



Core Integration Report: A Review, 2001 - 2005

Kaye Howe, Principal Investigator, UCAR
Dean Krafft, Principal Investigator, Cornell University
Kate Wittenberg, Principal Investigator, Columbia University

February 8, 2006



Supported by the National Science Foundation

Core Integration Report: A Review, 2001 - 2005

Introduction

This report outlines CI progress to date, lessons learned, and current plans and strategies. As with any maturing organization, our experience has led to a series of course corrections and reconceptualizations of NSDL and its directions. Some of our observations may serve as valuable input to NSF as it structures other large and distributed efforts.

CI is charged with integrating the collections, services, and targeted research of the NSDL projects into a coherent library, along with engaging the community, providing technology, and operating core services. The CI team is a partnership among the University Corporation for Atmospheric Research (UCAR), Columbia University, and Cornell University. UCAR has been responsible for central office functions, outreach and community support, and technical support for NSDL.org. Cornell has provided core library infrastructure development, operations, and communications. Columbia has engaged publishers in building the library, and developed access management solutions.

The early and continuing management challenge of CI was to build a collaborative work environment and to divide the work among groups that originally applied to do the whole project on their own. The resulting collaboration included the three primary partners that we have today (UCAR, Cornell, Columbia) and a small group of subcontracts. We immediately discovered the challenges of distributed work teams that included part-time staff who held other appointments and we wrestled with the problem of institutional vs. NSDL identity.

Over time, the structure of Core Integration has evolved in significant ways:

- We eliminated some early subcontracts and reallocated funds to outreach, communications, evaluation, technical development, Pathways support, iVia and the prototypical ENC Middle School portal in Math and Science. In some instances (e.g., search) we cancelled an unsatisfactory subcontract and pulled that work and resources back to Core Integration.
- We eliminated a senior position at UCAR and reallocated those funds to a project manager position. This position oversees the distributed work teams and has begun building a repeatable software delivery process.
- A full-time PI was appointed at Cornell to oversee technical development.
- We assigned core staff to support Pathways integration.
- We developed working teams across the three Core Integration institutions.
- We expend time and resources on face-to-face meetings, as well as frequent conference calls, to make certain that we coordinate our intensely collaborative work.

It is fair to say that Core Integration today is less a set of projects and more an operational organization (Appendix A).

Integrating Collections, Services, and Research

Project Integration

There have been 202 projects funded in the collections, services, targeted research, Pathways, and Core Integration tracks (Appendix F). The current CI team was initially funded in FY 01. Of the 66 unique collection projects¹ funded, 47 have been integrated (71% integration), with 34 providing item-level metadata. Access to item-level metadata was important early on to help users search for specific resources in the library, rather than being directed to a home page or other intermediate page on the collection provider's site. The number of integrated collection projects will continue to increase: eight collections with collection record entries are actively working on providing metadata harvesting services (Open Archives Initiative—OAI), another eight are still working on their collections, and four collections are online, but still need to submit records.

Several collections represented an integration challenge in that they do not provide resources per se, but complex data that can only be accessed through their services. Collections such as *eSkeletons*, the *Digital Archive Network for Anthropology*, and the *Digital Library of Microstructures* provide visualization software to access and manipulate complex data sets. Although they provide collection records that essentially point to their services, the structure of their data and services make it difficult to provide Dublin Core metadata to comply with the original metadata-centric repository architecture. These collections are more appropriately treated as services for the purpose of integration.

The 59 unique service projects represent a wide range of human-mediated services (such as workshops) and computer-mediated services (such as collection building and workflow tools). Six services have submitted collection records to the repository and of those, two have provided item-level metadata. Two services have been integrated into the core infrastructure: *AskNSDL* and *Instructional Architect*. Of software projects whose work is complete, nine are represented in the NSDL Library Builders' tool page² and four have hosted workshops. Eleven services represent work that enhanced the specific project offerings of collections and services.

From the beginning CI was challenged with the unintended consequences of grant funding and its timing. By FY 01 (the first year the current CI was funded), 29 projects had already completed their first year of work in building collections and services, or performing DL research. Of those 29, six were the initial Core Integration track projects. Several of the FY 01 projects were established collection providers. Because the order of projects was not phased from the beginning, there were lost opportunities as projects either ran out of money before they could be integrated or created services that were built before CI had an architecture to support them. These consequences of phasing resulted in some community frustration with our progress in library building. CI is currently reviewing these earlier projects to see what can be recovered and integrated. The "thousand flowers" approach produced by well intentioned panels of peer reviewers did not have as their primary goal and responsibility the production of coherent collections, services, and research in service of the larger goal of building a coherent STEM education library.

¹ Of the 202 projects, several were collaborations that resulted in only one unique collection or service, and in later years, some projects were continuation funding that extended existing work.

² Library builders tools can be found at http://nsdl.org/resources_for/library_builders/tools.php?pager=tools

The National Visiting Committee (NVC) understood this situation and recommended a more focused approach for NSDL. In response (and much to its credit), NSF initiated the Pathways track in an FY 04 solicitation. Pathways represent a good focusing strategy because they: 1) help NSDL connect to specific learner communities and their cultures; 2) provide in-depth knowledge of educational needs of that community, and 3) help with integration of other projects and collections. The entire ensemble of Pathways provides a locus of collaboration that will benefit the whole STEM education community (Appendix B). In addition, a signed Memorandum of Understanding (MOU) between CI and the Pathways addresses issues of deliverables and operational schedules as well as driving increasingly compelling collaborations (Appendices C and D). The MOU and the agreements it details are evolving, even as the work between and among us evolves. Most of all, the MOU makes it clear that an NSDL Pathways grant is not simply an infusion of dollars for a worthy stand-alone project, but a key building block for NSDL as a major STEM education resource.

In retrospect, more attention to the phasing of the elements of NSDL would have fostered better and earlier integration. On the other hand, NSDL is a classic start-up operation that must, almost necessarily, learn by doing. In going from the more disparate Collections track to the Pathways approach, NSF corrected deficiencies and inefficiencies and made clear the obligation and responsibility to participate in building NSDL through collaboration.

Collection Building and Quality

From the beginning, CI laid out Dublin Core and OAI standards to provide a common framework for metadata ingest and aggregation. This strategy acknowledged that existing collection providers were heavily invested in their own metadata frameworks, and CI provided a low barrier mechanism to ingest their information. When NSDL.org first launched in 2002, there were already 118 collections available.

CI continues to grow the NSDL collection with 580 collections representing well over one million resources. We are also pursuing partnerships to provide highly edited content from professional societies and other publishers to NSDL users. John Saylor, the NSDL Collection Development Director, has created an initial cadre of 29 Selectors to look actively for materials and add them directly into NSDL, using the automated NSDL Resource Recommendation System (RSS). He has solicited librarians, academics, and teachers and asked for their help in more than 50 disciplinary and interest areas (Appendix E). This group will continue to grow.

To provide additional metadata from non-NSDL collections, we have subcontracted with the University of California at Riverside for their iVia software, which crawls selected web sites and creates standard metadata for the resources it finds. This software will become an important tool for Pathways and other collection providers to help them more easily create collections and submit standard metadata to NSDL. ComPADRE and MatDL Pathways are currently pursuing the use of iVia to expand their collections. iVia is a very promising tool that is being developed, in part, by our use of it. CI staff are currently using iVia to crawl additional collections so that these resources can be accessed through NSDL search.

We have established the Collection Development Advisory Board (Appendix E), whose role is to advise the Director of Collection Development on all issues relating to building and maintaining the overall NSDL collection, including policy. The members represent several library associations and university and school libraries across the country, including, most recently, the NSF Librarian. CI has also asked

Rick Luce, Research Library Director at Los Alamos Laboratory, to chair a small Library Review and Planning Committee, which will review the overall operations and processes of the library and make recommendations for improvements. This review comes at a particularly efficacious moment, as John Saylor will be returning to his regular position as Engineering Librarian for Cornell in October, 2006. We expect good advice and perspective from the Review Committee on the job scope and focus for the position that will be left open by Saylor's return to the Cornell Library.

