BINGHAMTON UNIVERSITY" and "LIBRARIES" separated by a vertical line. Below the header, the page title is "Rosetta CSV NFS Ingest". A message states: "Directory **anewtest** already exists, [start over](#) or upload your CSV to it below." Below this, it says "Select CSV to upload:" followed by a file selection button labeled "Choose File" and the text "No file chosen". There is a dark green button labeled "Upload CSV". Below that, the word "OR" is displayed. Underneath, there is a checkbox labeled "Binghamton University Photograph Ingest (Automatic CSV creation)". At the bottom, there is a dark green button labeled "Submit".





# Testing the concept...

## Rosetta CSV NFS Ingest (BU Photography)

Select sample image to upload to generate CSV file preview:

Choose File No file chosen

Upload Image

### Raw JSON:

```
[{"SourceFile": "/exlibris/incoming/newtest/20100513_lammert02_jwc.tif", "FileName": "20100513_lammert02_jwc.tif", "FileSize": "55 MB", "DateTimeOriginal": "2010.05.13 01:26:54", "Format": "image/tiff", "City": "VESTAL", "Province-State": "NY", "Sub-location": "-", "Description": "JONATHAN COHEN/BINGHAMTON UNIVERSITY \nBinghamton University graduate student Dawn Lammert, photographed outside Science II, Thursday, May 13, 2010.", "Creator": "JONATHAN COHEN", "Keywords": "[2010,'student life','portrait','graduate','science II','s2']" }
```

### PHP Array VAR\_DUMP:

```
array(1) { [0] => array(11) { ["SourceFile"] => string(54) "/exlibris/incoming/newtest/20100513_lammert02_jwc.tif" ["FileName"] => string(26) "20100513_lammert02_jwc.tif" ["FileSize"] => string(5) "55 MB" ["DateTimeOriginal"] => string(19) "2010.05.13 01:26:54" ["Format"] => string(10) "image/tiff" ["City"] => string(6) "VESTAL" ["Province-State"] => string(2) "NY" ["Sub-location"] => string(1) "-" ["Description"] => string(147) "JONATHAN COHEN/BINGHAMTON UNIVERSITY \nBinghamton University graduate student Dawn Lammert, photographed outside Science II, Thursday, May 13, 2010." ["Creator"] => string(14) "JONATHAN COHEN" ["Keywords"] => array(6) { [0] => int(2010) [1] => string(12) "student life" [2] => string(8) "portrait" [3] => string(8) "graduate" [4] => string(10) "science II" [5] => string(2) "s2" } }
```

### Test array parsing:

```
SourceFile: /exlibris/incoming/newtest/20100513_lammert02_jwc.tif
FileName: 20100513_lammert02_jwc.tif
FileSize: 55 MB
DateTimeOriginal: 2010.05.13 01:26:54
Format: image/tiff
City: VESTAL
Province-State: NY
Sub-location: -
Description: JONATHAN COHEN/BINGHAMTON UNIVERSITY Binghamton University graduate student Dawn Lammert, photographed outside Science II, Thursday, May 13, 2010.
Creator: JONATHAN COHEN
Keywords:
```

### PHP Array Parsed:

```
SourceFile -- "/exlibris/incoming/newtest/20100513_lammert02_jwc.tif"
FileName -- "20100513_lammert02_jwc.tif"
FileSize -- "55 MB"
DateTimeOriginal -- "2010.05.13 01:26:54"
Format -- "image/tiff"
City -- "VESTAL"
Province-State -- "NY"
Sub-location -- "-"
Description -- "JONATHAN COHEN/BINGHAMTON UNIVERSITY Binghamton University graduate student Dawn Lammert, photographed outside Science II, Thursday, May 13, 2010."
Creator -- "JONATHAN COHEN"
0 -- "2010"
1 -- "student life"
2 -- "portrait"
3 -- "graduate"
4 -- "science II"
5 -- "s2"
```



# Proof of concept

- Can automate ingest CSV creation, partially from static information, partially from extracted EXIF data
- Next step: Decide how to handle photo metadata, properly format an ingest CSV with it

