FLOW: Co-constructing Low Barrier Repository Infrastructure in Support of Heterogeneous Knowledge Collection(s)

Karen S. Baker
Scripps Institution of Oceanography
kbaker@ucsd.edu

Anna K. Gold
UCSD Libraries
agold@ucsd.edu

Frank Sudholt
San Diego Supercomputer Center
fsudholt@sdsc.edu

Abstract:

Institutional repositories are being constructed today to address the needs of scholarly communication in a digital environment [1, 2]. The success of such institutional infrastructures as knowledge collections depends in part on offering low barriers for participation and on supporting heterogeneous knowledge inputs and outputs. The San Diego Supercomputer Center (SDSC) in partnership with CERN (European Center for Nuclear Research), the Scripps Institution of Oceanography (SIO), and the University of California, San Diego (UCSD) Science & Engineering Library, has modified CERN's CDSware software to initiate the process of creating a local low barrier repository.

The SDSC prototype of CDSware is a work in progress, a modified package of CDSware's open source software for managing an institutional repository of references and documents. Such short-term/local approaches are not only compatible with long-term federation strategies but are critical to initiating information flow, contributing to knowledge diversity, and ensuring reflexivity in the design and development processes.

Given the social aspects of documents in particular and artifacts in general, building successful institutional repositories requires exploring how to create new infrastructures that place documents, people, and organizations into unified and expressive relationships [3]. CDSware as implemented at SDSC adapts an OAI-compliant document and citation management system to more fully express the relationships between documents, people, and the organizational resources to which both are related [4]. By starting with a document-oriented view and making the social relations of documents explicit (to individuals, research groups, administrative reports and organizational metrics, sources of funding and research instruments), this project seeks to expand the knowledge management role of a document repository [5]. This work also suggests that personal document collections provide a mechanism to extend informal communications *if* collection-making tools are available [6]. We seek to create a process of learning and informing involving participants by providing mechanisms that enable the work of individuals and organizations. The project name FLOW is a purposeful metaphor calling to mind how rivulets shaped by the local landscape join a river of information, contributing to heterogeneous pools of knowledge.

CDSware encompasses both the referatory concept of collecting citations and links, as well as the repository concept of archiving the digital objects cited along with local classification and relationship information about the objects. Challenges include *technical* hurdles (ongoing resource support, open design, implementation strategies), *social* barriers (the need for a critical mass of participation to supply the activation energy for the system), and *conceptual* difficulties (articulating associations between individuals, materials and organizations in order to capture complex interdependencies). Our design approach is to start with a single research program within the SDSC - with an eye to federation, both across the institution via closely-managed institutional data tables - and across a looser network of affiliated research partners. For coordination at the institution level (SDSC) an ORACLE-based management data pool was linked to CDSware. To extend this to include associates, a second tier was implemented with MySQL, used by the CERN distribution of CDSware. (Project support has included funding from SIO, SDSC (Integrative Computational Sciences), and NSF Grants DBI-01-11544 and OPP-02-17282.)

References:

[1] R. Crow, "The Case for Institutional Repositories," 2002, http://www.arl.org/sparc/IR/ir.html.

[2] C. A. Lynch, "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age," February 2003, *ARL Bimonthly Report* 226. http://www.arl.org/newsltr/226/ir.html.

[3] S. J. McMillan, "Exploring Models of Interactivity from Multiple Research Traditions: Users, Documents, and System," *The Handbook of New Media*, L. Lievrouw and S. Livingstone (eds.), p. 163-182, SAGE Publications, London (2002).

[4] A. K. Gold, K. S. Baker, K. K. Baldridge, and J. LeMeur. "Building FLOW: Federating libraries on the Web," *ACM-IEEE Joint Conference on Digital Libraries*, Portland Oregon, July 2002, http://doi.acm.org/10.1145/544220.544286.

[5] A. K. Gold, "The Roles of Documents in Knowledge Management," unpublished paper, December 18, 2002.

[6] K. S. Baker, A. K. Gold, F. Sudholt, "Collection Repository: Personal to Project to Organizational," Technical report http://pal.lternet.edu/projects/semantics (2003).