National Science* Digital Library:

NSF, Community, Cornell & Collection Development

John M. Saylor
Cornell University
jms1@cornell.edu
8/29/01

National Science Foundation

"The National Science Foundation (NSF), established in 1950, is the federal government's only agency dedicated to the support of education and fundamental research in all scientific and engineering disciplines. Its purpose is to ensure that the United States maintains leadership in discovery, learning and innovation across science, mathematics and engineering."

-http://www.nsf.gov/od/lpa/nsf50/

National Science Foundation

Important DL Infrastructure Accomplishments:

□□·□□1986 NSFnet created

·1991 NSFNET allows for-profit traffic and thus becomes the global internet we know today

·1993 Research supported by NSF created the "killer application" credited with sparking the Internet explosion.

NSF's Science Digital Library Program

- *1996* Vision articulated by NSF's DUE
  - http://www.nsf.gov/cgi-bin/getpub?nsf9882
- *1997* NRC workshop
  - http://www.nap.edu/readingroom/records/0309059771.html
- *1998* Preliminary grants through DLI- 2
  - http://www.ehr.nsf.gov/ehr/due/programs/nsdl/projects.asp

NSF's Science Digital Library Program

- *1998* **SMETE-Lib workshop**
  - **http://www.dlib.org/smete/public/report.html**
- *1999* **NSDL Solicitation**
  - **http://www.nsf.gov/pubs/2000/nsf0044/nsf0044.htm**
- *2000* **6 Core Integration System projects + 23 others funded**

Zia, Lee. "The NSF National Science, Mathematics, Engineering, and Technology Education Digital Library NSDL Program: A Progress Report."
http://www.dlib.org/dlib/october00/zia/10zia.html

NSF's Science Digital Library Program

- *2001 Proposal* - 1 very large Core Integration System project to be funded
  - http://www.ehr.nsf.gov/ehr/due/programs/nsdl/

- Fall 2002 - NSDL goes live with a "comprehensive" collection and services.

NSF's Science Digital Library Goals

- Stimulate and sustain continual improvements in the quality of *Science*, Mathematics, Engineering, and Technology *Education.*

- *Comprehensive* library of the digital resources and services that are available for a broad audience of *learners* (students, instructors, researchers, etc) in both formal and informal settings for *K to higher education and the life long curious.*

Criteria to meet Goals

- Breadth and depth of coverage
- *Useability*
- Scientific and educational quality

NSDL 2000 Core Integration (CI) System Projects

- University of California at Berkeley
  - Materials for education in engineering and other disciplines
- Cornell University
  - Comprehensive scientific collections and services
- University of Missouri – Columbia
  - National biology digital library
- University Center for Atmospheric Research
  - Federated NSDL for earth system education -DLESE
- Eastern Michigan University
  - Digital community and collections for teacher education
- Columbia University
  - Columbia Pubscape – NSDL Publishing Center

NSDL 2000 Collections Projects

- Biology Education Online – An Interactive Electronic Journal
- Bioscience Education Net
- Digital Library for Earth Systems Education
- Atmospheric Visualization Collection
- MATHDL – Online Learning Materials in Mathematics
- A Digital Library Network for Engineering and Technology
- Mathematics, Science, and Technology Teacher Preparation
- Geoscience (Solid Earth) Data Sets
- The Alsos Digital Library

NSDL 2000 Services Projects

- Prioritizing Content Creation in Digital Libraries
- Peer Review of Digital Learning Materials
- E-Journal of Earth System Science Education Resources
- Breaking the Metadata Generation Bottleneck
- Discovering, Recommending, and Combining Learning Objects
- Information Pathways through NSDL Video
- Component Repository & Environment for Teaching Environments

NSDL 2000 Working Groups (est. 9/00)

- Coordinating Committee
- Collections Development
- Community Development, Pedagogy, and Education
- Evaluation and User Studies
- Shared Governance and Organization
- Research Issues
- Services and Tools
- Standards and Metadata
- Intellectual Property and Sustainability
- Technical Infrastructure and Interoperability

- http://www.smete.org/nsdl/workgroups/index.html

FOR MORE INFO...

