Newspapers in Hydra (Hydra Connect unconference session)

When: Thursday October 6, 2016, 11:00-11:50 AM  
Where: Boston Public Library, Orientation Room  
Facilitator: Eben English  
Notetaker: Eben English

(While this was not an official meeting of the Newspapers Interest Group, many members were in attendance, and it seems logical to post the discussion notes under the auspices of this group.)

Replacement for current systems

Many people expressed a desire for a locally-hosted, open-source platform to replace vendor-provided systems, such as Veridian and CONTENTdm. Support for migrating from these systems would be highly desired.

Existing work

- Indiana: Hydra-based 'Pumpkin' application builds off the paged-media functionality in Princeton's 'Plum'.  
- National Library of Wales: Reportedly has a nice system that implements IIIF manifests at the article level.  
- Maryland: Not presently using Hydra, but have newspaper content in Fedora+Solr-based system
  - PCDM-based modeling, as shown in poster session  
  - METS-ALTO stored as blob/datastream in Fedora  
  - word-coordinate data indexed in Solr to facilitate search highlighting on page image  
- open-oni, based on Library of Congress software developed for Chronicling America project

BPL and Utah Grant update

Application status

- Seeking 250K from IMLS  
- Initial preliminary proposal submitted in February 2016, but was not selected for full proposal application  
- A revised preliminary proposal was submitted in September 2016, hoping to hear results by early November  
- Full proposal due January 13, 2017  
- Would receive notice of funding in April 2017  
- Project would begin in June 2017

Project goals

- Develop a set of gems that could be integrated into Hydra based applications (as opposed to a stand-alone system such as GeoBlacklight or Avalon)
  - Admin - would provide ingest, management functionality  
  - Display - would provide discovery and display functionality  
    - Would likely use IIIF APIs for search and display, so that this gem would be fairly agnostic about back-end implementation  
- Implementers could choose which gem(s) they want to use.

Newspaper IG possible activities

- Resource Lists  
  One possible initial group activity would be to create a list of existing digitized newspapers resources.  
  - Systems currently in use by libraries  
  - Tools for working with newspaper content  
  - Other parallel groups/activities
    - Open-oni community  
    - IIIF Newspapers Interest Group  
      - meets monthly  
      - meetings/notes announced on IIIF-Discuss Google group.
- Requirements  
  Come up with a list of features for newspaper content ingest, management, discovery, and display
  - Hydra-in-a-Box integration is highly desired, since that system may be widely implemented  
  - Sufia / CurationConcerns integration  
  - NDNP XML data format batch ingest  
- PCDM modeling  
  - Need models at several levels: edition(?), volume(?), issue, page, article  
  - Might be possible to create edition or volume access using descriptive metadata rather than actual modelled 'objects'  
  - Newspapers can have multiple articles per page/image, and articles can span multiple pages/images
Newspaper IG meetings

Eben will initiate this, most likely a Doodle poll to figure out a workable time.