

The Distributed Library: OAI for Digital Library Aggregation

Abstract

The Digital Library Federation proposes a 2-year project to research, design, and prototype a “second generation” OAI finding system, capitalizing on the lessons learned from the first wave of OAI harvesting and using as its raw material collections drawn from across the DLF membership. Our research here builds on the digital objects, motivated scholarly users, and high-level OAI expertise that we have across our 36-member organization, and is informed by ongoing research into metadata creation and service building at UIUC, Michigan, Emory, and elsewhere, including our colleagues in the NSF’s (OAI-based) National Science Digital Library.

The Open Archives Initiative (OAI) has proven itself as a protocol that allows basic metadata records to be created at many “providers” and then gathered up by “harvesters” who use those records to create library services (www.oaister.org, for a single example). In the act of using it over several years in library settings, however, a range of issues have come to light that need research and development if OAI is going to mature into its full potential: collections as well as item records need further development, and we need richer mechanisms of creating dialog between harvesters and providers; the hurdles to adoption need careful study, particularly how to embed the very idea of creating public, harvestable metadata as a routine step in our digitizing workflows, and how to speed up the feedback loop from harvester to a community of providers such as exists in the library world, who typically respond positively to such “good practice” guidance.

The aim we have clearly in mind is to foster better teaching and scholarship through easier, more relevant discovery of digital resources, and a much greater ability for libraries to build more responsive local services on top of a distributed metadata platform. To this end, a team of scholars – alongside digital library experts -- will be assembled at the early stages of the project and their input will inform and challenge our assumptions about the nature and design of our prototype service throughout the grant period.

The DLF is an ideal community in which to prototype collection-level and item-level metadata that is explicitly designed for use in digital library aggregation services. We are large enough to give confidence that what we discover can be generally applied; we have rich intellectual resources in our librarians and a proven track record for effective collaborative work; and we have in our number some of the most active OAI harvesting service providers. We have a good sense of the barriers to providing OAI records routinely, and the training and consultancy component of our work seeks to overcome those hurdles; related work such as the recent DLF *Survey of Digital Aggregation Services* provides rich feedback on the strengths and weaknesses of the first wave of OAI services – lessons we will take to heart in this “second wave” research.