UNESCO

UNITED NATIONS EDUCATIONAL,
SCIENTIFIC AND CULTURAL ORGANIZATION

Analytical Survey

DIGITAL LIBRARIES IN EDUCATION

IITE

UNESCO INSTITUTE
FOR INFORMATION TECHNOLOGIES IN EDUCATION

MOSCOW 2003

UNESCO
UNESCO Institute for Information Technologies in Education (IITE)

COORDINATING AUTHOR
Leonid Kalinichenko
(Russian Federation, Institute of Informatics Problems of the Russian Academy of Sciences, Moscow)

IITE PROJECT COORDINATION
Sergei Christochevsky

ANALYTICAL SURVEY WORKING GROUP
Daniel Atkins (USA, University of Michigan)
Edward Fox (USA, Virginia Polytechnic Institute and State University)
David Fulker (USA, University Corporation for Atmospheric Research, Boulder, Colorado)
Yannis Ioannidis (Greece, University of Athens)
Stephan Körnig (Germany, Darmstadt University of Technology)
Bertram Ludäscher (USA, San Diego Supercomputing Center, University of California)
Mary Marlino (USA, University Corporation for Atmospheric Research, Boulder, Colorado)
Pasquale Savino (Italy, Istituto di Elaborazione della Informazione, Consiglio Nazionale delle Ricerche, Pisa)
Narasimhiah Seshagiri (India, Supercomputer Education and Research Centre, Bangalore)
Tamara R. Sumner (USA, University of Colorado, Boulder)
Alexey Ushakov (USA, University of California, Santa Barbara)
Michael Wright (USA, University Corporation for Atmospheric Research, Boulder, Colorado)

Digital Libraries in Education. Analytical Survey
This analytical survey provides an overview of current technologies for Digital Libraries in Education (DLEs) and their anticipated evolution through consideration of several DLE projects that have been carefully selected to provide for the identification of typical DLE features, characterization of the current state of DLE technologies, clarification of the social and organizational issues surrounding DLE development, and prediction of the DLE impact on education. Advanced frameworks and methodologies related to DLEs are described to forecast the further evolution of DLE technologies and applications. The survey is not intended to be an exhaustive coverage of DLEs, nor is it an aim to meld all the perspectives into one coherent picture, but rather to present them all. This information could serve a ground for planning the forthcoming phases of future DLE projects.

FOR FURTHER INFORMATION PLEASE CONTACT:
UNESCO Institute for Information Technologies in Education
8 Kedrova St. (Bld. 3), Moscow, 117292, Russian Federation
Tel.: 7 095 129 2990
Fax: 7 095 129 1225
E-mail: info@iite.ru
Web: www.iite-unesco.org

ISBN 5-902116-07-4 (Publishing House “Education-Service”)

Printed in the Russian Federation

Table of contents

1. Introduction

The project aimed at Digital Libraries in Education (DLEs) is being developed in accordance with strategic goals of the UNESCO Institute for Information Technologies in Education (IITE). UNESCO intends to foster new forms of networking between teacher-training institutions and teachers, using digital libraries (DLs) as well as production and deployment of digital educational materials. This project started with the international expert meeting held in June 2001 in Moscow. The information materials *Digital Libraries in Education: State-of-the-Art Report* had been prepared as a basis for discussion during that meeting [IMEM01]. These materials contained a brief analysis of (1) current trends leading to quite rapid changes of learning environments in contemporary society and of (2) the most notable existing programmes for the development of DLEs. One of the recommendations of the expert meeting was to continue investigation of current experience with DLEs, under the IITE framework, and to prepare an analytical survey of the area, to be widely disseminated. This analytical survey is the current stage of the IITE project on DLEs.

The content of the analytical survey reflects the results of joint work and meetings with several US groups: the NSDL (National Science Digital Library) Policy Committee, UCAR (University Corporation for Atmospheric Research, Boulder, Colorado), SDSC (San Diego Supercomputer Center, University of California, San Diego), and the University of Michigan (Ann Arbor). The conversations involved Edward Fox, David Fulker, Mary Marlino, Timothy Spangler, Tamara Sumner, Michael Wright, Bertram Ludäscher, Michael Freeston, Alexey Ushakov, Daniel Atkins, and their colleagues. The analytical survey also incorporates information collected at these meetings and centres. Additionally the survey includes excerpts of [IMEM01] that were discussed in 2001 with Yannis Ioannidis (Greece, University of Athens), Stephan Körnig (Germany, Darmstadt University of Technology), Pasquale Savino (Italy, IEI CNR, Pisa), Narasimhiah Seshagiri (India, SERC, Bangalore), and other participants of the international expert meeting at the IITE in June 2001.

The analytical survey is a step in procedures for the promotion of international cooperation, globalization of products and methodologies, and their adoption in communities. It focuses on the application of information technologies in education. Through consideration of some carefully selected DLE projects, the survey provides a detailed analysis of current DLE technologies and their anticipated evolution, thus forming a base for planning the forthcoming phases of future projects for DLE development.

The survey does not emphasize hybrid libraries, which draw together online and physical collections and services, presenting them to the user in a seamless and integrated manner, supported by middleware that handles aspects, such as authentication and cross-searching. In universities and colleges, libraries offer their catalogues online, many through a standard Web browser. University libraries also offer access to increasing numbers of electronic journals and other online information sources, including those provided internally or from remote locations. Virtual versions of library services, such as reservations, registration, and reference enquiries, are also starting to be offered, particularly to distance learners.

The survey does not attempt to analyse every known DLE-related project. Instead, several DLE projects have been carefully selected to provide for identification of typical DLE features, characterization of the current state of DLE technologies, clarification of the social and organizational issues surrounding DLE development, and prediction of the DLE impact on education. Advanced frameworks and methodologies related to DLEs are described to forecast the further evolution of DLE technologies and applications.

The survey does not consider every educational discipline, instead concentrating on education in the natural sciences and engineering. Thus the specificity of DLEs for many other disciplines needs to be investigated further. Geographically, the report is based on information produced mostly in the USA and Europe. Collecting information about the state-of-the-art in other regions is important for UNESCO. Several significant issues (such as sustainability and economic issues, DLE globalization versus national or regional development) are not sufficiently analysed in this document and will require separate discussion.

As the area of DLEs is too broad a subject to cover exhaustively in one survey, and in view of the constraints on the survey mentioned in the previous few paragraphs, the survey also outlines subjects for further analysis.

The survey is the result of the joint effort of discussions with, and information provided by, the Working Group:

Daniel Atkins (USA, University of Michigan)
Edward Fox (USA, Virginia Polytechnic Institute and State University)
David Fulker (USA, University Corporation for Atmospheric Research, Boulder, Colorado)
Yannis Ioannidis (Greece, University of Athens)
Leonid Kalinichenko (Russian Federation, Institute of Informatics Problems, Russian Academy of Sciences, Moscow), coordinating author
Stephan Körnig (Germany, Darmstadt University of Technology)
Bertram Ludäscher (USA, San Diego Supercomputing Centre, University of California)
Mary Marlino (USA, University Corporation for Atmospheric Research, Boulder, Colorado)
Pasquale Savino (Italy, Istituto di Elaborazione della Informazione, Consiglio Nazionale delle Ricerche, Pisa)
Narasimhiah Seshagiri (India, Supercomputer Education and Research Centre, Bangalore)
Tamara R. Sumner (USA, University of Colorado)
Alexey Ushakov (USA, University of California, Santa Barbara)
Michael Wright (USA, University Corporation for Atmospheric Research, Boulder, Colorado)

The survey is structured as follows: Section 2 contains a general discussion of the role of digital libraries as an integral part of the rapidly changing educational environment. Section 3 presents the wide range of understanding of the digital library concept. Section 4 gives a general picture of DLEs as repositories of educational resources, with services. Section 5, *Integrated Learning Environments,* presents definitions and information about Managed Learning Environments (MLEs) and Virtual Learning Environments (VLEs) and interactions between them. The related pedagogic models for VLEs are also discussed there. Section 6 lists criteria for digital library quality within the learning environment. This section also contains a brief description of the proposal for a joint project between the UK Joint Information Systems Committee (JISC) and the US National Science Foundation (NSF) to incorporate VLEs and digital libraries into the learning process. Section 7 surveys activities to standardize educational metadata.

Sections 8 and 9 survey current DLE technologies. Several American projects (NSDL, DLESE, CITIDEL, and NDLTD) and several European projects (DNER, Scholnet, and Cyclades) were selected to demonstrate the current state-of-the-art and the planned evolution of technologies. (See those sections for the full names of these projects.) To show the anticipated evolution of DLE frameworks, from current ones based on the conventional library metaphor to more knowledge-based systems, Section 10 collects information on several advanced frameworks and methodologies related to DLEs. Issues in instructional course development with reuse of preexisting learning objects are discussed in Subsections 10.1 and 10.2. Work on extension of the information content of DLs for education and research with specialized educational resources of data (including real time data) is reflected in Subsections 10.3 and 10.4. The use of advanced infrastructures (cyberinfrastructure and data grid) as a possible basis for future DLEs is considered in Subsections 10.5 and 10.6. Research on knowledge-based approaches for DLE frameworks is included in Subsections 10.7 (navigation and search interface for NSDL based on science literacy benchmarks) and 10.8 (preparing a course of physical geography in an ADEPT concept-based architecture). Section 11 summarizes the survey. A reference list of publications used for compiling the survey is given at the end of the document.

2. Transforming the way to learn

2.1. Transformation of the educational system is inevitable

We are on the threshold of a revolution that is making the world's accumulated information and knowledge accessible to individuals everywhere. Digital technology will not only transform the intellectual activities of universities, it is likely to cause restructuring of the current higher-education enterprise into a global "knowledge and learning" industry.

A radical transformation of the educational system is coming under growing pressure in the school systems themselves, changing society into one in which knowledge work becomes ever more important, and developing information and communication technologies (ICTs) which are transforming the economies. The following trends indicate that the transformation is inevitable [DLTO01]:

- The number of students is still growing, as health and population increase.
- Different types of students are asking for education; different students bring different experiences with them.
- Increasingly, work and study are combined, and that leads to a need for more flexible learning arrangements in which the campus or school building is no longer central to the educational process.
- More generally, there is a trend towards lifelong learning.
- Lifelong learning leads to an emphasis on "learning to learn". Knowledge[1] becomes obsolete at an ever-increasing rate in a knowledge economy, and knowledge workers need to be able to refresh their knowledge on a regular basis.
- Because of the differences among students, there is a need to accommodate different learning styles, to provide customization and alternative learning routes. Courses have to take better account of the different experiences and backgrounds of students.
- Higher education institutions have long had a monopoly in providing education, but increasingly, companies and public bodies possess knowledge that can be reused for educational purposes, partly for in-house training (knowledge management) but also to offer to external markets.
- Education is under constant budget pressure; thus there is a need for more efficient and effective education.
- Students increasingly are behaving like consumers, and want to make informed choices about how and where they want to be educated, which implies that students are no longer committed to one institution.
- Teaching staff will exhibit more job-hopping behavior than they did in the past.
- There are too many dropouts in the current educational system.

A global learning environment of the future is envisaged as:

- being student-centred,
- being interactive and dynamic,
- enabling group work on real world problems,
- enabling students to determine their own learning routes,
- emphasizing competencies like information literacy to support lifelong learning.

Active learning implies that students do not limit themselves to resources supplied by their instructors, but also search for and organize new materials themselves in order to solve problems and to develop their competencies continuously.

2.2. Information users in the changing educational environment

It is important to analyse the characteristics of users who will make significant choices within this new educational environment. Starting with the students, one of the first and most important characteristics to note is that of preferring to be self-sufficient in their information gathering. It seems that users are beginning to perceive the traditional library as something used at the end, or at best the middle, of their information search. This has important implications for education programmes, as well as for understanding how those users who come to the library decide to do so. They do not come to the library first for the problem definition and information gathering phases. They prefer to dive into the problem alone first rather than coming to the common space.

Users want control of their own information environment. It is important to them to have some items owned for convenient personal consultation. They prefer to use a private good rather than a common good if they can. Users do

[1] The word "knowledge" is an overloaded term. This word is used here to emphasize that such phrases can often be found in the DL context. Of course, fundamental knowledge has not changed so rapidly.

not want to be dependent on anyone else if they can afford (in terms of both time and money) not to be. The convenience factor and the value added by the functionality of the service itself will be key in how choices about service providers are made.

The amount of information used in one's professional work that is available openly on the Internet is dramatically greater than it was in 1990. Users have the impression that they will find useful information free on the Web. If students do need to ask a person for information help, they go to a friend or co-worker because that person already has an understanding of the student's context for either the problem or their level of understanding.

Because of the characteristics described above, people are becoming ready to pay for ubiquitous, convenient, fast, and customized information access (this may not apply in developing countries). Faculty also wants control of their information environments, especially for teaching. Course management software packages are proliferating on campuses as part of a larger academic agenda to address the need to support information technology in both distance education and campus-based learning. Conventional libraries have an uncertain role in Web-based learning environments.

A second change in learning environments is that of a greater emphasis on the public scrutiny of teaching and learning. The quality of someone's teaching is no longer a personal matter or departmental matter. Faculty behavior in the classroom is brought out of the dimension of a personal contribution and is considered a service that is evaluated for its quality, just like other services.

2.3. Digital Libraries in Education as a way of restructuring the current higher–education enterprise into a global "knowledge and learning" industry

One of the natural responses to the above challenges consists in introducing the DLEs as a core of networks of learning environments and resources [SLRM01], that is:

- Designed to meet the needs of learners, in both individual and collaborative settings;
- Constructed to enable dynamic use of a broad array of materials for learning, primarily in digital formats;
- Managed actively to promote reliable anytime, anywhere access to quality collections and services, available both within and outside the network.

The digital library must not be seen as merely a digitized collection of information objects plus related management tools, but as an environment bringing together collections, services, and people to support the full cycle of creation, dissemination, discussion, collaboration, use, new authoring, and preservation of data, information, and knowledge. The challenges and opportunities that motivate advanced DL initiatives are associated with this view of the digital library environment. Work on digital libraries aims to help in generating, sharing, and using knowledge so that communities become more efficient and productive and the benefits of collaboration are maximized. It seeks to aid existing communities and to facilitate the emergence of new communities of research and education.

Introducing digital libraries into the education process has been underway in distance education for a number of years [DELO00]. With the Internet and the Web, distance education programmes can mount sets of materials on Web servers to support a course. The range of materials that currently are in digital form is great. In some disciplines enough materials are available with open access so that students already have access to broad collections. Digital libraries can provide adequately broad library services to local and remote students. One of the basic ideas is to join learning materials on various topics and written by many teachers in a digital library of courseware [DELI97]. Such a DL provides a basis for creating courses on specific topics.

