NewspaperWorks:

It's How Samvera Does Newspapers

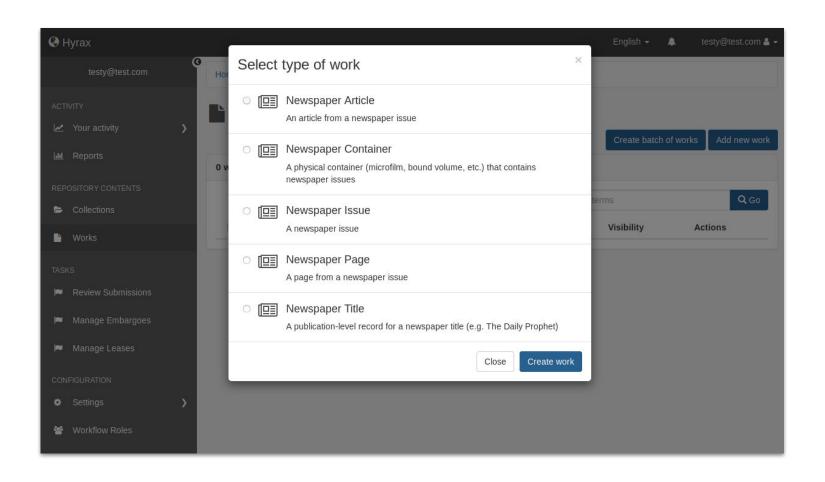`https://github.com/samvera-labs/newspaper_works`

NewspaperWorks is a gem

Provides content models, batch ingest workflows,
and front-end interaction for digitized newspaper content

- Works with any Hyrax 2.5.*-based repository app

- Easy to install and configure

```
$ vim Gemfile
gem 'newspaper_works'
$ bundle install
$ rails generate newspaper_works:install
```
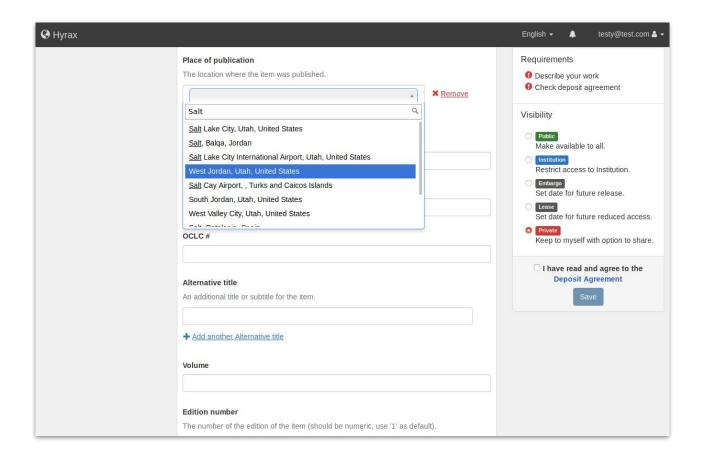
Newspaper work types & relationships

