Embedding the Managed Repository in National Science Digital Library (NSDL) Semantic Library Services

Carol Minton Morris, Elly Cramer, NSDL Core Integration, Cornell University

Abstract

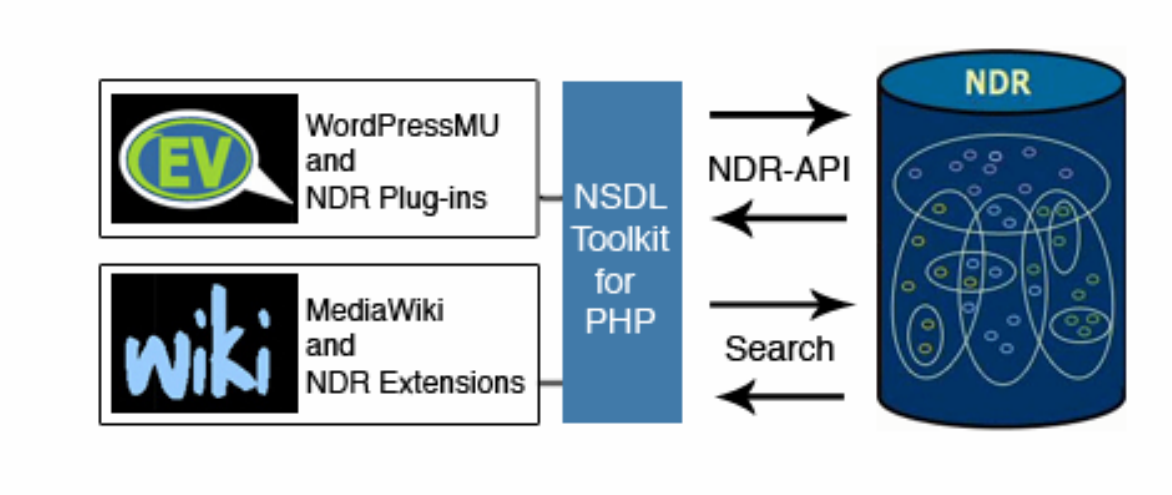
Managed institutional digital assets are fixed commodities, while serendipitously created web communications about those assets are often hard to pin down. The National Science Digital Library's (NSDL) community of users contributes two kinds of contextual communications about library resources that are often interrelated—short news and information items and mid-sized, editable articles. This paper presents one solution to the "semantic synthesis challenge," or how to extract warranted knowledge from these two types of casual context by embedding the managed repository using web application plug-ins that interact with the NSDL Data repository.

Introduction

Tim Berners-Lee and James Hendler suggest that the semantic web will change the way resources and related information about resources are mixed and matched.

> The very notion of a journal of medicine separate from a journal of bioinformatics, separate from the writings of physicists, chemists, psychologists and even kindergarten teachers, will someday become as out of date as the print journal is becoming to our graduate students.[1]

Expert Voices, NSDL's blogosphere, and NSDLWiki are collaborative services that allow users to create, annotate, and organize information that can be used by both humans and machines. Users contribute adjunct knowledge and expertise to NSDL[2] by adding information about resources and links to resources in the repository, as well as resources themselves to a collection in the managed repository through blog posts and wiki articles. The blog and wiki applications interact with the repository via NSDL Data Repository (NDR) plug-ins.



NSDL Data Repository and communication tools interaction.

The users of NSDL services run the gamut from teachers who need just-in-time resources aligned with state and national standards coupled with ideas for classroom use, to researchers engaged in collaborative scientific publishing and at-large science buffs with personal research interests. Focused and casual-use cases are supported by these open source NDR plug-ins. This paper explains how Expert Voices and NSDL Wiki embed user contributions to the repository in production processes and workflows that increase usage and participation while extending the impact of the Library's managed collections.

Expert Voices

One of the communication services, Expert Voices (http://expertvoices.nsdl.org/), is based on the WordPress Multiuser (MU) blogging system. WordPress is a popular Web open source technology that supports a "website where entries are written in chronological order and commonly displayed in reverse chronological order."[1] Many search engines are "tuned" to detect WordPress content because they scan for "fresh" content. RSS generated by this widely-used software is partially responsible for the more than 106 million blogs currently available on the Web.[3]

An Alpha version of Expert Voices was launched in the spring of 2005 with the expectation that NSDL users and collection contributors would become bloggers based on the Expert Voices content model of a discovery team approach to problem-solving. The original blogging discovery teams included recruited teachers, experts, librarians and others who shared insights and research around science, technology, engineering, and mathematics (STEM) topics in a blog. Additional Library news and information content models also emerged. The NSDL blogosphere has continued to grow as the system was released into Beta, increasing its content four fold.

Expert Voices Index	Alpha Release (4/06 – 12/06)	Beta Release (2007)
Number of blogs	17	43
Number of registered users	95	414
Number of Americans who say they keep blogs[4]	12 million	
Total number of posts	92	894
Total number of comments	47	633
Percentage of bloggers nationwide who post new material every few weeks.[4]	28%	

Visits to Expert Voices accounted for a majority of visits to NSDL.org in the spring and summer 2007. NSDL Core Integration team members Brad Edmondson, Elly Cramer, and Dean Eckstrom investigated evidence that search engines were weighting the value of RSS citations in NSDL blog communications and in turn boosting repository resource use[5].

