NSDL Technical Network Services:
A Cyberinfrastructure Platform for STEM Education

Since 2001, the project team at Cornell University has been developing, implementing, and supporting the central tools and services of the National Science Digital Library (NSDL). Cornell and its partners have a proven record of delivering reliable, production NSDL services; providing innovative educational solutions to the science, technology, engineering and mathematics (STEM) community; and supporting a wide range of NSF grantees and outside partners to deliver valuable educational tools and services to their own communities. This project will:

- Engage the NSDL and NSF STEM education communities to develop a common vision for NSDL and its technical infrastructure and foster broad-based community contributions
- Enable the library as a shared, collaborative, contributory space supporting the creation and display of context around library resources to enhance discovery, use, and understanding
- Support and extend the NSDL Core infrastructure (NCore) as a cyberinfrastructure platform for implementing and developing new educational applications and digital library tools for audiences across the full range of STEM research, learning, and education
- Ensure that the library and its accompanying tools and services are a robust, reliable, and highly available resource for all its users
- Create a business model that can support this effort well beyond the term of this grant

Intellectual Merit

The intellectual merit of this effort lies in technical and organizational contributions to innovative educational cyberinfrastructure. NCore will provide a robust, extensible, state-of-the-art cyberinfrastructure platform for STEM education and beyond. It offers a unique and potentially transformative opportunity to integrate a wide range of tools, services, content, and context from many different institutions and projects into a common framework of knowledge organization and discovery, enabling other projects to accomplish far more than they could have on their own. The project's work on integrating resources, description, context and contribution from a huge range of organizations, projects, and individuals, and on building tools to explore and understand this knowledge space, could transform digital libraries and processes of STEM education. The TNSP organizational model with its embedded formal community engagement process will contribute to understanding scalable social processes for fostering community innovation and contributions to evolving cyberinfrastructure. Through business and sustainability planning, new models for sustainable cyberinfrastructure will be investigated and defined.

Broader Impacts

This project will significantly enhance the nation's infrastructure for research and education, enabling STEM scientists and educators to develop and disseminate digital learning resources and learning applications to new and larger audiences of science learners, including K-16 students, pre- and in-service science teachers, and higher education faculty. It will establish a broad-based community engagement process to foster creativity, innovation, and contributions to the design and use of next-generation educational cyberinfrastructure and interactive learning applications. It will provide technical tools and services to a large number of NSDL and other NSF grantees, helping them to better serve their specific audiences while focusing less on recreating technical developments and more on serving as sites for teaching, mentoring, and supporting research of large numbers of STEM students and teachers. Finally, it will stimulate new forms of cyberinfrastructure-based education and learning science research by instrumenting the NSDL infrastructure to collect detailed (anonymized) usage and behavioral data, providing the basis for next-generation research into personalized learning, educationally-informed recommendation engines, educational social networks formation and evolution, and other areas where large volumes of user-generated actions and activities are needed.

NSDL Technical Network Services:
A Cyberinfrastructure Platform for STEM Education

Since 2001 [11, 45], the project team at Cornell University has been developing, implementing, and supporting the central tools and services of the National Science Digital Library [79]. Cornell has a proven record of delivering reliable, production NSDL services; providing innovative educational solutions to the STEM community; and supporting a wide range of NSF grantees and outside partners in delivering highly valuable educational tools and services to their communities. Together with our project partners – Digital Learning Sciences and Columbia University – we are uniquely positioned in terms of know-how, experience and desire to fulfill the role of NSDL Technical Network Services Provider (TNSP), delivering the best possible tools, services and support to NSDL grantees and the STEM education communities served by NSDL. As the TNSP, we will pursue a clear set of goals:

- Provide the best possible educational cyberinfrastructure and support for communities across the full range of science, technology, engineering, and mathematics research, learning, and education
- Engage the NSDL and NSF STEM education communities in developing a shared vision and encouraging their contributions to this infrastructure
- Enable the library as a shared, collaborative, contributory space by incorporating collaboration and communication tools, and enabling user-contributed content
- Support the creation and display of context around library resources to enhance discovery, use, and understanding
- Put the library in the path of the user, enabling users to easily and comfortably integrate the library into their teaching and learning workflows
- Ensure that the library and all its accompanying tools and services are a robust, reliable, and highly available resource for all its users
- Support and extend the NSDL Core infrastructure (NCore) [44] as a platform for implementing and developing new educational applications and digital library tools [55]
- Develop and implement a business model that can support the continued operation of the central NSDL Technical Network Services Provider beyond the term of this grant

Our mission over the next four years will be to work with our customers – Resource Center, Pathways, other NSDL grantees, other NSF-funded projects, and outside partners – to create a new kind of digital library that integrates the best of traditional libraries, the powerful tools of Web 2.0, the organizational capabilities of the NSDL Pathways and partner societies, and the collective intelligence of all those who use and contribute to the library. Fulfilling this mission will create an intellectual commons, empowering scholars, scientists, teachers, librarians and students with the ability to work together to create knowledge and understanding.

The NCore Platform

NCore is an open-source suite of tools and services for creating production digital libraries, repository systems, and content-rich STEM learning environments. The PI team, led by Cornell University, has been the primary developer of and contributor to this infrastructure. Our work on the NCore platform and our support for NSDL demonstrate a clear, proven track record of success in advancing STEM education into the cyberinfrastructure age. A core function of the TNSP is to support, maintain, and extend this platform to meet the evolving needs of the NSDL and STEM communities.

In implementing NCore we were guided by a number of key architectural design principles that we would carry forward as the Technical Network Services provider:

- Build on robust, open-source technologies and employ open vendor-neutral standards, protocols and architectures
- Follow a web-services based architecture to enable library tools and services to be easily combined and embedded in end-users' teaching and learning workflows, and used by third parties to create innovative libraries, portals, and learning applications
- Employ modular design to ensure that tools and services can be easily added, independently modified, and removed as needed without impacting other operations
- Listen to the specific needs of community members in order to identify general solutions of broad utility that can be configured to meet a wide range of educational needs across the full sweep of disciplines and audiences that make up NSDL.

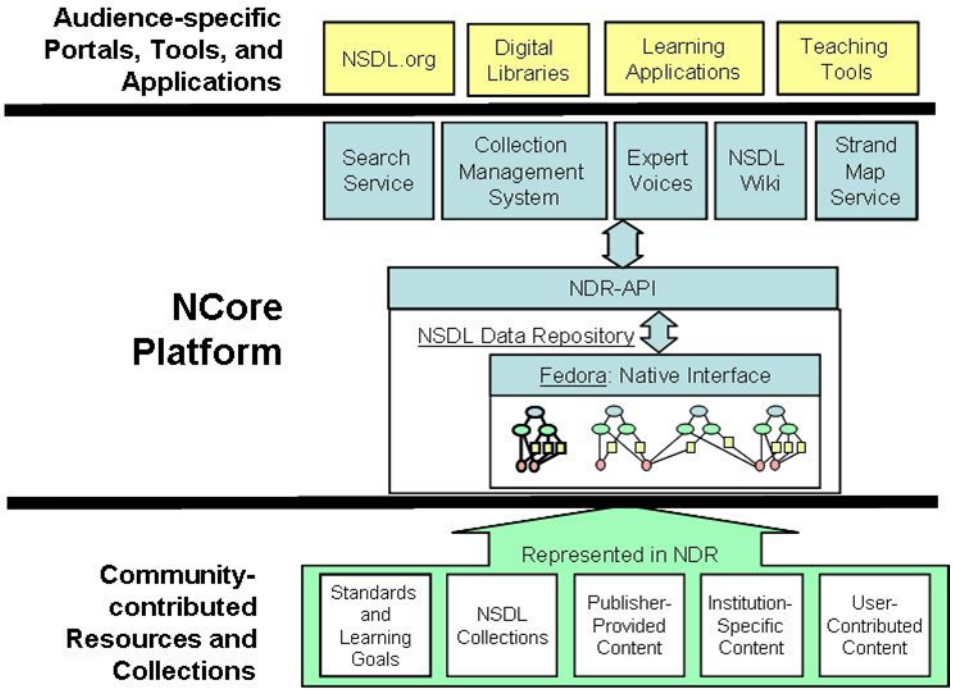

Figure 1. The NSDL Technical Network Service Provider Architecture and Existing Infrastructure

The role of the NCore platform in the three tier NSDL technical infrastructure as a mediator between community-contributed collections and audience-specific portals, tools and learning applications is shown in Figure 1. The NSDL Data Repository forms the heart of NCore; the platform also consists of a number of tools and services for digital librarians and end-users, including search services, the NSDL Collection System, the Strand Map Service, the Expert Voices blogging system, and a repository-integrated wiki for collaborative authoring around NSDL resources. All of these components are released on SourceForge under open sources licenses [43].

