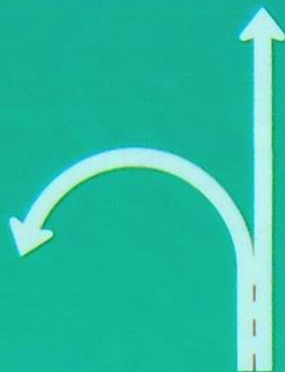


ALL

THE
ROADS



FIGGY



A
Tour

OF

Ingest
Workflows

Presented



Anna Headley &
Trey Pendragon



Princeton University



What is Figgy?

Figgy is our digital repository, with a staff-facing UI for ingesting and managing digital objects. The kinds of things staff do in Figgy include:

- Create resources
- Pull metadata from another system (for some resource types)
- Describe a resource using forms (for other resource types)
- Organize the files that comprise a resource, and track its progress along a workflow
- Create collections that can be published as spotlight exhibits

Ingest

Figgy has several ingest workflows. We add them either to:

- Accommodate how users are already working, or want to work.
- Ingest material from different vendors.

We're going to run through a bunch of ingest user stories and show how Figgy accommodates them.

Digitization Studio

As a Digital Imaging Technician I want to ingest a set of images in as few steps as possible in order to focus on photographing library materials. I want to create resources with complex hierarchical arrangements as part of the ingest process in order to reduce the need for structural editing later on. I need to easily understand whether the ingest was successful in order to complete the scanning project.



Save and Ingest

Workflow

1. Place images on a mount of our high-volume storage array
2. Name the directory after a bibliographic ID or finding aid component ID
3. Enter that same ID on the resource creation page in Figgy
4. Wait a moment for Figgy to discover that the directory is there
5. Click “Save and Ingest”

New scanned resource

Metadata

Title

Required if Source Metadata ID is blank

[+ Add another](#)

Source Metadata ID

Bib ID (1234567) or PULFA ID (C123 or C123_c456)

☐ Refresh metadata
from
PULFA/Voyager

☐ Set visibility by
Date Created

Collections

required Rights Statement

Rights Note

Save Work

Requirements

✓ Describe your work

Visibility

- ☒ **OPEN** Open to the world. Anyone can view.
- ☐ **PRINCETON** Restricted access. Only users logged in with a Princeton University NetID can view.
- ☐ **ON CAMPUS** Restricted access. Users can only view on campus or VPN.
- ☐ **READING ROOM** Restricted access. Only users who have been granted reading room permission in this application and are physically in an RBSC reading room can view.
- ☐ **PRIVATE** Only privileged users of this application can view.

Save

Searching...

New scanned resource

Metadata

Title

Required if Source Metadata ID is blank

[+ Add another](#)

Source Metadata ID

Bib ID (1234567) or PULFA ID (C123 or C123_c456)

☐ Refresh metadata
from
PULFA/Voyager

☐ Set visibility by
Date Created

Collections

required Rights Statement

Rights Note

Save Work

Requirements

✓ Describe your work

Visibility

- ☒ **OPEN** Open to the world. Anyone can view.
- ☐ **PRINCETON** Restricted access. Only users logged in with a Princeton University NetID can view.
- ☐ **ON CAMPUS** Restricted access. Users can only view on campus or VPN.
- ☐ **READING ROOM** Restricted access. Only users who have been granted reading room permission in this application and are physically in an RBSC reading room can view.
- ☐ **PRIVATE** Only privileged users of this application can view.