The NSDL Expert Voices "Search Popup" operates from within the application. Blog contributors search the library and add links to resources in their posts. It's not only a convenience to provide access and insertion of the link into a post, but the process also provides a second link to the NSDL information page for that resource.

Expert Voices “NDR Publish” allows authorized users to add individual resources and associated metadata to the NDR, making it discoverable in NSDL.org's search results. The resource may be the post itself or valuable links referenced in the post that the contributor would like to see in the library.

By contributing a post into an Expert Voices blog, adding it to the library, and adding a link to a resource that is also in the library, relationships are established within the repository. These relationships will then appear in the managed repository metadata for both resources.

An example of a successful NSDL Expert Voices blog is the NSDL Middle School Portal Pathway at Ohio State University, which posts a weekly annotated news item for science teachers entitled “Connecting News with National Science Standards”[6]. This blog makes it easy to blend current events into the regular curriculum. Project staff create links from current news articles to related teaching resources that connect specific content standards to an event. Articles may be appropriate for students to read in the blog, or they may serve as background knowledge for teachers and inspiration for subsequent instructional activities. The articles aim at a grade 5-8 audience and include links to lessons, reference articles, and definitions. Currently this blog hosts the most visitors to NSDL’s blogosphere. The NSDL icons that track back to NSDL’s managed repository from the blog increase usage, and add context for teachers as they go about the daily business of synthesizing knowledge for their students.

NSDL Wiki

NSDL Wiki (http://wiki.nsdl.org) provides a collaborative online environment based on MediaWiki, another popular open source web-based software. It allows users to organize, create, and annotate resources. Vetted articles and referenced resources can then be added back to the library for search and discovery on nsdl.org with the same type of extensions that exist on Expert Voices.

> Wikis are often used to create collaborative websites, power community websites, and are increasingly being installed by businesses to provide affordable and effective Intranets or for use in Knowledge Management. Ward Cunningham, developer of the first wiki, WikiWikiWeb, originally described it as "the simplest online database that could possibly work".[7]

An interdisciplinary team from Ohio State University (OSU), College of Education and Human Ecology; the Ohio Resource Center (ORC) for Mathematics, Science, and Reading; the Byrd Polar Research Center; COSI (Center for Science and Industry) Columbus; and the National Science Digital Library (NSDL) Core Integration team at Cornell University and University Corporation for Atmospheric Research (UCAR) are using the NSDL Wiki to plan and manage the Beyond Penguins and Polar Bears Project[8] to develop and document project deliverables that include creating 20 issues of a multimedia “e-zine” with a combined focus on inquiry-based science and content-rich literacy learning based on International Polar Year research. As this project moves forward a new degree of knowledge synthesis is possible as repository citations are automatically linked to pre-publication resources and tracked back to the managed repository.

Conclusion

After working with Expert Voices for almost two years and for several months with NSDL Wiki, some workflow observations have emerged. User communities are often enthusiastic about the idea of contributing on-the-fly communications to a melting pot of evolving knowledge. The contributors can become easily discouraged with authentication and work flow processes, in spite of our attempts to decrease the barriers. If they are passionate about the topic or if their participation increases their community's web presence, they will persist and be successful at contributing blog posts and wiki articles and interacting with the library content. The reality is that jumpstarting a blogosphere or a wikidom takes editorial planning and often requires some incentives for "what's in it for me?", in addition to usable web interfaces.

RSS has become widely used[9], but we found that the non-technical users still want email notifications to stay involved and are intimidated by the concepts and terminology. We're interested in looking at how these different forms of communication (i.e. email, blogging, and wikis) affect building online communities and how these communities affect the growth and success of the library.

The next round of development will look at integrating more tagging functionality like Connotea[10], creating custom collections or aggregations; and making better use of audio and video podcasts and other multimedia. As NSDL Wiki becomes more active, we will be studying whether it receives the same attention and exposure in the search engines as the blogging software has demonstrated.

Acknowledgements

This paper is based on prior work by the Cornell NCore Team that includes Dean Krafft, Aaron Birkland, Jim Blake, and Lynette Rayle.

This material is based upon work supported by the National Science Foundation under Grants No 0227648, 0227656, and 0227888. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

1. Berners-Lee, T., Hendler, J., Scientific Publishing on the "Semantic" Web, Nature Web Debates, April 5, 2001 http://www.nature.com/nature/debates/e-access/Articles/bernerslee.htm

2. Lagoze, C. et al, Representing Contextualized Information in the NSDL, Proceedings of the Euorpean Conference on Digital Libraries, September 2006. Alicante, Spain. http://arxiv.org/abs/cs.DL/0603024

3. "About." Technorati http://technorati.com/about/

4. Lenhart, A., Fox S. "Bloggers: A Portrait of America's New Storytellers." Pew Internet and American Life Project Reports. July 19, 2006. http://www.pewinternet.org/pdfs/PIP Bloggers Report July 19 2006.pdf

5. Edmondson, B., “How to Improve Rankings With Blog Posts, Or, How to Get Cute Bots to Notice You.” NSDL Wiki. September 2007, http://wiki.nsdl.org/index.php/How_To_Improve_Rankings_With_Blog_Posts

6. http://expertvoices.nsdl.org/connectingnews/

7. http://en.wikipedia.org/wiki/Wiki

8. http://wiki.nsdl.org/index.php/Community:Polar

9. Grossnickle, J. “RSS -- Crossing into the Mainstream” 2005. http://publisher.yahoo.com/rss/RSS_whitePaper1004.pdf

10. http://www.connotea.org/