NSDL Data Repository (NDR) and API: NCore consists of a central repository (the NDR) and a web services interface (NDR API) that enables community members to develop new tools and services. The NDR houses and manages library content. It currently contains metadata descriptions of over 2 million STEM resources in over 130 collections. These collections are developed, curated, and contributed by the STEM community, including NSDL Pathways, Selections Services and other grantees, as well as external organizations such as universities, federal agencies, and publishers. The NDR was designed to support the highly distributed nature of NSDL by providing a purpose-built data model for integrating collections created and managed by different institutions. The NDR is implemented in the open source Fedora repository platform, which is a mature, scalable platform that can handle large volumes of content and

users [63, 67]. The NDR leverages the Fedora data model, using Resource Description Framework (RDF) triples [36] to support sophisticated relationship modeling between resources, metadata, annotations, and standards. This approach enhances the pedagogical value of digital education collections by enabling rich semantic relationships among resources to be incrementally and flexibly added by collection developers, teachers, learners, and other NSDL users [50].

Search Services: The NCore platform includes two REST-based search services for library metadata records and text content, built on the open source packages Lucene and Nutch [1]. The general search service allows fielded search of NSDL resources using Dublin Core metadata records in the NDR, as well as full-text searches of the crawled content of available library resources. Fielded search enables users to locate resources by grade level, format, subject, and collection. Searching by K-12 educational standards is being released on nsdl.org shortly. This search service is optimized to support scalable searching over millions of records. It is used by the nsdl.org web site and currently handles five million queries per month. The Digital Discovery System search service is optimized to support the rapid construction of audience-specific portals and applications and can be flexibly configured to search over any XML schema structure. It also supports geospatial searching and can be integrated with Web 2.0 applications such as Google Maps. Both search services are used by NSDL Pathways, Services, and other grantees.

Collection Management System: This system enables collection developers to efficiently and flexibly create and manage collections of metadata describing educational resources, such as learning objects, user-provided comments and reviews, scientific data sets, and other curricular components. This system supports NSDL Qualified Dublin Core and the IEEE Learning Object Model, and can dynamically generate cataloging interfaces for a range of metadata frameworks and structured text collections given an XML schema representation. It also supports the as well as the assignment of science education standards and geospatial information to resources.

Expert Voices: The NSDL blogosphere was created by building plug-in extensions to the WordPress MultiUser blogging system. It is fully integrated into the library infrastructure. Blog authors can search NSDL for resources, which they can then link to in their blog posts. The posts themselves can be published as NSDL resources, and the linkages are captured in RDF triples in the NDR. Within Expert Voices, authors can add metadata to existing NSDL resources and create metadata for new resources. Author and editor rights in Expert Voices are controlled on a per user and per blog basis. Expert Voices is currently operating in production with over 46 individual blogs and 749 registered users.

NSDL Wiki: The NSDL Wiki is an implementation of MediaWiki, the system underpinning Wikipedia, using plug-in extensions to fully integrate the wiki system into the library. Wiki pages can reference NSDL resources, add metadata to those resources, and create new external resource references in the library. In addition, wiki pages themselves can be published as new NSDL resources. An example of the educational value of the NSDL Wiki are the recently released "Classic Articles in Context"[31] providing new educational context for classic scientific articles in a growing set of fields. Editing rights for a set of related pages in the NSDL Wiki can be administered on a per user basis.

Strand Map Service: The Strand Map Service provides a web service protocol supporting the construction of interactive knowledge map interfaces based on the learning goals articulated in *Benchmarks for Science Literacy* [5] and the learning progressions and strand maps published in the *Atlas of Science Literacy* [3, 4]. Software developers use the web service protocol to construct knowledge map interfaces appropriate to the needs of their specific audiences using dynamically generated visualizations provided by the Service [71]. The resulting interactive map interfaces enable educators and learners to: (1) discover educational resources supporting learning goals, (2) browse learning goals and their interconnections, and (3) enhance their own content knowledge by examining related learning goals, the National Science Education Standards, and prior research on student conceptions[21]. This service is being used to support the NSDL Science Literacy Maps (http://strandmaps.nsdl.org/), the Educational Benchmarks Collection, and other applications.

Other NCore components (not shown in Figure 1), many developed by community partners and integrated into the infrastructure, include:

Community Sign-On: In collaboration with our Core Integration partners at Columbia University, we implemented an NSDL-wide single sign-on technology based on Shibboleth [57]. This component is designed to provide users with seamless access across distributed NSDL Pathways and other sites. The NSDL Shibboleth Identity Provider service is currently operated by Columbia and would continue to be operated by them as a subcontract under this proposal. Shibboleth Service Provider clients have been implemented for the nsdl.org web site, the Expert Voices blogging service, and the NSDL Wiki.

K-12 Science Education Standards: NCore uses Achievement Standards Network (ASN) identifiers as the mechanism for representing educational standards. These identifiers support the unambiguous referencing and interchange of educational standards across services, collections, and digital libraries. This approach, which maintains a mapping between standards and a unique, resolvable URI, is the result of an NSDL Services grant (PI: S. Sutton, Univ. of Washington, NSF Award DUE-0121717). We currently host the Content Assignment Tool (CAT) and Standards Alignment Tool (SAT), also resulting from a prior NSDL award (PI: A. Diekema, Syracuse University, NSF Award DUE-0435339). CAT helps collection developers to assign education standards to resources. It uses natural language processing algorithms to analyze a target digital library resource and suggest a ranked list of potential standards from a specified standards framework. SAT uses similar algorithms to automatically map standards in a target framework to corresponding standards in a destination framework. For instance, given a specific National Science Education Standard, SAT can return a list of corresponding Colorado State Standards. Currently these tools are available through web page and web service interfaces. The next release of the NSDL Collection System will embed the Content Assignment Tool into the cataloging workflow, enabling collection developers to easily incorporate standards assignment into their cataloging processes.

OAI-PMH Harvest Aggregation System: NCore includes a large-scale OAI-PMH harvesting management and metadata aggregation system, which controls the automated scheduling, performing, and reporting on collection harvests [46]. This system currently harvests an average of 150,000 metadata/resource record updates from over 130 collections each month. These records are represented using the NSDL Qualified Dublin Core metadata framework.

OnRamp/OnFire Multimedia Content Management System: OnRamp/OnFire supports the creation of multiple modes of dissemination of STEM education content from a Fedora/Fez-based repository. It has recently been put in production in support of "Beyond Penguins and Polar Bears: Integrating Literacy and IPY in the K-5 Classroom" (PI: K. Lightle, Ohio State University, NSF Award DRL-0733024) where it supports creation of the multimedia cyberzine. It is also used to store and manage the dissemination of other NSDL content, such as the monthly NSDL Whiteboard newsletter.

TNSP Project Roadmap

The TNSP has responsibility for providing a wide range of tools, services, and processes in support of NSDL grantees and the broader STEM community. The TNSP is also charged with engaging the NSDL community to develop the NSDL technical infrastructure, tools, and services over the period of the grant and beyond. In this project roadmap we lay out our planned activities in three periods. The first section describes the operational capabilities we will provide the community from day one of this grant. This is the critical infrastructure that the entire NSDL community relies on. In the second section, we describe new development and process activities we will undertake in the first six to twelve months. These tasks will advance the current infrastructure in important ways while we ramp up our Community Engagement Process (described later). In the third section, we present strategic directions for the mid- to long-term development of the NSDL technical infrastructure. This roadmap will serve as a starting point as we engage with the community to develop a full, multi-year roadmap for NSDL technical network services development.

NSDL Technical Network Services – Day One

We will provide continued operation, ongoing maintenance, and customer support for the following:

NCore Services and Tools described above: The continued operation and support of the NCore platform that we developed under the previous NSDL Core Integration grant is critical to the entire NSDL community, and it is explicitly called out in a number of places in the NSDL solicitation. We will continue to: (1) harvest metadata from existing NSDL collections; (2) work with new Pathways and collection providers to identify and incorporate appropriate new collections into NSDL in order to keep the NSDL content growing and up-to-date; (3) operate and manage search indexing and content crawling; (4) provide support for the creation of new NSDL blogs and wikis and the management of comments and spam; (5) manage the infrastructure to support the nsdl.org domain and subdomains, email addresses, and mailing lists; and (6) ensure integrity of NSDL resources and metadata by working to minimize bad links, ensure schema valid metadata,.

NSDL.org Web Site: The solicitation implicitly divides responsibility for the NSDL.org web site between the TNSP and the Resource Center. The TNSP is explicitly charged with "maintaining and upgrading the infrastructure underpinning NSDL and its primary website, nsdl.org". It also states that "The Technical network services team will also maintain NSDL.org's high standards of stability and reliability, to ensure that end-users can interact effectively with the site." On the other hand, the Resource Center is charged with evaluation, including "webmetric analysis", as well as outreach and "promot[ing] both national and international dissemination of the research and development contributions of the NSDL program." Given this division, the TNSP will manage the underlying code, databases, the look & feel, navigation, and layout of the site. The TNSP will also instrument NSDL.org, and TNSP-supported tools and services, to support evaluation and research; basic forms of instrumentation are already in place including collection of usage data and query logs. We will also use the site to showcase existing NSDL tools and services (e.g. search, Science Literacy Maps, Instructional Architect) as well as new tools and services being developed by current and future NSDL grantees and outside projects. We will provide technical support for the Resource Center to present content and tools relevant to its outreach and dissemination missions within NSDL.org (e.g. news, reports, and links to other projects). We will work closely with the Resource Center and other NSDL community members to ensure that the capabilities of the site fully meet the needs of the Resource Center and the NSDL community.