CI has relied on the metadata quality standards of each NSF-funded project. As we started to ingest metadata from collections, inconsistent quality standards between collections resulted in uneven metadata quality, which offered challenges in presenting coherent search results to end-users. Metadata experts are expensive and in short supply; the range of expertise in the projects in creating metadata resulted in uneven metadata quality. Our course correction included actions that take the best advantage of current tools and practices in the community to provide better service to the end user. This includes using automatic processes such as iVia, and community/end-user-based processes to annotate or add to metadata, such as the Middle School Pathways work done by ENC to improve K-12 resource metadata.

An improved understanding of how resources can be designed for reuse has helped highlight areas that we now know require further research, including definition of the criteria required to constitute good learning resources and the granularity needed to support reuse of resources.³ Issues such as digital rights management and tracking provenance will be key issues for the future as end users find creative ways to reuse NSDL materials.

Engaging the Community

Building the Developer Community

The first NSDL PI meeting was held in 2000, prior to the creation of the current CI team. This group defined key issues and priorities and developed a process for establishing a governance structure, which led to the *Pathways to Progress* document⁴ in 2001. By the 2002 Annual Meeting, it was clear grantees were interested in building a community with a common vision of the future. Survey results from successive meetings underscore the importance of the Annual Meeting in support of community collaboration and networking, and the leveraged development that collaboration makes possible.

We established a project liaison position (Susan Jesuroga) in our first year to provide ongoing communications and integration assistance. We provide assistance to the Policy Committee (PC) for face-to-face meetings and regular conference calls, developing policy, and performing research as needed. The Policy Committee created three initial NSDL policies on key issues of process, privacy, and metadata⁵, and established a set of standing committees in key interest areas: Content, Community Services, Evaluation, Sustainability, and Technology. Each standing committee has a CI liaison who helps outline and refine issues for community input. CI has also supported—often in partnership with a standing committee—community workshops on sustainability, publisher and user engagement, web analytics, and evaluation. The

³ Reusability is a complex issue, which includes metadata, context, intellectual property (IP), and interoperability challenges. Foundational work was completed by the NSDL project on reusability. <http://www.reusablelearning.org>

⁴ Pathways to Progress can be found at <http://doelib.comm.nsdlib.org/PathwaysToProgress.pdf>.

⁵ Current policies can be found at <http://policy.comm.nsdlib.org/docs/>.

feedback and discussions provided through committees and workshops have been invaluable to both CI and the community as a whole (Appendix H).

While the PC had early success in developing specific policies for NSDL, the challenge of creating and implementing an effective governance structure has proven more difficult, and the committee struggled to define its role in the NSDL organizational structure. Policy Committee members differed as to whether the committee was advisory to CI or had authority over it and our day-to-day operations. That lack of clarity—ultimately resolved on the side of an advisory role—made for significant stress and misunderstanding. The Standing Committees varied widely in finding appropriate work that fit the fact that they were volunteers with other significant responsibilities and little time. Despite that, work that has mattered greatly to the development of NSDL was and continues to be accomplished.

As NSDL work moves forward, it is clear that committees and working groups are the appropriate location for defining broad community needs (such as helping to define web metrics and evaluation requirements) and representing community-wide issues and concerns (reflected through policies), while CI focuses on the development and operations of core services, the transfer of research and policies to operations, and our most fundamental mission—the support of community integration and collaboration.

As new projects have been added to NSDL, we have worked to facilitate direct project-to-project integration in addition to supporting project integration with the core infrastructure. As an example, CI facilitated the collaborative efforts of the Achievement Standards Network (ASN) project with other groups, including Syracuse with their Content Alignment Tool (CAT) and the Teachers' Domain Pathway, to solve the issue of dealing with educational standards. CI plans to connect to the ASN database and provide the CAT on NSDL.org. Other non-funded stakeholders have joined the community in support of the STEM education focus. Textbook and journal publishers are contributing metadata and selected content, while other groups like EPIC (Engaging People In Cyberinfrastructure) are helping to promote NSDL in the spirit of seeking synergies to achieve common goals. Still others see the library as an education and public outreach opportunity—the Institute of Electrical and Electronics Engineers (IEEE) helped recruit *AskNSDL* experts for ESTME Week (Excellence in Science, Technology, and Mathematics Education) in 2004, resulting in the participation of 48 IBM employees.

Communications and Outreach to NSDL Users

Outreach

Engaging a broad base of stakeholders to develop and use the library is critical to NSDL's long-term success. Current outreach and communications activities cultivate connections with a wide range of groups to build awareness, increase usage, grow library resources, create new advocates, and aid sustainability. Outreach and communications efforts have steadily expanded and diversified since the library's early stages. As of the initial release of NSDL.org in December 2002, CI had a single brochure and a visionary multimedia presentation on CD-ROM. Our visibility was minimal outside of higher education and our presence at professional meetings was limited to a handful of papers delivered at digital library conferences.

When CI was formed, a first and necessary task was to build both the NSDL infrastructure and a community of developers. Those necessary activities led us to an inward focus in the initial versions of NSDL.org that had the unintended consequence

of keeping end users at a distance. Feedback from our target users and professional colleagues underscored the problem of focusing on the promise of NSDL before services and tools were fully developed: our valued user audiences—teachers and librarians—came to a site that confused them about its intent and use. At the first several annual meetings, we were appropriately criticized for seeming to focus on technical issues to the exclusion of our educational mission. We also assumed a level of user knowledge about digital libraries and their uses that user testing would have revealed as inaccurate.

NSDL sought to address these issues by focusing on building general awareness about the library, its activities and resources, and to attract a base of regular users in varying sectors of the STEM education community. Tactics included increasing the presence of NSDL at major professional conferences where potential users and contributors gather such as the American Association for the Advancement of Science (AAAS), the National Science Teachers Association (NSTA), the National Council of Teachers of Mathematics (NCTM), the Association of Science-Technology Centers (ASTC), and the American Library Association (ALA), as well as the broad dissemination of promotional materials to target audiences. Concurrently, Core Integration staff networked with organizations in K-12 and higher education as well as the library and museum communities. These cross-organizational conversations focused largely on awareness-building and exploring common goals.

CI has developed 18 targeted brochures and fliers, and distributed an estimated 10,000 copies in 2005. NSDL has exhibited at 26 conferences, reaching over 200,000 people involved in higher education, libraries and museums, K-12 schools, and scientific research institutions. CI also coordinates and supports NSDL project staff in their own conference and workshop activities. Coordination is a key element in ensuring consistent branding and message. We organized well-attended briefings on Capitol Hill, on the Senate side in May, 2005 and on the House side in September 2005. The September briefing was sponsored by the House STEM caucus and NSDL speakers were introduced by Congressmen Ehlers and Udall, co-chairs of the caucus. The newly appointed NSF Deputy Director, Kathie Olsen, also spoke and commented on the significance of NSDL in the NSF education portfolio.

Outreach work now extends beyond awareness-building to include training users on broader information literacy and digital resource integration skills. CI offered three pilot online seminars in 2005 (two for K-12 teachers and one for university faculty). Plans are currently in development to offer monthly online seminars throughout 2006 in conjunction with Pathways and other NSDL projects as part of a partnership with the National Science Teachers Association. Outreach Director Susan Van Gundy is facilitating an effort of the Community Services Standing Committee to present a series of online community discussions that showcase tools and services emerging from NSDL projects. The first of these online sessions, dubbed *Tool Times*, was held in November 2005. CI is providing use of the web conferencing platform and a session moderator, as well as assistance in planning and promotion. A list of upcoming *Tool Times* and archives of previous sessions, will be made available at <http://commserv.comm.nsd.org>.

