

## U. S. Government Use of the OAI-PMH

Michael L. Nelson  
Old Dominion University  
Department of Computer Science  
Norfolk VA 23529  
mln@cs.odu.edu

The Open Archives Initiative (OAI) is an international consortium focused on furthering the interoperability of digital libraries (DLs) through the use of "metadata harvesting". The metadata harvesting approach is defined in the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) (Lagoze, Van de Sompel, Nelson & Warner, 2002). Metadata harvesting is in contrast to the "distributed searching" approach of many previous DL interoperability projects for federating different DLs into a single service. While feasible for small numbers of nodes (e.g., < 20), large-scale distributed searching has proven difficult in an Internet environment for large numbers of nodes (e.g., > 100) (Nelson, 2001).

Key to the philosophy of the OAI is the separation of responsibilities into "service providers" and "data providers". In practice, a DL act as both a data provider and a service provider, so it is important to understand the distinction. A data provider is simply a repository, a collection of metadata records. A service provider offers value added services (e.g., searching, browsing) on the metadata extracted from one or more data providers. More detailed reviews of the OAI-PMH, its origins, version history and applications can be found in (Van de Sompel & Lagoze, 2000; Lagoze & Van de Sompel, 2001; Van de Sompel & Lagoze, 2002; Lynch, 2001).

Despite its "eprint" origins, the United States Government has been heavily involved with the OAI for Scientific and Technical Information (STI) purposes. After a brief review of the OAI-PMH and some of the associated issues, we will examine the development and status of service providers and data providers from NASA, the Department of Energy and the Department of Defense. We will also relate these efforts to similar efforts taking place in the U.K., such as the MAGiC project (Sidwell, Needham & Harrington, 2000). We will also introduce the emerging role of OAI-PMH aggregators (service providers that re-expose harvested metadata from data providers) and the issue of community building in the OAI.

Lagoze, C. L. & Van de Sompel, H. (2001). The Open Archives Initiative: Building a low barrier interoperability framework. Proceedings of the 1<sup>st</sup> ACM/IEEE Joint Conference on Digital Libraries, pp. 54-62. Available at <http://www.openarchives.org/documents/oai.pdf>

Lagoze, C. L., Van de Sompel, H., Nelson, M. L. & Warner, S. (2002). Open Archives Initiative Protocol for Metadata Harvesting – Version 2.0. Available at <http://www.openarchives.org/OAI/openarchivesprotocol.html>

- Lynch, C. A. (2001). Metadata harvesting and the Open Archives Initiative. ARL Bimonthly Report 217, 2001. Available at <http://www.arl.org/newsltr/217/mhp.html>
- Nelson, M. L. (2001). Better interoperability through the Open Archives Initiative, *New Review of Information Networking*, 7, 133-145.
- Sidwell, C. A., Needham, P. A. S. & Harrington, J. D. (2000). Lightening grey literature: Making the invisible visible. *New Review of Information Networking*, 6, 121-136.
- Van de Sompel, H. & Lagoze, C. (2000). The Santa Fe Convention of the Open Archives Initiative. *D-Lib Magazine*, 6(2). Available at <http://www.dlib.org/dlib/february00/vandesompel-oai/02vandesompel-oai.html>
- Van de Sompel, H. & Lagoze, C. (2002). Notes from the interoperability front: A progress report on the Open Archives Initiative. *Proceedings of the 6<sup>th</sup> European Conference on Digital Libraries*, pp. 144-157. Available at <http://www.openarchives.org/documents/ecdl-oai.pdf>