NCore Toolkit: In support of several ongoing efforts, including the recently funded "SGER: Increasing the Educational Impact of NSF-funded Research" (PI: D. Krafft, Cornell University, NSF Award DUE-0829759), a toolkit supporting the easy integration of externally developed tools, services and applications into the NCore platform will be provided. The toolkit, available in both PHP and Java, greatly simplifies the use of the NDR API, making it easier to create lightweight plug-ins and extensions for existing applications or for newly developed tools and services.

RSS/Atom Feed Registration and Ingest: Cornell is completing work enabling authorized users to register RSS feeds into NSDL. This system can be used as a lightweight alternative to OAI-PMH ingest. It also will advance the vision of the library as a shared collaborative space by enabling user-developed bookmarks, resource recommendations, and folksonomies to be easily incorporated into NSDL.

Initial Development Plan

While much of the development effort under this grant will be guided by the Community Engagement Process, our initial development plan is focused on strengthening the NCore platform and TNSP services for cost-effective business operations; transitioning from a centrally managed core integration effort to a more distributed model for NSDL; and enhancing developer and end-user access to the advanced educational cyberinfrastructure capabilities of the Fedora-based NSDL Data Repository.

Enhancing Reliability, Integrity and Efficiency: Although we have released a number of independent components of NCore on SourceForge, there are still parts of NCore that are not publicly released.

Moreover, we do not yet provide a complete, installable software package and build system. Finally, based on the experience of the DLESE Program Center, we believe that there are significant opportunities for streamlining the code and the processes that underlie the library. We plan to refactor and improve all aspects of the NCore systems, create a fully automated and integrated build process for NCore, and improve the efficiencies of workflows for the management of NCore tools, other NSDL systems, and technical collections integration. In addition to reducing costs and improving the central operation of NSDL, this should make it significantly easier for Pathways and other NSDL grantees to integrate with and build on top of the NCore platform.

Integrating Resource Vitality Checking and Archived Resource Access: A significant issue for NSDL is dead links. Resources can be moved or taken off-line, leading to significant user frustration with library contents and search results. Frequently, an archived copy of the resource will be available in either the NSDL archive at the San Diego Supercomputer Center or the Internet Archive. We plan to take several steps to address these issues: (1) building and running automated vitality checking to identify dead links; (2) providing users with the ability to flag suspect links and resources; (3) creation of an "Archived Collection" for materials no longer available at their original URLs, but only available as archived copies; (4) enhancing search to allow including or excluding resources in the Archived Collection, and linking directly from the search results list to the archived copy for resources whose primary link is dead.

Enabling Distributed Collection Management: In the Core Integration (CI) model of NSDL, primary responsibility for overall NSDL collection development and management resided with CI. With the new solicitation, this responsibility is widely distributed among NSDL grantees. To support this new model, we plan to extend the tools and infrastructure to distribute responsibility for management of resources and collections to authorized community members. For example, when a user flags a resource as suspect, the system must support distributed review by a group of authorized editors across the NSDL Pathways and grantees. Existing tools would be extended to allow both distributed action and distributed management of authorization and access.

Enhancing Search and Browse: We will continue to expand mechanisms for putting NSDL in the path of the user, and expanding the context around resources to support discovery and use. Specific initial development tasks include: (1) improving how NSDL resources are indexed and accessed by commercial search engines, which provide an important gateway and significant traffic to NSDL.org,; (2) improving mechanisms for associating educational resources with learning goals in the NSDL Science Literacy Maps; (3) providing increased support for searching and browsing by K-12 science education standards; (4) providing tools and services for geo-referencing NSDL resources, conducting geospatial and map-based searches, and integrating geo-referenced NSDL resources into third party applications such as Google Maps; and (5) supporting explicit and implicit user feedback and quality rankings of NSDL resources through user-provided ratings of resources (explicit feedback) and click-through statistics aggregating searching and browsing behaviors (implicit feedback).

Strategic Directions

The information and web landscape in which NSDL exists is constantly evolving. The read-only document-based "Web 1.0" has transformed into the highly participatory interplay of documents, social networks, and semantic relationships of "Web 2.0". The evolution of the NSDL technical infrastructure reflects this web transformation, moving from a metadata-based union catalog paradigm to one that is resource centric, participation centric, and relationship centric. The result is an information environment that integrates with the way people work on the web now, rather than one that resembles the libraries of a previous era. The NSDL technical infrastructure must continue to evolve over the this performance period, both to fully exploit work done under the previous core integration project and to keep NSDL current with the evolving Web. The remainder of this section briefly summarizes the proposed work.

Search and Browse: The NDR, with its rich and extensible relationship-based data model, provides the foundation for a new generation of search and browse services. Metadata-based catalog searches will continue to be an important component of the NSDL service structure. However, we will leverage semantic relationships modeled in the NDR to support search-in-context and serendipitous browsing, enabling users to explore the context of resources. We will work closely with the NSDL community to design and develop these new searching and browsing capabilities through two efforts:

> ***Enhanced query API:*** We plan to enhance the existing NDR search API to allow clients to issue expressive semantic queries against the contents of the NDR. We will use the emerging standard for semantics queries, SPARQL [23], a technology developed in the semantic web community and embodied in a W3C recommendation. Support of SPARQL in the query API will facilitate the development of applications that fully leverage the graph structure of the NDR. For instance, the NSDL Science Literacy Maps could be extended to enable users to query the relationship structure of the maps to locate resources related to prerequisite learning goals.
>
> ***Semantic exploration:*** We plan to work with our partners to implement visual browsing tools that expose directly to the user the semantic relationships among resources that we now store in the NDR and the ontology structure upon which they are based. This will make use of techniques for semantic graph visualization developed in the semantic web community [58]. Implementation of this capability will build on the API implementation described above.

Object Reuse and Exchange: The NSDL collections consist of resources of various levels of complexity. Some resources are simple, atomistic web pages, but many are complex aggregations of learning objects consisting of multiple, multiple-genre (text, images, video, etc.) resources. The exposure and utility of these rich NSDL resources to the rapidly expanding suite of Web 2.0 applications (e.g., mash-ups, social network applications) depends on the ability to both identify them in the web context, via URIs, and describe them in manner that allows applications to access their constituent components. This has been the focus of work by Open Archives Initiative Object Reuse and Exchange (OAI-ORE) [51]. For instance, digital curriculum overlay models integrating learning goals, library resources developed by multiple institutions, and lesson plans [41, 42] illustrate the types of complex educational aggregations that can be described and shared via OAI-ORE. The OAI-ORE standards, developed with separate funding by Co-PI Lagoze and colleagues, form an integral part of the cyberinfrastructure essential for future NSDL work. Our technical plan includes both implementing the current OAI-ORE standards and undertaking further development of the standards in a manner that will benefit the NSDL and other digital libraries. One aspect of this work will address the scalability of ORE methods to allow description of large aggregations. We will investigate the utility of URI patterns for doing this, instead of relying on current methods where all components of an aggregation need to be explicitly enumerated.

Datanet and eScience: Fedora and OAI-ORE are being adopted as the infrastructure of choice in many eScience projects, both nationally and internationally. Submissions to the NSF's recent Datanet solicitation are proposing to build upon these technologies to support access, integration, and preservation of scientific data [2]. Through our involvement with Fedora Commons and emerging Datanet initiatives, the PI Team will promote considerable cross-fertilization between NSDL and eScience projects, ensuring interoperability and sharing across education and research technical infrastructures.

Advanced Educational Context: Advances in machine learning and natural language processing are providing new, scalable methods supporting the creation and display of context around resources to support discovery, use and learning. Two areas demonstrating strong potential include topic discovery and conceptual summarization. Topic discovery uses sophisticated algorithms to analyze large-scale collections of inter-linked documents to identify emerging subject areas and track their evolution over time [40]. These algorithms are also being used to support repository management workflows to automatically identify resources that are out-of-scope (off topic) for particular collections (PI: J. Martin, University of Colorado at Boulder, NSF Award CISE/IIS-0741326). As these algorithms become more

robust, we can develop automated tools to improve the quality of metadata within NSDL [46] and provide users with innovative interfaces for browsing related resources. Conceptual summarization algorithms analyze the text of digital library resources to automatically construct concept inventories describing key ideas and learning goals [25]. These inventories can be used in library interfaces to assist learners in selecting and evaluating learning materials, and they can underpin new forms of personalized resource selection services [26]. Recent studies demonstrated that such services can have a positive impact on learners' cognitive processes and outcomes [22]. We will continue to track research in these areas and deploy the results in tools available to the NSDL user community.