Applying digital libraries in education has the potential to drastically change fundamental aspects of the classroom [WTSM99] in ways that could have an enormous impact on teaching and learning. The DLEs can be seen as an *information space* in which students are moving around intellectually, encountering new information, and working with the teacher and other students to make sense of what they encounter. A traditional library for education, a school library for example, typically includes textbooks, curriculum materials, artifacts (such as charts, physical samples, and equipment), enrichment books, and the teacher's own personal collection of teaching tools (as well as resources shared by other educators and learners). Taken generically, this space has been constituted over the decades to include the content required for teaching, for example, a sanctioned subject to high school students. The texts may be carefully designed and written, and the entire space is designed to bring students into the discipline that the class represents. Thus students can be exposed to only the best products of the human mind. The content of education can be carefully evaluated and filtered to include only the most worthy. Such a point of view may lead to a controversy with the point of view that emphasizes expanding the number of educational resources and avoiding being prescriptive and confining.

Educational applications of digital libraries range from primary schools through graduate schools and across all disciplines. DLs change the possibilities for the education information space. Boundaries are expanded to include not only canonical versions of the subject, but other products put into DLs. However, the DL content can be controlled better than that of the Web. Although some would argue that content on the Web is motivating and interesting to students, it is varied and unpredictable in its design, rather than carefully designed to help students learn. Web sites can be complex and confusing or deep and significant. The quality, quantity, and substance of information available in the classroom teaching and learning space may become vastly different once the Web is included. Clearly, there are both positive and negative features of the new space opened up by the Web. In fact, positive and negative characteristics of the traditional text-based classroom and of the Web can be seen as mirror images of each other. DLE offers a middle ground between the overly constrained information space of traditional classroom resources and the overly open and undependable information space of the Web. Controls on content allow accountability and dependability while open submission provides opportunities for a wealth of materials beyond what an individual teacher or school system would envision and select.

New pedagogical methods should accompany DLs as an emerging technology to reach the compelling vision of education expressed in [ACRA98]:

"Any individual can participate in online education programmes regardless of geographic location, age, physical limitation, or personal schedule. Everyone can access repositories of educational materials, easily recalling past lessons, updating skills, or selecting from among different teaching methods in order to discover the most effective style for that individual. Educational programmes can be customized to each individual's needs, so that our information revolution reaches everyone and no one gets left behind".

Describing the situation in the USA, [ACRA98] reported that "in education, information technology is already changing how we teach, learn, and conduct research, but important research challenges remain. In addition to research to meet the scalability and reliability requirements for information infrastructure, improvements are needed in the software technologies to enable development of educational materials quickly and easily and to support their modification and maintenance. We know too little about how best to use computing and communications technology for effective teaching and learning. We need to better understand what aspects of learning can be effectively facilitated by technology and which aspects require traditional classroom interactions with the accompanying social and interactive contexts. We also need to determine how best to teach our citizens the powers and limitations of the new technologies and how to use these technologies effectively in their personal and professional lives".

"Access to and use of IT, particularly in educational settings (K-12 [primary and secondary] as well as higher education), is a prerequisite to building the skills base that will allow our citizens to function productively in the information society of the next century. It is also a critical stepping stone for instilling interest and developing the skills of the budding IT researchers who will be essential to sustaining our national research capabilities".

It was predicted that "the Nation [is] facing an impending crisis in preparing workers to be productive in an economy that is increasingly dependent on IT. Although the use of computers in education is increasing at all levels, and computer literacy is increasing dramatically across the country, too little percentage of the population are entering or receiving necessary re-training in the computing, information, and communications professions. Market forces alone will not correct the problem. The government must do more to help educate and re-train people in these crucial fields and to bolster the academic pipeline from elementary school to post-graduate study".

This and other analyses gave rise to various research and development programmes for DL technologies in education around the world and resulted in planning specific research areas, including [NSF996]:
- preservation and archiving of digital scholarly information, including technology and procedures for long-term information asset management;
- utilization of digital libraries in educational technology at all levels of instruction, electronic publishing, and scholarly communication technology, including
- collaboratories, online repositories, and new methods of organizing scientific knowledge distribution.

3. The digital library concept: A common term with many interpretations

Digital libraries are becoming an integral part of digital learning environments. At the same time, the notion of "digital library" is subject to a broad range of definitions. Different audiences associated with a digital library have different interpretations; they evaluate a digital library differently and use different terminologies. On one end of the range, DLs are considered to be related to physical libraries performing similar functions, thus creating a *hybrid library* (combining traditional and electronic resources). On the other end, DLs are considered to be knowledge repositories, and services, organized as complex information systems. For example, global information repository projects are devoted to the accumulation of digital forms of information related to the Earth, universe, art, environment, or humans.

Such diversity of definitions is reflected in the different communities related to DLs. One such community is the research community formed in the United States by the Digital Libraries Initiatives (DLI) of NSF. DLI did not attempt to define "digital library" strictly. In order to incorporate a wide range of possible approaches and domains, the concept is treated broadly and vaguely. Thus the projects cover a wide range of topics, stretching the possible meaning of "digital library". This is perfectly acceptable for research – frontiers need to be stretched. On the other hand, the practicing community, the majority of whom reside in operational libraries, concentrate on building operational digital libraries, maintaining and operating them, and providing services to users. As a result, thousands of digital libraries have emerged worldwide. The efforts are diverse. Numerous types of collections and media are included and processed in many different ways. Some of them are located in physical libraries, creating a hybrid library, while others are not bound to a library at all. JISC in the UK (in contrast to the US NSFs DLI effort) promotes the innovative application and use of information systems and information technology, and emphasizes the development of content and new technologies that would be widely applicable and not just of benefit to the participating institutions. In 1999 NSF and JISC established a joint initiative to bring together the best elements of the two funding bodies.

To show the difference of understanding of the notion of DLs, several definitions of "digital library" follow. The broad (unrestricted) definition may be considered closest to the approach taken by the research community [DLCO00]:

"Digital libraries are organized collections of digital information. They combine the structure and gathering of information, which libraries and archives have always done, with the digital representation that computers have made possible" [PDLB97].

The following definition may serve as a bridge between the research and practicing communities:

"Digital libraries are a set of electronic resources and associated technical capabilities for creating, searching, and using information; they are an extension and enhancement of information storage and retrieval systems that manipulate digital data in any medium. The content of digital libraries includes data and metadata. Digital libraries are constructed, collected, and organized by (and for) a community of users, and their functional capabilities support the information needs and uses of that community" [COVB99].

The Digital Library Federation (USA) agreed in 1999 on the following working definition of a digital library representing a definition of the practicing community. In this definition the emphasis is on an organizational or institutional setting for the collection of digital works and aspects related to its functioning in the larger context of service:

"Digital libraries are organizations that provide the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities" [DLFS99].

The United Nations Task Force on Digital Libraries gives the following definition of digital libraries:

"Digital libraries are organized collections of information resources in digital or electronic format along with the services designed to help users identify and use those collections. Digital libraries promise to provide more effective information services than has been possible in the past, by offering the following advantages: faster delivery, a wider audience, greater availability, more timely information, more comprehensive”.

The Technical Committee on Digital Libraries of the IEEE (Institute of Electrical and Electronics Engineers) Computer Society (IEEE-CS) uses the more general term “(digital) collective memory” to emphasize the convergence of libraries, museums, archives, and collections of all kinds including those of private citizens. Collective memory development faces challenges in several areas, including storage, classification, and indexing; user interfaces; information retrieval; content delivery; presentation, administration; and preservation.

Diversity of understanding of what digital libraries are leads to a wide range of possible visions for DL frameworks and methodologies of use: from the conventional library metaphor to knowledge-based systems. This survey focuses on evaluation of digital libraries in learning environments.

4. Digital libraries of educational resources and services

"The network is the library" [SLRM01]: In a library, be it digital or analog, the essential transaction is the same: a user interacts with content. But richer interaction is possible within the digital environment, not only as more content is put within reach of the user, but also as more tools and services are put directly in the hands of the user. These include the abilities to search, refer, validate, integrate, create, customize, publish, share, notify, and collaborate, to name but a few. Students, teachers, faculty, and those pursuing continuing education will "connect to learn"; but they will also "learn to connect", as they leverage their participation with other users of the library and its resources.

By networking users and content with tools, the digital library enables three chains of support. First, *users supported by profiles* are able to form *learning communities*. These can be communities of one or they may be communities of thousands; they may be short-lived communities born of immediate needs, or they may grow into persistent communities. However, an important concern to acknowledge is the potential loss of privacy, which must be balanced against the potential gain in personalization of a user's experience. A second chain of support closely related to the first is that *content supported by metadata enables the formation of customizable collections of educational objects and learning materials*. These collections may target an individual or they may target a community; and they may learn and adapt to the behavior of their users. Finally, *tools supported by common protocols or standards enable* the development of varied application services that enhance the value of the library's content for the learner.

The following long-range objectives for DLEs were formulated by Tom Kalil (The White House) and further articulated in an article by Zia [SLRM01]:
- Lifelong learning.
- Learning anytime anywhere.
- Distance learning demonstration programme.
- Government as "model user" of technology-based training.

For these objectives, a number of intermediate goals are formulated, such as:
- Improve student performance.
- Get more students excited about science.
- Increase the quantity, quality, and comprehensiveness of Internet-based science educational resources.
- Make these resources easy to discover and retrieve for students, parents, and teachers.
- Ensure that these resources are available over time.

Studies show [SLWSJ98] that the Internet has the potential to transform the highest level of education, but only a fraction of that potential is now being realized. Some of this gap lies in the maturation process that is part of any transition, but a larger part is the result of fragmentation. Resources of great value are not being used because students and faculty do not know about them, or do not know how to use them.

While great efforts have been placed on creating materials, less attention has been given to organizing them, maintaining them in the long term, helping people find them, and training people how to use them. For example, a faculty member who is planning a course has only the most rudimentary tools to discover what materials are available or whether they have proved effective in other courses. A student who is researching a topic is forced to choose between general-purpose Web search services and commercial databases designed for scientific and technical research. Neither faculty nor students can safely rely on resources that might be withdrawn without notice, or change subtly overnight.

A DLE is envisaged as a *comprehensive library of the digital resources and services* that are available for education in science, mathematics, engineering, technology, and other disciplines. The key word here is "comprehensive". Faculty are very specific in wanting a single place where they and their students can discover, use, and possibly contribute a wide range of materials.

A DLE is considered to be a *federation of library services and collections* that function together to create a digital learning community. Organizationally, a DLE will consist of a small central operation with a wide range of partners. Some of the services and collections are already well organized; for these, the DLE will act as a gateway. Others exist but are poorly organized; for these, the DLE will stimulate the creation of specific services. Some materials are fragmented, unorganized, or hard to find; in these cases the DLEs will build library services and may even manage specific collections. Across all these areas, the DLEs will provide tools to help faculty and students find and use materials, with services to assist them in evaluating quality and appropriateness.

DLEs will take a broad view of science and technology, and of scientific education. The primary audience is faculty and undergraduate students, but there is no hard distinction between the needs of high school students, undergraduates, and graduate students, nor between students in formal programmes, independent learners, and the general public.

DLEs should have a variety of *financial models* for access to the materials; some content will be free of charge while other materials will be available on a fee basis.

The range includes curricula and courseware materials, lectures, lesson plans, computer programmes, models and simulations, intelligent tutoring systems, access to remote scientific instruments, project-based learning, tools, the results of educational research, scientific research reported both formally in journals and informally in web sites, raw data for student activities, and multimedia (image, audio, or video) banks. DLEs should provide services for authors and instructors, such as annotation, evaluation, and peer review of donated materials. For students and faculty, they will offer the capability to search for desired information by subject area, to have access to scientific data sets, to interact with peers, and to provide archiving, location-independent naming, recommender systems, selective dissemination of information, and copyright management. Faculty, students, and other clients, such as independent learners, will be able to participate in forums. Interdisciplinary activities, lifelong learning, and the process of education will all benefit. In this way, the DLE will be much more than the sum of its parts, and will promote change and innovation in scientific and technical education at all levels.

The following are guiding principles that DLEs should follow:
1. Be driven by educational and scientific needs.
2. Facilitate educational innovations.
3. Be stable, reliable, and permanent.
4. Be accessible to all (though not all materials will be free).
5. Build on, and leverage, past and current work in courseware libraries, digital library research, and successful commercial sites.
6. Be adaptable to new technologies.
7. Support the decentralized creation of services.
8. Provide tools and organizational background for the integration of resources.

DLEs are intended to encourage *the dissemination of research in educational methods*. They will also facilitate the involvement of industry and government laboratories in the educational process. Whereas some universities benefit from guest speakers from industry or government in the classroom, not all schools are able to arrange such visits. The digital library, enabled by new information technologies, would provide a forum for real time video or voice communication to a wider range of learners. These *virtual lectures and discussions* could be captured and then added to the library for later access.

DLEs will also facilitate *cross-institutional sharing* of educational resources, including all types of courseware, as well as materials for distance and self-learning. The ultimate goal is the development of a community of science and technology educators who use the library for cross-disciplinary and cross-institutional collaboration. Access and discussions with authors and prior users would be possible, along with an archive of past reviews and discussion of materials in the DLE. The collections could be annotated and linked to these discussions and reviews.

The digital library also opens the opportunity for students at different institutions to work on *joint projects or experiments*, perhaps sharing and adding to the same data set and its analysis. This would also promote *physical resource sharing*, as students and instructors may have varying access to high-end instrumentation, computational capabilities, data collections, and technology.

The following rationale for DLEs in science and math has been expressed [SLWSJ98]:
1. Student performance in math and science is poor and needs to be improved.
2. Today's Internet lacks the cataloguing, organization, archiving, collections management, etc., of a library.
3. The effort to connect every classroom to the Internet will be of limited value without high-quality content.
4. A digital library can be a resource for the entire population (marginal cost of dissemination is almost zero).

One of the methods of determining the success of digital libraries in improvement of student learning is to examine whether they are helping to achieve pedagogical objectives. Development of scientific thinking in students might be one of the criteria. Examples of skills that are to be developed in students by educators are asking questions, acquiring information, organizing information, analyzing information, and answering questions in certain scientific disciplines.

Accessibility [SLWSJ98] is an important property of DLEs that requires a two-part strategy. The first is that the library should be realistic in its technical expectations. Since a range of factors, including network bandwidth, availability of computers, and costs, can limit accessibility, the library must be designed to accommodate a wide range of users and be realistic about the technology that they use. However, not all DLE services need be limited to the lowest common denominator of the current capabilities of computers, networks, students, and faculty. Technology is improving rapidly, and the library must grow with it.

The second part of the strategy is that the DLE should work vigorously with concerned individuals and organizations, including federal and local agencies, to ensure that all students and faculty have good Internet access. Modern scientific and technical education requires that all faculty and students have computers and telecommunications, with the training to use them effectively.

Sustainability [SLRM01] is another important property of DLEs. There are strong arguments for the national DLE to be considered a “national treasure” and supported as a public good; indeed the frequent calls for open, free access to content are rooted in this view. An attractive scenario for the long-term management of the digital library is to place responsibility in the hands of a non-profit organization.

