The Case for OAI

This handout reviews (1) the background and development of the OAI protocol, (2) past and present motivations for its use, (3) the benefits of increasing the quantity and quality of shared metadata through OAI implementation, and (4) current OAI-based discovery systems.

Responding to scholarly communities' demands for greater access to pre-publication articles and other scholarly resources, the Open Archives Initiative (OAI) developed an interoperable, scalable protocol for sharing metadata: the OAI-PMH, or Open Archives Initiative Protocol for Metadata Harvesting, also simply referred to as the OAI protocol. Scholars' continuing desire for greater access to or awareness of institutional collection holdings, combined with their increased use of online search engines which cannot locate these materials, points to a greater need for institutions to make their holdings more easily searchable. As a scalable, interoperable format, the OAI-PMH offers a cost-effective and low-barrier means for institutional metadata sharing. On-going projects using the OAI protocol demonstrate the benefits to researchers and institutions and the areas for further improvement.

SEE ALSO: **Project Abstract** (more in-depth background on this IMLS-funded OAI project).

Development of the OAI Protocol

Early demands for electronic repositories of scholarly information came from the sciences, where scholars desired faster means for sharing and accessing pre-print publications. The Santa Fe Convention of the Open Archives Initiative established initial interoperability agreements for sharing these publications across institutions through its development of the OAI-PMH, a protocol for harvesting metadata. As the OAI protocol was being successfully implemented in the scholarly scientific community, libraries were investigating how they might expand their own services into this realm. In 2000, the Andrew W. Mellon Foundation awarded seven grants funding efforts to develop metadata harvesting services in digital libraries; OAI was selected as a protocol that could be easily adapted for this project.

Need & Demand

Current popular search engines still cannot locate resources within searchable databases (like library catalogs)—an area of the web that contains documents and resources highly valuable to learning communities. As online searching continues to be a first stop for researchers (and sometimes the only stop), resource-sharing institutions must find new ways to make their collections visible and accessible. At the same time, the reality of software and hardware costs and the potential time, effort, and personnel involved in reformatting metadata discourages many institutions from implementing new digital-library infrastructures. Efforts towards implementing new systems must consider the long-term value as well as immediate demands: what offers the greatest long-term use with minimum expenditures, and can also be easily integrated with existing systems?

Benefits

The OAI protocol works with multiple systems and users, and its use of Dublin Core (DC) establishes a low-barrier means for metadata sharing. Many solutions currently exist for implementing OAI on top of current digital-library infrastructures, often through the use of open-source software available online for free download. OAI's interoperability makes it a viable alternative for institutions seeking to quickly and inexpensively build their collection online, and helps ensure that learning communities, libraries, and other resource-sharing institutions can access collection metadata. Exposing richer metadata formats in addition to DC further increases the usability of metadata and enhances awareness of collections.

Examples

The following projects demonstrate some of the services and benefits that the OAI protocol makes possible. These are just a sampling of the OAI-based discovery systems currently available. The Experimental OAI Registry at UIUC (http://gita.grainger.uiuc.edu/registry/) and the Open Archives Initiative's community registry (http://www.openarchives.org/community/index.html) offer more complete lists of OAI projects.

- **_AmericanSouth_**   **http://americansouth.org/**
  AmericanSouth, a project of Emory University with initial funding from The Andrew W. Mellon Foundation, is a topically focused digital library whose inter-institutional collection of resources all relate to Southern culture and history. AmericanSouth currently uses the PKP Open Archives Harvester for gathering, searching, indexing, and browsing metadata from these collections.

- **_CIC Metadata Portal_**   **http://cicharvest.grainger.uiuc.edu/index.asp**
  This UIUC project allows researchers to locate and view the digitized resources available through the twelve member institutions of the CIC (Committee on Institutional Cooperation). Keyword and focused searching is available to all users; geographic browses are currently limited to registered users. Results may be sorted by collection, institution, type, and by materials available online.

- **_NSDL (National Science Digital Library)_**   **http://www.nsdl.org**
  This digital library for science education, initiated by the National Science Foundation, collects resources and services directed primarily towards undergraduate populations but also serving learners of all ages. It represents "a complex network of libraries within libraries" (Brogan, 2003), with the NSDL search engine and directory page as main entry points.

- **_OAIster_**   **http://www.oaister.org**
  With funding from the Andrew W. Mellon Foundation, The University of Michigan developed this OAI search engine to uncover digital resources in the "hidden web." OAIster locates OAI-exposed metadata and web sites, but retrieves only those accompanied by the digital resources they describe, thus providing learning communities with a "no dead ends" or "one-stop shopping" approach to research.

REFERENCES

Brogan, Martha L. 2003. "A Survey of Digital Library Aggregation Services." Washington, DC: The Digital Library Federation, Council on Library and Information Resources. (Internal report. 2005 report forthcoming)

Digital Library Federation. 2004, December 28. "DLF Evaluation of the Open Archives Initiative." http://www.diglib.org/architectures/testbed.htm.

Hagedorn, Kat. 2003. "OAIster: A 'no dead ends' OAI service provider." *Library Hi Tech 21*(2), 170-171.

Lagoze, Carl, & Herbert Van de Sompel. 2001. "The Open Archives Initiative: Building a low-barrier interoperability framework." Proceedings of the ACM/IEEE Joint Conference on Digital Libraries, Roanoke VA, June 24-28, 2001, pp. 54-62. http://www.openarchives.org/documents/jcdl2001-oai.pdf.

Lagoze, Carl, William Arms, Stoney Gan, Diane Hillmann, Christopher Ingram, Dean Krafft, Richard Marisa, Jon Phipps, John Saylor, Carol Terrizzi, Walter Hoehn, David Millman, James Allan, Sergio Guzman-Lara, & Tom Kalt. 2002. "Core Services in the Architecture of the National Digital Library for Science Education (NSDL)." http://arxiv.org/abs/cs.DL/0201025.

MetaScholar Initiative. (2004, March 1). *AmericanSouth Project Final Report: A Collaborative Project to Improve Access to Digital Resources on Southern History and Culture, January 2002-January 2004.* http://www.metascholar.org/pdfs/AmSouth-FINAL-REPORT.pdf

MetaScholar Initiative. (2004, March 1). *MetaArchive Project Final Report: A Project for the Development and Evaluation of an Open-Source OAI-Compliant Infrastructure for Surfacing Hidden Scholarly Resources.* http://www.metascholar.org/pdfs/MetaArchive-FINAL-REPORT.pdf

Shreeves, Sarah L., Joanne S. Kaczmarek, & Timothy W. Cole. 2003. "Harvesting cultural heritage metadata using the OAI protocol." *Library Hi Tech 21*(2), 159-169.

Sompel, Herbert Van de, & Carl Lagoze. 2000, February. "The Santa Fe Convention of the Open Archives Initiative." *D-Lib Magazine 6*(2). http://www.dlib.org/dlib/february00/vandesompel-oai/02vandesompel-oai.html.