Save

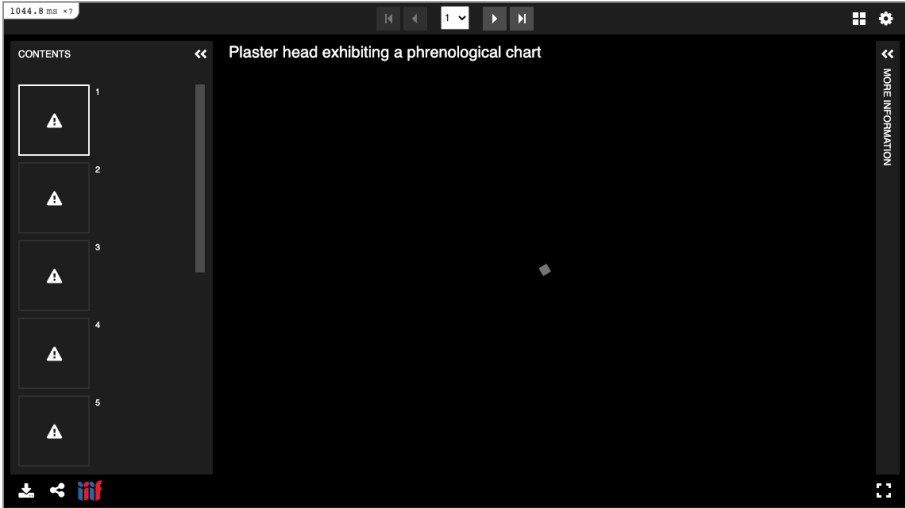
Save and Ingest

Ingest 10 files from ashley/10634526

Plaster head exhibiting a phrenological chart

Claim

Files Processed: 0 / 10

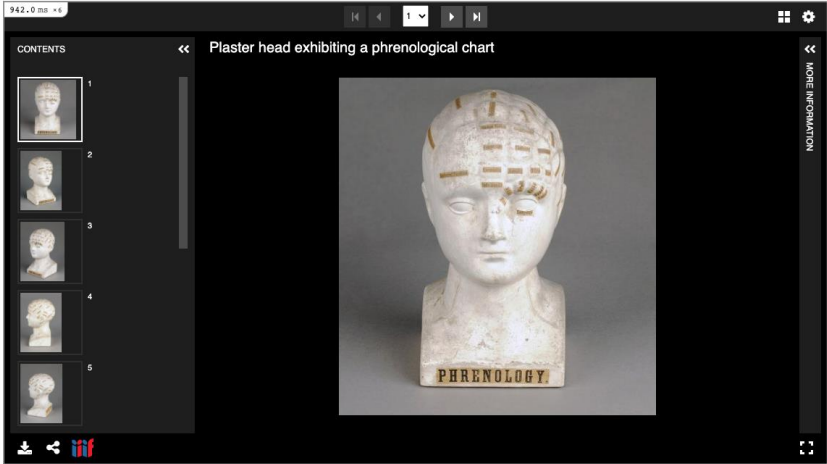


Imported Attributes

Attribute Name	Values
References	See Museum Objects Description File.
Extent	1 item
Title	Plaster head exhibiting a phrenological chart
Content Type	Visual material

Plaster head exhibiting a phrenological chart

Claim



Imported Attributes

Attribute Name	Values
References	See Museum Objects Description File.
Extent	1 item
Title	Plaster head exhibiting a phrenological chart
Content Type	Visual material
Description	Unidentified Artist [manufacturer] Medium: Plaster

Save and Ingest

Technical overview

- When form field is populated javascript hits a Figgy endpoint with the id; Figgy looks for directory and if found the button is activated
- At save time, Figgy finds the directory again to add files to the resource
- Relies on directory being named after the external id
- Directory contents must align to specified structure - details in the next use case (Bulk Ingest)
- Figgy's remote metadata services pull metadata from ILS bibliographic record or from finding aid to populate imported metadata fields.

Digitization Studio

As a digital imaging technician I want to shoot a full set of items and ingest them all at once in order to photograph more items at once without stopping to individually ingest each one into Figgy.



Bulk Ingest

Workflow

- Prepare files according to required directory structure
- Select directory in Figgy UI

Example directory structures

Single volume work:

- ingest_dir
 - bib_id
 - 1.tiff

Multi-volume work:






- ingest_dir
 - bib_id
 - vol1
 - 1.tiff
 - vol2
 - 1.tiff

Bulk Ingest Scanned Resources

Ingest one or more individual, multi-volume, or mixed resources. Files selected must be contained in a parent directory.