Support of community workshop efforts can be as simple and direct as providing print and presentation materials, promoting community workshops on NSDL.org, or soliciting promotion within specific user communities for workshops hosted by CI. Other efforts involve more in-depth coordination, such as the Pathways' user workshops, where we are managing a series of workshops for Pathway audiences

with the primary goal of increasing use of the library. While each Pathway is determining the specific content, participant recruitment plans, and appropriate venue for their event, all workshops will share strategies for evaluation and targeting of participants who will disseminate information to additional users. For example, the MatDL pathway will convene teams of researchers and educators from Materials Research Science and Engineering Centers (MRSECs), while the CSERD pathway will be targeting teams from among NSF's Math and Science Partnerships awardees. By coordinating workshops at various levels across projects, NSDL strengthens its reputation with user communities and collects participant data that can be later analyzed to assess impact. Our shared goal is to give users a better understanding of the interdisciplinary breadth and depth of NSDL. Seminars and workshops are also an expression of our maturing partnership strategy. As the richness of library resources, tools, and services develop, so too does its usefulness to educators.

Communications

Before the release of NSDL.org, CI laid out plans for increased communications about NSDL and its mission. NSDL has several email lists used for communications including a general list (nsdl-all) with over 700 subscribers and a Whiteboard Report subscribers list with around 450. Whiteboard Report (originated and edited by Carol Minton Morris) started in 2000 as a monthly publication about NSDL activities, largely aimed at the developer community, and is now published bi-weekly and available through links and search at NSDL.org. Distribution of the Report extends further than the list, including circulation through the Cornell Library and Outreach lists, the Humanities, Arts, Science and Technology Advanced Collaboratory (HASTAC), and several NSDL projects, along with cross-posting of news items with the Electronic Privacy Information Center (EPIC) *CyberInfo Beat Newsletter* (<http://www.eotepic.org/>).

As a service to NSDL users, NSDL headline news and events have been highlighted on the NSDL.org site since the beginning of the project. The editorial process and scope of content have grown over time, as the site has changed to provide highly edited content areas for specific audience groups. The hand-crafted content and hand-selected resources present users with a tailored view into NSDL. They include:

- News - news about the library, events, and collections of interest to NSDL users
- Top Picks - a stable set of exemplary resources from NSDL that have been judged as most useful for an audience group
- Resources of Interest - a rotating set of resources selected from NSDL that highlight the breadth of NSDL
- Using NSDL - resources from NSDL that can be put to use immediately: a lesson, event or an activity
- Research Articles - research articles of interest to an audience group
- NewsFeeds - external RSS news sources from the STEM community

In 2003, with the help of the Educational Impact and Evaluation Standing Committee, CI published the first NSDL Annual Report and continues publishing Annual Reports in conjunction with the Annual Meeting. Over 3,000 copies have been distributed at national meetings and conferences, through projects and collaborators each year since it was first published. The report is also available online at NSDL.org, and in 2005 we created the first "NSDL Annual Report Supersize" version, which offers expanded project profiles online.

Several other advancements in outreach and communications are in development. To enhance targeted communications to different stakeholder audiences, we are migrating from manually assembling content for each communications instance to an automated framework that will facilitate the repurposing of content in multiple publishing outlets. This new publishing system, called *On Ramp*, will greatly increase our communications options, and will push information to target audiences. We are in the midst of developing *Expert Voices*, an educational blog that will allow students and educators to read about and engage scientists in continuing discussions.

NSDL Publisher Partnerships: Goals and Strategies

In its work, the Core Integration team has proceeded from the premise that in order for the NSDL to become a resource of choice, used frequently by a broad range of teachers and students on a national scale, it is necessary to engage the interest and participation the scientific publishing community. This community includes both non-profit and for-profit organizations that control a substantial percentage of the high-quality educational science content used by researchers, teachers and students.

To this end Kate Wittenberg and Mike Luby, of the CI team, undertook steps to engage this community in a collaborative and productive manner, so as to insure that the NSDL becomes a strong and valued partner rather than a competitor to the traditional science publishing community. Science publishers possess assets that are critical to the success of the NSDL, including an efficient and stable mechanism for acquiring and peer-reviewing high quality content from scientists and science teachers; an effective system for editorial development, design, and production of this content; excellent market research and evaluation mechanisms, established models for contracts, licenses, copyright, and intellectual property management, and a reliable system for marketing and sustainability. In addition, many of these publishers work with vendors who provide technical infrastructure and support for schools.

Through its publisher relations efforts, the CI team has established an effective means to engage the science publishing community, including mechanisms for harvesting metadata and enabling controlled access to publishers' content. These activities will ensure that NSDL reaches its full potential as a functional, valued, and highly used resource, and will serve as a model for partnerships with other collaborators. Appendix G provides an annotated list of current publisher agreements.

Providing Technology and Core Services

Library Infrastructure

In the beginning, the CI team designed the infrastructure to support three functions: selecting web-based STEM resources, searching across them, and facilitating access to them. The resulting architecture represents a union catalog, which was implemented as a metadata repository (MR) containing Dublin Core records for resources held by NSDL collection projects and other contributing organizations. The metadata records, which include URLs to the corresponding digital resources, are ingested into the MR via OAI. During ingest, these records are processed to normalize dates and various controlled vocabulary elements.

The union catalog approach was a reasonable decision, given the state of the digital library world at the time: metadata standards were well-developed and OAI was an accepted protocol for metadata transmission. Collections on the leading edge of

digital library technology were able to supply metadata to NSDL without much assistance.

This approach revealed a number of implementation problems over time. The most outstanding of these relate to uneven metadata quality across collections and problems with OAI validity, especially XML-schema compliance. Even though standards and protocols were well-documented and widely used in the DL community, some collection projects did not have the cataloging expertise to create quality metadata or technical help to deploy an OAI server. Busy building and deploying infrastructure, we did not have the resources to provide extensive hands-on help for those collections. As a result, the administrative costs of maintaining the MR have been unexpectedly high.

Even for collections that were successfully harvested initially, we would sometimes encounter new problems with their metadata and/or OAI service as their collections changed over time. The requirement of strict adherence to XML schemas, UTF-8 character formats, and OAI protocol proved a recurring challenge. Collection providers are provided automated emails detailing harvest issues, but the OAI harvest success rate has remained in the 50-75% range. This despite the fact many of the collections have been in production for over a year, requiring CI staff to work with the collections to identify and fix their problems before attempting repeat harvests.

The early emphasis on metadata also meant we postponed creating a structure to support routine service integration. Any service integration into the core infrastructure required hand-crafted solutions that were not scalable. In addition, some collections are more service-like in nature: they have data resources, but need a user-friendly service on top of the data to be of use. Finally, we could not ingest and make use of other forms of information, such as annotations and reviews, or actual content.

As a library, NSDL needs to point to the most useful resources for STEM education. To support learning, it needs to give those resources a learning context to help the end user meet specific educational goals, including the information and services needed to reuse resources. This combination of early integration challenges and educational needs led to the decision to build a Fedora-based repository, called the NSDL Data Repository (NDR), which allows us to gather and use information beyond metadata, including content.

Fedora is the Flexible and Extensible Digital Object and Repository Architecture—an information model and repository architecture for encapsulating and securing multimedia content. In non-technical terms, it provides the means to ingest, manage, and provide access to complex and dynamic information. The Fedora-based NDR will support MR functions such as OAI ingest and exposure of Dublin Core metadata while extending the types of information that can be managed. Fedora increases the options available for building user services, including resource discovery, access, repurposing, and reuse.

Fedora work began in 1997 as a DARPA and NSF-funded research project at Cornell University. It grew into an open-source project in collaboration with the University of Virginia, funded through the Mellon Foundation. Work is focused on creating a production-quality content repository system and the CI team has worked closely with the Fedora team in addressing scalability issues exposed by the NDR

development. The Fedora software has been deployed in a variety of applications including institutional repositories, archives, museums, and commercial library projects besides NSDL.

The NDR allows us to manage relationships among objects in the repository, so we can model a resource-centric, as opposed to a metadata-centric view of the online world. The metadata-only view limited our ability to represent multiple references to a single resource in search results, and restricted resource information to that provided by a single collection provider. Managing resource-centric views allows us to provide myriad information to the end user, including metadata, reviews, annotations, educational standards, collection provider background, and services that can act on those resources. With this wide ranging information, we can start to provide a broad context about resources to users.

