Catholic Research Resource Initiative (CRRI)
Digital Access Committee
Minutes, Thursday, March 13, 2008

The committee held its meeting via conference call. Present were Tom Leonhardt, Chair; Susan Leister, Eric Lease Morgan, Marta Deyrup, Maura Kenny, Kevin Cawley, and Ruth Bogan.

Agenda: Questions for discussion
1. Is it our responsibility to provide online forms to enable libraries/archives to do cataloging for their own institutions?
2. Are we a portal, i.e., a directory, or are we a digital depository? What would it take to become a central metadata depository?
3. What is the search platform? What is a good search experience?
4. Can we bore into a finding aid? That is, can we provide access to specifics below the collection level? What content do we want to make searchable through the portal?
5. What would not show up in a Google search?
6. Can we eliminate what is not pertinent?

Note: the questions were all addressed, however, not necessarily in the order posed, and discussion ranged across related questions.

1. Eric's to-do list includes a June deadline for creation of an online form that outputs an EAD record. His understanding is that the form will a) gather necessary information, which will be b) formatted as an EAD document, which can be c) ingested into the portal, d) indexed, and e) displayed. The ideal configuration would be to have one form only with one underlying format, probably EAD. Data could be crosswalked to other formats as needed, for example, Dublin Core. The form is not envisioned to be a tool for editing existing EAD documents. In Eric's assessment, online creation of records is easy, however, online editing of existing documents is more difficult.

A question was raised about making things "visible to Google" vs. putting data into a database where it would not be visible. A followup question was asked, "What is our goal?" Is it to expose information about collections to search engines? Capabilities to expose data in multiple ways, e.g., to search engines, to OAI harvesters, and others seems ideal. Another question was posed about what the portal does to add value, i.e., what will it do that Google won't do? Making collection information accessible to search engines does not guarantee a place at the top of a hit list.

2. Currently, the CRRA utility is a "portal with potential." It is currently an indexer rather than a digital repository. The goal is to be a repository, so that searchers would not just find a description of a chalice (indexer) but would be able to seamlessly access the digitized image of the chalice, which would be stored centrally (repository). It is likely that searchers will increasingly ask, "What can I do with the digital object?" It was

suggested that some future enhancements might be demonstrated with existing objects, i.e. build a trial form of a fully functioning repository for demonstration purposes.

3. The search platform has the following components and capabilities:
- It is an index
- All words in finding aids are extracted to lists
- All words are associated with their host documents
- Words have pointers to the exact places they are found, allowing for specific or fielded searching
- Search engine looks at lists
- Search results are displayed to searcher, with some context
- SRU interface in front of the indexing and searching software allows a change to underlying software without a corresponding change to the user interface

Steering Committee has suggested a preference for a Google search experience, meaning the default search screen is simple with the option of more complex searching in the background. In addition, the Google search experience presents the user with his search term(s) highlighted.

6. To address what is "pertinent" to a user, we need to identify the target user. We assume this to be a researcher. Several methods for making searching more specific were discussed:
- Drop down menus
- Guided searches (Eric thinks these are not often used and that people prefer a single search box where they will enter few search terms)
Intermediate examination of searchers' queries would allow developers to embed search expertise into system, e.g.:
- Ranking of search results with phrase matches being favored in results lists
- Threshold limits that generate feedback to users when exceeded (too many hits generates an option to limit) or unreached (too few hits generates suggestions for other search options)
- Links to thesauri

The current search interface supports search syntax.

4. The question of boring into a finding aid depends on the depth of the metadata provided. Institutions must provide the fullest possible description. The online form will likely not generate exhaustive EAD files.

There followed a discussion about the guidelines, whose creation Marta has overseen. The consensus was that we are ready to put the guidelines out, in their current form, and begin to get feedback. We know we will need to flesh out the guidelines for subject terms when the online form is completed.