## PHP Array Parsed:

```
SourceFile -- "/exlibris/incoming/aneptest/20100513_lammert02_jwc.tif"
FileName -- "20100513_lammert02_jwc.tif"
FileSize -- "55 MB"
DateTimeOriginal -- "2010:05:13 01:26:54"
Format -- "image/tiff"
City -- "VESTAL"
Province-State -- "NY"
Sub-location -- ""
Description -- "JONATHAN COHEN/BINGHAMTON UNIVERSITY Binghamton University graduate student Dawn Lammert, photographed outside Science II, Thursday, May 13, 2010."
Creator -- "JONATHAN COHEN"
0 -- "2010"
1 -- "student life"
2 -- "portrait"
3 -- "graduate"
4 -- "science II"
5 -- "s2"
```

## CSV preview:

Object Type	Title (DC)	Description (DC)	Creator (DC)	Subject - LCSH (DC)	Created (DCTERMS)	Physical Description (dc:description)	Type (DC)	Language (DC)	Medium (dcterms)	Identifier (DC)	Publisher (DC)	Modified (DCTERMS)	Is Part Of (DCTERMS)	Extent (dcterms)	Rights (DC)	Techni (DC)
SIP	NameOfSIP															
IE	Title	Description	Creator	Subject	Created	Physical	Type	Language	Medium	Identifier	Publisher	Modified	IsPartOf	Extent	Rights	Techni
/exlibris/incoming/aneptest/20100513_lammert02_jwc.tif	20100513_lammert02_jwc.tif	55 MB	2010:05:13 01:26:54	image/tiff	VESTAL	NY	-	JONATHAN COHEN/BINGHAMTON UNIVERSITY Binghamton University graduate student Dawn Lammert, photographed outside Science II, Thursday, May 13, 2010.	JONATHAN COHEN	2010	student life	portrait	graduate	science II	s2	



## Diary of Maurice Leyden, 1871



### Preservation Master

- BNG\_MLD\_01\_0008.jpg
- BNG\_MLD\_01\_0009.jpg
- BNG\_MLD\_01\_0010.jpg
- BNG\_MLD\_01\_0011.jpg
- BNG\_MLD\_01\_0012.jpg
- BNG\_MLD\_01\_0013.jpg
- BNG\_MLD\_01\_0014.jpg
- BNG\_MLD\_01\_0015.jpg
- BNG\_MLD\_01\_0016.jpg
- BNG\_MLD\_01\_0017.jpg
- BNG\_MLD\_01\_0018.jpg
- BNG\_MLD\_01\_0019.jpg
- BNG\_MLD\_01\_0020.jpg
- BNG\_MLD\_01\_0021.jpg
- BNG\_MLD\_01\_0022.jpg
- BNG\_MLD\_01\_0023.jpg**
- BNG\_MLD\_01\_0024.jpg
- BNG\_MLD\_01\_0025.jpg
- BNG\_MLD\_01\_0026.jpg
- BNG\_MLD\_01\_0027.jpg
- BNG\_MLD\_01\_0028.jpg



BNG\_MLD\_01\_0008.jpg



Maurice Leyden Collection  
Binghamton University Special Collections

### Descriptive Information

#### Diary of Maurice Leyden, 1871

**Creator :** Leyden, Maurice

**Description :** The Maurice Leyden Collection includes diaries kept by Maurice Leyden, a Civil War Veteran and dentist. Throughout his 1871 diary, Leyden discusses the weather, his dental practice with partner George Davis, local politics in Monroe County, Rochester, trips to nearby cities, and riding with his wife, Maggie. In addition to these topics, Maurice Leyden also discusses the various talks he went to hear with his wife, Maggie, who was one of fourteen women who voted with Susan B. Anthony in the 1872 Presidential election. For example, Maurice and Maggie went to hear Anne E. Dickinson, an advocate for the abolition of slavery and a proponent of women's rights, speak alluding to

# References

## Kitodo references

- <https://github.com/Kitodo>
- <https://uknowledge.uky.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1003&context=ifla-news-media>
- <https://drive.google.com/file/d/14rPII7vDIXZQ8Lqzmvk3lIM32T1bX2ZMdfmNcAcp0RFHDVVK2bxOsxJiYyBcBrb0rwqKY2Tbmkad-1H/view?usp=sharing>