In addition, we are currently working with the Ockham Project⁶ in NSDL to further refine and implement their digital library service registry. The registry is a directory of Internet-available digital library services, which enables any service provider to discover and interface to other online services. We will participate as an Ockham node, which enables the online discovery of NSDL services like search and OAI harvest by the broader digital library community. We will also send staff to an upcoming workshop on service registration in which participants will begin to lay out a plan to create a service registry model that can be adopted and used by the broader digital library community.

Library Services

Early NSDL library services included search, browse, and limited exhibits, which consisted of hand-selected information from the library (e.g. new in the library, available tools and services). Initial CI subcontracts provided archiving and search services, along with support for *AskNSDL* and research into handling geo-spatial data. Over time, CI management decided to consolidate search work at Cornell, using open source software, and reallocate the geospatial resources to other library services and other development efforts within CI.

NSDL.org went online in December 2002. In late 2005, we had 580 collections representing over a million resources and are doing routine OAI harvests of metadata. The site supports search and browse of resources, archiving of publicly available resources, news and events, and a reference desk service (*AskNSDL*). In November 2005, the site received over 13,000 visitors, an increase of almost 50% over April 2005. Pathways web metrics have been introduced only relatively recently, and continue to be introduced. Preliminary analysis based on ten weeks of traffic between mid-November 2005 and the end of January 2006 suggests, very approximately, that traffic across nine NSDL sites (including NSDL.org) approximates 100,000 visitors per month and is trending up. The number of referrers (i.e. locations from which visitors come to the site) has also grown. In fact, there is a very "long tail"⁷ to the distribution of referrals. Twice as many referrals come from locations other than search engines including:

⁶ More information on OCKHAM can be found at <http://www.ockham.org/>.

⁷ In 2004, Chris Anderson first argued that products that are in low demand or have low sales volume can collectively make up a market share that rivals or exceeds the relatively few current bestsellers and blockbusters, if the store or distribution channel is large enough. "The theory of the Long Tail is that our culture and economy is increasingly shifting away from a focus on a relatively small number of 'hits' (mainstream products and markets) at the head of the demand curve and toward a huge number of niches in the tail." <http://www.wired.com/wired/archive/12.10/tail.html>

- Previously bookmarked NSDL.org pages, and/or directly from URLs typed directly into the browser navigation bar
- Links embedded in non-search engine pages

Currently, basic services in NSDL.org include login, search and browse, archiving, news and events, audience-specific content, and *AskNSDL*. The login service is based on Shibboleth, an Internet2 protocol. Search now uses the Lucene open source software and current work focuses on presenting resource-centric search results. The archiving service, a subcontract with the San Diego Supercomputing Center, routinely stores 6 terabytes of publicly available NSDL metadata, resources, and related links. News and content are handled through Scout Portal Toolkit (from Internet Scout at the University of Wisconsin) and PHP programming. *AskNSDL* uses software developed at Syracuse. Outside of NSDL.org, CI also provides a site crawl and metadata generation function through its contract with the University of California at Riverside, for the iVia software.

In providing context for specific audiences, CI shifted some development resources to two key areas: *On Ramp* and *Expert Voices* (discussed previously). Both are under development, and take advantage of the NDR to associate additional content and information with library resources. For example, if a scientist is blogging about black holes and references a resource in NSDL, the blog entry will be associated with that resource, providing wider context than the resource alone. In publishing NSDL news items, if an article talks about how a teacher has used a particular resource in the classroom, the related information from the article will also be available with the resource.

As *Expert Voices* comes online, CI will move away from *AskNSDL*. The reference desk service seemed a necessary part of the initial site, but usage is consistently low in spite of sustained outreach, and there is a tendency of students to use it as a homework hotline. With *Expert Voices*, NSDL can better model scientific inquiry and discussion for its users. With *On Ramp*, CI will move away from Scout Portal Toolkit to create and support the distribution of news, events and audience-specific content in multiple venues.

In an effort to improve the user experience, CI is devoting resources to creating a community sign-on environment, which builds on the Shibboleth work done so far. Shibboleth leverages existing campus identity and access management infrastructures, but CI discovered that many projects did not have the resources needed to implement Shibboleth, although several wanted to support an authentication system at their site.

Starting with the Pathways, CI is creating a community log-in, so that a user registered and logged in at one NSDL site can seamlessly move to another NSDL site without additional log-ins. This concept was first tested between the NSDL.org site and *AskNSDL*, and we are working on ways to standardize the user attributes that are shared between sites, and provide programming support and software. Community sign-on will surface both policy and community issues, (such as privacy issues and defining the range of information about users that sites agree to share), that will need to be resolved by discussions among CI, the Pathways, and the Policy Committee.

Looking forward to new services, CI will begin discussing issues of preservation with the help of the Pathways. Archiving helps to retrieve missing content in the short

term, but we recognize the challenges of providing continued access to resources given ongoing changes in operating systems and software. Also, with the new NDR, we can also consider hosting content for collections whose resources might otherwise be lost. As in providing any new service, cost-benefit analyses will be necessary to understanding how best to expand library services.

Significant issues for our next report

Evaluation

Evaluation efforts were initially undertaken in the Educational Impact and Evaluation Standing Committee (EIESC). Their early work focused on community building and later moved on to helping projects do internal evaluation. This included a pilot project on web metrics that attempted to collect and correlate site statistics over several NSDL sites. The EIESC had many talented participants with an ambitious agenda, but faced the reality that volunteers do not have time to carry out complex evaluation studies or the capacity to review the data coming out of such studies. The EIESC and others in CI recognized that NSDL also needed an evaluation plan that focused on the end users of NSDL services. CI had hired HCI (human-computer interaction) staff and contractors to improve site usability, but the question of overall evaluation did not move forward in the early years of NSDL.

In recent months, we hired an evaluation specialist (Mick Khoo) to start basic program assessment. To date, Khoo has created a meta-framework for evaluation that focuses on the library resource development lifecycle, which can guide data-driven operational decisions and will eventually shape longer term research questions. In addition, early analysis of the web metrics data is beginning to yield a picture of overall resource use. This is a key issue for NSDL because overall usage is more than the number of visitors to NSDL.org; it also includes access to the resources on the distributed sites that make up NSDL and usage data on the whole network of NSDL sites, most importantly, the Pathways.

Sustainability

NSDL is currently entirely NSF-funded. CI is working closely with the Sustainability Standing Committee to identify successful sustainability scenarios for NSDL projects. CI has the responsibility for developing business and organizational models for the NSDL program itself. There is considerable overlap with the project sustainability work.

We currently have reports done by Dan Greenstein and his staff of the California Digital Library (with an NSDL grant⁸) as well as work done by Brad Edmondson and Stuart Takeuchi on models, strengths and weaknesses, and, in Edmondson's report, interviews with significant NSDL community members.

It seems clear that the strong center of NSDL as it goes forward is the collaborative partnership between and among Core Integration and the Pathways, now nine in number plus the Middle School Math and Science prototype. That center is importantly supplemented by other significant NSDL collections and services, increasingly integrated, as well as non-NSDL funded partners who find value in being part of NSDL. Realistic plans for sustainability must focus on that strength. We

⁸ Greenstein project: *Adding Value to the NSDL by Integrating it into Academic Libraries: A Business Proposition and a Service Enhancement*; NSF Award #333710.

expect to consult extensively on business models with the NVC as well as the Policy and Sustainability committees.

We will further address both these issues in our next report.

Conclusion

Digital libraries offer universal access to online resources and hold the promise of widespread reuse of digital resources for education. In *What is a Digital Library Anymore, Anyway? Beyond Search and Access in the NSDL*, Carl Lagoze, of the Core Integration team, writes:

In the age of Google, what is a digital library anymore, anyway? Just asking the question is bound to raise passions. Despite our zealous defense of the successful work of the digital library community over the past decade, the amazing success of commercial web search engines has changed the playing field. Search and access over a set of resources, while important to any digital library, are not sufficient. Digital libraries need to distinguish themselves from web search engines in the manner that they add value to web resources. This added value consists of establishing context around those resources, enriching them with new information and relationships that express the usage patterns and knowledge of the library community. The digital library then becomes a context for information collaboration and accumulation—much more than just a place to find information and access it.⁹

A review of education technology literature suggests that to add value, the NSDL must collect and manage context about resources to ensure adoption and successful reuse in the classroom. NSDL architecture must support not only resources and descriptive metadata, but also the relationships among these resources and a wide range of contextual information about learners, their needs, their learning environments and communities.

