

Metadata Process Proposal

Developing Best Practices for Metadata in the NSDL Data Repository

Background:

NSDL's focus in 2007 and beyond is centered in the rich contextualization around resource-centric discovery that the Fedora-based NSDL Data Repository (NDR) enables:

- efficient, effective search and discovery of high quality STEM resources and related information
- the provision of supporting tools and services, based on and driven by attention to the needs, practices, and behaviors of its users

This focus on user-centered design and fulfilling the needs of unique communities of practice is also at the heart of NSDL's partnership building efforts with multiple stakeholders engaged in educational endeavors, and ultimately helps to promote long-term sustainability for the library. These same goals are reflected in the priorities of the Pathways, the primary partners in constructing NSDL: building and expanding collections, efficient discovery of quality resources, and better understanding and ability to meet user needs (Jesuroga 2006).¹

A number of circumstances have converged in the last half of 2006 that allow NSDL to capitalize on the experience, skills, and contributions of its community in synergistic ways:

1. NSDL will operationally transition from the current Metadata Repository (MR) to the NDR in early 2007, allowing for relationships, linkages, and behaviors around resources and services in the NDR, in addition to rich metadata expression. This transition is accompanied by ongoing work in Core Integration to
 - improve and expand capabilities of the NSDL search service
 - metadata enhancement activities by San Diego Supercomputing Center with respect to grade level for existing NSDL resources, and exploration of use of iVia to further enhance metadata
 - the need to address a recognized gap in NSDL service relating to collection contribution (accessioning)
 - the creation and documentation of best practices across all aspects of NSDL development
2. The National Science Foundation has specified in several of its solicitations that awarded projects should utilize NSDL as the recommended dissemination mechanism for project output. This means that awarded projects from the Teacher Professional Continuum (TPC), the Course, Curriculum, and Laboratory Improvement (CCLI) initiative, as well as the NSDL solicitation in 2007, will be seeking resource and/or collection accessioning support from NSDL in the future. NSDL's NSF program officer (Lee Zia) has requested that NSDL establish a web site providing information and support to potential new contributors to NSDL from these initiatives, and that this site be made available in early 2007, with guidance, recommended best practices for resource and collection contribution, and points of contact for assistance in the process. NSF will include this proposed URL in selected upcoming solicitations and direct potential grantees to it. Sample language as follows:

¹ Jesuroga, Susan. (2006). *Proposal to Core Integration and the Pathways regarding metadata*. Boulder, Colorado: National Science Digital Library (NSDL), University Corporation for Atmospheric Research (UCAR). Retrieved August 2006 from http://comm.nsd.org/download.php/873/MetadataIssues-1_final.pdf

Metadata Process Proposal

Projects that plan to use the World Wide Web as a component of their overall dissemination strategy should connect the project's website to the National Science Digital Library (NSDL). The proposal should describe how the web pages will be tagged with descriptive metadata (see <http://dublincore.org>) so that the material becomes part of the NSDL. The following website <http://nsdl.org/contribute> provides information and instructions for connecting the project's website.

3. NSDL is collaborating with the Digital Library for Earth System Education (DLESE) to integrate and adapt a selection of DLESE tools and services into the NDR. This work is paving the way for integration of Pathways resources, tools, and services into the NDR, enabling adaptation to NSDL and wider distribution of education standards-supporting tools like Syracuse's Content Assignment Tool (CAT) along with the Achievement Standards Network (ASN) database of state educational standards, a standards-supporting cataloging system and associated services, and concept-mapping capabilities.
4. Additional collaborations in the community are supporting NSDL adaptation and customization of the CWIS cataloging system, and wider adoption of iVia's metadata generation capability.
5. At the October 2006 Pathways meeting in Washington DC, there was substantial discussion about ways to address standardization within NSDL and Pathways for a basic set of NSDL Dublin Core metadata elements, and the development of recommended best practices for same. Formation of a small, nimble working group of representatives from Core Integration and Pathways was suggested as a way to work productively and iteratively to achieve agreement on metadata issues, based on the recommendations in the Jesuroga 2006 white paper (distributed at the August 2006 Pathways meeting; see footnote reference and link).
6. The NSDL-funded Metadata Registry project (Sutton and Hillman, PIs) is maturing and will be ramping up to register metadata schemas and controlled vocabularies within the community in the coming year.