This vision of DLEs still begs the question of how creators will be compensated for their efforts. For contributions of “fine-grained” content (e.g. short Applet tutorials or simulators) the digital library can offer recognition from peers, which would be suitable and important “compensation”. Digital rights management technologies also hold promise for identifying usage of, and then appropriately providing compensation for, content. This would allow the creators and purveyors of content to differentially price and/or repackage portions of “coarse-grained” material that has been disaggregated (some publishers have begun to offer custom runs of selected textbook chapters to professors). It has been observed that reconceptualizing information as a service rather than a good offers the opportunity for new revenue streams that can be directed back towards content creators. This view suggests interesting possibilities for the development of new services for users that could be available, for example, individually or through affiliation with existing organizations, such as professional societies. More generally, these considerations may lead to rethinking of reward systems, such as promotion and tenure, to reflect the importance of developing, sharing, and using educational resources in DLs.

The role of digital libraries in the learning environment will be clarified further. First, models of learning environments will be characterized.

5. Integrated learning environments

5.1. The Virtual Learning Environment and its interaction with the Managed Learning Environment

Major changes are happening across the educational community, affecting all aspects of learning, teaching, and administration activities. These present opportunities for huge benefits to colleges and universities from new learning environments and management technologies.

To discuss the involvement of digital libraries in the process of education, a model of learning environments is required. A JISC model of a Managed Learning Environment (MLE) and a Virtual Learning Environment (VLE) [MLEB02, MLEJIS, MLEVLE, VLEREQ, VLEGRE, VLEPED] will be used here. JISC has described an MLE as follows:

"MLE is the taking advantage of the potential of new technology based learning environments to integrate information systems around the learner. These learners may be working in different modes at different times, on campus, off campus, part time or full time. To support this, information systems will need to be student centred and fully accessible from multiple locations. They will need to be integrated at multiple levels, ensuring interoperability between administrative and financial systems, learning support and learning environments, and between collaborating institutions".

Using such a definition, educational institutions need to, first, rethink their educational and organizational processes as an essential part of MLE development, and second, make student-centred approaches central to that development. All MLE development should combine two processes: the design and implementation of appropriate secure and robust technical systems, and the enabling of effective educational and organizational integration towards an improved student experience. Of course, rather than being automatically or necessarily aligned, these two processes often operate with varying degrees of separation and lack of integration.

A VLE [MLEVLE, JISCMLE, VLEREQ] consists of the components through which learners and tutors participate in online interactions of various kinds, including online learning. The principal functions that the complete VLE needs to deliver are:

- Controlled access to curriculum that has been mapped to elements (or "chunks") that can be separately assessed and recorded. (Computer Curricula 2001, http://www.acm.org/sigcse/cc2001/, the new curriculum for computing developed by the Association for Computing Machinery, ACM, and IEEE-CS, is a fresh document waiting for the application of such an approach.)
- Tracking student activity and achievement against these elements using simple processes for course administration and student tracking that make it possible for tutors to define and set up a course with accompanying materials and activities to direct, guide, and monitor learner progress.
- Support of online learning, including access to learning resources, assessment, and guidance. The learning resources may vary from self-developed to professionally authored and purchased materials that can be imported and made available for use by learners.
- Communication between the learner, the tutor, and other learning support specialists to provide direct support and feedback for learners, as well as peer-group communications that build a sense of group identity and community of interest.
- Links to other administrative systems, both in-house and externally.

As shown in Figure 1, the VLE will act as a "portal" to online curriculum mapping, assessment, communication, delivery, tutor support, and tracking facilities.

The VLE makes up only one part of a college's overall systems (both computerized and noncomputerized). Interfacing between these systems is possible by "connecting up" the constituent parts through the use of interoperability standards such as the IMS (Instructional Management System). Examples of these interfaces are ones between the student record system and the VLE, and between learning resources (or content) and the VLE.

VLEs are learning management software systems that synthesize the functionality of computer-mediated communications software (email, bulletin boards, newsgroups, etc.) and online methods of delivering course materials (e.g. the Web). Several systems are emerging for the management of online learning, but none is currently able to deliver

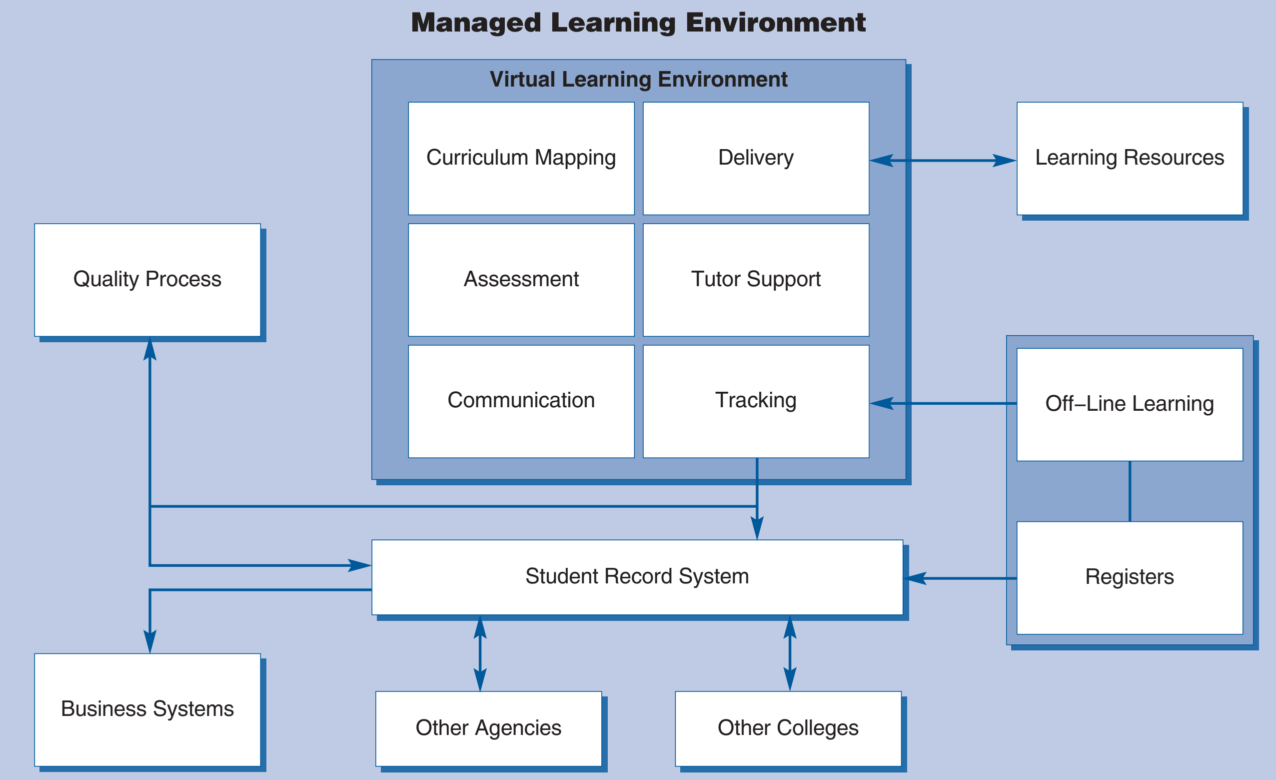

Figure 1. Schematic of a Managed Learning Environment

the full set of functions and linkages shown in Figure 1. Reports of evaluations carried out on particular VLEs, indicating the various pedagogical assumptions that developers may make in implementing VLEs, can be found at http://www.jisc.ac.uk/jciel/mlesg/.

Most VLE systems are intended not simply to reproduce the classroom environment online, but to use the technology to provide learners with new tools to facilitate their learning. They aim to accommodate a wider range of learning styles and goals, to encourage collaborative and resource-based learning, and to allow greater sharing and reuse of resources. A summary of the features of a number of current systems is provided in [VLETLS,VLEGRE]. In the MLE/VLE particular attention is given to the issues of interoperability, both between VLEs and various administrative systems and between VLEs and various providers of learning content.

5.2. Integrated learning environments in different countries

The e-learning situation worldwide is driven by similar factors, namely, technological developments increasing and changing the expectations of learners, changes in society resulting in changes in the nature of the student population (including globalization of learning), and new developments and understandings of what learning is and how it can best be accomplished.

The US higher education system is not as driven by central government as it is in other countries (e.g. in the UK). In addition to the traditional, privately run, large universities, there is growing use of commercial contracting out of instructional responsibilities rather than using tenured faculty, and increasingly private firms, such as Microsoft[2], are collaborating in higher education or running their own universities. These last developments are already becoming factors in the UK and are likely to increase in coming years, once again driven by the globalization of the educational marketplace.

[2] See: http://research.microsoft.com/programs/. The Jones Education Company aims to "get the cost of real estate out of education" and uses cable television to deliver six certificates and 11 degrees in conjunction with 14 institutions. For these and more see: Michael Thorne. *Universities of the Future*. PowerPoint Presentation at SeSDL videoconference seminar series 2000/2001; scroll down page at: http://www.sesdl.scotcit.ac.uk:8082/seminars/index.html for link.

The terms "VLE" and "MLE" are not used in the US; distance learning is frequently referred to as remote learning, and the terms "virtual classroom" and "Web-based instruction" are examples of terms used in a broader sense than VLE. There is no generic term for a networked learning environment encompassing both campus- and distance-based online instruction. This is in spite of the fact that the systems or platforms referred to as VLEs and MLEs are widely developed and used in the US, by both commercial firms and educational institutions. The difference in language probably indicates a difference in conception or priorities.

5.3. Alternative pedagogic models for evaluating Virtual Learning Environments

Two different models (one from education, the other from systems modeling) have been explored in [VLEEVA] as a basis for constructing a pedagogical evaluation methodology for VLEs. The educational model was developed and applied to the use of learning technology in higher education by Laurillard [CONV93] as a Conversational Framework. The Organizational Model is drawn from the Viable Systems Model for modeling organizational systems proposed by Beer [ORGM81]. It has previously been suggested that this organizational systems approach may be applicable in a pedagogical context [VLEEVA]. Basic ideas of these frameworks are presented in Figures 2 and 3.

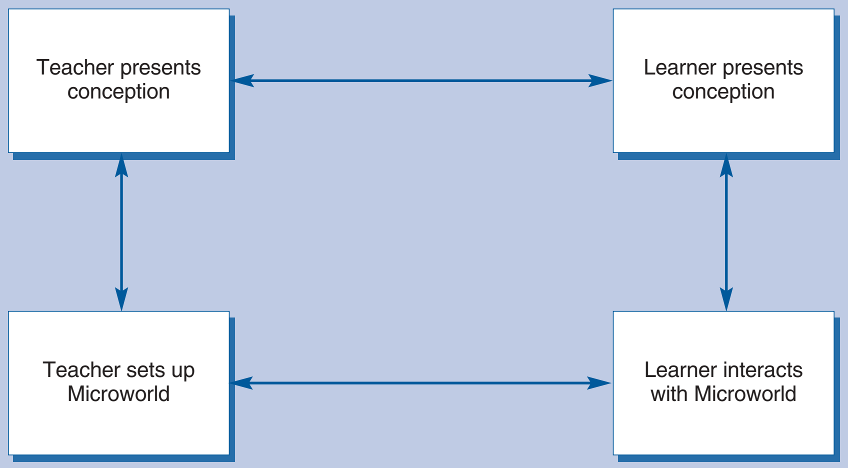

Figure 2. Conversational learning

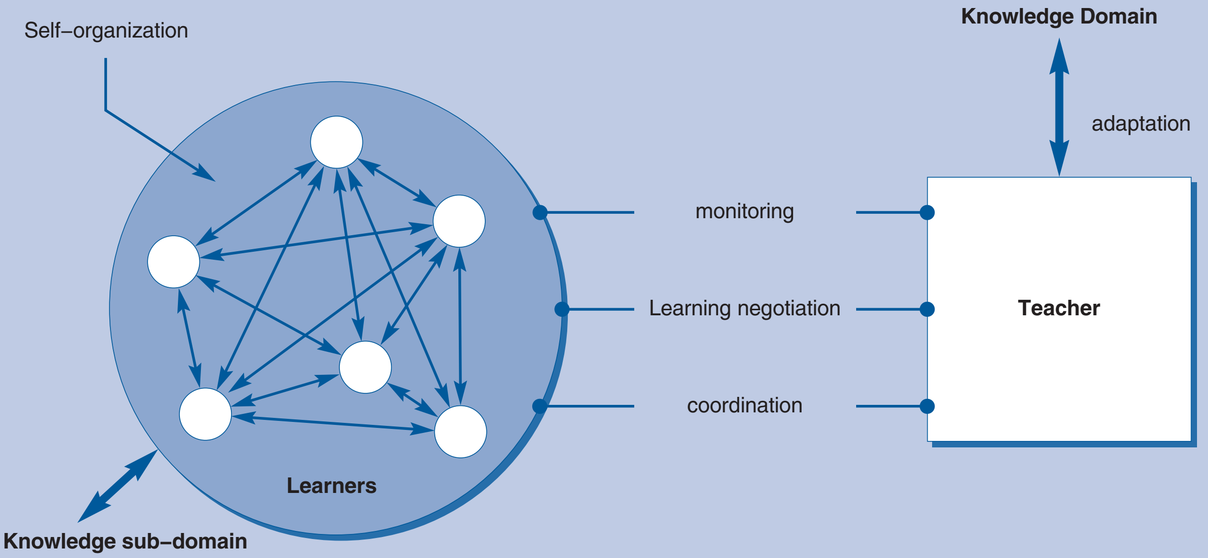

Figure 3. Organizational Model for a learning environment

The US Department of Defense's Advanced Distributed Learning Initiative (ADL) summarizes many approaches to solving the scalability or "teacher bandwidth" problem in a model close to the conversational one:

- A one-to-one instructional model in which a teacher tailors instruction to individual student needs is preferable to other instructional models.
- Human (teacher-student) interaction in large scale learning environments is not economically feasible; therefore,
- Automating feedback and other learning support via intelligent instructional systems is the only viable solution to providing scalable online learning.

But automated instructional systems completely lack human interaction and social negotiation, which learning theorists are increasingly stressing as crucial to supporting meaningful learning. Highly decontextualized learning objects are reusable in the greatest number of learning contexts, but they are also the most expensive and difficult for instructional designers to reuse, creating a "reusability paradox". Fortunately, however, when educators have been trained and motivated, they can easily identify and share very small knowledge resources that can be readily reused by others, once discovered. This process can be facilitated by tools that support the synthesis and construction of larger resources out of a number of small resources. Examples of such tools include the Walden's Path software from the digital library research group at Texas A&M University, USA, the instructional architect from Utah State University, USA, and colleagues, and the VIADUCT tool that is part of the CITIDEL (www.citidel.org) effort at the Virginia Polytechnic Institute and State University, USA.

Some ideas of learners' organization and interaction with a knowledge subdomain of the Organizational Model are reflected in the online self-organizing social systems (OSOSS) [OSOSS]. The most significant departure of the OSOSS approach from conventional learning objects approaches is that it relies on human beings to locate, assemble, and contextualize the resources. The OSOSS provides a conceptual framework for a new method of indexing, discovering, combining, using, and evaluating digital educational resources:

Indexing and discovery: Learning objects are not catalogued with metadata and submitted to a central repository. Community members know of existing resources and local resource collections. Learners gather information from a variety of sources. Individual resources are discovered through "community queries" in which community members respond with pointers to resources they know personally.