Title, Container, Issue, Page, Article

Newspaper metadata fields

Publisher, place of publication, ISSN, LCCN, edition, volume, issue, section, page, geographic coverage, text direction, etc.
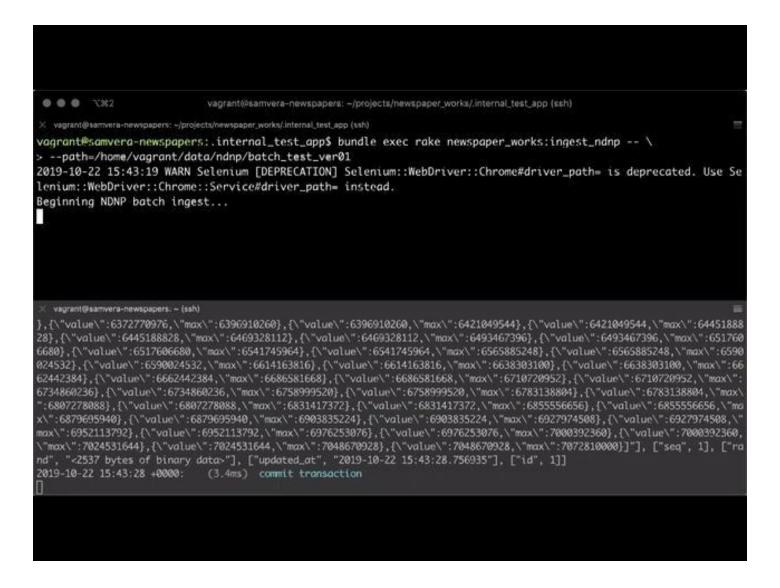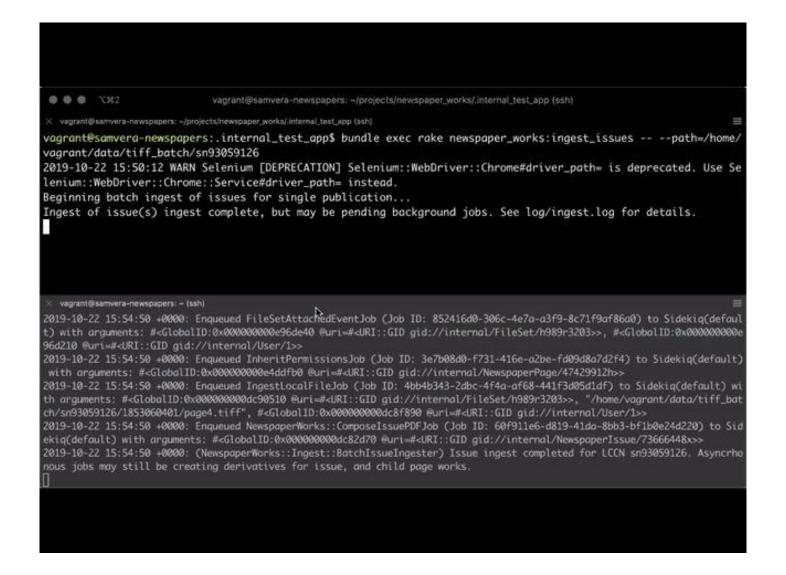
Batch ingest workflows

- National Digital Newspaper Program (NDNP) spec
  - with or without TIFF files

- PDF issues
  - splits PDF into component pages

- TIFF or JP2 page images

```
$ rake newspaper_works:ingest_ndnp -- \
    --path=/path/to/your/ndnp_batch \
    --admin_set=admin_set/default \
    --depositor=admin_user@example.com \
    --visibility=open
```
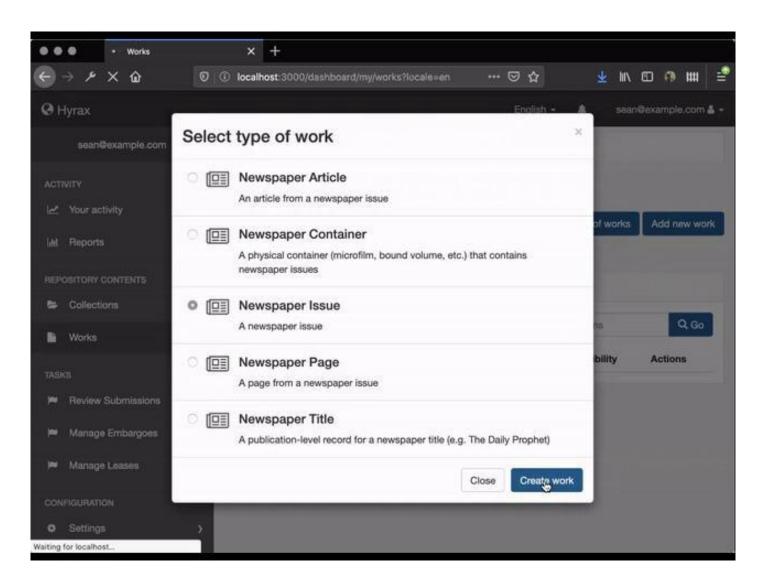
Batch ingest in action: NDNP

