The Mellon Metadata Harvesting Initiative

Major Findings from Participating Projects

DLF Fall 2002 Forum
Seattle, Washington
Tuesday 5-Nov-2002

Session Agenda

- Overview of the Mellon Metadata Harvesting Initiative (brief)
- Reports from Projects (~ 50 min)
- Discussion (~ 30 min)

Notes:

(Session will not review OAI-PMH)

(Goal of discussion)

The Mellon Metadata Harvesting Initiative

- In 2000 the Andrew W. Mellon Foundation sought to explore how the OAI-PMH could be used by libraries and other repositories of scholarly information to make metadata about scholarly collections more visible/useful to Internet users
- In conjunction with the DLF, Mellon hosted a series of planning meetings and eventually awarded 7 grants to fund exploratory projects
- Aimed at ***"uncovering the hidden web"***

The Seven Institutions

1. Emory University (MetaArchive.Org)
2. SOLINET / ASERL (AmericanSouth.Org)
3. The Research Libraries Group
4. The University of Michigan (OAIster)
5. University of Virginia
6. University of Illinois at Urbana-Champaign
7. (Woodrow Wilson International Center for Scholars at the Smithsonian)

The MetaScholar Projects

- The MetaArchive and AmericanSouth projects merged to form the ***MetaScholar Initiative***
- $600K two-year effort (now at mid-point):
  - Creating an extended metadata aggregation network encompassing some two dozen academic libraries, archives, and museums
  - Building subject portals / online communities developed by scholars, archivists, and librarians to organize access to the metadata aggregated

Partner Institutions

MetaArchive

- Univ. Georgia
- Emory Univ.
- Washington & Lee
- Southwestern
- Univ. Richmond
- Sewanee
- Davidson College
- Atlanta History Center
- (others)

AmericanSouth

- Univ. Georgia
- Emory Univ.
- Univ. Tennessee
- UNC Chapel Hill
- LSU
- Auburn Univ.
- Univ. Florida
- Univ. Kentucky / KVL
- Vanderbilt

Research Areas of the MetaScholar Initiative

- Metadata Aggregation Networks
  - OAI-PMH programming services
  - Open source tools and systems
- Online Communities
  - Crafting and organizing coherent subject domains of related resources
  - Evaluation of new tools for scholarly discourse / content creation built on top of metadata services

MetaScholar Highlights

- Open sourced and re-implemented Old-Dominion ARC and Virginia Tech ODL software environments
- Established extended OAI networks and central harvester/portal infrastructure
- Scholarly Design Team studying new forms of online scholarly discourse
- SPWG studying metadata organization

Metadata Gardening

Some broad observations and initial findings from the MetaScholar Projects regarding the OAI-PMH and what it enables

Wave of the Future

- Metadata is the foundation of online communities
- Roles for libraries and archives
- Need to understand new services based on traditional strengths

Deciding what to Cultivate

- Case studies & focus groups to identify needs & formulate subject domains
- Align with your organizational mission and priorities
- Assemble resources

Planting Seeds

- Fostering adoption of the PMH
- Local / regional support infrastructures
- Working alliances with other institutions

Working the Garden

- Policies (minimal elements, authority control?)
- Key tools (providers, harvesters, transformers)
- Metadata format collisions (ex. UDC vs. EAD)

Harvesting

- Which formats
- Scheduling
- Filtering mechanisms for out of scope metadata
- Possibility of combining with other kinds of metadata feeds

Welcoming Others

- Raw metadata must be organized and contextualized
- Directly engage the online community of users in commentary and analysis

Enhancing the Garden

- Develop the virtual collection through ongoing additions to ensure that the metadata base continues to grow
- Engage community in this activity

Questions and Answers

- Contact:
  - Martin Halbert
  - Director for Library Systems, Emory University
  - 404-727-2204
- Websites for more information:
  - http://MetaScholar.Org
  - http://MetaArchive.Org
  - http://AmericanSouth.Org