Combination: Learning objects are not automatically populated into one of many instructional templates. Without the direction of any single grand architect, peers contribute relevant resources and descriptions of how they might be employed within the context of the initiator's problem.

Use: Learners do not sit through a temporal sequencing of resources and assessments linked to decontextualized instructional objectives. They employ resources provided by peers as mediational means in the solution of a self-selected problem or accomplishment of another self-selected goal.

Evaluation: Learning objects are not critiqued out of an instructional context with a summative quality rating. Learners evaluate the relevance and suitability of resources within a specific learning context.

In an OSOSS learners are provided with meaningful learning support "anytime anywhere", yet the support is reached with human-to-human interaction. Learning objects are successfully embedded in a meaningful learning context, but the objects are discovered and contextualized by humans – again without scalability's becoming an issue.

Potential problems with the OSOSS approach:

- A standard curriculum may be difficult to impose on individuals in an OSOSS.
- Assessment of individuals may be difficult to carry out in an OSOSS.
- Required feedback may not be immediate in an OSOSS.

6. Digital libraries and Virtual Learning Environments

6.1. Tasks of digital libraries in the learning environments

Within the context of changes in society, technology, and education in recent years, there have been two key developments relating to e-learning infrastructure in UK universities and colleges:

- The adoption of virtual learning environments and managed learning environments.
- The implementation of digital and hybrid libraries.

VLEs are tools which support e-learning through the provision and integration of Web-based materials, including learning materials, links to other resources, online communication tools (such as electronic bulletin boards), and assessment tools. When such VLEs are integrated with other information systems and processes of the institution, e.g. student records, the resultant system is generally referred to as an MLE.

For VLEs, truly digital libraries are required with all resources and services available online. Some of the candidate tasks that DLEs could support include [EDLL00]:

1. Highly directed uses, such as lab exercises to reinforce a specific disciplinary concept.
2. Instructional modules that introduce concepts in an incremental manner and can be customized and extended by faculty for use in lectures.
3. Free form exploration conducted by students preparing term papers or faculty putting together a lecture that might include personal manipulation of data sets, information visualization, and the integration of new information or data sets to augment existing content.
4. Collaborative applications that might be used by students doing team projects or faculty and teaching assistants who are team teaching.
5. Discipline- or domain-specific methods of building knowledge that support specific information seeking and use processes.

The key characteristic of learners with regard to the linkage of VLEs with digital libraries is their diversity. More and more learners are learning from home, from their workplace, part-time, or from a geographical distance to their course. They are coming from all age groups, and are learning throughout their lives. They are coming to the university expecting more, based on their experiences with the Internet and other information and communication technologies. There is no longer a typical "higher education" learner. Where library and information resource support to teaching was once comfortably housed in a library building, that support must now be provided to all students regardless of the medium or location of their learning.

6.2. General criteria for digital library quality within the learning environment

Quality of the resources to be discovered in the library: There is a great deal of discussion and divergence — some libraries focus on quantity as in the public library model; some focus on quality, as in specialized collections that might be found in a public library.

Seamless access: This includes seamless integration between the learning environment and the library or information resources at any point in the VLE and within one user's portal across different courses, departments, or even institutions. The most important aspect of this was the single sign-on; one authentication procedure, regardless of whether the user is accessing the VLE from on- or off-campus.

Warning notes that were sounded included potential problems with seamless cross-searching of different databases, indexes, and other information resources.[3] Lack of interoperability of search vocabularies, and a lack of awareness of and strategies to deal with this in course design, could lead to confusing, ineffective resource discovery experiences for learners.

All library functions online: Concerns about this include the potential diminishment of two important educational functions of traditional libraries: serendipitous browsing (finding the book you need right next to the one you were actually searching for); and their social function as a place to meet fellow students and discuss sources of information, etc.

[3] See JISC/RSLP funded project HILT (High Level Thesaurus) for a full analysis of this issue, at: http://hilt.cdlr.strath.ac.uk/

Individualization for the learner: This concept includes such ideas as the student portal, which could cross institutions and be available throughout a learner's life; the Amazon.com idea of tailoring resources and notifying the user about relevant resources; the ability to save and share searches; the ability to take and embed notes with information resources, and to share resources; and settings for "level", such as undergraduate, third-year, etc., with options to adjust upwards if the user wishes.

Flexibility for the teacher: Teachers would like to be able to adapt or update courses easily, including the information resources embedded in or linked to them, from anywhere. Flexibility in terms of being able to design the course according to their own pedagogical approach, rather than having it dictated by the system, was also seen as extremely important, and vital for bringing academics on board with e-learning. Finally, the system should have the capability to feed back data to the teacher about what information resources and services are being used.

Universal accessibility: Universal accessibility includes accessibility for users with differing physical abilities, adaptability to differing learning styles; availability on- and off-campus (an issue with regard to certain subscription library materials), equitable access for distance learners abroad (usually the biggest problems are access to hard copy resources and time zone problems with communications), equitable access for the economically disadvantaged (those who have to wait in line at a computer centre versus those with a PC or laptop of their own), and usability on any platform or hardware.

6.3. Example of a project incorporating a VLE and digital libraries into the learning process

Incorporation of VLEs into the learning process and interaction of VLEs with digital libraries can be illustrated by the framework planned according to a recent JISC/NSF proposal [GEOFRA]. The framework shows how the courses, content, and delivery mechanisms at the participating institutions (University of Southampton, University of Leeds, University of California at Santa Barbara, Pennsylvania State University) are linked to a range of digital resources through the application of VLE and DL technologies. As courses and modules cover particular topics, students will use links to digital resources that include a geospatial classification, within each of the following areas: human geography, geomorphology, geographical information science, and Earth observation.

Layer 1 in Figure 4 outlines the existing courses offered at partner institutions. Within each course, smaller components of learning can be identified that might involve one lecture and seminar in a week, or one practical class or one field visit, which are denoted as student learning nuggets, and which form *Layer 2* in the framework. *Layers 3* and 4 represent the integration of VLE and DL technology, where *Layer 4* shows a DL middleware technology for managing collections of resources, such as those shown in *Layer 5*. The instances of the middleware can interoperate, so that all the resources in the distributed collection are available to the VLE users.

The Alexandria Digital Earth Prototype (ADEPT) [ALEXBI, ADEPRO, ADEDLE, ADEPTS, ADEPTA, EGEODL] project has developed distributed digital libraries for heterogeneous geo-referenced information. In ADEPT, *libraries* are sets of collections. Libraries expose a single standard set of interfaces to all their collections, making it possible to issue a single query against multiple collections. (By contrast, the interfaces to collections are not standardized; instead, a library has standard mechanisms for adapting itself to whatever interfaces the collection exposes.). A library is, in effect, a "collection broker", mediating standardized access to its.

ADEPT incorporates a *bucket framework*: a canonical, simplified representation of the source metadata of heterogeneous collections, allowing uniform querying across all the collections.

The ADEPT architecture has recently been extended [ADECOA] to support an innovative form of VLE, based on the hypothesis that learning should proceed from a formal presentation of concepts and their relationships within a domain of knowledge. The architecture supports a representation of a domain ontology, or concept space, linked to a collection of learning objects. The granularity of these objects is more appropriate to their use in VLEs than traditional "information containers", such as books.

The proposal [GEOFRA] that has been accepted aims at the resource-based learning that involves active participation with multiple resources. Students are motivated to learn about a topic by trying to search for and evaluate authentic information. This learning experience mimics real life in targeting the learner as the routine information hunter and interpreter who constructs knowledge by problem solving with information tools. The advantages to this approach include:

- a student-centred approach to learning;
- promotion of the development of thinking skills (such as problem solving, reasoning, and critical evaluation);
- improving student research skills, which supports the research-led mission of all four partners;
- flexible and adaptable resources and materials for different learning styles and strategies;
- integration of key skills and competences within the academic framework.

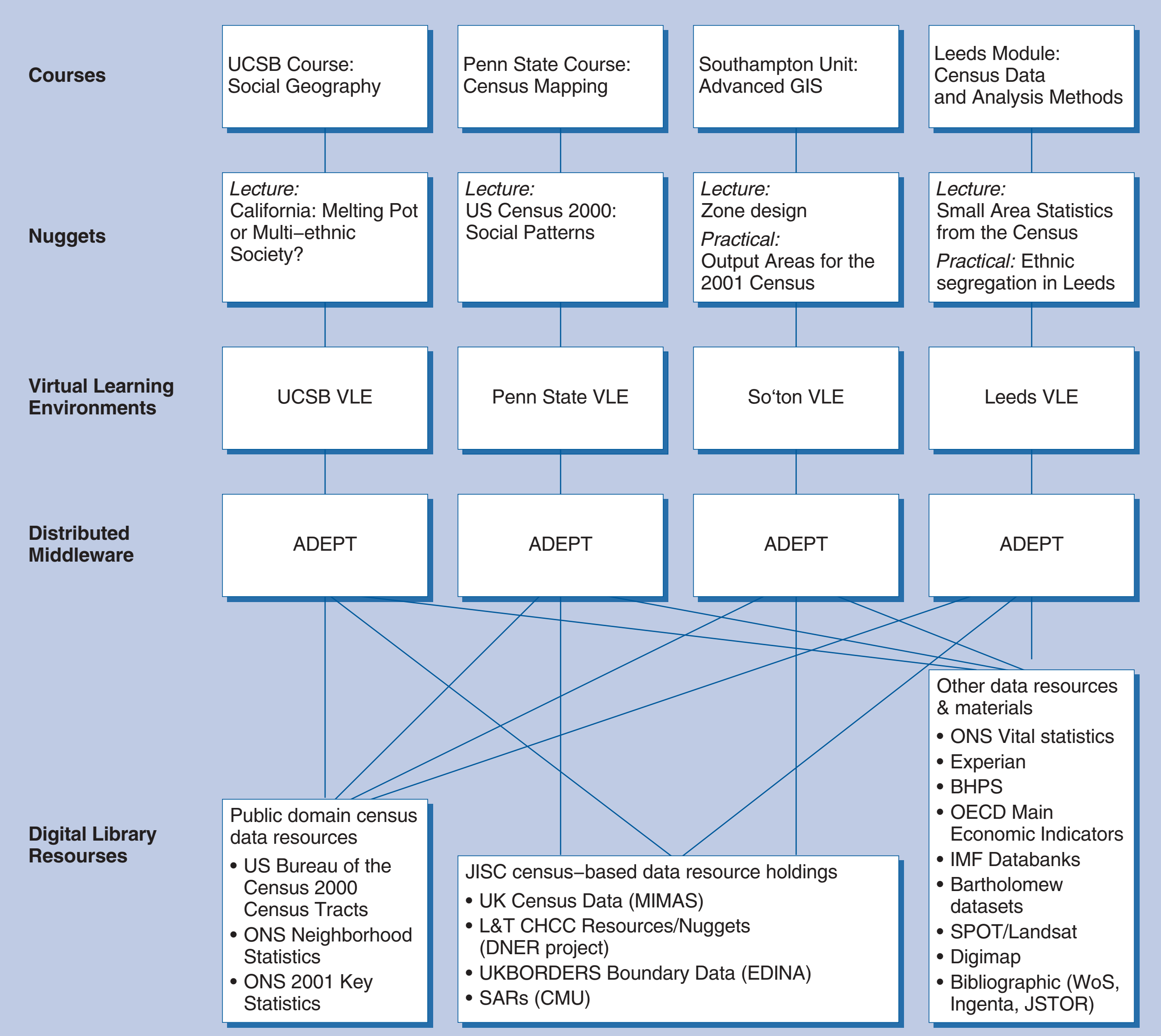

Figure 4. Infrastructure embedding digital resources across multiple institutions

The rest of this survey is devoted to the current state and anticipated evolution of DLEs that eventually should meet the requirements discussed above. The survey shows that the existing technology and services constitute a step in the evolution to true DLEs.

7. Taking a common view of educational metadata

Instructional Management Systems (IMS) Project: Designers and developers of online learning materials have an enormous variety of software tools at their disposal for creating learning resources. These tools range from simple presentation software packages to more complex authoring environments. They can be very useful in allowing developers the opportunity to create learning resources that might otherwise require extensive programming skills. Unfortunately, the wide variety of software tools available from a wide variety of vendors produce instructional materials that do not share a common mechanism for finding and using these resources. Descriptive labels can be used to index learning resources to make them easier to find and use. Such metadata specification makes the process of finding and using a resource more efficient by providing a structure of defined elements that describe, or catalogue, the learning resource, along with requirements about how the elements are to be used and represented.

In 1997, the IMS Project, part of the nonprofit EDUCOM consortium (now EDUCAUSE) of US institutions of higher education and their vendor partners, established an effort to develop open, market-based standards for online learning, including specifications for learning content metadata [LRBP11]. Also in 1997, groups within the US National Institute for Standards and Technology (NIST) and the IEEE P.1484 study group (now the IEEE Learning Technology Standards Committee, LTSC [LTSA01]) began similar efforts. The NIST effort merged with the IMS effort, and the IMS began collaborating with the ARIADNE Project (www.ukoln.ac.uk/services/elib/projects/ariadne), a European project with an active metadata definition effort. In 1998, IMS and ARIADNE submitted a joint proposal specification for standardization to IEEE, which formed the basis for the current IEEE learning object metadata (LOM) standard. IMS publicized the IEEE work through the IMS community in the US, UK, Europe, Australia, and Singapore during 1999 and brought the resulting feedback into the ongoing specification development process.

The IEEE LOM base document [LOMN00] defines a set of metadata elements that can be used to describe learning resources. This includes the element names, definitions, data types, and field lengths. The specification also defines a conceptual structure for the metadata. The specification includes conformance statements for how metadata documents must be organized and how applications must behave in order to be considered IEEE-conforming. The IEEE base document is intended to support consistent definition of metadata elements across multiple implementations. The number of items defined within the IEEE base document was large and many participating organizations within the IMS community recommended that a select core of elements be identified to simplify initial implementation efforts. The IMS developed a representation of the metadata in Extensible Markup Language (XML) [LRXB11] and surveyed its member institutions around the world to identify the core elements. The IMS *Metadata Best Practice and Implementation Guide* [LRBP11] provides general guidance about how an application may use the core and extended metadata elements.

Many metadata implementers were initially optimistic that their participation in the IMS consortium would help produce a relatively small but well defined and agreed upon set of metadata elements. This optimism soured as the set of proposed metadata elements grew increasingly larger.

Learning object metadata [LOMN00]: This standard specifies a conceptual data schema that defines the structure of a metadata instance for a learning object. For this standard, a learning object is defined as any entity, digital or nondigital, that may be used for learning, education, or training. For this standard, a metadata instance for a learning object describes relevant characteristics of the learning object to which it applies. Such characteristics can be regrouped in general, educational, technical, and classification categories. The conceptual data schema specified in this standard will allow for linguistic diversity of both learning objects and the metadata instances that describe them. This standard will be referenced by other standards that will define the implementation descriptions of the data schema so that a metadata instance for a learning object can be used by a learning technology system to manage, locate, evaluate, or exchange learning objects.