CI has taken substantial steps towards building an environment rich with context by:

- Creating the NSDL Data Repository, which supports relationships among diverse pieces of information about a resource
- Developing *Expert Voices*, which supports scientific communication and discussion
- Integrating and collaborating with the Pathways, who bring experience and knowledge of specific learner audiences as well as providing a focus for technical and educational collaboration
- Facilitating service integration of several education standards efforts to provide the end user resources aligned to state and national educational standards
- Improving the NSDL.org communications through *On Ramp* and providing content tailored to specific audiences
- Reaching out to users through professional development and partnerships

Over the next year, in addition to completing and implementing the work detailed in this report, we want to deepen the collaborative work with the Pathways projects. Fedora, iVia, the standards work from ASN and CAT, joint work on increasing NSDL

⁹ Lagoze, C., Krafft, D., Payette, S., and Jesuroga, S. What is a Digital Library Anymore, Anyway? Beyond Search and Access in the NSDL. *D-Lib Magazine*, 11 (11). Available online at <http://www.dlib.org/dlib/november05/lagoze/11lagoze.html>

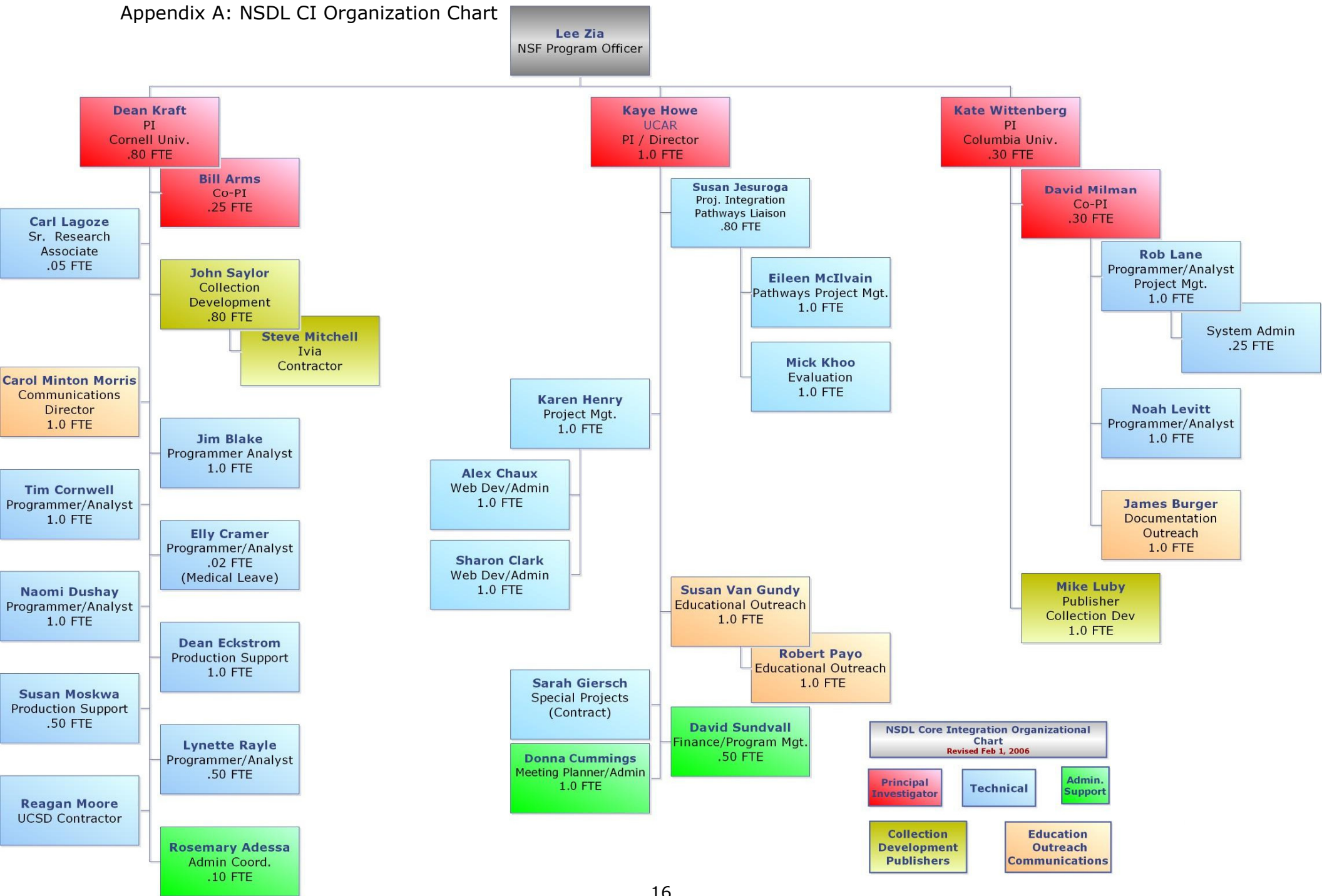
usage through strategic partnerships, community sign-on, direct outreach to the educational community, all of these and more are areas for collaborative work.

The Pathways are at very different stages of development and even those that are made up of mature organizations (e.g. the Math Gateway) face enormous technical and cultural challenges of their own. The Pathways hold a rich array of technical and educational elements which must be brought together to serve the STEM educational community. The multiple challenges of collaborative integration, continuous technical and educational development, while maintaining an already useful STEM educational digital library, are worth the talents of everyone involved.

Appendices

- A. Organization Chart - NSDL CI**
- B. NSDL Pathways Projects**
- C. MOU 2004**
- D. MOU 2005**
- E. Collection Development Advisory Council (CDAC) and Selectors list**
- F. NSDL Projects by State, 2000-2005**
- G. NSDL Publisher agreements report**
- H. NSDL Workshops report**
- I. NSDL CI Work Tracker, as of February 3, 2006**

Appendix A: NSDL CI Organization Chart



Appendix B: NSDL Pathways Projects

[AMSER](#)

NSDL APPLIED MATHEMATICS AND SCIENCE EDUCATION PATHWAY

To Resources and Services for Community and Technical Colleges

AMSER will consist of a focused metadata repository and a variety of integrated services designed specifically to enhance the learning experience of the community college students and the teaching capabilities of instructors at those institutions. AMSER is in development. For more information about AMSER, please e-mail info@amser.org

[Ben: BioSciEdNet](#)

NSDL BIOLOGICAL SCIENCES PATHWAY

To Digital Libraries for Teaching and Learning in the Biological Sciences

Ben Collaborative partner organizations provide access to over 3,800 reviewed resources covering 76 biological sciences topics. Ben is managed by the American Association for the Advancement of Science (AAAS). Registration is required.

[ComPADRE](#)

NSDL PHYSICS AND ASTRONOMY PATHWAY

To Physics and Astronomy Education Resources

Through a partnership of authors and organizations ComPADRE acts as a steward for the educational resources used by broad communities in physics and astronomy by creating and sustaining a network of collections that provide learning resources and interactive learning environments. ComPADRE resources positively influence physics and astronomy students and their teachers in both individual and collaborative settings.

[Computational Science Education Reference Desk](#)

NSDL COMPUTATIONAL SCIENCE, MODELING, AND TECHNOLOGY PATHWAY

To Computational Science Education

Shodor provides curriculum development materials that include interactive activities and instructional materials for students, educators, and parents; workshops, online courses, and educational resources geared towards educators, and; internships and workshops that provide students with resources in computational science.

[Digital Library Projects at the Ohio State University](#)

NSDL EDUCATION PATHWAY

To Middle School Teacher Resources

The NSDL Middle School Portal provides access to selected online resources for instruction and professional development from the National Science Digital Library. Subject pathways in mathematics, science, and technology present topic lists that take an in-depth look at teachable concepts in science, math, or technology.