Show example directory structure

Visibility (1)

- ☒  **OPEN** Open to the world. Anyone can view.
- ☐  **PRINCETON** Restricted access. Only users logged in with a Princeton University NetID can view.
- ☐  **ON CAMPUS** Restricted access. Users can only view on campus or VPN.
- ☐  **READING ROOM** Restricted access. Only users who have been granted reading room permission in this application and are physically in an RBSC reading room can view.
- ☐  **PRIVATE** Only privileged users of this application can view.

Files (4)

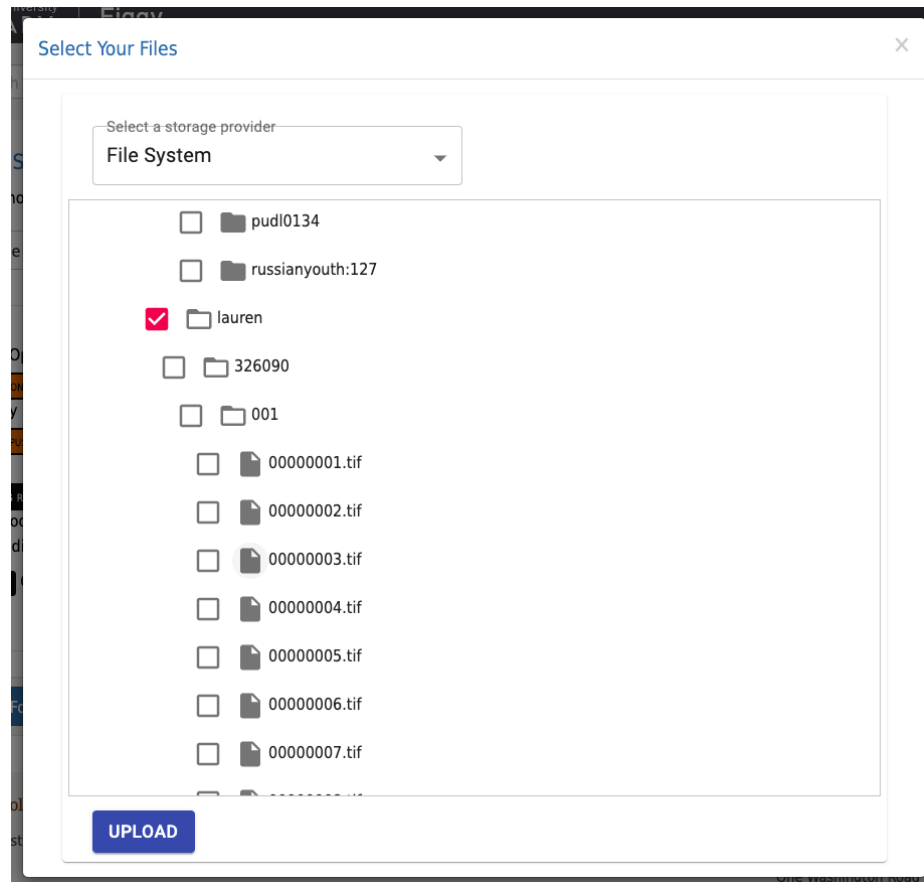
Select Folder on Server

Workflow State (2)

pending

Collections (3)

Nothing selected



Bulk Ingest

Technical Overview

- Browse everything provides UI
- Browse everything integration also provides directory name to pull all files and arrange into resources
- Use Figgy services to pull metadata from ILS or Finding Aids

Request Fulfillment

As an archives specialist I want to store and preserve digitized materials when I fulfill requests in order to increase their availability and avoid re-scanning when they are needed again. I want to integrate this preservation into my existing delivery workflow in order to keep document delivery straightforward for patrons.