TNSP Organizational Design

The organizational structure of the TNSP must be carefully designed to balance short-term responsiveness to community needs and interests with the long-term goal of providing the NSF STEM research and education community with the best possible educational cyberinfrastructure. Additionally, over this four year period of performance, the organization must progressively adapt to changing financial models, as fixed NSF support declines and is replace by variable amounts of support provided by NSDL grantees in the form of 15% back charges on their individual grants. The TNSP also must plan and prepare for a new business model, capable of supporting its technical services and infrastructure beyond the four year period of performance. This is clearly a complex undertaking with few success models and precedents to draw upon. Our organizational design rests on four foundations: (1) an experienced senior management team, (2) an organizational model with clear roles and responsibilities, (3) agile project management processes based on industry best-practices, and (4) an independent Advisory Board.

Key Personnel

We have an experienced senior management team – Krafft, Lagoze, Sumner, and Wright – that together have the combined expertise to address these multi-faceted demands. The PIs have significant experience developing, operating, and supporting digital libraries and cyberinfrastructure for STEM education and research. As part of the NSDL Core Integration Team since its inception, Krafft and Lagoze (Cornell University) have provided technical leadership for the NSDL program, responsible for developing, operating, and supporting the technical and operational infrastructure underpinning NSDL. Sumner and Wright (Digital Learning Sciences) have provided technical leadership to the Digital Library for Earth System Education and NSDL communities since 1999. They have worked closely with the Cornell team to integrate tools, services, processes, and know-how resulting from prior projects (see Results from Prior NSF Support) into the NSDL infrastructure and to provide technical collections support to the NSDL community. In this grant, we propose to formalize this productive partnership to enhance and extend the technical network services available to the NSDL and larger NSF communities. This team has a track record of successful collaboration and has demonstrated their ability to deliver robust NSDL infrastructure and services, while adhering to project resource constraints and milestones.

Organizational Model

We have developed an organizational model designed to balance short-term services and enhancements with long-term strategic initiatives, which puts community services and engagement at the center of program planning (Figure 2). Specific members of the senior management team have primary responsibility for directing each of these functional areas:

Platform Development and Library Operations (Krafft): The primary responsibilities of this functional area are to maintain the technical integrity of the NSDL Data Repository, to maintain and extend the NCore technical platform, to maintain and extend the library portal (NSDL.org), and to operate library services and infrastructure.

Krafft (PI: Cornell University) is the senior manager responsible for this functional area. Krafft is extremely experienced and qualified to fulfill this role; he has managed NSDL technical development and operations on behalf of the Core Integration team for over five years. Under his leadership, NSDL

technical infrastructure and NSDL.org have a track record and reputation for reliable, robust, and efficient operations and innovative technical developments.

Community Services and Engagement (Wright): The primary responsibilities of this functional area are to provide technical collections integration support to enable the community to share collections into the NSDL Data Repository, to support NSDL grantees in utilizing and integrating NCore tools and services into their products and applications, and to establish and facilitate the community engagement process enabling distributed contributions to library directions and technologies.

Wright (Co-PI: Digital Learning Sciences/UCAR) is the senior manager responsible for this area. Wright has significant prior experience in leading distributed community members in digital library development in both the NSDL and DLESE programs (DLESE: Digital Library for Earth Systems Education). He served as Co-Chair of the Technology Standing Committee from 2003-2004, facilitating dialog between community members and the Core Integration team. In DLESE, as Director of Technology and Operations, his core job was providing technical community services and supporting distributed community engagement in library directions and technologies [76]. He worked with the DLESE Technology Committee, led DLESE Developer Workshops, and established technical plans responding to community ideas and feedback generated by the 200+ attendees at DLESE annual meetings.

Strategic Initiatives (Lagoze): The primary responsibility of this functional area is to scan the emerging research landscape and to anticipate long-term trends and needs in educational cyberinfrastructure. Goals such as "putting the library in the path of the user" and "supporting the creation and display of context around library resources" will require new advances that respond to the evolving eScience repository landscape, eLearning landscape, and the World Wide Web.

Lagoze (Co-PI: Cornell University) is an acknowledged thought leader in digital library architectures, eScience repositories, and the design of interoperability protocols for cyberinfrastructure. He has been a key team member of the Cornell NSDL Core Integration team. Lagoze has led the development of technical standards currently in use in NSDL and world wide (OAI-PMH) and is actively leading international efforts to develop models and mechanisms for representing compound, multiple media eScience and eLearning objects (OAI-ORE). Supporting the incorporation of these emerging scientific and learning objects into NSDL will be an important strategic initiative.

Evaluation (Sumner): The primary responsibility of this functional area is to develop and collect measures and metrics that will enable the team to assess the value and efficacy of the TNSP technologies and services to NSF, NSDL grantees, and the larger STEM research and education community. We will conduct ongoing formative evaluations that help the team to continuously refine their processes, products, and services, and help inform the business planning process.

Sumner (Co-PI: Digital Learning Sciences/UCAR) is the senior manager responsible for this area.

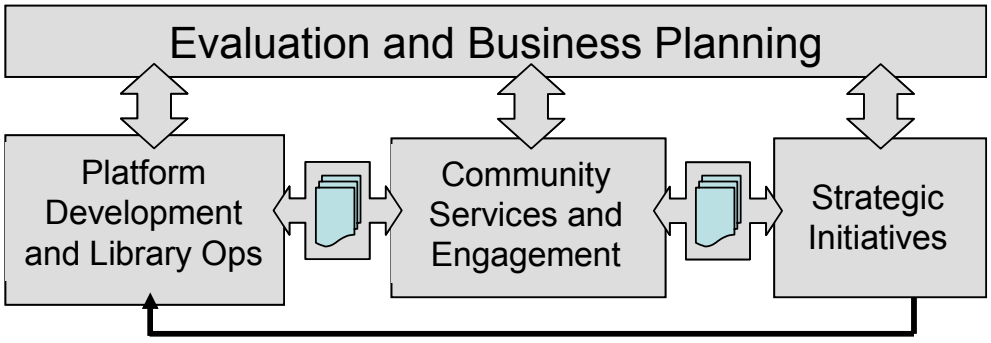

Figure 2. Organizational Model of the Technical Network Services Provider. "Boundary objects" such as Project Roadmaps will structure and mediate communication between different functional areas.

Sumner has significant experience in user-centered design and evaluation methods, the development and evaluation of educational technologies and digital libraries, and cognitive and learning science. She was Chair of the NSDL Educational Impact and Evaluation Standing Committee from 2002-2003, where she led a collaborative community effort to collect metrics on library collections and growth, library usage, and community engagement and governance. As a co-PI on the DLESE Program Center, she was responsible for ensuring that research and evaluation results fed back into library operations and design.

Business Planning (PI Team): The primary responsibility of this functional area is to investigate, evaluate, select, and implement a financial and operational model for sustaining the TNSP beyond the four year performance period. This effort will be guided by an independent Advisory Board composed of members with relevant library, education, and business experience.

The entire PI team is responsible for this critical functional area. Sumner and Wright recently completed a similar process on behalf of DLESE and have overseen transfer of the library to a new operating host. Wright is currently working with GLOBE, a NASA-funded K-12 education program, to develop a long-term business model for post-NASA funding. Lagoze has significant experience with establishing open source-style community initiatives to sustain technical developments (Open Archives Initiative).

Project Management Process

We are explicitly adopting an organizational learning model, where ongoing formative evaluations, business planning, and consultation with the NSF and the NSDL community will be used to inform and refine our organizational structures and management processes. This will be applied to an initial project management model based on agile development methods, which we have successfully employed over the past two years and which form the basis of industry best practices in software development firms. These agile methods advocate the value of frequent communication and short development cycles, in contrast to formal specification processes and long planning cycles [13, 14, 17].

The Senior Management Team is responsible for (1) developing the TNSP Project Roadmap in conjunction with the NSDL community, the NSDL Resource Center, and Advisory Board input, (2) analyzing the Project Roadmap to identify priorities and deliverables over the next short-term development period, and (3) managing teams in their functional area to complete deliverables. We have found that three short-term development periods per year, roughly corresponding to semesters, work well for synchronizing with our primarily academic community members.

"Lightweight" teams composed of technical staff from across participating institutions are the fundamental work units. Teams are convened as needed to address specific development objectives, for instance "Support OAI-ORE." Other teams, such as "Library Operations" are long-term. Cornell, Digital Learning Sciences, and Columbia personnel are all experienced working with each other in these team structures at a distance. Bi-weekly meetings, consisting of team leaders and senior managers, are conducted to resolve dependencies and conflicts between team remits and workloads, and to track progress. It is vital to the success of the NSDL program for the TNSP and the Resource Center to work together effectively and efficiently. Towards this end, we will encourage the Resource Center to nominate a team lead representative to participate in these meetings.

Advisory Board

TNSP will recruit and convene an independent advisory board composed of 6 recognized experts in STEM education, cyberinfrastructure, and sustainability and business planning. Participants will be selected based on their expertise and reputation, their contribution to the overall knowledge profile of the board, and their financial independence from the NSDL program. Program independence of advisory board members emerged as a critical factor during the sustainability planning of DLESE. Without this program independence there may be conflicts of interests between community members interested in increased and improved services, and the need to reduce and control costs to support long-term sustainable and scalable operations [52-54].