NSDL as a Collaborative Venture

- How to ensure that the projects become partners, not competitors?
- Build connections to partners that are not NSF funded - e.g.IMLS, ACS, etc.
- Professional society involvement needed.

NSDL Infrastructure Challenge

- the key is an infrastructure that will support a library with the widest possible scope coupled with initial collections and services that will have a more focused and manageable immediate impact.

CI Infrastucture Goals

- Collection Building
  - Help contributors integrate their resources into NSDL with selectable levels of interoperability
- Services
  - Build coherent services for users that unify resources provided by the individual collections
- Portals
  - to meet requirements of different users
- Science* education quality
  - Embodied in structure of the library

CI Categories of Effort

Objectives	Engaging the Community	Providing Technology	Operating Core Services
Diverse Collections	Seek out Collections Integrate Collections Consult on Metadata	Cataloging Tools Middleware	Collection Registry Metadata Repository
User Interfaces & Library Services	Integrate Service Projects Consult on Portals	Customized Portals Protocol Library Client Toolkit	Primary Portal Specialized Portals Information Discovery
Educational Excellence	Outreach to Disciplines Workshops and Training Consult on Evaluation	Evaluation Toolkit Tools for Collaboration	Community Portal Usage Statistics
Access & Content Management	Dialogue w' Community Models for Access Consult on IP	Authentication Toolkit	User Database Authentication
Framework for Evolution	Plan for Sustainability Develop Business Model Consult on Software	Open Interfaces Open Implementations	Reference Architecture

Cornell NSDL Project Team

William Arms	Overall coordination
Elly Cramer	Programming
Donna Bergmark	Programming
Diane Hillmann	Metadata Goddess
Dean Krafft	Systems
Carl Lagoze	Architecture
Rich Marisa	Architecture and engineering
John Saylor	Collection development
Carol Terrizzi	Design and communications
Sarah Thomas	Cornell University Library

Cornell's First Year Contributions

- Work with the NSF and other NSDL awardees to establish an *organizational framework* for the NSDL.
- Coordinate the task of identifying digital library *collections and services* in science, mathematics and engineering, particularly those that are most suitable for education. This includes working with the NSDL partners, and recruiting and integrating materials from independent collections.
- Provide a *prototype* for the underlying *systems technology* for the NSDL. This will be based on the Dienst architecture (developed at Cornell) and the Open Archives Initiative, (http://www.openarchives.org/) which uses the Dienst protocol.

Cornell's First Year Contributions

- Develop *guidelines* and *standards* for the collections and services in the NSDL. These include metadata standards for information discovery, quality standards for the NSDL and the process standards for achieving them.
- Develop and integrate an initial set of *services* for users of the NSDL, in particular a metadata repository that will index and provide information discovery and access to all the collections and services.
- Develop *prototype portals*, connecting a substantial number of NSDL partners and independent collections, tailored for targeted groups of users.

A technology Model for a National Science Digital Library that will exist exclusively to serve the public good.

Components

-A tightly coupled federation of autonomous digital libraries, interacting via agreed standards to function as a single coherent library.

-A loosely coupled network of collections and digital libraries, combined with (independent) services that mediate among their differences and provide certain common capabilities, such as for system-wide user identification.

-A large aggregation of collections with a central metadata repository on which all information discovery services are built.

