



Catholic Research Resources Alliance

Catholic Newspapers Program Meeting at St. Joseph's University

Monday, Jan. 27, 2014

2:00-4:00pm

Saint Joseph's University

5600 City Ave, Philadelphia, PA 19131

We will meet in the [Post Learning Commons/Francis A. Drexel Library](#), located at the corner of Lapsley and City Ave

[Directions to SJU campus](#)

Getting to or from the AIRPORT: The best options are by taxi or train. If folks choose to go by train, they can orient themselves on the [Amtrak](#) hot link. Using this option, get off at the Overbrook stop and walk northeast up the hill on City Ave. to Lapsley (the 2nd light).

Meeting Agenda and Goals

Purposes: to bring together individuals from the various groups who have participated in the Catholic Newspapers Program (CNP) in a discussion of the three primary aspects of the program: the directory (ICON), digitization, and the repository with an emphasis on critical questions and next steps for implementation.

Goals:

1. Bring people and perspectives together
2. Put together the pieces of the project: the directory (ICON), digitization, the repository (facilitate common understanding of the whole and how the pieces interrelate)
3. Identify critical questions and next steps

Agenda

- Background and status report
- Big picture
- Directory (ICON)
 - Amy Woods: The ICON registry is open for searching by CRRA members, CRL members, others
 - Amy Woods: Workflow to be developed. Will address what is needed for "permission age" for putting newspapers location and holdings data received from Pat into the

queue. Also will address with Pat if possible or desirable for CRRRA members to submit data directly to Amy for inputting in ICON registry. ICON staff do all direct entry of data into ICON registry. Registry the preferred term, instead of directory. ICON is an authoritative registry of titles, locations and holdings with issue level holdings data. Developed as a collection management tool. Individual institutions and/or consortia can use to manage retention of materials, to set priorities for digitization, etc.

- Amy Woods: Next step is to add data on commercially filmed or digitized newspapers to ICON to provide complete picture of where newspapers are located, in what formats, what holdings, etc. This will enable libraries and archives to set their own priorities for access, physical preservation, digitization and digital preservation.

-

- Digitization

- Laurie Arp: digitization process with Lyasis is coordinated, single point of contact, an A to Z process from initial cost estimate through ordering, digitization and return of images and materials. Brings solid track record and support in fundraising.
- Laurie Arp: Deliverables include specifications, formats, OCR – ABBY (fine reader) with option for Standard METS/ALTO or with article level segmentation and headline correction.
- METS is Metadata Encoding and Transmission Standard for encoding descriptive, administrative and structural metadata for objects in a digital library. Widely used standard.
- ALTO provides latitude and longitude of letters in text so that functionally later, can highlight keywords within the context of an article. ALTO (Analyzed Layout and Text Object) is a XML Schema that details technical metadata for describing the layout and content of physical text resources, such as pages of a book or a newspaper. It most commonly serves as an extension schema used within the [Metadata Encoding and Transmission Schema \(METS\) administrative metadata section](#). However, ALTO instances can also exist as a standalone document used independently of METS.
- Shawn Weldon: PAHRC has extensive newspapers collections, including those of American Catholic Historical Society. Willing to let any/all of these newspapers be used in digitizing. PAHRC would require assistance in preparing, shipping, reviewing materials for digitizing. PAHRC has no staff for this. Possible to request support in grant request?
- Pat Lawton: What are the desired specifications and standards that should be used in CRRRA project?
- Betsy: these are the guidelines for how materials are handled, what the vendor does and how.
- Kate Nevins, others: Set up small group to work out guidelines for CRRRA project. All agree will use national standards, etc. but need to define the specifications and standards so as to achieve desired level of functionality and user experience in CRRRA project.
- Betsy: What if a funder requires a competitive bidding process?

- Kate/Laurie: Lyrasis can do a bidding process if needed but prefers to start with assessing quality (high) of work done by its current digitization vendors.
- Marta Deyrup: Which copies and issues would be used in digitizing?
- Pat: Assessing now in discussions among digitizing partners. Intent is to find best copy and complete runs. Will need to use issues from more than one location in some, many cases.
- Any: ICON registry can help. Includes issue level correlation sheet so can make notes about individual issues, missing pages, misnumbering, etc.
- Laurie: Some of this information could be helpful to have in digitizing so know e.g., if misnumbering is ok.
- Kate: Lyrasis is seeking funds to assist smaller institutions in digitizing and to get the digital content accessible via DPLA (Digital Public Library of America)
- Darren Poley: Villanova bases its digitizing agreement on Creative Commons license using the broadest license as the default. Villanova asks for perpetual access via the Villanova Digital Catholic Library, indemnification against copyright violation (low risk in most cases) and owner insurance to cover damages or loss of materials. Villanova retains a copy in exchange for doing the work; will not sell digital copy but does provide access to it in perpetuity. Content is on Villanova site (with wrapper and owner logo around it) because most partners don't have capacity to do so on their own sites.
- Shawn: What tracking capability and/or use data is there? Google analytics.
- Evelyn Minick: University attorney questioned "in perpetuity" at Villanova.
- Amy Woods: ICON registry records (as sent by CRRA) will have much of the metadata needed for packing lists with bibliographic and holdings data accompanying material when sent for digitization.
- Pat Lawton: See the parallel activities of entering CRRA newspapers data – location, titles, holdings – into ICON registry and digitization activities – support creation of packing lists, for future rounds, to set digitizing priorities.
- Repository
 - Betsy Post: The Repository TF defined functional specifications and has narrowed choices to 2 options for a hosted repository. TF believes important to set functional specifications at high level to support maximum user experience and to look for hosted repository from vendors who have experience and expertise in working with digital newspapers content.
 - Betsy Post: Need to consider how to handle options for correcting OCR mistakes at the back end in the repository, perhaps crowdsourcing?
 - Will have to determine what metadata goes into repository. Title? Holdings? Other? More?

Next steps

Considerations

Critical questions

What's missing

What do we need to know

Attendees

1. Amy Wood (Center for Research Libraries)
2. Betsy Post (Boston College) Directory/Repository group and Digitizing Partner
3. Cait Kokolus (St. Charles Borromeo) Newspapers TF and Digitizing Partner
4. Darren Poley (Villanova) Newspapers TF and Digitizing Partner
5. Demian Katz (Villanova) Digital Access Committee (DAC), Directory/Repository group
6. Evelyn Minick (SJU) Board of Directors, Membership Committee
7. Jennifer Younger (CRRA)
8. Kate Nevins (Lyrisis)
9. Kristin Mudrick (SJU) (Liaison)
10. Laurie Gemill Arp (Lyrisis)
11. Marian Courtney (SJU)
12. Marta Deyrup (Seton Hall) DAC
13. Michele Woods (Xavier New Orleans) Digitizing Partner (Unable to attend)
14. Pat Lawton (CRRA)
15. Shana McDanold (Georgetown) DAC (Unable to attend)
16. Shawn Weldon (PAHRC) Newspapers TF, Digitizing Partner