The Advisory Board charge will be to provide guidance on the development, operations, and long-term vitality of the Technical Network Services Provider, advising the PI and senior management team on strategic directions, tactics and strategy implementation, plans and financials, sustainability and business planning, and marketing and partnership development. The Board will meet annually face-to-face, with supplemental telecons at mid-year and as needed.

TNSP's Community Engagement Process

The success of the TNSP hinges on the ability to create and manage an open and transparent process for: (1) harnessing community creativity and innovation, (2) efficiently gathering broad-based customer input to inform TNSP's annual work plan, and (3) communicating work plan decisions and their rationale. We will devote significant time and careful thought, particularly in the first two years of the project, to formulating and carrying out a community engagement process, and progressively refining and enhancing it based on evaluative data and community feedback.

We will build on the high-level process schematic shown in Figure 3. This process is based on discussions with consultants at Ithaka (www.ithaka.org: experts in developing non-profit sustainability models for higher education), who highlighted the need for structured small-group discussions encouraging deep participation, multiple opportunities for feedback in the process, and clear reporting of decisions and plans. It also follows prior social science research on the role of "boundary objects" in facilitating multi-disciplinary project communication [19]. Social science researchers have repeatedly demonstrated the critical role of structured representations or boundary objects in supporting productive dialogue and the development of shared understanding in multi-disciplinary and distributed design teams [18]. Many successful open source projects use "roadmaps" as a boundary object to support project planning and communication with distributed participants.

We will conduct this process annually over a 6-8 week period with the community input phase commencing at the NSDL Annual Meeting, which occurs shortly after new awards are announced.

- **Project Roadmap (Alpha):** The senior management team will develop a preliminary version of the project roadmap to seed community discussions and planning. We will work closely with the NSDL Resource Center to ensure that this roadmap is consistent with their planned outreach and coordination activities. This roadmap will communicate TNSP's planned service offerings for the coming year (library operations, NSDL.org, NCore platform support, collections integration), planned advances for the coming year (NCore enhancements, community-developed tools and services to be integrated and supported), and selected strategic initiatives and planned progress milestones towards them over the coming year. Roadmap choices will be guided by "business case" criteria developed with guidance from our Advisory Board. Initial criteria used to formulate the roadmap in this proposal include anticipated operating costs, utility to the broad community, adherence to architectural design principles, and strategic importance to the vitality of NSDL and TNSP. In Year 1, the alpha roadmap will be based on the objectives presented earlier in this proposal. In subsequent years, the alpha roadmap will logically extend the previous year's roadmap, incorporating evaluation results, emerging community needs, and the needs of newly funded NSDL projects. The alpha roadmap will be posted on an NSDL Project Roadmap wiki and available for community input.
- **Community Input**: While anyone can respond and make contributions to the wiki, we will conduct a series of formal sessions with constituents within NSDL to ensure systematic and broad-based input. At the NSDL Annual Meeting, we will review the roadmap with the NSDL Technology Committee and also with new grantees as part of an orientation session. This orientation session will provide us with an opportunity to learn about the specific technical needs and project goals of new grantees. After the NSDL Annual Meeting, we anticipate holding two to three focus group telecons for NSDL Pathways and other grantees. Participant comments, suggestions, and concerns discussed from each session will be recorded on the Project Roadmap Wiki.

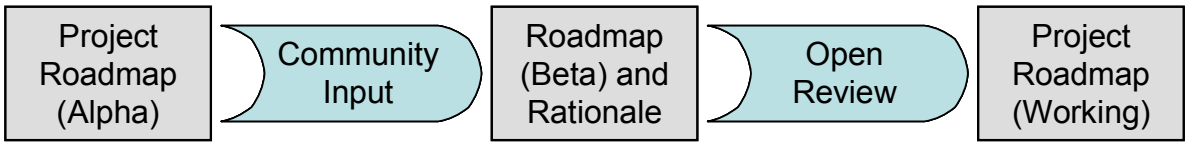

Figure 3. High-level schematic of community engagement process

- **Roadmap (Beta) and Rationale**: The senior management team will produce a revised roadmap selecting community ideas and contributions using our business case criteria. The senior management team will also provide brief rationale for roadmap changes resulting from this criteria-based analysis.
- **Open Review**: The NSDL community will be invited to comment on the beta version using the wiki for an open period of 2 to 3 weeks and these comments will inform the "Working" version.
- **Project Roadmap (Working):** The senior managers will produce the "Working" version of the roadmap, based on feedback gathered during the open review period. This version will be used to guide TNSP project management and to assess and track project progress. Mid-year adjustments to the Working Roadmap will be posted on the wiki and announced via the NSDL Whiteboard.

It is important to reiterate that all responsibility for formulating the TNSP roadmap rests with the senior management team, who must also ensure that the roadmap is aligned to the TNSP's emerging business sustainability model. This process is a systematic means to engage in dialog with our customers to ensure that our planned TNSP services and tools address community needs; it is not a voting process.

Other Communication Tools. We will continue to support publication and dissemination of the NSDL Whiteboard, the monthly newsletter outlining grantee accomplishment and developments, new tools and services, and developments in STEM education salient to NSDL. We will also promote broader use of NSDL Labs, an area in NSDL.org for the TNSP and other NSDL grantees to post announcements about prototypes and beta releases of new tools and services for experimental use and feedback.

Dissemination to other NSF grantees. McArthur and Zia outline the potential of NSDL and NCore to become a platform for next-generation education research and advanced learning applications [55]. We will work with the NSF to encourage investigators responding to EHR, OCI, and appropriate solicitations in other directorates to build upon the NSDL technical infrastructure and to incorporate NSDL tools, services, and data into their research objectives and project designs.

Evaluation

The objective of our evaluation plan is to systematically gather formative data to inform improvements to TNSP infrastructure, operations, and services; to ensure that TNSP tools and services are in sync with community needs; and to inform our business planning. Towards these ends, our evaluation questions and mixed-method protocols are designed to gather continuous feedback from our primary customers – NSDL grantees – about the value, utility, and perceived quality of TNSP infrastructure, operations, and services (Table 1). Our plan includes anticipated points of coordination with other NSDL entities. We will facilitate research and evaluation by providing instrumentation data to the NSDL Resource Center and other grantees funded to evaluate the impact of TNSP-supported infrastructure, tools, and services.

Table 1. Key components of TNSL evaluation plan: questions and methods to be employed

Evaluation Questions	**Methods to be Employed**
<u>NSDL Grantee Satisfaction with TNSP:</u> ▪ Operations – Perceptions of reliability and robustness of library operations ▪ Services and Tools – Use of and satisfaction with TNSP services and NCore platform/tools	Annual Web-based Community Survey ▪ Likert-scale and open-ended response questions ▪ Deploy to all PIs, Co-PIs, and Senior Personnel listed on awarded NSDL grants

(both overall and individual services/tools) ▪ Community Engagement – degrees and venues of participation, perceptions of personal engagement and quality of community process, overall satisfaction with influence on process and annual roadmap outcomes ▪ Agreement and satisfaction with TNSP long-term strategy and vision	▪ Coordinate with the NSDL Resource Center and the Educational Impact and Evaluation Standing Committee (EIESC) to jointly design and deploy only one community survey annually to reduce potential survey fatigue amongst participants
NCore Usage and Support Services: What are the uptake and usage of NCore and its constituent tools by NSDL grantees? What triggers and barriers are grantees experiencing with respect to NCore use, technical collections integration, or accessing technical support services? (Triggers are positive experiences promoting adoption and use.)	▪ Questions on Community Survey about usage of NCore and individual components ▪ Annual semi-structured phone interviews with sample of NSDL Pathways and other grantees
NSDL.org: How are site components being utilized and by whom? What tasks are users trying to accomplish? How satisfied are users with site features, tools, and services?	▪ Instrument NSDL.org, NCore tools and TNSP services to augment web metrics analysis performed by Resource Center ▪ Coordinate with the Resource Center to deploy site usage survey at regular intervals
NDR Collections Management: How are collections growing and changing over time? How available are resources managed in the NDR (link checking)? How does resource availability and technical collection integrity change over time?	▪ Instrument NDR to track collection growth and collect resource availability metrics ▪ Develop scripts for regular, automatic reporting of collection-level resource availability and harvesting metrics for evaluation purposes and to aid grantees in collection curation
Viability of Alternate Business Models: Probe perceptions of potential uptake and perceived value of alternative business or pricing models	Additional questions as needed on Community Survey and semi-structured phone interviews

Business Planning and Sustainability

Our approach to promoting and sustaining the TNSP beyond the grant period will involve three, intertwined activities: (1) business planning, (2) identifying TNSP services that provide unique value to STEM audiences and should be sustained or expanded, and (3) quantifying and lowering our operating costs. The latter will be an important activity in Years 1 and 2, after which the TNSP's fixed funding from the NSF is halved. We will also ensure that the NSDL technical architecture is engineered to easily support the elimination or out-sourcing of tools and services that are too expensive to operate or do not provide enough added value to justify ongoing support. This three-pronged approach is based on our prior experiences with sustainability planning for both the DLESE [52-54] and GLOBE programs.

