Facet Forward: Faceted Navigation of Federated Search Results for Cultural Heritage Materials

Mark E. Phillips
LITA National Forum 2007
Denver, Co
OAI-PMH

data-provider - (institutions)
service-provider – (UNT Libraries')

Harvester
OAI-PMH Repository

6 verbs
  GetRecord
  Identify
  ListIdentifiers
  ListMetadataFormats
  ListRecords
  ListSets
Getting Metadata Records

Harvesters:
Many available at openarchives.org

Customized version of OCLC's one page harvester.
Indexing OAI Records

Python conversion scripts
Convert from harvested temp format to Solr format
Send to Solr for indexing

Solr Index
  Dublin Core based document structure
  MD5 hash of OAI Identifier as doc id
  Fulltext and several sorting fields
mod_python based SRU access
CQL conversion to Lucene query was challenging
Less than 1 second from 10 simultaneous requests.
Challenges

Thumbnail support
Simultaneous SRU requests.
Getting repositories to support OAI-PMH
Customization of OAI-PMH output