With this background in mind, Core Integration has identified a highly qualified individual to facilitate metadata issues. NSDL's collaboration with DLESE is enabling Core Integration to utilize the expertise of Katy Ginger, to accomplish the following:

- in the short-term, assist in the development of collection accessioning best practices and support for potential new projects (NSF request)
- in the long-term, help define and provide core documentation of NSDL best practices in multiple areas
- facilitate and support the transition to the anticipated hiring of an Editor-in-Chief and additional metadata specialist, to be located at Columbia University Core Integration, in 2007

Qualifications: Katy has been the collections integrator, metadata architect, and controlled vocabulary specialist for DLESE for the past six years. She has participated in the NSDL controlled vocabulary working group and workshops, and worked with the Dublin Core Education Working Group in developing the metadata application profile for the education domain. She will be working with the Metadata Registry project as well in the upcoming year. She has domain expertise in science, as she is also an atmospheric scientist. Katy has experience with numerous metadata schemas, including Dublin Core, IMS, and ADN

Metadata Process Proposal

(DLESE native metadata schema, adapted from IMS. The ADN schema for geoscience metadata was developed jointly by the Alexandria Digital Library, DLESE, and NASA).

Draft - Process Proposal: Develop Best Practices to address the following:

- Suggested required metadata for an initial set of metadata fields: ed level, audience, subject, resource type, title, URL, description
- Controlled vocabulary support for metadata fields
- Outline and make recommendations for formalizing accessioning and deaccessioning processes/policies
- Develop transition plans for Pathways to adopt and use these Best Practices

Benefits: Best Practices will

- Improve metadata quality, library search/discovery and the display of resource information in the user interface
- Improve the development and provision of library services that depend on consistent and reliable resource information (e.g. metadata)
- Increase the selection of and reuse of educational content across NSDL and all Pathways
- Describe how to contribute and manage content in the library with accessioning and deaccessioning policies
- Address NSF concerns of how to contribute to NSDL for new projects in forthcoming NSF solicitations
- Establish the groundwork so that greater educational contexts around resources can flourish. That is, after a resource is described on a basic level, it is easier to gather information on how to use and teach with a resource and what educational standards may be associated with the resource

How: A small working group of not more than five people develops these Best Practices, makes recommendations, and assists Pathways with transition plans to adopt and use them. With iterative Pathway review and agreement, a desired goal is that these Best Practices

- become formal policies
- have XML support
- are adopted as NSDL expectation and practice

Essentially, the Best Practices and the small working group serve to move NSDL towards this goal in a timely manner and by developing the necessary materials and processes. It is projected that by March 1, 2007 the working group will: define communication processes, define review and approval processes by Pathways projects and key personnel, and develop an initial set of recommendations for education level and determine and recommend the next priority metadata fields for attention. The expectation is that PW PIs or designated PW representatives will contribute to the working group as needed, to resolve unique issues that might arise for individual Pathways, after Pathways have had the opportunity to review the recommendations emergent from the working group. At the same time, or shortly thereafter, some Pathways are involved in adopting and using the emerging Best Practices, reporting back, and resulting in iterative development. The expectation for involvement level of participants in the small group is that initial weekly phone calls will be the norm for a period of time, most likely moving to biweekly status later in the process. Beyond the conference calls, it is anticipated that several hours a week may be required by working group members for reviewing and iteration of best practices recommendation for PW review.

Skills required: In order to facilitate the development of these Best Practices, the following skills are necessary within the small working group:

- Expert-level understanding and knowledge in metadata issues, cataloging, and controlled vocabularies

Metadata Process Proposal

- Participation in large and small scale collection building efforts from both a technical and social perspective
- Experience in developing accessioning/deaccessioning policies and best practices as applied to digital libraries
- Working knowledge of XML

Who:

Katy Ginger will head up the working group. CI is in the process of identifying 3-4 additional individuals with the appropriate background and experience and will advise everyone within a week to 10 days.

Contact and CI Lead: Katy Ginger (ginger@ucar.edu)

Further information, updating, and documentation of the proposed process will take place on the Medata Management wiki (NSDL Community portal):

http://metamanagement.comm.nsdlib.org/cgi-bin/wiki.pl?Metadata_Management