Features include lively text and graphics along with background for teachers, interactive online activities, data analyses, and links to related topics.

[Materials Digital Library](#)

NSDL MATERIALS SCIENCE PATHWAY

To Effective Delivery of Materials Science Information

The NSDL Materials Digital Library (MatDL) brings materials science research and education closer together. MatDL is exploring the various roles digital libraries can serve in the materials science community including: 1) supporting a virtual lab, 2) developing markup language applications, and 3) building tools for metadata capture. MatDL is being integrated into an MIT virtual laboratory experience.

[Mathematical Sciences Digital Library](#)

NSDL MATHEMATICS PATHWAY

To Mathematics Resources for Teachers and Students

MathDL is an online resource published by the Mathematical Association of America. The site provides online resources for both teachers and students of mathematics that include: *Convergence*, a magazine that teaches mathematics using its history; the *Journal of Online Mathematics and its Applications (JOMA)*, a scholarly journal, plus; free classroom-tested and peer reviewed digital classroom resources.

[Needs](#)

NSDL ENGINEERING PATHWAY

To Online Learning Materials in Engineering and Related Areas of Science and Math

NEEDS provides web-based access to a database of learning resources where the user (whether they be learners or instructors) can search for, locate, download, and comment on resources to aid their learning or teaching process. In addition NEEDS supports a multi-tier evaluation system from our national award competition with sponsors from industry to user-based reviews of individual learning resources.

[Teachers' Domain](#)

NSDL EDUCATION PATHWAY

To Multimedia Resources for the Classroom and Professional Development

Teachers' Domain is a multimedia digital library for the classroom that provides learning experiences in ways no textbook can. This ever-expanding library currently includes collections on Science (Life Science, Physical Science, Engineering) and Social Studies (The Civil Rights Movement, Brown v. Board of Education). Registration is required.

Appendix C: Memorandum of Understanding between NSDL Core Integration and Pathways Projects

10/18/2004

“The success of the NSDL program will depend, to a large extent, on the development of a collective sense of identity and common cause by all projects”

- from the 2004 NSF-NSDL Solicitation

1. Introduction

This Memorandum of Understanding (MOU) between NSDL Core Integration (CI) and the FY04 Pathways Projects seeks to clarify the administrative and development activities needed to support a successful partnership in building the National Science Digital Library (NSDL). The CI group has been charged by NSF to work with new and existing projects to develop technical standards, infrastructure, and organizational processes for integrating collections and services into NSDL. Pathways Projects have been charged to provide a stewardship role for the educational content and services needed by a broad community of learners. The Pathways concept is a critical and central strategy to help build NSDL in a way that is useful for defined communities of users.

In order to move from a group of collected projects to a coherent NSDL organization, this collaboration between the Pathways Projects and CI will provide a leadership role in building a more comprehensive, recognized and trusted on-line science, technology, engineering and mathematics (STEM) library to serve learners at all levels of education. Through this collaboration, NSDL and NSF will gain a better understanding of technical and administrative frameworks required to sustain a long term NSDL organization, and learn what it takes to support additional Pathways Projects in future years.

2.0 General Responsibilities

CI staff is responsible for coordinating and managing the core library, main library portal at NSDL.org and underlying technical infrastructure. This includes the metadata repository architecture and metadata services, search, authentication services, archiving, *AskNSDL*, and the main portal. CI also adopts and support standard protocols and technologies in response to CI and community needs. In addition to technical responsibilities, CI coordinates and engages the NSDL community, which includes library developers, educators, and users of the library, through an overall NSDL Outreach and Communications Plan. CI staff also support the NSDL governance process and run the NSDL Annual Meeting.

As laid out in the NSDL Solicitation, the Pathways Projects are responsible for maintaining criteria and mechanisms to identify, select, annotate, and generate metadata for high-quality and relevant digital content as it continues to become available; providing all item-level metadata to the NSDL core metadata repository; sustaining the currency of the project resources, either by acquiring/linking, or deaccessioning; active archiving that preserves the usability of digital content as the

underlying information technology systems evolve; and anticipating and providing value-added services that may be specific to the targeted learning community.

The CI and the Pathway Project Partners will make all good faith efforts to fulfill the conditions of this MOU within their time resources and budgetary constraints. The Pathway Projects are being funded with NSF support and their primary obligation is to deliver the products to which they are committed by the terms of their project descriptions.

3.0 Specific Collaborative Activities in Year One

Following an initial meeting in Boulder on August 18-19, 2004, Pathways and CI staff discussed the initial list of priorities and tasks to commence Year One work. The tasks listed below are of equal priority: major areas of work can run concurrently, although specific tasks within each area may require phased activities. A brief description of each collaborative task, the primary contact (first name listed in Contacts), and staff are listed for each task. Where specific staff names are not identified, the type of staff needed is indicated and marked TBD. Due dates are broadly listed as fiscal year quarters (e.g. Q1 starting October 1, 2004 and ending December 31, 2004), with the understanding that milestones may be changed by agreement of the PIs as the work commences.

3.1 Technical

3.1.1 Identity Management

The goal is to support single sign on and central management of common NSDL identity attributes.

Task	Contact(s)	Due
IM-1 Determine attributes required across projects, that can be managed centrally	Karen Henry, Pathways technical staff (TBD)	Q1
IM-2 Define each attribute	David Millman, Diane Hillmann, Kim Lightle, Joe Brandt, Dave Joiner, Ed Almasy	Q2
IM-3 Discuss login options with Math Gateway's CMS developer Lucidea Corporation –report back to PIs	David Millman, Math Gateway staff	Q1
IM-4 Collect and post current/required vocabularies (role, education level, etc) related to user registration from each Pathway – share results with Content SC	Kim Lightle, with help from Pathways staff	Q2
IM-5 Assemble and analyze current Privacy Policies	Susan Jesuroga, with help from Pathways staff	Q1-2
IM-6 Create ground rules to verify identity and level of confidence	David Millman, Ed Almasy	Q2
IM-7 Background work: Improve Shibboleth software packaging; Identify alternative identity systems; Confirm "cookies" restrictions; Detail common	David Millman	Q2

approaches and services		
IM-8 Decision Point: When to roll out NSDL login across portals	CI and Pathways PIs, on advice of technical team	Q3
IM-9 Begin implementation of NSDL login on portals	Pathways technical staff, with help from CI	TBD in IM-8

3.1.2 Data Management

The goal is to understand each Pathways metadata and resource sharing strategies to ensure the FEDORA repository can meet Pathways' needs.

Task	Contact(s)	Due
DM-1 Gather Pathways requirements for metadata representation in FEDORA system <ul style="list-style-type: none"> - What is the context by which your users find resources? - What is the granularity of those resources? - What do we mean by "resource"? - How do we deal with vocabulary differences? 	Dean Krafft, Pathways technical staff (TBD)	Q1
DM-2 Decision Point: Priorities for FEDORA implementation beyond March 2005 release	CI and Pathways PIs, on advice of the technical team	Q2
DM-3 Collect/post vocabularies from Pathways, especially those used to enable discovery (education level, type)	Susan Jesuroga	Q1
DM-4 Research other possible vocabularies	Blythe Bennett	Q2
DM-5 Provide regular status to Pathways about CI work, including FEDORA conversion	Karen Henry	Monthly to PIs
DM-6 Supply metadata and vocabulary information to NSDL repository	Pathways technical staff, when available.	When available
DM-7 Set up K-12 standards sub-group <ul style="list-style-type: none"> - Share results of Plato meeting - Review JES&Co NSDL project - Make recommendations for future work, including if work is to be done by CI, Pathways and/or or by engaging the community 	Karen Cariani, Susan Van Gundy, Mike Luby, Kim Lightle	Q2
DM-8 Discuss resource reusability, in context of Robby Robson's work <ul style="list-style-type: none"> - Make recommendations for future work, including if work is to be done by CI, Pathways and/or or by engaging the community 	CI and Pathways technical leads (TBD)	Q3

3.1.3 Other Technical Work

Task	Contact(s)	Due
OW-1 Provide DNS resolution services for the Pathways URL using XXX.NSDL.org	Dean Krafft	Q2
OW-2 Discuss each groups' MyLibrary work at Annual Meeting to determine points of collaboration/need - Make recommendations for future work, including if work is to be done by CI, Pathways and/or or by engaging the community	CI and Pathways technical leads (TBD)	Q1
OW-3 Provide improved archiving documentation	Dean Krafft	Q2
OW-4 Create minimum Pathways/CI accessibility guidelines	Rachael Bower, Ed Almasy, Casey Jones, other Pathways UI representatives (TBD)	Q3
OW-5 Participate in Webmetrics Pilot	Casey Jones, Peter Pinch, Pathways technical staff (TBD)	Q2-3

3.2 Outreach and Communications

Goals include coordinated outreach and communications efforts, and development of co-branding strategies. Specific goals include naming strategies for NSDL.org domains, project titles, and tags to ensure consistency in communications from project-to-project, and key NSDL messages to be conveyed by Pathways and CI when talking about individual and combined efforts.