The purpose of this standard is to facilitate search, evaluation, acquisition, use, sharing, and exchange of learning objects, for instance by learners or instructors. Catalogues and inventories may take into account the diversity of cultural and linguistic contexts in which the learning objects and their metadata will be exploited. By specifying a common conceptual data schema, bindings of learning object metadata will have a high degree of semantic interoperability.

Basic metadata structure: A description of a learning object consists of data elements. The latter are grouped into *categories*. The base scheme consists of nine such categories:

a) The *General* category groups the general information that describes the resource as a whole.
b) The *Lifecycle* category groups the features related to the history and current state of this resource and those who have affected this resource during its evolution.
c) The *Meta-metadata* category groups information about this metadata record itself (rather than the resource that this record describes).
d) The *Technical* category groups the technical requirements and characteristics of the resource.
e) The *Educational* category groups the educational and pedagogic characteristics of the resource.
f) The *Rights* category groups the intellectual property rights and conditions of use for the resource.
g) The *Relation* category groups features that define the relationship between this resource and other targeted resources.
h) The *Annotation* category provides comments on the educational use of the resource and information on when and by whom the comments were created.
i) The *Classification* category describes where this resource falls within a particular subject classification system.

Data elements: Categories contain data elements. For each element, the base scheme defines:

- *Name:* the name by which the data element shall be referenced.
- *Explanation:* the definition of the element.
- *Size:* the number of values allowed.
- *Order:* whether the order of the values is significant (only applicable for elements with multiple values).
- *Value space:* the set of allowed values for the data element — typically in the form of a vocabulary or a reference to another standard.
- *Data type:* a set of distinct values.

The following data types are included: list values, vocabularies, minimum-maximum values, and character sets. For each of the data elements, the specification includes the data type from which it derives its values, such as LongString or Date, etc. These will be defined separately, and will be implemented in a particular way in a particular system. In order to maximize interoperability, future work may define a common representation for these data types. In the absence of such a common representation, an exchange format, such as XML, would allow systems with different representations to achieve interoperability through a conversion process.

DELOS Working Group on Metadata Registries: Thomas Baker (Fraunhofer Institute of the German National Research Center for Information Technology - GMD), who formed the DELOS Working Group on Metadata Registries, identified problems of harmonization of various metadata standards in the following manner.

Standards organizations and metadata implementers could in principle link their standards (as well as any application-specific extensions based on them) by cross-referencing their schemas over the Internet. However, there are currently a variety of models for defining the nature and attributes of metadata entities ("elements"). One important model is the ISO 11179 set of standards for defining the attributes of data elements and the architecture of registries. Another is the somewhat simpler model being used by the Dublin Core Metadata Initiative in its prototype schema registry, which is based on RDF schemas. Other terms are used for the IEEE learning objects metadata model and in XML Schemas. Yet others are used for web services. An appropriate level of harmonization between these various approaches would help ensure a degree of interoperability as such initiatives deploy networked registries.

8. Current Digital Libraries for Education: American projects

8.1. National Science Digital Library (USA)

NSDL is a large and comprehensive DLE initiative. It now involves 118 projects, funded by the US National Science Foundation, but additional organizations are engaging as partners as well — as the organization takes shape, it may evolve into a nonprofit institution, NSDL, Inc., or become an official sub-branch of the federal government. Current operations relate to four funding tracks, connecting with a parallel volunteer structure for governance.

The core integration track aims to provide coherence and production-level operational support, yielding a central portal to services and collections. The collections track focuses on content, with different groups working on varied horizontal or vertical slices of the STEM field. The services track emphasizes both general and specialized services, including those needed in any digital library and those appropriate to support teaching and learning. Examples discussed below include the strand map service built on the strand maps developed by the American Association for the Advancement of Science's (AAAS) Project 2061 and the concept map tool in the GetSmart project. The research track encourages innovation, new approaches, and evaluation/assessment.

The following subsections explore various aspects of NSDL, which had its first glimmerings over a decade ago, and which now is in a crucial phase of initial testing and launch. Later in this report are discussions of DLESE (Section 8.2), which is connected with NSDL, and CITIDEL (Section 8.3), one of the NSDL collection projects, to provide more details regarding how the NSDL deals with particular areas in science, technology, engineering, and mathematics, in particular Earth systems (DLESE), and computing and information technology (CITIDEL).

8.1.1. NSDL features

We concentrate first on a macro way of looking at digital library support for learning environments and on learning resources in general. The most visible representative of this approach is the US NSF National Science Digital Library (NSDL) programme [SLRM01, NSDLAR, ICADL02]. This programme seeks to bring together a vast collection of learning resources supporting all possible kinds of education, ranging from primary to graduate and lifelong learning, into one big library for the USA — and even beyond. As such, the NSDL approach is consistent with the large scale of many of the other NSF digital libraries projects.

The goal of the NSDL programme is to enhance all aspects of scientific and technical education in America. The NSDL library will be a distributed information environment for accessing quality-assured digital resources from many sources. NSDL resources span a nearly unlimited range of materials with educational value, including web pages of all sorts, digital objects such as geospatial images, proxies for physical objects, such as specimens, and threaded discussions.

Beyond its digital resources, the NSDL environment will include an extensible set of services to enhance the experience of library use. These will offer, for example: interfaces for browsing and discovering NSDL resources; tailored views of NSDL; means to annotate resources (augmenting owner-supplied metadata); support for social interactions among NSDL users; and managed access to resources by various groupings of end-users (i.e. by enforcement of usage policies). NSDL services and content eventually may alter basic pedagogic and academic practices in science, technology, engineering, and mathematics education.

NSDL is envisioned as an integrated information environment, constructed in a highly distributed effort. The goal is for end-users to interact with NSDL mainly as a coherent whole rather than as a set of individual collections and services. These ideas for coherence parallel, in several ways, the technical architecture for the Distributed National Electronic Resource (DNER) under development in the United Kingdom [DNERLP, DNERAV]. Planned NSDL features are characterized below.

End-Users: NSDL end-users are viewed primarily in their roles as educators and learners, though many simultaneously are library builders or content providers. The NSDL information environment can be characterized as supporting four high-level activities by such end-users:

- *Discovery:* NSDL facilitates discovery of content and services corresponding to end-user needs and interests.
- *Access:* NSDL enables and manages access to (discovered) content and services, potentially resulting in use for educational purposes. Use often results in further cycles of discovery, access, and use.
- *Tailoring:* Educators and learners tailor NSDL for personal purposes and for use by specified groups, such as classes of students.
- *Social interaction:* End-users may enrich their NSDL information environment through social interaction and community discourse.

Content: NSDL content is typically made available in the form of collections, defined to be any aggregation of one or more items. There will be collections of metadata about other collections. NSDL content will be characterized (and discovered) mainly via "metadata records" that describe content at the collection or item level. The accessible resources include:
- digital collections, managed by dispersed institutions;
- primary and derived information;
- data and metadata;
- diverse services, some embedded with human roles;
- policy-controlled and non-policy-controlled items.

The content available through NSDL includes (though not exclusively) the following types, with or without policy-controlled access:
- Web pages — such as lesson plans, teacher guides, monographs, abstracts, manuscripts, scholarly journals, and still images — that are accessible and usable via conventional browsers.
- Digital items used outside the browser environment or with special plug-ins (usually after downloading), or requiring specialized access protocols. Examples include numeric data, geospatial images, moving pictures, sound collections, music scores, learning objects, and computer simulations (e.g. simulations of real-world objects and processes).
- Discussions on special topics archived from community discourse (e.g. email threads).
- Digital proxies for physical items, such as textbooks, lab supplies, and specimens.
- Thematically organized collections of the above (potentially nested).

Most content-specific aspects of the library are addressed via library services. For example, the initial release of NSDL will include a content-based search service restricted to textual documents in common formats.[4]

Services: Services for displaying, processing, and analyzing images, maps, and other scientific data, as well as specialized portals, may be characterized and "discovered" somewhat like other library content. Non content-based services (such as help desks or community forums) will focus on some form of social interaction, though recorded discourse from such interaction may well become NSDL content. Some services are compound, i.e. they depend on other services and therefore must interoperate via matching or brokered protocols. Some services may provide real physical objects (e.g. books or specimens) that correspond to digital proxies. As with content, some NSDL services will have policies that constrain usage; typically, such policies will be part of a metadata record describing the service.

Discovery: The act of discovering NSDL content (or services) generally entails using a general-purpose Web browser to call upon one or more services designed for searching, browsing, or querying NSDL collections. Nonbrowser contexts for using NSDL discovery capabilities also are envisioned. For example, the user of a geographic information system (GIS) may wish to search NSDL for maps, images, or gazetteer services without ever leaving the GIS environment.

Access: Access to (discovered) content within NSDL is often very simple, such as when the desired resource is openly available on the Web. Access becomes more complex when use of the resource is controlled by policy or when the resource requires more than a browser for effective use.

Whenever access entails the use of services, or when the desired resource is itself a service, then it becomes necessary to interface such services (applying protocol matching or brokering) to the end-user's information environment. A minimal analysis shows use of the resource to include the following:

[4] NSDL is being developed by a large federation of groups working on collections and services well integrated. NSDL now is up to 118 projects. One key track is for collections and one track is for services. Examples of collections and services will be provided later.

- unpacking the resource (as may be typical with learning objects and software packages);
- viewing it (e.g. visualizing a large data set) or listening to it;
- processing it (e.g. loading the resource into a spreadsheet or computer model);
- incorporating or assembling it into other (new) resources;
- storing, sharing, or publishing it for use by others.

Tailoring: User adaptations of NSDL (e.g. to match personal needs or to support a specified group of students) generally take one of two forms. The first is the creation of constrained views of NSDL, such as portals that are designed for specific education levels. A second form of tailoring is the creation of individual or group “profiles” that trigger automatic setting of user-interface parameters.

Social interaction: NSDL will continue the long historical tradition of libraries as centres of scientific discourse. NSDL supports the following activities by users:

- launching and joining electronic discussion groups;
- reviewing, editing, or annotating content developed by others;
- posing or answering questions at a human-mediated “help desk”;
- posting messages and announcements that reach a target audience;
- participating in collaborative educational or library development.

NSDL cannot be characterized as a one-way flow of information from providers to users. Users are both recipients and creators of primary content, secondary content, and metadata, especially as they give shape to the NSDL social context.

NSDL will create and manage a registry of services. Service descriptions are expected to include the nature of the service (human, programmatic, or both); a broad categorization (e.g. marking, discovery, content manipulation, as above); interfacing technology (e.g. SOAP, Z39.50); and perhaps semantic metadata, such as subject area and educational level.

8.1.2. NSDL core integration

The core integration effort has concentrated on building the central framework for the NSDL, with key services. Work to integrate the first NSDL collections into this framework has been done for the initial release in December 2002. In the next phase, the focus will be on maximizing the impact of the library on education. This will require enormous growth in the scale of the library, and partnerships with all facets of the science education community.

The core integration team provides the organizational and technical glue that binds the NSDL projects into a single coherent whole. Computers and networking can significantly enhance learning, when combined with pedagogy that is informed by current understandings about cognition, knowledge creation (i.e. constructivism), and the dynamics of collaborative learning [LTCOGN]. Thus a central NSDL challenge is to create not just a rich repository but also an *intellectual commons*, where students and educators interact and are stimulated to change the way they teach and learn.

Networks of institutions and individuals (educators, librarians, learners, publishers, parents, etc.) are to be stimulated to become stakeholders in NSDL, utilizing its resources, enhancing its quality, and finding it to be a place for fruitful discourse on education and learning.

The library is considered to be a cognitive tool, fostering the active creation of knowledge by both teachers and students. In the case of teachers, the knowledge gained will pertain to effective modes of pedagogy, in addition to subject knowledge about the worlds of science, technology, engineering, and mathematics.

Fitting NSDL to the needs of different user groups is made possible by the technical strategy of *one library, many portals*.

Over the next five years NSDL is expected to serve millions of users and provide access to tens of millions of digital resources [NSDLCO].

The architectural design (see Figure 5) for the initial NSDL phase is based on sharing of human and machine-generated metadata and exploitation of that metadata for the deployment of core services (e.g. search and discovery).

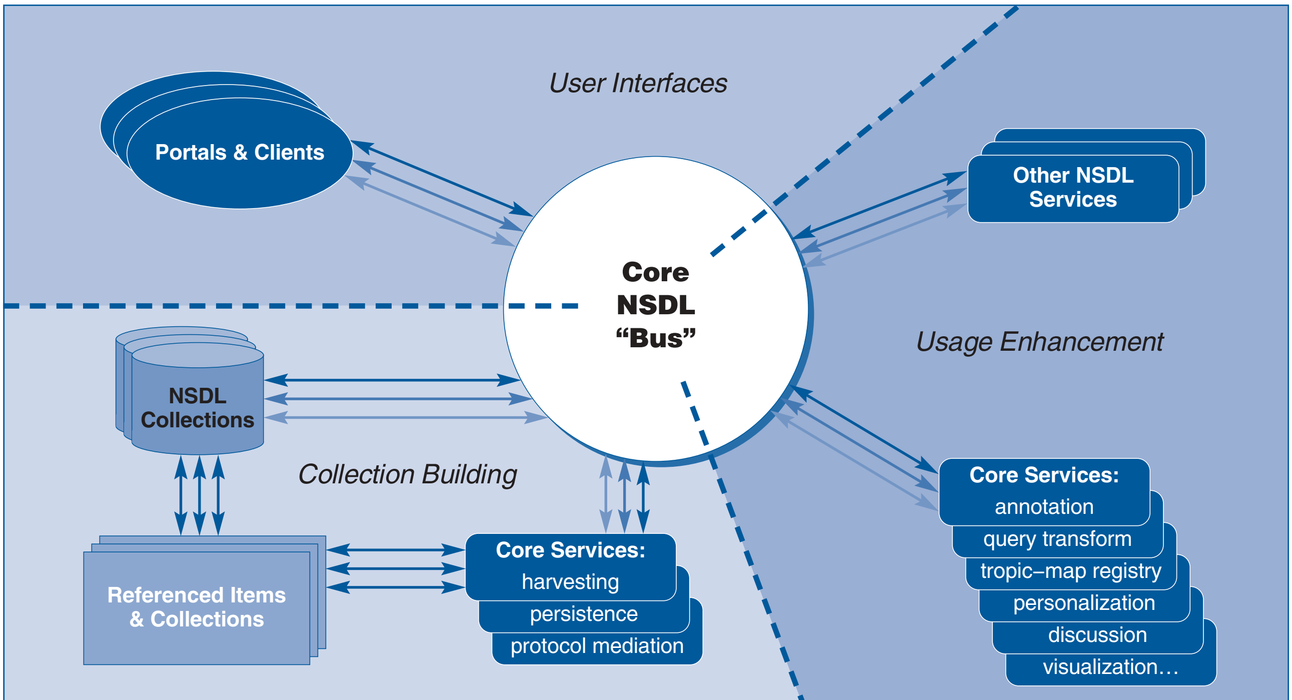

Figure 5. NSDL Architecture

8.1.3. Components of NSDL core architecture

As is illustrated in Figure 5, the core architecture of NSDL involves a number of integrated components.