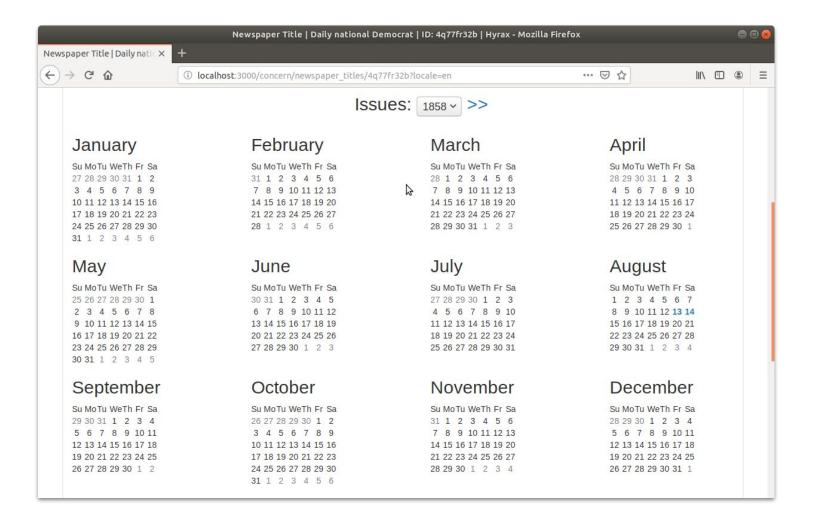# Batch ingest in action: TIFF

Single item ingest: PDF issue

Batch ingest actions
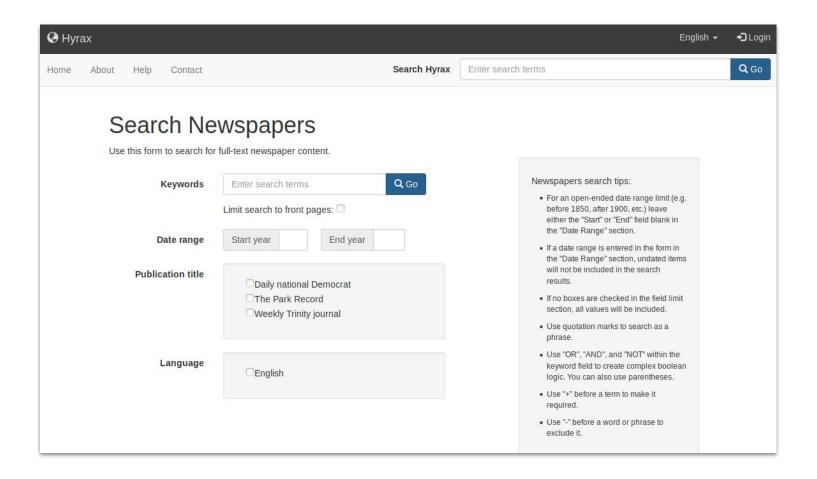
Command-line interface (rake task)

- Parses directories and files

- Creates repository objects
  - Title, Issue, Page

- Adds available metadata
  - publication title, date, place of publication

- Attaches files and creates derivatives
  - PDF, JP2, ALTO XML, JSON word-coordinates