Task	Contact(s)	Due
OC-1 Joint press release announcing Pathways	Carol Minton Morris, Deb Burns, Bob Panoff, Rachael Bower, Lang Moore, Len Simutis	Q1
OC-2 Determine XXX.NSDL.org names for each Pathways Portal	Susan Van Gundy, Carol Minton Morris, Deb Burns, Bob Panoff, Rachael Bower, Lang Moore, Len Simutis	Q1
OC-3 Determine appropriate level of co-branding for web and print material - lightweight NSDL brand for web sites - templates and standard descriptions for promotional materials	CI OutComm group, working with Pathways design staff (TBD)	Q1
OC-4 Create and start implementing combined Outreach & Communication plan	Susan Van Gundy, Carol Minton Morris, Deb Burns, Bob Panoff, Rachael Bower, Lang Moore, Len Simutis	Q2
OC-5 Pursue <i>AskNSDL</i> as outreach mechanism, including figuring out how to work with established help services like Ask Dr. Math	Susan Van Gundy, Carol Minton Morris, Deb Burns, Bob Panoff, Rachael Bower, Lang Moore, Len Simutis	Q2-4
OC-6 Develop statement of quality as it	John Saylor, Ed Almasy	Q2

relates to branding across CI/Pathways		
OC-7 Combined access to marketing lists/identify marketing group	Sarah Giersch	Q2
OC-8 Form advocacy subcommittee	Kaye Howe, Pathways PIs	Q2

3.3 Community

The goal is to engage and work with the NSDL community.

Task	Contact(s)	Due
C-1 Participate in the Pathways Panel presentation at Annual Meeting: Review of ENC work, overview of each Pathways work, and how they will engage the community.	Susan Jesuroga, Rachael Bower, Lang Moore, Bob Panoff, Ted Sicker, Len Simutis	Q1
C-2 Submit posters explaining project for Annual Meeting	Pathways staff (TBD)	Q1
C-3 Develop Pathways FAQ to answer NSDL community questions about the projects' goals and plans	Outreach working group	Q2
C-4 WBR article about Pathways	Carol Minton Morris	Q1
C-5 Join NSDL standing committees	Pathways staff (TBD)	As needed

3.4 Strategic Planning

The goal is to track and prioritize future needs not represented in Year One work.

Task	Contact(s)	Due
SP-1 Create quarterly goals and work assignments	Karen Henry, CI and Pathways teams	Quarterly
SP-2 Determine Year Two priorities	CI and Pathways PIs	Q3-4
SP-3 Rights management	CI and Pathways PIs	Q3
SP-4 Evaluation of collaborative process and resulting products	CI and Pathways PIs	Q4
SP-5 Sustainability	CI and Pathways PIs	Q3-4

4.0 Administration

Locus of collaborative planning and decision-making will be a group of the CI PIs, CI Team Leads, Pathways PIs and designated principals. Implementation of those plans will be detailed to work groups of CI and Pathways staff as required by the nature of the work. CI and Pathways PIs will meet via monthly conference calls to discuss work progress and make decisions with regards to priorities, and agendas will be prepared by CI staff at UCAR. CI and Pathways PIs will meet face-to-face twice yearly. The first meeting took place in August, 2004, and the second will be scheduled in late winter or early spring, 2005. The PIs, team leads, and principals will also meet at the NSDL Annual Meeting.

Susan Jesuroga will be the CI liaison to Pathways Projects, and Karen Henry will produce and track CI-Pathways priorities through a managed work tracker system. Both will coordinate the above tasks by working through the primary contact listed for each task. Any project management issues (missed deadlines, conflicting priorities, etc) will be brought to the CI and Pathways PIs for resolution. CI will maintain a private workspace to manage email lists, the work tracker, common documents and other communications required by this group. Work groups will establish their own processes, workspaces, and meeting schedules according to need.

5.0 Relationship to NSDL Governance

There are two governance bodies of importance to the CI-Pathways collaboration. The National Visiting Committee is appointed by the National Science Foundation to provide high-level strategic advice to CI, and continuing evaluation of the project to NSF. The committee also serves as an advocate of the library to the larger public. The Policy Committee advises the CI team, other NSDL grantees, and NSF on operational strategies, policies and implementation priorities for NSDL.

Pathways and CI PIs will attend the winter 2005 NVC meeting and participate in conference calls as needed to brief NVC members on issues of priority and matters requiring strategic advice. (CI will work to make the PI face-to-face and NVC meetings coincide – the meeting will most likely to be held in Washington D.C. or New York City.) Pathways and CI staff will advise the Policy Committee of policy and standards issues coming out of this joint work, such as Privacy Policy changes necessitated by the identity management work. All groups will be expected to participate in Standing Committee meetings as necessary.

6.0 MOU Updates and Revisions

Updates and revisions to the MOU may be proposed by any of the CI-Pathways partners and must be agreed to by a consensus of that group. Records of the updates and revisions will be kept by the CI staff at UCAR and posted to the Pathways workspace (<http://pathways.comm.nsd.org>).

Appendix D: Memorandum of Understanding between NSDL Core Integration and Pathways Projects

Version 1.2

8/30/2005

“Through its distributed digital library building strategy, the NSDL program is fostering the creation and development of a comprehensive cyberlearning infrastructure.”

- from the 2005 NSF-NSDL Solicitation

1. Introduction

The National Science Digital Library (NSDL) mission, as defined by the National Science Foundation, provides a clear focus on education and the need to encourage and sustain continual improvements in the quality of STEM education. To be a successful catalyst for education, NSDL must address two basic priorities. As a library, the first is discovering and organizing the most useful resources for STEM education. To support learning, the second is giving users a coherent learning environment and set of tools that help meet specific educational goals.

This Memorandum of Understanding (MOU) between NSDL Core Integration (CI) and the Pathways Projects seeks to clarify the administrative and development activities needed to support a successful partnership in building the NSDL. The CI group has been charged by NSF to work with new and existing projects to develop technical standards, infrastructure, and organizational processes for integrating collections and services into NSDL. Pathways Projects have been charged to provide a stewardship role for the educational content and services needed by a broad community of learners.

The Pathways strategy is a key element to NSDL success through their focused efforts to organize a specific domain area and provide users the environment needed to meet their educational goals. Each Pathway project has specific community needs to meet, so CI and the Pathways must ensure that combined development outcomes are integrated to not only meet the NSDL mission, but also create a library that appears to be coherent, organized and coordinated for the broader NSDL user community.

This collaboration between the Pathways Projects and CI provides a leadership role in building a more comprehensive, recognized and trusted on-line science, technology, engineering and mathematics (STEM) library to serve learners at all levels of education. Through this collaboration, NSDL and NSF will continue to build the technical and administrative frameworks required to sustain an NSDL organization for the future.