Metadata repository

By combining metadata from many collections, the metadata repository can be considered to be a generalization of the concept of a union catalog, i.e. a catalog that combines records from many libraries. For metadata NSDL is adopting the following strategy:

- Collect item metadata from cooperating collections in any of eight supported "native" formats [NSDLME].
- When appropriate, automatically crosswalk native metadata to qualified Dublin Core [DCQUAL], which will provide a lingua franca for interoperability.
- When item-level metadata does not exist and where possible, process content and generate metadata automatically [METAGE].
- Accept that item-level metadata will not always exist but mandate that collection-level metadata always exists. Concentrate limited human effort on the creation of this collection-level metadata.
- Assemble all metadata, core and native, item and collection, in the central metadata repository.
- Expose metadata records in the repository for service providers to harvest.

The NSDL Standards and Metadata Workgroup, whose members represent all NSDL projects, identified the following list of preferred metadata element sets:

- Dublin Core,
- Dublin Core with DC-Ed extensions,
- LTSC (IMS),
- ADL (SCORM),
- MARC 21,
- Content Standard for Digital Geospatial Metadata (FGDC),
- Global Information Locator Service (GILS),
- Encoded Archival Description (EAD).

Three categories of metadata are available and can be stored in a metadata repository:

1. *Collection-level metadata:* The repository contains a registry of the collections that are known to the NSDL. There is a metadata record for each collection. It uses qualified Dublin Core, including descriptive information about the

material in the collection (e.g. courseware for high school biology) and, where available, technical information about protocols, formats, etc., and access restrictions and authentication requirements.
2. *Original item-level records:* Whenever a collection is able to supply item-level metadata that follows one of the preferred standards, the original records as supplied by the collection are stored in the metadata repository.
3. *Normalized item-level records:* The preferred metadata standards differ considerably in their syntax and semantics. Many collections implement only part of a standard and others have no standardized metadata. To enable the delivery of services that require consistent information about items from such diverse collections, all item-level metadata records are normalized by metadata crosswalks to Dublin Core.

By providing item-level metadata both in its original form and also as normalized metadata, the repository offers service providers a choice. Services that are able to make use of the original metadata can use it. Others can use the simpler, normalized records.

Qualified Dublin Core (with DC-ED extensions [DCEDEX]) is the normalized format for both item and collection metadata. Eight native metadata formats are supported. The metadata repository stores both the native metadata and the DC metadata.

The following approaches for adding metadata to the metadata repository are supported: metadata ingest via OAI (Open Archive Initiative), metadata ingest via FTP, email, or Web-upload (batch); metadata ingest by direct entry or metadata ingest by gathering.

Both the Dublin Core and native metadata records in the repository are made available to services through the OAI protocol. No restriction is placed on access to the Dublin Core records, making it possible for any party, NSDL-funded or external, to create a service building on the data in those records. In most cases, the contents of the native metadata records are also open-access.

Search services

Generally speaking, any item that has metadata in the repository is accessible via the search and discovery component. Where possible, the search and discovery component also allows search by the actual content of the resources corresponding to a record in the repository. The content is accessed using open network protocols (e.g. HTTP or FTP) linked via the identifier in the metadata record. In the first phase of NSDL, for content to be available via search, the content must be freely accessible over the Internet and it must be stored in one of a small set of textual formats: open formats such as ASCII text and HTML and a handful of proprietary formats, such as PostScript, PDF, and Microsoft Word.

Content-based search initially supports only textual queries, meaning that nontext items (e.g. images, sound recordings) are accessible via metadata only. Thus, the discoverability of nontextual items depends on adequate descriptive metadata: e.g. in the DC title or description elements. Search and discovery services are available using metadata, content, and any combination of the both. The query language that supports these services is independent of the architecture. However, the service provides a language modeled after Z39.50 type 102 ranked list queries [Z203RLQ]. The search engine interacts with its clients — portals, not people — using the SDLIP protocol [SDLIP].

Authentication and authorization

The core access management system relies on standard (e.g. Kerberos [KERBER], LDAP [LDAP05]) or emerging protocols (e.g. Shibboleth [SHIB01]) to distribute identity verification (authentication) and cohort membership (authentication) to the administrators of distinct communities of users.

User profile server: The NSDL core architecture includes a profile server, which holds attributes associated with a user. Portal interfaces, or other services, may use the Profile Server to store and retrieve information to customize a user's experience. For example, a portal may store search preferences and histories, disability information, grade level, etc., in the profile Server and adjust its interface behavior when the user next visits the portal.

Rights management broker: The core architecture also includes a rights management broker service. The rights broker enforces access decisions for items in the library based on the characteristics of the user and of the item.

8.1.4. Levels of interoperability in NSDL

Three levels of digital library interoperability are identified [NSDINT]: federation, harvesting, and gathering.

Federation: Federation can be considered the conventional approach to interoperability. In a federation, a group of organizations agree that their services will conform to certain specifications (which are often selected from formal standards). The libraries that share online catalog records using Z39.50 are an example of a federation. Another federation is the ADEPT project for geospatial materials, led by the University of California at Santa Barbara, USA; one of the NSDL partners in the core integration production team [ALEXBI].

Harvesting: The Open Archives Initiative is based on the concept of metadata harvesting [OAI001]. The metadata about collection items can be harvested by service providers and built into services such as information discovery or reference linking. While services built by metadata harvesting may be less powerful than those provided by federations, the burden of participating is much less. As a result, many more organizations are likely to join and keep their systems current.

Gathering: Even if the various organizations do not cooperate in any formal manner, a base level of interoperability is still possible by gathering openly accessible information using a Web crawler. The premier examples of this approach are the Web search engines. Because there is no cost to the collections, gathering can provide services that embrace large numbers of digital libraries, but the services are of poorer quality than can be achieved by partners who cooperate directly.

Some of the most interesting Web research at present can be thought of as adding extra function to the base level, which will lead to better interoperability, even among totally noncooperating organizations. Even though the concept of a fully semantic Web is still only a vision, it is reasonable to expect that the level of services that can be provided by gathering will improve steadily. ResearchIndex (also known as CiteSeer) is an example of a digital library built automatically by gathering publicly accessible information. Thanks to the assistance of NEC, both the data and software used in ResearchIndex are being integrated into CITIDEL (www.citidel.org), one of the NSDL collection projects.

8.1.5. Initial release of NSDL

The initial release of NSDL will be in December 2002, involving mostly the results of funding provided to over 60 projects. The discussion below sketches key aspects of that release, and then proceeds to consider future growth plans, which will involve over 30 more additional projects funded in the fall of 2002. Section 8.1.6 outlines the plan for continued evolution. Section 8.1.7 contains a discussion of the NSDL community, which will expand both in terms of size and engagement as NSDL content and services expand. Section 8.1.8 sketches the organization of NSDL, with support by NSF, guidance by the NSF-selected National Visiting Committee, steering efforts of the Policy Committee, and focused efforts by Standing Committees.

Providing technology: An architectural framework has been established [NSDLNSF] to support the core technology and to integrate resources created by others, including:

- Building a metadata repository, which holds native and standardized metadata records for each collection and item known to the NSDL;
- Implementing interfaces (primarily OAI protocols) by which the metadata repository is populated and by which its contents may be accessed to construct various library services;
- Interfacing and testing fundamental library services;
- Building the primary portal from user-interface components that can be reused in a broad array of additional portals.

Operating core services: The first library-wide services include the main NSDL portal, a comprehensive search and discovery service, and an initial authentication and authorization service.

NSDL emphasis in the next phase will be on growth: expanding the collections, adding new services and partners, and above all encouraging use for education.

Growing the collections: A major goal is to make the NSDL a comprehensive library, covering all areas of scientific, technical, engineering, and mathematics education. This requires a vast expansion in the number of collections, far beyond those funded by the NSF. NSDL strategy has three collection-building strategies: working with other digital libraries, partnering with publishers, and building collections automatically from the Web.

Ultimately, the success of the NSDL will be judged by its impact on STEM education. A variety of reforms is required, including the adoption of inquiry-based science learning, and emphasis on the process of doing science and the integration of scientific research into education. Digital libraries promise to be a powerful tool in realizing the goals of these reform initiatives.

The NSDL goal is to be the resource of choice for disciplinary, curricular, and pedagogical issues. Hence, NSDL should be a partner and major asset for every significant initiative in education for science, technology, engineering, and mathematics.

NSDL focuses its efforts on developing strategic partnerships with a key set of professional organizations, including the National Science Teachers Association and the American Library Association (ALA). Within the ALA, NSDL aims to develop partnerships with the American Association of School Librarians, the Association for Library Services to Children, and the Public Library Association, as well as the Association for Library Collections and Technical Services. The outcome of these partnerships for NSDL will be an intimate knowledge, through consultation and feedback, of the pedagogical and resource needs of teachers and librarians as they face the daily realities of American classrooms and schools.[5]

Many of the collection track projects of NSDL are actively engaged in collaborations with professional societies related to their specific subject areas. These collaborations lead to increased access to extensive collections of materials. An example is the CITIDEL (Computer and Information Technology Interactive Digital Education Library) project, a collections project that covers the entire realm of computing materials. CITIDEL has entered into an accord with the Association for Computing Machinery (ACM) by which the metadata of the entire ACM digital library is accessible for search through CITIDEL. In addition, ACM has agreed to provide free access to specific entries in the ACM DL that are particularly relevant for educational purposes. NSDL's aim is to engage a group of such partners, who cover all disciplines and all levels of science education. The societies can help in understanding how adequately NSDL resources represent the most current resources and research in their fields.[6]

Project 2061 [ATLAS] (see discussion later in this survey) supports "the development of new tools for teachers, curriculum developers, and textbook authors and publishers". As the national library for scientific education, NSDL intends to become the distributor of these educational resources.

8.1.6. Planned NSDL evolution

Extending the architecture

Describing and integrating services: A multi-service framework requires standards for describing services, to permit interactions among them and access to them by users and agents [NSDLNSF]. It is planned to leverage recent developments in Web services, including the Web Services Description Language (WSDL) work within the W3C and perhaps the .net work from Microsoft.

Creation and management of federated ontologies: The NSDL will encompass resources from many disciplinary domains. Providing a coherent end-user view will require developing and establishing conceptual relationships among distributed resources.

Creating, packaging, and accessing complex content: The resources available in the NSDL will be rich and complex. Dynamic and multimedia content requires access methods and digital object structuring standards that extend far beyond those now available via Web standards.

Scientific data: Of particular value in science education are numerical data sets and other digital objects that are of little use except in conjunction with appropriate tools.

Preservation of digital resources: The utility of the NSDL depends on the integrity of the resources, notably their longevity and stability. There is no one solution to the problems of digital preservation and longevity.

[5] This orientation to get a bottom-up feedback may provide an average pattern of requirements. How to lead to excellence based on the average is an open issue.

[6] To make such partnership efficient, the resource representation in DLE should be more focused, more structured, more curriculum or subject oriented.

Data provenance: Another factor related to the integrity of digital resources is their record of origin. This is especially true for scientific data, where origin and derivation are critical to determining the veracity and utility of the data.

Annotation and review frameworks: The metadata repository has flexible means for annotating resources. This is intended to support educational services by linking reviews of resource quality and appropriateness, and encouraging collaborative reviews.

Authentication of digital objects: Communities of authors, publishers, librarians, and consumers have consistently emphasized the importance of document authenticity. In the digital realm, comprehensive document authentication is in only the early stages of development, but considerable progress can be made with simple tools.

Accounting: For authors, an important value in NSDL will be to understand how their work is being used (e.g. by whom and how often), so that they receive the same level of academic credit for digital citations as for print citations. The planned NSDL access management infrastructure is uniquely positioned to supply this information to an author without compromising the privacy of the user. Work on logging standards [GONCAL] across NSDL may also have broad impact.

Collaborative learning: Supporting collaborative learning is of central importance to NSDL. NSDL plans to work with the collaborative learning community to begin the process of identifying interoperability standards for integrating collaborative learning systems with digital library substrates.

Large-scale growth

Version 2 of NSDL is scheduled for summer 2004 [NSDLNSF]. By the end of this phase, NSDL should be a very large, production-quality library, with substantial richness of educational resources, and a rapidly growing community of users and contributors. The community support and evaluation activities will be continued.

In parallel with this growth in resources, a major emphasis on educational outreach is planned, building partnerships with groups of educational users, including discipline-specific groups, sharing resources with them, providing services via NSDL, understanding their needs, and seeking resources that meet them.

Fundamental to this growth, in Version 3 and beyond, is an increased emphasis on creating communities of educators and learners, and feedback mechanisms to facilitate parent-teacher-student interaction. One of the important directions is a technical framework, for sharing and managing semantic information, to support the sharing of learning concepts across diverse communities and the creation of new learning communities within the NSDL rubric.

8.1.7. NSDL education community

NSDL provides opportunities to develop a new science, technology, engineering, and mathematics education community [PATHWAY] that is interdisciplinary in nature. There is intrinsic value in recognizing and accentuating the connections among the knowledge bases, skills, and methodologies employed by the contributing disciplinary communities. NSDL can make a substantive contribution towards bridging current disciplinary boundaries. An NSDL that effectively integrates concepts, knowledge, and methods across the STEM disciplines will be a much greater resource than a collection of discipline-specific libraries.

The NSDL community is an aggregate of many existing and intersecting communities, including:
- disciplinary groups,
- educational group,
- technology and information science group,
- special interest groups (e.g. policy makers, journalists, commercial sector),
- learners of all kinds — students and citizens-at-large.

8.1.8. NSDL governance

The NSDL governance model [NSDGOS, NSDLGOI] must:
- create a public sphere of influence that promotes partnering and shared vision and balances interests which sometimes differ,

- establish a framework for accountability,
- reflect the special nature of key players.

To achieve its purpose, the NSDL governance model (Figure 6) is built around committees, subcommittees, task forces, and interest groups that provide several domains of guidance. In decreasing order of authority (and formality), the advisory entities in this model are as shown in Table 1.