Google Drive Integration

Workflow

- Fulfiller photographs material, creates a PDF, and uploads it to Google Drive
- Create a Resource via a form, add a component ID or BibID, and then use the “Upload Files from Server” button to select the PDF from their Drive

Technical Overview

- We use Browse Everything

Princeton University

LIBRARY

Enter search

File Manager

Focus

Total

1

Release Notes

trend

Select Your Files

Select a storage provider

Google Drive









AUTHORIZED

Select a file

Google Drive

Q

A-Z

Name	Owner	Last modified
 test_ingest	me	Apr 28, 2020
 Test Files	me	Nov 19, 2019
 Work on Valkyrie, an Update - Samvera 2020	me	Oct 22, 2020
 2020 Valkyrie Hackathon Working Document	me	Oct 22, 2020
 Valkyrie Hackathon 2020	me	Oct 22, 2020
 Technical Interview Brainstorming	me	Oct 22, 2020
 Samvera Slide Template	me	Oct 21, 2020
 Interview Pair Exercise	me	Oct 20, 2020

Select

Cancel

Drag / Drop Upload

Workflow

- Fulfiller photographs a file, creates a PDF or set of files, and saves it on their own computer.
- Drag and Drop the file from their machine into the target area.

Technical Overview

- We use vue2-dropzone, which utilizes dropzone.js.
- Dropzone.js is configured to directly PUT to the resource's edit action to append a file.

File Manager

[Board Member Positions](#) / File Manager

Upload Files from Server

Drag and drop files here to upload.

Total files: 1

Scanning Station

As an access and lending specialist I want to provide materials to off-site patrons during University and building closures in order to support their research. I want to use existing scanning stations and workflows, and avoid using new interfaces like Figgy directly, in order to minimize required staff training.



Mounted Ingest Directory

Workflow

1. Users copy a PDF made by the station to a given directory with the file named after a catalog ID, e.g. 123456.pdf
2. Team leaders receive hourly emails informing them of those which are done, and notify requesters.

Mounted Ingest Directory

Technical Overview

1. Automated process every hour copies all PDFs last modified > 1 hr ago into another directory and spawns an ingest job.
 - a. This long last modified window protects against ingesting files that are still in transit
2. Ingest job creates a resource with the right privacy/collection for these items, uses the catalog ID from the filename for metadata, splits the PDF into pages, creates derivatives, and keeps the PDF as the first member.
3. Every hour another process “complete”s all these objects which have had all their derivatives made and sends notification emails.

Ephemera Collections

As an ephemera collection manager I want to describe objects directly in Figgy as I prepare them for shipping to a digitization vendor in order to maintain data in a single location. When I receive the files from the vendor I want to bulk load them and have Figgy attach them to the resources I've already created.



Bulk Ephemera Ingest

Workflow

- Description-first
- For these items we provide a lot more form fields than we normally do, since they do not have an external metadata system

In Ephemera Box 2

required Barcode

required Folder number

required Title

Sort title

[+ Add another](#)

Alternative title

[+ Add another](#)

Transliterated title


[+ Add another](#)

required Language






Nothing selected

Save Work

Requirements

 Describe your work

Visibility

- ☒  **OPEN** Open to the world. Anyone can view.
- ☐  **PRINCETON** Restricted access. Only users logged in with a Princeton University NetID can view.
- ☐  **ON CAMPUS** Restricted access. Users can only view on campus or VPN.
- ☐  **READING ROOM** Restricted access. Only users who have been granted reading room permission in this application and are physically in an RBSC reading room can view.
- ☐  **PRIVATE** Only privileged users of this application can view.

Save

Save and Duplicate Metadata

Save Draft

Bulk Ephemera Ingest

Technical Overview

- When a digitization vendor sends us back a hard drive of a box of materials in a pre-defined folder format our digitization studio mounts it and runs a rake task in Figgy to automatically ingest and associate those images with pre-created records.
- The folders in the hard drive are named after the barcode on the physical folders in the box, which is in the metadata for each object.
- They spot check the files that get ingested and mark a full box as “complete.”