2.0 General Responsibilities

CI staff is responsible for coordinating and managing the core library, main library portal at NSDL.org and underlying technical infrastructure. This includes the metadata repository architecture and metadata services, search, authentication services, archiving, *AskNSDL*, and the main portal. CI also adopts and support

standard protocols and technologies in response to CI and community needs. In addition to technical responsibilities, CI coordinates and engages the NSDL community, which includes library developers, educators, and users of the library, through an overall NSDL Outreach and Communications Plan. CI staff also support the NSDL governance process and run the NSDL Annual Meeting.

As laid out in the NSDL Solicitation, the Pathways Projects are responsible for maintaining criteria and mechanisms to identify, select, annotate, and generate metadata for high-quality and relevant digital content as it continues to become available; providing all item-level metadata to the NSDL core metadata repository; sustaining the currency of the project resources, either by acquiring/linking, or deaccessioning; active archiving that preserves the usability of digital content as the underlying information technology systems evolve; and anticipating and providing value-added services that may be specific to the targeted learning community.

The CI and the Pathway Projects will make all good faith efforts to fulfill the conditions of this MOU within their time resources and budgetary constraints. The Pathway Projects are being funded with NSF support and their primary obligation is to deliver the products to which they are committed by the terms of their project descriptions.

3.0 Collaborative Activities

This MOU concentrates on the high-level areas of collaboration. Because the Pathways' focus on serving specific audience needs, the focus of this MOU defines collaborative areas of work, including technical infrastructure, outreach, policies, and procedures, which will be required to support a coherent, coordinated library for all NSDL users. Because the Pathways Projects are at different levels of maturity, specific development tasks and deliverables for CI and individual Pathways Projects will be outlined and tracked in a separate Pathways Work Tracker document. The following outline lists the key areas of collaboration required to sustain an operational NSDL.

3.1 Keep barriers to NSDL users as low as possible

- 3.1.1 Work towards implementing a single NSDL sign on
 - Develop an approach or policy to achieve goal
 - For those Pathways with existing logins, determine appropriate integration path consistent with policy
- 3.1.2 Share collection development policies online
 - Post collection development policies for all collections under each Pathway
 - Annotate policy with appropriate definitions/descriptions of important features, such as resource quality or how resources are selected
- 3.1.3 Create and implement plans for resource persistence
 - Include clear definition of what is saved, for how long, plans for migration
- 3.1.4 Define issues surrounding modification of resources

- Define user requirements that allow for successful modification, including rights management
- Define granularity of resources required

3.1.5 Commit to usability and accessibility standards

- Follow industry best practices
 - See AccessNSDL.org, <http://ncam.wgbh.org/accesslinks.html>, http://www.imslobal.org/accessibility/accmdv1p0/imsaccmd_o_viewv1p0.html
- Implement usability studies, focus groups, user testing, etc

3.2 Improve the NSDL collection

3.2.1 Share resource centric information, such as metadata, annotations, and related content

- Define common set of resource-centric features for Pathways
- Submit metadata to the NSDL Data Repository (NDR)
- Define potential issues surrounding metadata and sustainability beyond the period of NSF funding
- Define and develop web services for expressing resource-centric information and relationships across collections and the NDR

3.2.2 Provide resource recommendations

- Make recommendations where materials are out of individual collection scope, but within the broader NSDL scope
- Participate in using the iVia expert-guided crawl

3.3 Adopt standards and policies

3.3.1 Adhere to the NSDL Metadata Policy

- Online at <http://comm.nsd.org/download.php/83/PC3Final.pdf>

3.3.2 Participate in defining necessary policies on IP and rights management

3.3.3 Post privacy policy on portal site

- Help define minimum elements required for consistency across all sites
- Meet COPPA guidelines, where required

3.3.4 Use and share defined vocabularies and taxonomies, such as subject classifications, audience, and grade level

- Develop agreed-upon shared vocabularies
- Post vocabularies centrally

3.3.5 Use NSDL-wide services and information, where possible

- Achievement Standards Network, for education standards

3.4 Coordinate outreach and communications activities

3.4.1 Participate in co-branding NSDL and Pathways

- Web sites, promotional materials, press coverage and other related documents

- 3.4.2 Coordinate workshops and other user training opportunities
- 3.4.3 Coordinate conference presentations and exhibits
- 3.4.4 Facilitate networking and partnership building between NSDL and Pathways' target audiences
- 3.4.5 Share original content and domain expertise
 - Current events, articles, reports, and papers through *On Ramp*
 - News feeds, blogs, and other RSS feeds
- 3.4.4 Contribute to NSDL reports, briefings, and other presentations

3.5 Contribute to the overall evaluation of NSDL

- 3.5.1 Participate in web metrics evaluation efforts
 - Install Omniture tracking software, or other software as provided by CI
 - Provide regular summary reports, once installed
- 3.5.2 Participate in the larger NSDL program evaluation effort
 - Define yearly benchmarks and metrics
 - Provide metrics as defined above

3.6 Support NSDL operations

- 3.6.1 Identify primary technical contacts and backups
 - Contacts for daily operational questions, and emergency outage contact
 - Contacts for user questions and support
- 3.6.2 Meet site availability targets and provide contingencies for downtimes
 - Determine targets for each site
- 3.6.3 Perform routine checks and repairs for broken links

3.7 Enable collaboration

- 3.7.1 Provide web service interfaces for broadly applicable services
 - Participate in defining core central services, such as search, NDR access, and archiving
 - Identify and provide other agreed upon Pathways-enabled services
- 3.7.2 Set and manage co-development priorities through
 - Regular meetings
 - Document and project management
 - Clear schedules, deliverables and responsibilities
 - Document the impacts of collaboration in time and resources
 - Share lessons learned

4.0 Administration

Locus of collaborative planning and decision-making will be a group of the CI PIs, CI senior staff, Pathways PIs (or PI-designates) and designated principals. Implementation of those plans will be detailed to work groups of CI and Pathways staff as required by the nature of the work. CI and Pathways PIs (or PI-designates) will meet via monthly conference calls to discuss work progress and make decisions with regards to priorities, and agendas will be prepared by CI staff at UCAR. CI and Pathways PIs (or PI-designates) will meet face-to-face twice yearly. The first meeting took place in August, 2005, and the second will be scheduled in late winter or early spring, 2006. The PIs and staff will also meet at the NSDL Annual Meeting.

As first point of contact, Susan Jesuroga will be the CI liaison to Pathways Projects, and Karen Henry will produce and track Pathways Work Tracker. Both will coordinate development activities by working with the primary contact listed for each task in the Pathways Work Tracker. Any project management issues (missed deadlines, conflicting priorities, etc) will be brought to the CI and Pathways PIs (or PI-designates) for resolution. CI will maintain a private workspace to manage email lists, the work tracker, common documents and other communications required by this group. Work groups will establish their own processes, workspaces, and meeting schedules according to need.

5.0 Relationship to NSDL Governance

There are two governance bodies of importance to the CI-Pathways collaboration. The National Visiting Committee is appointed by the National Science Foundation to provide high-level strategic advice to CI, and continuing evaluation of the project to NSF. The committee also serves as an advocate of the library to the larger public. The Policy Committee advises the CI team, other NSDL grantees, and NSF on operational strategies, policies and implementation priorities for NSDL. CI will be the primary liaison to these committees and will ask the Pathways projects to provide briefings from time-to-time.

Pathways and CI PIs will attend the winter 2006 NVC meeting and participate in conference calls as needed to brief NVC members on issues of priority and matters requiring strategic advice. (CI will work to make the PI face-to-face and NVC meetings coincide – the meeting will most likely to be held in Washington D.C. or New York City.) Pathways and CI staff will consult with the Policy Committee about policy and standards issues coming out of this joint work, such as Privacy Policy changes necessitated by the identity management work. All groups will be expected to participate in Standing Committee meetings as necessary.

6.0 MOU Updates and Revisions

Updates and revisions to the MOU may be proposed by any of the CI-Pathways partners and must be agreed to by a consensus of that group. Records of the updates and revisions will be kept by the CI staff at UCAR and posted to the Pathways workspace (<http://pathways.comm.nsd.org>).

This MOU replaces the 2004 agreement between the Pathways Projects and CI, dated October 18, 2004.