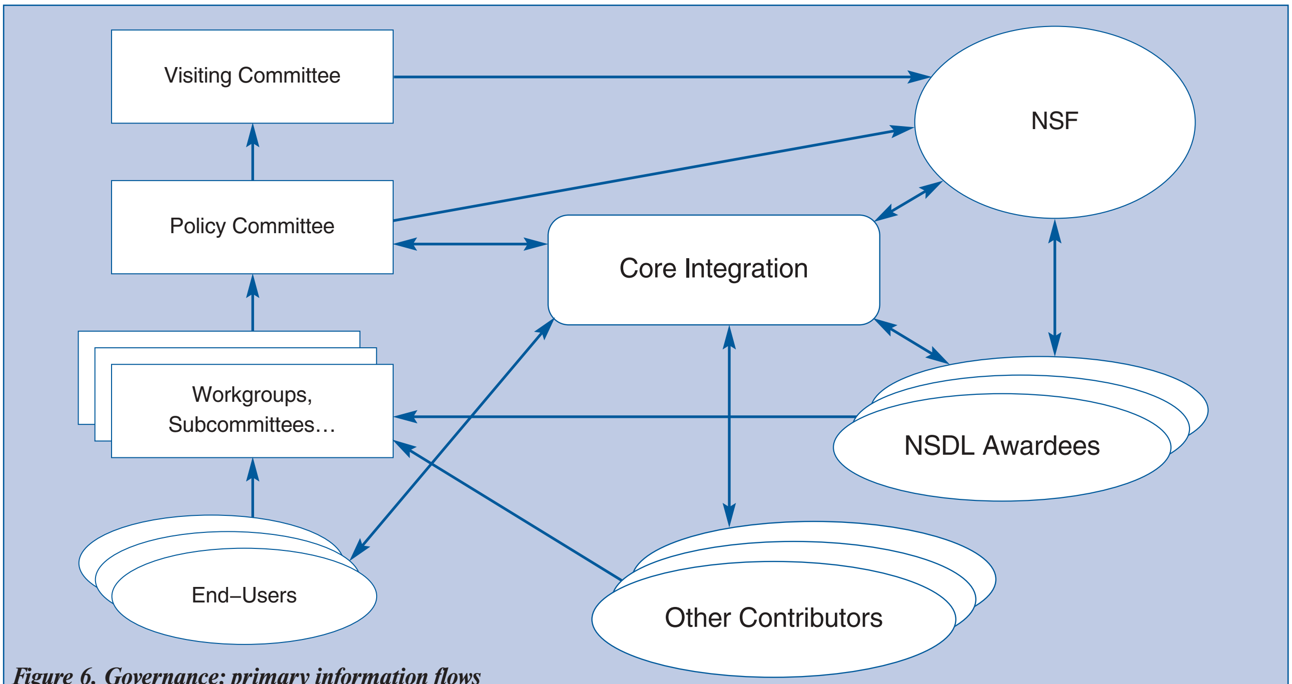

Figure 6. Governance: primary information flows

This table summarizes the key elements of the governance model, organized by the types of guidance and decisions that pertain:

Entity	Area of Decision Making and Guidance	Composition
National Visiting Committee	Accountability and top level strategy and feedback (to NSF and the CI team)	NSF appointed advisory group
Full Assembly	Adopts the governance model, elects the Policy Committee, and responds to polls when appropriate	One representative from each project (PI or designee)
Policy Committee	Empowered to act on behalf of the Assembly, articulating policies, strategies, and priorities	Elected by the Assembly from a slate prepared by a Nominating Committee
Topical Standing Committees	Each reports to the Policy Committee, which charters it to reflect a subdomain of responsibility	Defined by the Policy Committee
Subcommittees and Task Forces	Ad hoc groups may be chartered to address specialized needs of the NSDL such as a long–term governance, incorporation, and communications. Each reports to a specified committee	Defined by Policy or Standing Committees as needed
Spontaneous Interest Groups	Address ad hoc common interests from individual project members	Composed of individuals within and beyond the NSDL projects as appropriate
Project–specific	Operations, management, and budgets	Individual awardees

Table 1. Elements of NSDL governance

Topical Standing Committees set by the NSDL Policy Committee in spring 2002 include: Community and Services, Content, Educational Impact and Evaluation, Sustainability, and Technology.

The Community and Services Standing Committee is focused on library users and learners, and also on supporting other communities and views, such as teachers, the primary and secondary education communities, and other groups involved in education. Relative to earlier governance documents, it serves as a users committee. The Content Standing Committee serves as the primary group to recommend or adopt policies associated with the creation, development, and maintenance of individual collections within NSDL. It addresses issues of quality, in conjunction with other groups, such as Educational Impact and Evaluation. The purpose of the Educational Impact and Evaluation Standing Committee is to ensure that participatory and stakeholder evaluation principles are integrated into the development and implementation of NSDL. The Sustainability Standing Committee's mission is to facilitate the development of an NSDL entity for long-term sustainability of the NSDL collective through a diversified funding stream. Intellectual property issues and publisher relations are among the tasks of the Sustainability Subcommittee. The Technology Standing Committee is a forum for those involved in building the technical infrastructure of the NSDL. As such, the Committee serves as the voice of the Assembly in matters ranging from technical standards (such as metadata) through the technical integration of resources and services into the NSDL architecture.

8.2. Digital Library for Earth System Education (USA)

8.2.1. DLESE version 1.0

The Digital Library for Earth System Education (DLESE) [DLESNSF, DLESEC, DLESME] mission is to "improve the quality, quantity, and efficiency of teaching and learning about the Earth System, by developing, managing, and providing access to high-quality educational resources and supporting services through a community-based, distributed digital library". Over the past two years, DLESE has emerged to support the specific educational needs of the geoscience community within the larger NSDL network. In the tradition of community libraries, DLESE can fundamentally change how students learn, instructors teach, and researchers interact, by providing new ways of sharing information, tools, and services. DLESE serves as a vehicle for the geoscience community to respond to the challenges of systemic educational reform and the changing technological landscape.

Version 1.0 of DLESE is the operational library that was released in August 2001, with approximately 1,000 carefully selected resources in its initial collection. Version 1.0 provides educational discovery features that enable users to search by grade level, educational resource type, and keyword. This version also contains a resource cataloguer, and community oriented services, such as discussion forums for working groups and a community-posting tool. Version 1.0 creates an initial three-tier library architecture in which requests made in the user interface are mapped to operations over centrally held metadata through middleware services. To ensure interoperability with the NSDL, support for the Open Archives Initiative harvesting protocol has been implemented. This protocol is being used by DLESE to support transport of metadata between it's distributed collections.

The DLESE collections grow through community contributions from individuals or institutions. Contributions consist of individual resources or entire collections. The effectiveness of this collections model depends on having community members who are able to create and share their efforts to further common intellectual goals. The DLESE Program Center (DPC) enables the community to consciously and actively shape the intellectual framework of the DLESE collection by providing tools, components, and services that reflect DLESE policy, assure collection quality, and promote pedagogical innovation. Much of the DPC's prior work has focused on creating resource characterization frameworks (a metadata framework implemented in an XML schema), resource characterization tools, and services necessary to grow the collections and support resource discovery.

Version 1.0 supports resource discovery using keywords, grade level, and educational resource type descriptors (map, visualization, activity, lab, etc.). Discovery is provided through both basic and advanced search interfaces and a graphical browsing interface.

Library awareness and effective use of library services by educators and students have been promoted by DPC through a strong presence at geoscience professional meetings, community Working Groups (e.g. Diversity, Data Access, K-12), and through the DLESE annual meeting. These efforts have been targeted towards stimulating library development projects.

To evaluate user experiences, library collections and services, and educational effectiveness, the DPC has been focused on: (1) formative studies of current and potential library users to inform library design, (2) ethnographic studies of library building processes, (3) participant surveys to gauge the usefulness of each annual meeting and to provide broad input into library planning, and (4) online survey instruments for the beta test programme.

DLESE is a community-owned and governed digital library offering high-quality electronic resources that foster learning about the Earth at all educational levels. The DLESE Strategic Plan, which outlines the five-year vision for the next stage of library development, was formally approved in November 2001. When fully operational, DLESE will offer peer-reviewed teaching and learning resources, interfaces, and tools to allow exploration of Earth data; services to help users effectively create and use educational resources; and an "intellectual commons" facilitating sharing, collaboration, and excellence in Earth system education. DLESE users include learners and instructors in all venues, many of whom are also resource contributors, developing educational materials, providing scientific knowledge, and evaluating DLESE holdings.

8.2.2. DLESE evolution plan

During the next five-year performance period [DLESNSF], the DLESE Program Center (DPC), in conjunction with the larger DLESE community, will engage in a programme of work to achieve the following three goals:
1. Establishing DLESE as the premier, trusted source for high-quality geoscience educational resources.
2. Promoting sustainable library growth through community capacity-building and community participation in library governance, development, and operations.
3. Enabling the library to serve as a catalyst for geoscience educational reform.

These goals will be realized in stages and measured against two-year and five-year benchmarks outlined in the Strategic Plan. These benchmarks are articulated in two major library versions, Version 2.0 and Version 3.0 (Table 2).

DLESE version 2.0

In Version 2.0 (due for release in the summer of 2003), collections development services and resource discovery efforts will be extended to support science literacy standards and the Earth system science (ESS) perspective. Users will be able to search across multiple peer-reviewed collections, according to an Earth system perspective and a variety of science education benchmarks and standards. Community forums and library services supporting the effective use of resources and professional development will be available. Version 3.0 will represent a significant step towards supporting data use in geoscience education. Users will be able to search across spatially and temporally indexed resources, such as data, maps, images, and field guides. Gazetteer services will support user-centred geo-referenced discovery interfaces using place names and Earth system events. Integrated tools and services to assist with age-appropriate exploration of data will be available. Users will create and share a variety of personalized collections.

For Version 2.0 [DLESNSF], the current DLESE architecture will be extended to support distributed community collections and to interoperate with NSDL collections and services. Some of these services are already in place, in particular, a repository management system and Open Archives Initiative (OAI) data and service provider interfaces. These OAI interfaces support both DLESE's native ADEPT/DLESE/NASA (ADN) metadata framework, and the Dublin Core (DC) framework as recommended by NSDL. The ADEPT middleware, with its innovative search bucket architecture and mechanisms to work with geo-spatial searching, will be incorporated into the DLESE discovery system. The ADEPT bucket architecture enables collections to map their metadata framework onto common semantics (buckets) that can then be used by shared library services, such as discovery systems.

Thus, for Version 2.0, DPC efforts will focus on extending the resource metadata framework and resource characterization tools to enable contributors and collections developers to characterize their holdings according to a variety of national science literacy standards and an Earth system science perspective. In Version 2.0, DLESE discovery services will be extended to enable users to search and browse DLESE collections using the US national science education standards, the geography standards, and the Earth system science (ESS) vocabulary.

DLESE version 3.0

In Version 3.0 [DLESNSF] (the version to be delivered at the end of the five-year performance period), the architecture will be extended to support resource discovery using geo-referenced information, i.e. geospatial footprints (e.g. regions

	Version 1.0	Version 2.0	Version 3.0
Web Dimension			
Discovery	Grade level, educational resource type, key words	Educational discovery: standards, Benchmarks, ESS	Geo–referencing and Earth system events
Collections	Modest (about 1,000 resources)	Reviewed collections with broad topical breadth for primary, secondary, and higher education	Thematic and personalized collections
Community	Working groups, community postings	Community Forum, Prototype teaching and learning centre	Robust teaching and learning centre
Services & Support	Search tips available; support for resource contributors provided	Online end–user support (FAQs, Tutorials); community support personnel available	Robust community input mechanisms; email support provided
Data Dimension			
	Data sets and tools present in broad collection	Specially constructed thematic resources combining data, tools, and learning aids for primary, secondary, and higher education	Metadata extensions for integrated access to Earth data and learning resources
	DAWG convened; demonstrator projects under way	Specifications for data access and delivery developed with DAWG	Ability to create personalized inventories of data and tools
Use Dimension			
Target Audience	Contributors, faculty early adopters	Extended to mainstream primary, secondary, and higher education teachers/faculty	Extended to students and informal learners
Volume Capacity	1,000s of sessions/month	10,000s of sessions/month	100,000s of sessions/month
Building Dimension			
Governance	Committees established; some policies in place	All major library policies developed and in place	Mature community–based governance
Collections Development	Focus on object–level growth	Collection–level growth; collections developers toolkit	Brokering for persistent collections in place
Processes	Manual growth and maintenance processes	Collections management and harvesting processes automated	Automatic resource/collection cataloguing
Evaluation	Web metrics collected; beta testing is routine	Collection metrics collected; formative evaluations are routine	Summative evaluations of DLESE under way

Table 2. DLESE versions

on the surface of the Earth specified by latitude and longitude) and temporal footprints (e.g. a point or slice in time). The architecture also will be extended to support existing and emerging federated protocols necessary for interoperating with distributed DLESE and NSDL partners. For example, supporting the widely used Z39.50 protocol would enable DLESE to interoperate with existing libraries and bibliographic services.

Version 3.0 will also include support for geo-referenced discovery. Inquiry-oriented Earth system education makes use of many nontextual resources, such as maps, images, and data, to bring hands-on activities into the classroom. Indexing by temporal and spatial dimensions will be supported. Users will be able to locate such geo-referenced resources using a direct manipulation interface (e.g. selecting a region on a map) or via textual search terms, such as “New York”. To support these discovery features, the place name gazetteer developed by ADEPT and the Map and Imaging Laboratory at UCSB, USA, will be incorporated into the discovery system.

In Version 3.0, the discovery system will be extended to support the educational use of Earth data and data analysis tools. Educational use of real time and archived data is one of the foundations of the geoscience education reform movement and is critical for integrating geoscience research into educational practice. Tools appropriate to primary, secondary, and higher education learning contexts are needed for scientific data mining, analysis, and visualization.

Throughout the performance period 2003-2007, the DPC will continue to design library interfaces and services for broad accessibility. In future versions, partnerships with key professional societies that serve underrepresented populations will be established.

Metadata collaboration

ADEPT, DLESE, and NASA's Joined Digital Library are taking their current frameworks as the basis for developing a common metadata framework, the ADEPT/DLESE/NASA (ADN) Joint Metadata Framework. ADEPT, DLESE, and NASA are using XSL to transform the ADN XML catalogue records to Dublin Core records for use in NSDL. For the ADN Joint Metadata Framework, DLESE is responsible for schema development, maintenance, and documentation for required metadata and an overarching framework schema; ADEPT is responsible for geospatial and temporal components and model description schemas. All three organizations are responsible for collection level type information. A collection level metadata framework based on, using the Alexandria Digital Library collection level metadata is being investigated.

Since DLESE will be a federated system, search buckets might be used to facilitate searching [ADEPTA]. Search buckets are a small set of high-level metadata attributes that can be used to make collections appear to have the same searchable metadata. They are also designed explicitly for querying. The underlying metadata fields of a record do not map one-to-one to a bucket. Rather, groups of fields map to one bucket and fields may map to more than one bucket.

Collections policy

The scope of the DLESE collection is Earth system education, with particular emphasis on interdisciplinary areas. The collection will favour materials that bring the Earth into the classroom or other learning site, connecting the general with the specific, theory with evidence, and the global with the local. The collection will favour materials which are well documented, easy to use, bug-free, motivational for learners, pedagogically effective, scientifically accurate, and which foster mastery of significant understandings or skills. The types of resources and tools to be collected include research and education materials and sources of content, tutorials and learning resources, field trips, and tools.

Community collections

Community collections that collaborate with DLESE include NASA's Joined Digital Library (JDL), the Digital Water Education Library (DWEL) for primary and secondary education, the Electronic Encyclopedia of Earthquakes, the Geoscience Digital Image Library (GeoDIL), the Atmospheric Visualization Collection (AVC), and the COMET Multimedia Database.