A business consultant will be hired to structure and guide the planning process. The senior management team will work with the consultant to identify and explore alternative business models (both non-profit and for profit), and to conduct market research to identify long-term audiences and markets for TNSP technology and services. We will begin by analyzing the long-term viability of a fee-for-service model. As specified in the solicitation, the TNSP will begin operations under this model, with a significant portion of income derived from the 15% charges budgeted in other NSDL grants. As we quantify our different operating costs, we will investigate tiered pricing structures charging different fees for specific services or packages of services. Our evaluation activities are designed to inform this planning process by gathering feedback on NSDL grantee perceptions of TNSP services and their "value for money" and by

providing opportunities to probe the perceived value of alternative pricing structures and business models. The outcome of our planning process will be a business plan designed to sustain TNSP tools and services for a three-year period beyond the end of the grant. Our Advisory Board will provide strategic guidance to the business planning and business model selection process and the development of the TNSP Business Plan. They will also provide ongoing tactical guidance over the next four years to help us ensure that the TNSP is prepared and ready for new operating and financial conditions after the grant period.

Intellectual Merit

The intellectual merit of this effort lies in our technical and organizational contributions to innovative educational cyberinfrastructure. In terms of technical contributions, the NCore platform will embody state-of-the-art advances – semantic queries, OAI-ORE, and advanced educational contexts – and make these advances available to the STEM educational community in robust, cost-effective, and scalable ways. We will demonstrate the power of NCore and NSDL for promoting the development of innovative library interfaces, services, and learning applications capable of transforming teaching practices and learning experiences for a wide variety of STEM audiences. At the information level, this effort will yield a new form of knowledge organization, integrating resources, descriptions, context and contributions from a huge range of organizations, projects, and individuals, that could potentially transform our understanding of digital libraries and the processes of STEM education.

In terms of organizational contributions, the TNSP organizational model with its embedded formal community engagement process will contribute to our understanding of scalable social processes for fostering community innovation and contributions to evolving cyberinfrastructure. Through our business and sustainability planning, we will investigate and define new models for sustainable cyberinfrastructure.

Broader Impacts

Enhancing the nation's infrastructure for research and education is the core remit of the TNSP. As a result of the work proposed here, at the end of the four year period of performance we will have:

- Significantly enhanced the computational infrastructure – NCore, NSDL.org, and NSDL Data Repository collections – enabling STEM scientists and educators to develop and disseminate digital learning resources and learning applications to new and larger audiences of science learners, including K-16 students, pre- and in-service science teachers, and higher education faculty.
- Established a broad-based community engagement process to foster creativity, innovation, and contributions to the design and use of next-generation educational cyberinfrastructure and interactive learning applications. Our proposed process will help to formalize and structure critical collaborations and partnerships across STEM disciplines and across a range of institutions serving science research and education that are funded through the NSDL program (i.e., universities, museums, professional societies, public television stations, and other non-profit organizations)
- Provided technical tools and services to a large number of NSDL Pathways, Services, Targeted Research and other NSF grantees, helping them to better serve their specific audiences. By helping grantees to build on and customize our significant, shared technical infrastructure, we will enable grantees to focus less on recreating technical developments and more on serving as sites for teaching, mentoring, and supporting research of large numbers of STEM students and teachers
- Stimulated new forms of cyberinfrastructure-based education and learning science research by instrumenting the NCore platform, NSDL.org interfaces and other infrastructure-supported tools, processes, and services to collect detailed (anonymized) usage and behavioral data. Such data can provide the basis of next-generation research into personalized learning, educationally-informed recommendation engines, educational social networks formation and evolution, and other areas where large volumes of user-generated actions and activities are needed.

Results from Prior NSF Support

The three most relevant prior NSF awards of the PI Team are listed in Table 2. Major outcomes of these awards that directly relate to our proposed TNSP activities are:

- **Released and documented NCore:** Described earlier in the NCore section.
- **Populated and managed the NSDL Data Repository:** Also described in the NCore section.
- **Provided technical integration, support, and outreach services to the NSDL community:** Examples of integrating community-produced services were described in the NCore section. Examples of support and outreach services provided to NSDL grantees were described in TNSP Project Roadmap section, NSDL Technical Network Services – Day 1.
- **Hosted, supported, and developed NSDL.org and other audience-specific portals, tools, and applications:** Responsibility for NSDL.org, the main library portal, resided with the UCAR Core Integration team in the prior award (PI: K. Howe, UCAR, NSF Award DUE-0424671). Cornell staff worked closely with UCAR team members to design the library's flagship portal, which is hosted and operated by Cornell University. Additionally, we have hosted and maintained selected tools and applications developed by NSDL grantees, such as hosted application is Instructional Architect (PI: M. Recker, Utah State University, NSF Award DUE-0434892), which supports K-12 teachers to construct and share lessons plans composed of NSDL resources.
- **Operated the library portal, services, and technical infrastructure:** The Cornell University team has operated the NSDL technical infrastructure since its inception. We have made several infrastructure innovations to support reliable and scalable service. Working closely with the Fedora Commons group, we developed a robust, scalable repository architecture and implementation for NSDL [47-49]. We developed the MPTStore [16], a fast, highly-scalable RDF triplestore to support the roughly 300 million dynamically-updated RDF triples in the NDR. The repository architecture makes use of RDF triples to represent NDR aggregations [44], relationships between sourced metadata statements and the resources they describe, as well as contextual relationships among resources. To support reliable 24x7 operations, we created a new transaction journaling system for Fedora that supports running real-time production replicas of the full repository, with the ability to switch from leader to one of the two production followers in a matter of minutes. With this configuration we can manage zero-downtime upgrades of the repository servers and code.
- **Disseminated research and development outcomes to the STEM community:** Produced over 41 publications, presentations, and workshops as a direct result of prior NSF awards.

Table 2. Overview of most relevant prior NSF awards

NSF Award: DUE 0227648 Title: Collaborative Project: Core Integration – Leading NSDL towards Long-term Success Amount: $8,845,453 Investigators: **Dean Krafft**, William Arms, John Saylor, **Carl Lagoze**, Diane Hillmann Period: 1/1/03-12/31/07 Human resource development: Three graduate students Publications: See [9-12, 15, 20, 24, 27-30, 37-39, 44-49, 56, 59-62, 64-66, 74, 75, 78]	NSF Award: GEO 0639539 Title: DLESE Program Center 2007 Work Plan Amount: $999,457 Investigators: Mary Marlino, **Tamara Sumner, Michael Wright** Period: 11/1/06-10/31/08 Human resource development: Two graduate students Publications: See [41, 42, 52-54, 70, 77]	NSF Award: DUE 0226286 Title: Strand Maps as an Interactive Interface to NSDL Resources Amount: $472,538 Investigators: **Tamara Sumner**, Mary Marlino, Greg Janee, **Michael Wright** Period: 1/1/03-12/31/05 Human resource development: Two graduate students and one undergraduate researcher Publications: See [6-8, 21, 32-35, 68, 69, 71-73]