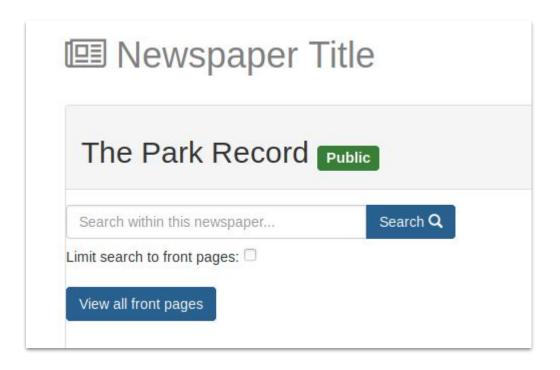- Can specify AdminSet, depositor, and visibility
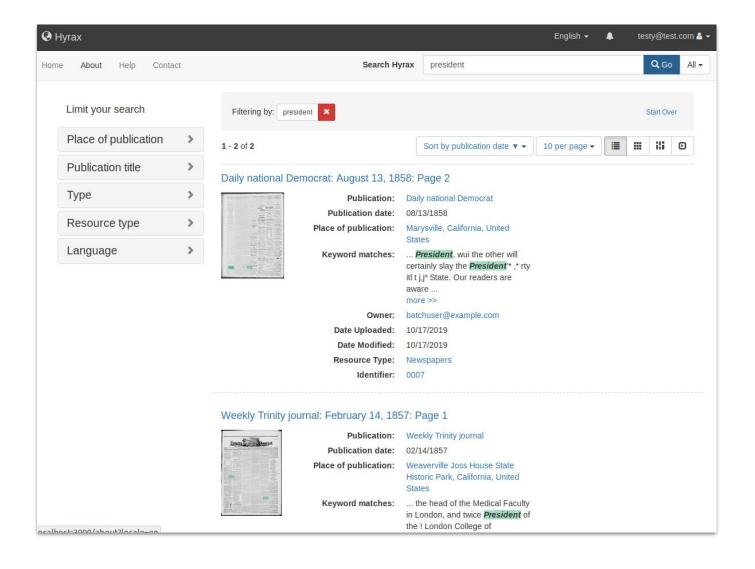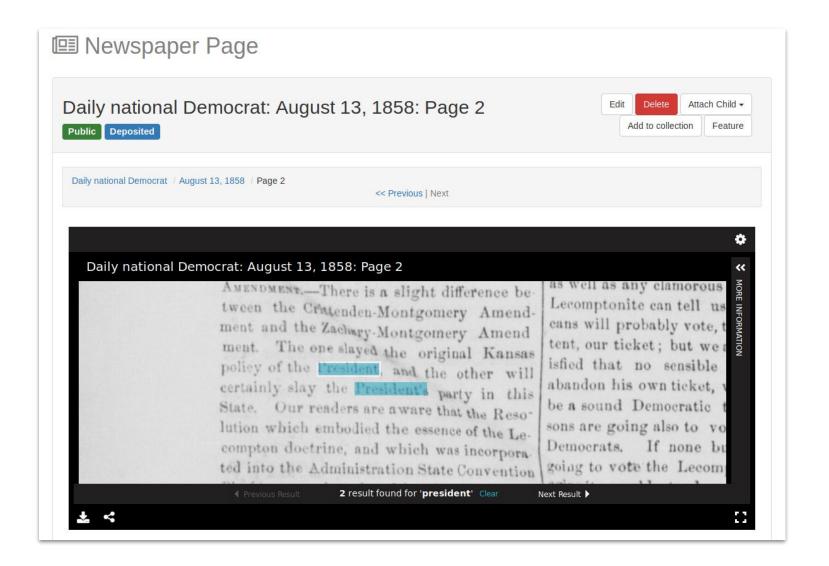
Calendar issue browsing