Archival Audio & Video

As an archivist I want to ingest a BagIt bag of digitized audio files, photographs, and xml metadata in order to preserve and provide access to fragile materials. I want digitization notes saved on the Figgy resource in order to inform researchers of quality issues. I want other descriptive metadata imported from the finding aid in order to avoid duplicate data.




BagIt Ingest

Workflow

- A set of A/V materials is sent to a vendor, they send back BagIt bags in a given format for every piece of media.
- Each bag includes a directory for each digitized material, each directory has a Preservation Master, an Intermediate File, a web-quality derivative, and a picture of the material along with some unique descriptive metadata in PBCore XML for each “side” of a tape.

New Archival Media Collection // Figgy

https://figgy-staging.princeton.edu/archival_media_collections/new120%Search

 **PRINCETON UNIVERSITY LIBRARY**

Figgy

Enter search terms

+

aheadley

New archival media collection

Metadata

required Source Metadata ID

C0662

☐ Refresh metadata from PULFA/Voyager

☐ Set visibility by Date Created

Bag path


/mnt/hydra_sources/archives_bd/bcadmin/mss/av_backlog/la_c0984_2017_05_bag


Save Work


Requirements

✓ Describe your work

Visibility

☐  **OPEN** Open to the world. Anyone can view.

☐  **PRINCETON** Restrict access to only users and/or groups from Princeton University Library.

☒  **PRIVATE** Only privileged users of this application can see the work.

Save

Research Tools

Data and Statistics

Databases

E-journals

Research Guides

Main Catalog

Search

Library Services

Article Express

Borrow Direct

Circulation

Course Reserves

Interlibrary Loan (ILL)

Library Privileges

About the Library

Collections and Collecting

Exhibitions




Library Locations

For Alumni


For Library Staff

Staff Directory

Princeton University Library

   **FRIENDS**

One Washington Road
Princeton, NJ 08544-2098 USA
(609) 258-1470 Phone

 **PRINCETON UNIVERSITY**

BagIt Ingest

Technical Overview

- An “Archival Media Collection” is created or appended to in order to group a bag’s materials.
- Background job goes through each directory and creates a work, attaches media files, extracts digitization metadata from XML, and attaches the XML file. The work is associated with the Archival Media Collection.
- Descriptive metadata is synced from finding aid.

Migrations and Other Special Requests

Services and rake tasks to:

- Ingest objects with METS metadata
- Ingest objects from custom-designed databases
- Ingest custom intermediate files (e.g. watermarked images)



Thank you!

Anna Headley

 aheadley@princeton.edu

 hackmaster.a

 @hackmasterA

Trey Pendragon

 tpendragon@princeton.edu

 tpendragon

 @pendragon_dt

Image credits

- Title slide

- Highway signage <https://www.flickr.com/photos/jjbers/32430317987/> (modified)
- Dog in sidecar <https://www.flickr.com/photos/121433567@N03/14346529683> (trimmed)
- Cats in taco truck used by permission from Marijane White (modified)
- Dog in white car <https://www.flickr.com/photos/joliexis/140252829> (trimmed)
- Background road / sky <https://www.flickr.com/photos/cogdog/19345141996>

- Slide headers

- Highway signage <https://www.flickr.com/photos/jjbers/45123552451> (modified)

Image credits

- Cat on motorcycle <https://www.flickr.com/photos/tchdavid/45375805824>
- Bunny on Truck <https://www.pxfuel.com/en/free-photo-omnaw>
- White Lab in Truck <https://www.flickr.com/photos/therubyring/42200499761>
- Terrier in rear view <https://www.flickr.com/photos/tedmurphy/4486409323>
- Poodles in volvo <https://www.flickr.com/photos/thomashawk/40335151142>
- Terrier steering <https://www.pikrepo.com/fajyy/dog-driving-vehicle>
- Stuffed Golden <https://www.flickr.com/photos/andreboeni/42230677825>