References

1. Lucene, Apache Lucene Project, 2008. Available at http://lucene.apache.org/
2. Sustainable Digital Data Preservation and Access Network Partners (DataNet), National Science Foundation, 2007. Available at http://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf07601
3. AAAS *Atlas of Science Literacy, Volume 2*. Project 2061, Association for the Advancement of Science (AAAS) and the National Science Teachers Association, Washington, DC, 2007.
4. AAAS *Atlas of Science Literacy, Volume I*. Project 2061, Association for the Advancement of Science (AAAS) and the National Science Teachers Association, Washington DC, 2001.
5. AAAS *Benchmarks for Science Literacy*. Project 2061, American Association for the Advancement of Science (AAAS), Oxford University Press, Washington DC, 1993.
6. Ahmad, F., Bhushan, S., Gu, Q. and Sumner, T., The concept space interchange protocol. in *Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2004)*, (Tucson, AZ, 2004), Association for Computing Machinery (ACM) Press, 382. Available at http://doi.acm.org/10.1145/996350.996448
7. Ahmad, F., Gu, Q. and Sumner, T., Concept space interchange protocol: A protocol for concept map based resource discovery in educational digital libraries. in *10th European Conference on Research and Advanced Technology For Digital Libraries (ECDL 2006)*, (Alicante, Spain, Poster presentation, 2006), 471-479.
8. Ahmad, F., Gu, Q. and Sumner, T., A technological infrastructure for developing curriculum based on learning progressions. in *Proceedings of the ED-MEDIA 2006 World Conference on Educational Multimedia, Hypermedia & Telecommunications*, (Orlando, FL, 2006), EDMEDIA, 3070-3077.
9. Arms, C. and Arms, W. Mixed Content and Mixed Metadata: Information Discovery in a Messy World. in Hillmann, D. and Westbrooks, E. eds. *Metadata in Practice*, ALA Editions, 2004.
10. Arms, W.Y., Dushay, N., Fulker, D.W. and Lagoze, C. A Case Study in Metadata Harvesting: the NSDL. *Library Hi Tech*, *21* (2), 2003.
11. Arms, W.Y., Hillmann, D., Lagoze, C., Krafft, D., Marisa, R., Saylor, J., Terrizzi, C. and Van de Sompel, H. A Spectrum of Interoperability: The Site for Science Prototype for the NSDL. *D-Lib Magazine*, *8* (1), 2002. Available at http://www.dlib.org/dlib/january02/arms/01arms.html
12. Bartolo, L., Lowe, C., Krafft, D. and Tandy, R., NSDL MatDL: Adding Context to Bridge Materials e-Research and e-Education. in *European Conference on Digital Libraries (ECDL)*, (Budapest, Hungary, 2007), Springer, 499-500. Available at http://dx.doi.org/10.1007/978-3-540-74851-9_51
13. Beck, K. *eXtreme Programming Explained: Embrace Change*. Addison-Wesley, Boston, MA, 2000.
14. Beedle, M. and Schwaber, K. *Agile Software Development with SCRUM*. Prentice Hall, Upper Saddle River, NJ, 2001.
15. Bergmark, D., Lagoze, C. and Sbityakov, A. Focused Crawels, Tunneling, and Digital Libraries. in *European Conference on Digital Libraries*, Rome, Italy, 2002.
16. Birkland, A. and Wilper, C. MPTStore: A Fast, Scalable, and Stable Resource Index *Open Repositories 2007*, San Antonio, Texas, 2007. Available at http://www.slideshare.net/cwilper/mptstore-a-fast-scalable-and-stable-resource-index
17. Boehm, B. and Turner, R. *Balancing Agility and Discipline: A Guide for the Perplexed*. Addison-Wesley Professional, Boston, MA, 2003.
18. Boland, R. and Tenkasi, R. Perspective making and perspective taking in communities of knowing. *Organization Science (Special issue on Electronic Communication and Changing Organizational Forms)*, *6* (4), 1995. 350-372.
19. Bowker, G. and Star, S.L. *Sorting Things Out: Classification and its Consequences*. MIT Press, Cambridge, MA, 1999.

20. Bruce, T.R. and Hillmann, D. The Continuum of Metadata Quality: Defining, Expressing, Exploiting. in Hillmann, D. ed. *Metadata in Practice*, American Library Association, Chicago, 2004.
21. Butcher, K., Bhushan, S. and Sumner, T. Multimedia displays for conceptual discovery: Information seeking with strand maps. *ACM Multimedia Systems Journal (Special Issue on Multimedia Technology for Education)*, *11* (3), 2006. 236-248.
22. Butcher, K. and de la Chica, S. Supporting student learning with adaptive technology: Personalized conceptual assessment and remediation. in Banich, M. and Caccamise, D. eds. *Generalization of Knowledge: Multidisciplinary Perspectives*, Taylor and Francis, London, England, in press.
23. Clark, K.G., Feigenbaum, L. and Torres, E. SPARQL Protocol for RDF, W3C, 2008. Available at http://www.w3.org/TR/rdf-sparql-protocol/
24. Cramer, E.J., Dushay, N., Krafft, D., Hillmann, D., Saylor, J. and Terrizzi, C. Contribution and Collaboration Strategies for the National Science Digital Library: Investigating Technological Solutions to Facilitate Social Evolution of a Collaborative Infrastructure *Joint Conference on Digital Libraries*, ACM, Houston, Texas, 2003.
25. de la Chica, S., Ahmad, F., Martin, J. and Sumner, T., Pedagogically useful extractive summaries for science education. in *Proceedings of the 22nd Meeting of the International Committee for Computational Linguistics (COLING 2008)*, (Manchester, UK, 2008 (to appear)).
26. de la Chica, S., Ahmad, F., Sumner, T., Martin, J.H. and Butcher, K.R. Computational foundations for personalizing instruction with digital libraries. *International Journal of Digital Libraries*, *Special Issue on Digital Libraries and Education*, 2008 (to appear).
27. Dmitriev, P. and Lagoze, C., Automatically Constructing Descriptive Site Maps. in *Eighth Asia Pacific Web Conference*, (Harbin, China, 2006).
28. Dmitriev, P., Lagoze, C. and Suchkov, B., As We May Perceive: Inferring Logical Documents from Hypertext. in *HT 2005 - Sixteenth ACM Conference on Hypertext and Hypermedia*, (Salzburg, Austria, 2005). Available at http://doi.acm.org/10.1145/1083356.1083370
29. Dushay, N. Visualizing Bibliographic Metadata - A Virtual (Book) Spine Viewer. *D-Lib Magazine*, *10* (10), 2004. Available at http://www.dlib.org/dlib/october04/dushay/10dushay.html
30. Dushay, N. and Hillmann, D., Analyzing Metadata for Effective Use and Re-Use. in *DCMI Metadata Conference and Workshop*, (Seattle, 2003).
31. Edmondson, B. Classic Articles in Context for Teachers *Expert Voices: NSDL Highlights*, 2008. Available at http://expertvoices.nsdl.org/highlights/2008/03/31/classic-articles-in-context-for-teachers/
32. Gu, Q., Ahmad, F., Molina, F. and Sumner, T., Dynamically generating conceptual browsing interfaces for digital libraries using SVG. in *3rd Annual Conference on Scalable Vector Graphics (SVG Open 2004)*, (Tokyo, Japan, 2004).
33. Gu, Q., Ahmad, F., Molina, F. and Sumner, T., Novel SVG visualizations for exploring distributed digital libraries. in *4th Annual Conference on Scalable Vector Graphics (SVG Open 2005)*, (Enschede, Netherlands, 2005).
34. Gu, Q., Ahmad, F. and Sumner, T. Resource discovery in distributed digital libraries through visual knowledge navigation. *Journal of Zhejiang University: Science*, *6A* (11), 2005. 1306-1311.
35. Gu, Q., Sumner, T., Bhushan, S. and Ahmad, F., Dynamically generating conceptual browsing interfaces for digital libraries. in *Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2004)*, (Tucson, AZ, 2004), Association for Computing Machinery (ACM) Press, 391. Available at http://doi.acm.org/10.1145/996350.996457
36. Hayes, P. RDF Semantics. McBride, B. ed., W3C, 2004. Available at http://www.w3.org/TR/rdf-mt/
37. Hillmann, D., Dushay, N. and Phipps, J., Improving Metadata Quality: Augmentation and Recombination. in *DC-2004*, (Shanghai, China, 2004). Available at http://students.washington.edu/jtennis/dcconf/Paper_21.pdf

38. Hillmann, D.I., Westbrooks, E.L. and American Library Association. *Metadata in practice*. American Library Association, Chicago, 2004.
39. Hitchcock, S., Bergmark, D., Brody, T., Gutteridge, C., Carr, L., Hall, W., Lagoze, C. and Harnad, S. Open Citation Linking: The Way Forward. *D-Lib Magazine*, *8* (10), 2002. Available at http://dlib.org/dlib/october02/hitchcock/10hitchcock.html
40. Jo, Y., Lagoze, C. and Giles, C.L. Detecting Research Topics via the Correlation between Graphs and Texts *The 13th International Conference on Knowledge Discovery and Data Mining (SIGKDD)*, ACM, San Jose, CA, 2007. Available at http://doi.acm.org/10.1145/1281192.1281234
41. Khan, H. and Maull, K., Realizing the role of digital repositories in educational applications: Supporting content and context. in *Second International Conference on Open Repositories*, (San Antonio, TX, 2007).
42. Khan, H., Maull, K. and Sumner, T., Curriculum Overlay Model for Embedding Digital Resources. in *8th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2008)*, (Pittsburgh, PA, 2008), Association for Computing Machinery (ACM) Press, 74-84. Available at http://doi.acm.org/10.1145/1378889.1378904
43. Krafft, D., Birkland, A., Blake, J., Cornwell, T., Cramer, E.J., Minton Morris, C., Rayle, E.L. and Weatherley, J. NCore Documentation, 2008. Available at http://ncore.nsdl.org
44. Krafft, D.B., Birkland, A. and Cramer, E.J. NCore: Architecture and Implementation of a Flexible, Collaborative Digital Library *Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries*, ACM, Pittsburgh, PA, 2008, 313-322. Available at http://arxiv.org/abs/0803.1500
45. Lagoze, C., Arms, W., Gan, S., Hillmann, D., Ingram, C., Krafft, D., Marisa, R., Phipps, J., Saylor, J., Terrizzi, C., Hoehn, W., Millman, D., Allan, J., Guzman-Lara, S. and Kalt, T., Core Services in the Architecture of the National Digital Library for Science Education (NSDL). in *Joint Conference on Digital Libraries*, (Portland, Oregon, 2002), ACM/IEEE. Available at http://arxiv.org/abs/cs.DL/0201025
46. Lagoze, C., Krafft, D., Cornwell, T., Dushay, N., Eckstrom, D. and Saylor, J., Metadata aggregation and "automated digital libraries": A retrospective on the NSDL experience. in *Joint Conference on Digital Libraries*, (Chapel Hill, NC, 2006), ACM. Available at http://arxiv.org/abs/cs.DL/0601125
47. Lagoze, C., Krafft, D., Cornwell, T., Eckstrom, D., Jesuroga, S. and Wilper, C., Representing Contextualized Information in the NSDL. in *ECDL2006*, (Alicante, Spain, 2006), Springer. Available at http://arxiv.org/abs/cs/0603024
48. Lagoze, C., Krafft, D., Jesuroga, S., Cornwell, T., Cramer, E. and Shin, E. An Information Network Overlay Architecture for the NSDL, Cornell University, 2005. Available at http://arxiv.org/abs/cs.DL/0501080
49. Lagoze, C., Krafft, D.B., Payette, S. and Jesuroga, S. What Is a Digital Library Anymore, Anyway? Beyond Search and Access in the NSDL. *D-Lib Magazine*, *11* (11), 2005. Available at http://dx.doi.org/10.1045%2Fnovember2005-lagoze
50. Lagoze, C., Kraft, D., Cornwell, T., Eckstrom, D., Jesuroga, S. and Wilper, C., Representing contextualized information in the NSDL. in *10th European Conference on Digital Libraries (ECDL 2006)*, (Alicante, Spain, 2006).
51. Lagoze, C., Van de Sompel, H., Johnston, P., Nelson, M., Sanderson, R. and Warner, S. Abstract Data Model *ORE Specification*, Open Archives Initiative, 2008. Available at http://www.openarchives.org/ore/0.9/datamodel
52. Marlino, M., Sumner, T., Kelly, K. and Wright, M. When the Music's Over. in Skinner, K., Halbert, M. and Halbert, M. eds. *Strategies for Sustaining Digital Libraries*, Emory University, Digital Library Publications, Atlanta, GA, 2008, 97-111.
53. Marlino, M.R., Sustainability Issues for Digital Libraries. in *Presentation to the Open Society Archives*, (Budapest, Hungary, 2007).