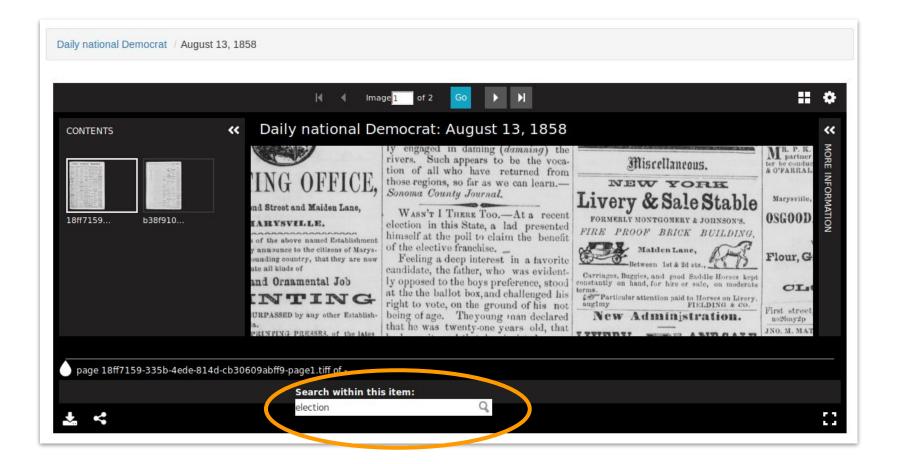# Newspapers advanced search

Search within title; browse front pages

Keyword matches in search results

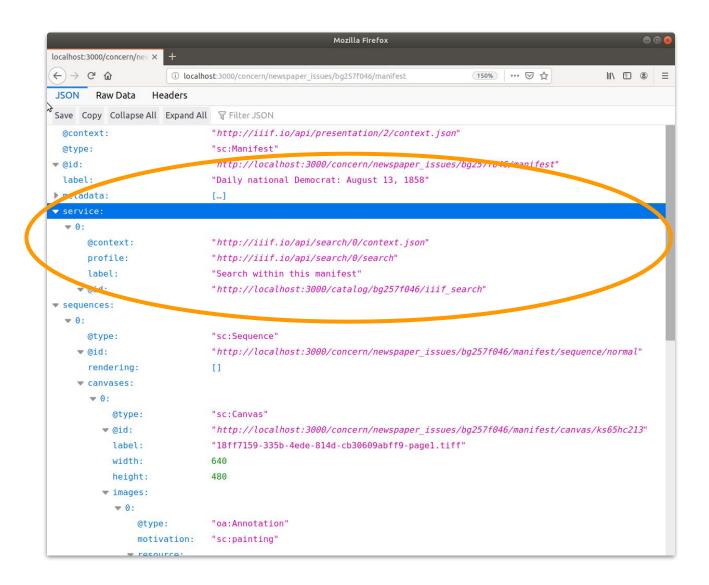# Zooming viewer

Full text search in viewer

IIIF: Presentation manifest + Content Search

Blacklight IIIF Search

Gem that provides an endpoint in your Blacklight app that will return a JSON response conforming to IIIF Content Search API v. 1.0.

Integrate the URL for this service into IIIF Presentation manifests to allow clients/viewers to provide functionality for searching within a resource and displaying results.

```
$ vim Gemfile
gem 'blacklight_iiif_search'
$ bundle install
$ rails generate blacklight_iiif_search:install
```

Hierarchical navigation

Semantic URLs

Predictable URL based on LCCN, pub date, edition, page

Publication date
  08/13/1858

Identifier
  0007

Resource type
  Newspapers

Persistent url
  http://localhost/newspapers/sn84038814/1858-08-13/ed-1/seq-2

Features in development

- Batch ingest of article-segmented content

- Download clipped image segment

- Geographic browsing & visualization

- Timeline browsing & visualization

- Hyrax 3.0 compatibility

How to get started

Wiki:

- Detailed docs on installation, setup, object modeling, metadata profile, batch ingest, and more
https://github.com/samvera-labs/newspaper_works/wiki

Test site:

- Hyrax repository application running latest release of gem
https://newspaperworks.digitalnewspapers.org/

Vagrant box:

- Installs all dependencies (including Fedora and Solr)
- Creates test app with NewspaperWorks installed
https://github.com/marriott-library/samvera-newspapers-vagrant

NewspaperWorks needs you!

Samvera Newspapers Interest Group:

- Meets bi-monthly on 1st Thursday 1 PM ET/11AM MT
  https://wiki.duraspace.org/display/samvera/Samvera+Newspapers+Interest+Group

Samvera Slack:

- #newspapers channel
  https://samvera.slack.com

Acknowledgements

- Institute of Museum & Library Services

- Grant advisory committee
    - Michael Appleby
    - Jenn Colt
    - Ryan Cordell
    - Michael Della Bitta
    - Karen Estlund
    - Tom Ewing
    - Michael Giarlo
    - John Herbert

- Newspapers Interest Group

Questions? Contact us

```
|----------------|------------------------------------|
| Eben English   | eenglish [at] bpl [dot] org        |
| Brian McBride  | brian.mcbride [at] utah [dot] edu  |
| Jacob Reed     | jacob.reed [at] utah [dot] edu     |
| Sean Upton     | sean.upton [at] utah [dot] edu     |
```