Features

-Library sections of interest to different kinds of visitors
-Easy navigation to essential services
-Exhibits of interest
-Collection tours
-Tools and services

-"My SITE," a portal personalization environment that guides users through creating, storing, retrieving, and editing up to four channels of information.
-"Add SITE to your Site," the first of a set of building tools that will give users the ability to add SITE design elements and active components to their own portals

NSDL - Collection Development

- ***The NSDL must have a very broad and comprehensive collection development policy (ie scope statement).***
  - **describes the types of resources that are suitable for inclusion in the library**
  - short and easy to understand
  - easy to apply quickly & consistently
- ***Examples:***
  - http://www.dlese.org/steering_committee/policy/collection_policy.html
  - http://www.geminfo.org/Participation/gov.html
  - http://elib.cs.berkeley.edu/calflora/about-database.html

NSDL Collections WG 11/2/00

- **Scope of Collection**
  - Any type of resource that can be used or re-purposed to facilitate learning in science, mathematics, engineering and technology.
- **Target Users and Providers**
  - Science, mathematics, engineering and technology educators and learners (which includes the research community) and content and service creators.

NSDL- Collection & Service Criteria-v1.0

- Relevance to Science* Education
- Conforms to NSDL Intellectual Property Policy *
- Basic Integrity criteria
  - (a) no blatant errors of fact,
  - (b) no blatant political, religious, or commercial message,
  - (c) it works; no blatant technical failures if the resource is digital

NSDL- Collecting the Collections

- Collections of digital materials that meet these broad filters above are being acquired by several means including:
- 1. Federated Collections - services will be built to certain specifications (which are often selected from formal standards). e.g. DLESE and NEEDS
- 2. Harvested Collections - Each digital library makes minimal metadata about its collections available in a simple exchange format. E.g. Open Archives Material
- 3. Gathered Collections - minimal or no metadata- material is selectively gathered automatically
- 4. User Contributed Collections - Collections contributed by the community of users that meet the basic filters mentioned above.

Challenges

- How do we apply criteria to the collection?
  - -authorized body or the resource creator?
- Creating incentives for independent digital libraries to adopt interoperability agreements.
- What effective services can we build with such minimal metadata as OAI?
- What will make the NSDL Library a permanent part of the educational landscape?

Collecting the Collections

- NSDL Collections first year strategy
  - Collaborate with
    - –NSDL funded collections (10 total)
    - –Cornell Collections
    - –Local School Science Teachers - Ithaca Science Zone
    - –Other science collections with interesting materials
  - Gather other interesting sites- ie "scrape" metadata

Collecting the Collections

- NSF Funded NSDL Collections partners and Collaborators
  - Merlot (Multimedia Educational Resource for Learning and Online Teaching)
  - Eisenhower National Clearinghouse
  - PubScience
  - Internet Scout Project (Science and Engineering)
  - CIESE (Center for Improved Engineering & Science Education)
  - NEEDS

Collecting the Collections

- Cornell partners
  - Building the Digital Earth (Atlas)
  - Plant Pathology
  - Paleontological Research Institute
  - Plantations
  - Laboratory of Ornithology
  - CUL - Math Books,
  - NCSTRL
  - Materials Science—'Ask a Scientist'
  - Cornell News Service

Collecting the Collections

- What's there?—collections from NSDL grantees who have materials with metadata, Cornell partners, some gathered materials

- What isn't?—the rest of the net—this will take time and community building

  Example: DLESE

- Read the Whitepaper (3/20/01):

http://www.smete.org/nsdl/workgroups/coordcomm/Whitepaper.doc

Collecting the Collections

Impediments to collaboration

- Lack of any metadata-
  - "We are pleased with the technical side…of the database and web access…but we are complete novices in terms of how to make our collection part of the digital library. I assume this hinges on appropriate metadata, but I am not sure exactly what kinds…"
- Newness of the Open Archives Initiative (http://www.openarchives.org)
- Proprietary and intellectual property concerns-
  - "Our collection is subscription based and so users who are not coming from subscribing institutions would not be able to access the content below the first screen. "
- Collection builders/providers are very busy & understaffed

How Can You Contribute?

- Join/observe an NSDL Working Group
  - http://www.smete.org/nsdl/workgroups/index.html
- Suggest collections for inclusion
  - http://www.siteforscience.org
- Report Bugs
  - http://nsdlib.nsdl.cornell.edu/cgi-bin/nsdl-bugs
- This presentation at
  - http://reuleaux.cornell.edu/nsdl/NSDLACScur.ppt