54. Marlino, M.R., Sumner, T., Kelly, K. and Wright, M., DLESE: A Case Study in Sustainability Planning. in *Open Repository 2008 Conference*, (Southampton, England, 2008).
55. McArthur, D., J. and Zia, L., L. From NSDL 1.0 to NSDL 2.0: towards a comprehensive cyberinfrastructure for teaching and learning *Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries*, ACM, Pittsburgh, PA, USA, 2008. Available at http://doi.acm.org/10.1145/1378889.1378902
56. Minton Morris, C. Telling Great Stories: An NSDL Content and Communications System for Aggregation, Display, and Distribution of News and Features *arXiv.org*, 2005. Available at http://www.arxiv.org/abs/cs.DL/0509094
57. Morgan, R.L., Cantor, S., Carmody, S., Hoehn, W. and Klingenstein, K. Federated Security: The Shibboleth Approach. *EDUCAUSE Quarterly*, *27* (4), 2004. Available at http://connect.educause.edu/Library/EDUCAUSE+Quarterly/FederatedSecurityTheShibb/39889
58. Mutton, P. and Golbeck, J. Visualization of Semantic Metadata and Ontologies *IV '03: Proceedings of the Seventh International Conference on Information Visualization*, IEEE Computer Society, London, UK, 2003. Available at http://www.cs.umd.edu/~golbeck/downloads/IV03.pdf
59. Pan, B., Gay, G., Saylor, J. and Hembrooke, H. One digital library, two undergraduate classes, and four learning modules: Uses of a digital library in classrooms. *Journal of the American Society for Information Science and Technology*, *57* (10), 2006. 1315-1325. Available at http://www.ota.cofc.edu/pan/new-Pan_JASIST_04128_2_submitted-revised-09-26.pdf
60. Pan, B., Gay, G., Saylor, J., Hembrooke, H. and Henderson, D., Usability, learning, and subjective experience: User evaluation of K-MODDL in an undergraduate class. in *Proceedings of the Fourth ACM/IEEE Joint Conference on Digital Libraries (JCDL' 04)*, (Tucson, Arizona, 2004), ACM. Available at http://www.ota.cofc.edu/pan/JCDL04.pdf
61. Pan, B., Hembrooke, H., Gay, G. and Gonsalves, G.C. Bridging the Gap: A conceptual model of the access of digital libraries. *Journal of Digital Information*, *7* (2), 2006. Available at http://www.ota.cofc.edu/pan/Pan-missingLinkPaper.pdf
62. Pan, B., Walker, K., Gay, G. and Saylor, J. Learning Objects in Classrooms. in Koohang, A. and Harman, K. eds. *Learning Object Applications & Future Directions*, Informing Science Institute, Santa Rosa, California, 2006.
63. Payette, S., Blanchi, C., Lagoze, C. and Overly, E.A. Interoperability for digital objects and repositories. *D-Lib Magazine*, *5* (May), 1999.
64. Paynter, G., Developing Practical Automatic Metadata Assignment and Evaluation Tools for Internet Resources. in *Fifth ACM/IEEE Joint Conference on Digital Libraries (JCDL 2005)*, (Denver, Colorado, 2005), ACM. Available at http://doi.acm.org/10.1145/1065385.1065454
65. Phipps, J., Hillmann, D. and Paynter, G., Orchestrating Metadata Enhancement Services: Introducing Lenny. in *DC-2005*, (Madrid, Spain, 2005). Available at http://arxiv.org/abs/cs.DL/0501083
66. Saylor, J. and Morris, C.M. The National Science Digital Library: An Update on Systems, Services and Collection Development. *Science and Technology Libraries*, *26* (3-4), 2007. 61-78.
67. Staples, T., Wayland, R. and Payette, S. The Fedora project: An open-source digital object repository system. *D-Lib Magazine*, *9* (4 (April)), 2003. Available at http://www.dlib.org/dlib/april03/staples/04staples.html
68. Sumner, T. From Digital Libraries to Educational Cyberinfrastructure, Invited Talk, Cultural Communication and Computing Research Institute, Sheffield Hallum University, England, 2005.
69. Sumner, T. The NSDL Strand Map Service: A Networked Knowledge Organization and Visualization System for K-12 Education *Invited Presentation on Next Generation Knowledge Organization Systems: Integration Challenges and Strategies*, The 7th Networked Knowledge Organization Systems (NKOS) Workshop, Held in Conjunction with JCDL, Denver, CO, 2005.

70. Sumner, T. Panel Title: Transforming Digital Content into Learning: The Potential and Challenges Facing Educational Digital Libraries, Panel Organizer and Participant, Educause Center for Applied Research Summer Symposium, Boulder, CO, 2007.
71. Sumner, T., Ahmad, F., Bhushan, S., Gu, Q., Molina, F., Willard, S., Wright, M., Davis, L. and Janee, G. Linking learning goals and educational resources through interactive concept map visualizations. *International Journal on Digital Libraries*, *5* (1), 2005. 18-24.
72. Sumner, T., Ahmad, F., Gu, Q., Molina, F., Willard, S., Wright, M., Davis, L., Bhushan, S. and Janee, G. A web service interface for creating concept browsing interfaces *D-Lib Magazine*, 2004. Available at http://www.dlib.org/dlib/november04/sumner/11sumner.html
73. Sumner, T., Bhushan, S., Ahmad, F. and Gu, Q., Designing a language for creating conceptual browsing interfaces for digital libraries. in *Proceedings of the 3rd ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2003)*, (Houston, TX, 2003), Association for Computing Machinery (ACM) Press, 258-260.
74. Terrizzi, C. Visual Arts and Communication at the Nexus of Usability in the National Science Digital Library User Interface *HCI, the Arts and the Humanities*, King's Manor, University of York, UK, 2003.
75. Terrizzi, C. and Cramer, E.J., Outside the box: Incorporating a Novel User Interface for the National Science Digital Library into Everyday Home Activity. in *Proceedings of the CHI Workshop: Designing Culturally Situated Technologies for the Home*, (Fort Lauderdale, Florida, 2003).
76. Wright, M., Marlino, M. and Sumner, T. Meta-Design of a Community Digital Library. *D-Lib Magazine*, *8* (5), 2002. Available at http://www.dlib.org/dlib/may02/wright/05wright.html
77. Wright, M., Sumner, T., Moore, R. and Koch, T. Introduction to Special Issue on Digital Libraries and eScience. *International Journal on Digital Libraries*, *7, Numbers 1-2*, 2007. 1-4.
78. Yookyung, J., Carl, L. and Giles, C.L. Detecting research topics via the correlation between graphs and texts *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, San Jose, California, USA, 2007.
79. Zia, L.L. The NSF National Science, Technology, Engineering, and Mathematics Education Digital Library (NSDL) Program. *D-Lib Magazine*, *8* (11), 2002. Available at http://dlib.org/dlib/november02/zia/11zia.html