Relationship to geoscience, science education, and information technology partners

DLESE has established relationships with science and science education professional societies, including the American Association for the Advancement of Science (AAAS), the American Geological Institute (AGI), the American Geophysical Union (AGU), the Incorporated Research Institutions for Seismology (IRIS), the National Science Teachers Association (NSTA), and the emerging Center for Ocean Sciences Education Excellence (COSEE) and Earthscope efforts. These partners provide outreach opportunities for DLESE through their conferences and workshops. AAAS library interfaces are under development to enable users to discover resources via the AAAS strand maps [ATLAS]. This would enable resources indexed via either the benchmarks or the standards to be queried and retrieved.[7] AAAS intends to integrate DLESE into their professional development offering for educators. DLESE has a codevelopment partnership with NASA and plans to extend its partnerships to other federal and state agencies during the performance period 2003-2007.

In the data arena, DLESE has two important partnerships, with the US NSF funded Unidata Program Center and the emerging Geoinformatics effort [SINHA00]. Both are developing mechanisms for describing Earth system data for discovery and use in educational and research settings. In the THREDDS (Thematic Realtime Environmental Distributed Data Services) project the Unidata Program Center is developing services to enable users to create personalized and sharable collections of data and data analysis tools.

[7] This is an absolutely different model and interface. Serious work on resources registration is required to achieve the indexing of resources via benchmarks.

8.3. Computer and Information Technology Interactive Digital Education Library (USA)

NSDL has many projects funded through the collection track. CITIDEL is a representative example, employing the latest in digital library technology. As such it is a microcosm of NSDL, with aspects related to content, services, and community building.

8.3.1. CITIDEL foundation

CITIDEL is built on a strong foundation of prior work and engages a dynamic and active community of computing educators. The CITIDEL goals are to:

- provide the greatest possible access to high-quality resources that contribute to learning about computing at any level;
- provide the greatest possible visibility to the products of talented educators and developers who produce materials that support student learning about computing;
- provide access to materials that may have value within specific contexts, without undue barriers;
- support open exchange of materials and comments on materials discovered;
- support community building and mutual support of teachers and learners with common interests;
- support the use of discovered materials in learning activities.

Foundation work to address these goals includes the use of the Computer Science Teaching Center (CSTC, http://www.cstc.org), a resource submission, review, and sharing resource funded by NSF in 1997. CSTC allows easy submission of materials for wide dissemination by simple upload procedures. Volunteers register to review materials, thus providing a level of assurance to resource users that the materials have been evaluated and found to be working and sound. Materials are categorized to support searching by subject area and type of resource, as well as by author name or resource title and keywords. CSTC was extended with multimedia-related content as a result of the US NSF-funded Curriculum Resources in Interactive Multimedia project [CRIM] that emphasized knowledge modules to aid in education [MCCKM].

In addition to CSTC, CITIDEL incorporates access to resources collected by other projects, including JERIC (the ACM *Journal of Educational Resources in Computing*, http://www.acm.org/pubs/jeric/ [CASSEL]), ResearchIndex (http://www.researchindex.org), the CS Virtual History Museum, the Networked Computer Science Technical Reference Library (NCSTRL, www.ncstrl.org), PlanetMath (a collaborative mathematics encyclopedia, http://planetmath.org/), and the ACM Digital Library (http://www.acm.org/dl). Efforts are under way to include metadata from the IEEE Computer Society, DBLP (the Computer Science Bibliography at http://dblp.uni-trier.de/), and relevant parts of the Networked Digital Library of Theses and Dissertations (NDLTD) collection, so that by 2003 there may be roughly a million records. The number of partners participating in CITIDEL grows regularly, just as the number of collections projects within NSDL grows. One of the challenges, at both the level of individual collections and the level of NSDL overall, is to assure the continuing accessibility and the status of maintenance of member collections. As is evident from the list of partners, CITIDEL includes resources of many types. Some are created specifically for supporting computing education. They include demonstration modules and simulators that display the behavior of computing components or algorithm performance, for example. Other resources are general-purpose tools or publications including refereed journal articles. Each of these has its place in supporting learning.

CITIDEL permits direct submission of materials to the library, but provides access to resources supported by others, too. CITIDEL also obtains materials by crawling the Web in search of candidate materials; with the permission of the author these materials are submitted to CSTC for evaluation and thus become part of CITIDEL. CITIDEL is really an example of NSDL on a smaller scale, bringing together a diverse set of resources and presenting them to the learning community as an integrated collection.

Digital libraries have differing policies regarding the evaluation of resources presented to its users. The approach of CITIDEL is to have a variety of types of materials, but to have them clearly distinguished for the user's convenience. Thus, CITIDEL searches will yield refereed resources from ACM journals and casual products of an individual's creativity. The searcher will know the source of the resource and can apply judgment with regard to the amount of confidence to place in the resource.

CITIDEL intends to serve the entire community of computing educators in the United States, with the understanding that access to the rest of the world is available as well. To reach the largest possible population, the resources of

CITIDEL must be available in both English and Spanish immediately, and in other languages in time. Spanish language access is a central component of the initial design effort.

Overall, the CITIDEL project presents materials with characteristics in three dimensions as shown in Figure 7. Note that quality is dependent upon source and corresponding approaches to add value.

Figure 7. The three dimensions of CITIDEL resources

8.3.2. CITIDEL growth

A key aspect of building a comprehensive digital library by integrating diverse resources is a plan for effective metadata harvesting. CITIDEL is a participant in the Open Archives Initiative. All member collections of CITIDEL are OAI sites and CITIDEL regularly harvests metadata from these sites in order to keep search results current with the collection revisions.

From Figure 8, we see that CITIDEL is both a harvester and a provider of OAI data. CITIDEL provides its metadata to the overall NSDL OAI harvester, contributing to the umbrella library, while serving the specialized needs of the computing community.

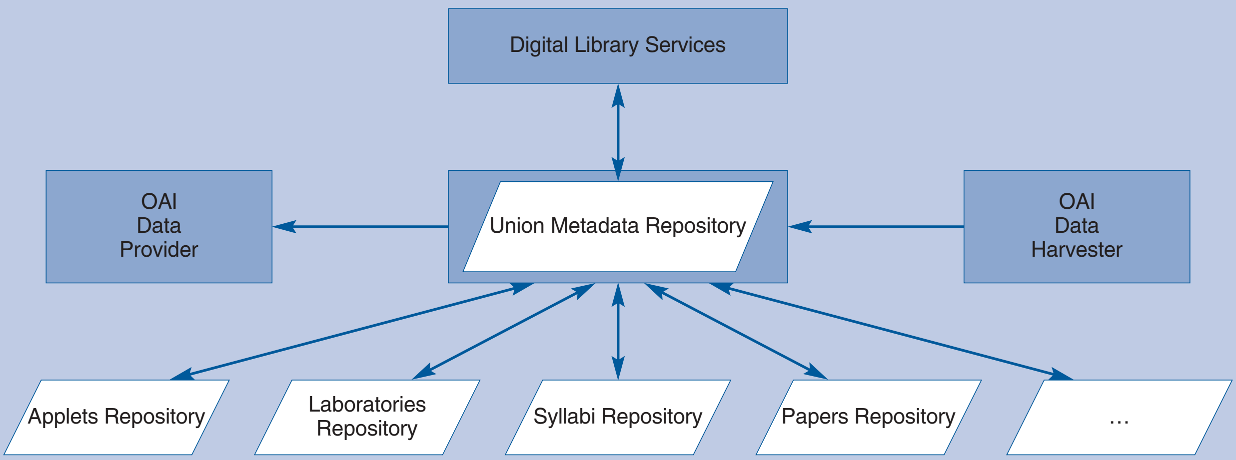

Figure 8. Metadata collection and provision in CITIDEL

CITIDEL provides a separate view for each of its user communities, including educators, learners, and administrators of the resource. Services are tailored to the needs of each group as shown in Figure 9. Filtering profiles allow a user to specify qualifications for the resources to be returned in response to searches. For example, some users may choose not to see materials that have not been peer-reviewed; other users may choose to limit the resources they see based on their education level or amount of familiarity with the topic. These profile characteristics will certainly change over time.

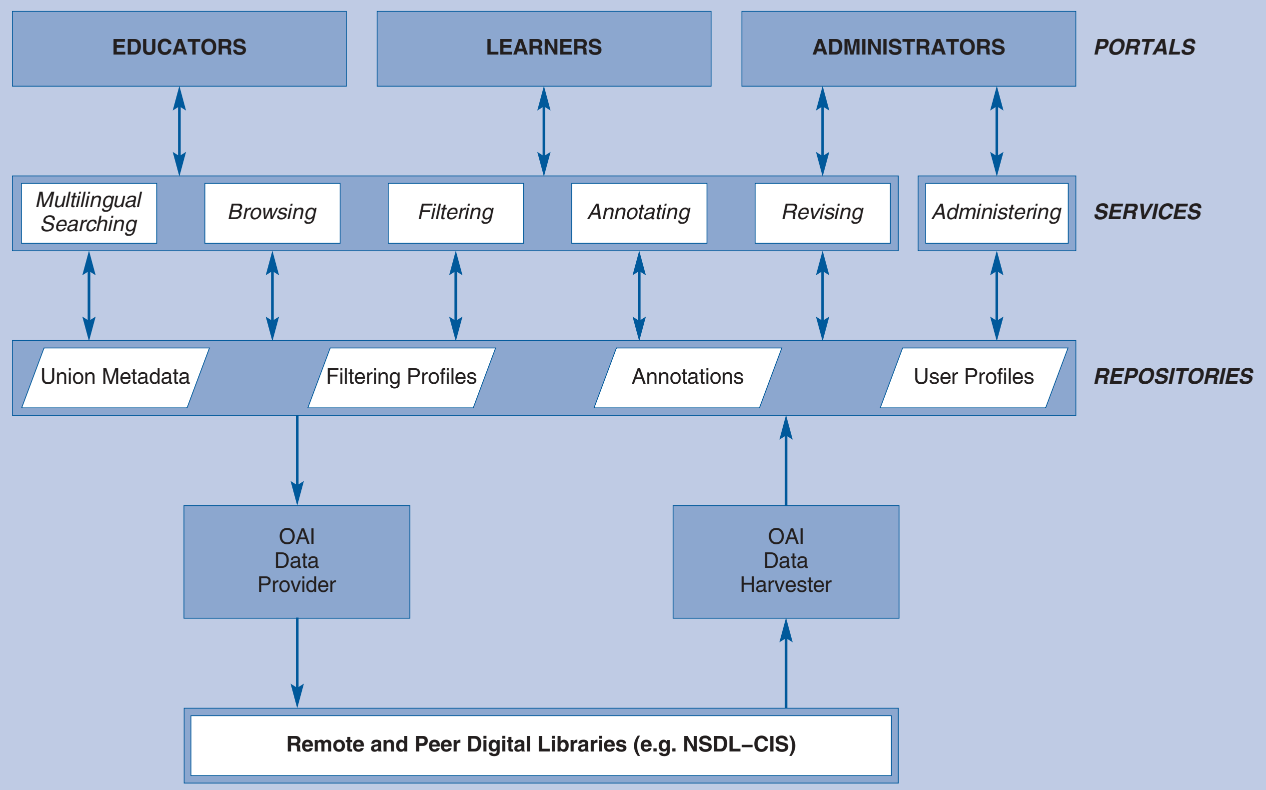

Figure 9. Architecture for local and interoperable CITIDEL services

CITIDEL additions

CITIDEL is more than a collection of resources and more than a collection of collections of resources. In addition to the usual DL services, such as browsing, filtering, annotating, revising, and administration, and the more unusual multilingual support, CITIDEL includes a number of innovative service activities. An example is the VIADUCT (Virginia Instructional Architect for Digital Undergraduate Computing Teaching) project.

VIADUCT is a supplement to the CITIDEL search and browse facilities that allows a user to group a subcollection of resources into a meaningful and useful teaching facility. The premise is that gathering resources is only the first step in using them effectively. The resulting collection of materials must be organized and presented in the context of overall learning goals. VIADUCT allows a CITIDEL user to gather materials and insert them into a syllabus, which defines a unit of a course as developed by an individual teacher. The resulting syllabus is a vehicle for the teacher to interact with students using resources discovered in CITIDEL, a concrete representation of that course unit for future use and revision, and a form of that course unit design that can be shared with other teachers for their use and their further development. Syllabus components constructed and populated with VIADUCT can easily be published in CITIDEL, becoming a new generation of resources for sharing in the computing educational community.

8.3.3. CITIDEL connections

CITIDEL exists to serve the community of computing educators. Thus, its connections to that community are critical. One of the foundation stones of CITIDEL, CSTC was created with funding from the US National Science Foundation, with support from the ACM Education Board. CSTC has benefited by close collaboration with the ACM Special Interest Group in Computer Science Education (SIGCSE, http://www.acm.org/sigcse/). Through the generosity of ACM, the contents of the *Journal on Educational Resources in Computing* and the publications of SIGCSE are freely available through CITIDEL. Other ACM digital library content is searchable through CITIDEL, with access to the full content limited by the usual subscription requirements. Like the overall NSDL, CITIDEL must address the issues of recognizing a subscriber to a member service in order to support access and to provide fee-based access to nonsubscribers.

Because CITIDEL is both a product of and a service to the computing community, its design and implementation provide illustrations for many of the issues that other DLs will have to address as they mature. CITIDEL is accessible at www.citidel.org.

Related projects

The CITIDEL project is connected with two NSDL services projects. One, led by Hsinchun Chen and Ann Lally at the University of Arizona, has developed a system supporting concept maps, integrated with searching, summarization, and visualization. Thus, CITIDEL and other collections can be searched and the results can be used to build concept maps whose nodes link to resources or other lower-level maps. The other project, GetSmart (http://collab.dlib.vt.edu/runwiki/wiki.pl?GetSmart), has been deployed in autumn 2002 in computing and information technology courses at Virginia Tech and the University of Arizona. While related to efforts described below (e.g. strand maps in Project 2061), GetSmart focuses on empowering students to create their own concept maps, as an aid to understanding readings and course materials, and to developing their term projects. In addition to supporting study and documentation efforts, GetSmart helps with class presentations and group discussion (e.g. with group and class-level maps that result from comparison of the maps of individuals).

In addition, the “DL in a box” effort (http://dlbox.nudl.org/), led by Su-Shing Chen at University of Florida, and involving Joe Futrelle at the National Center for Supercomputing Applications (NCSA), builds upon the Open Digital Library (ODL) effort launched at Virginia Tech (http://oai.dlib.vt.edu/odl/). The idea behind ODL [SULEMA1, SULEMA2] is to have a collection of components that conform to a set of lightweight protocols that minimally extend the protocol used in the Open Archives Initiative. Some of these components are key elements of the CITIDEL software. Recently, the ODL component for annotation was integrated into CITIDEL, demonstrating how easy it is to add services in the ODL framework. The DL in a box effort aims to extend, package, and document the current set of components so that NSDL and other efforts that require digital library software support can expand upon this framework, building their core technology around well-tested and modular routines that can be easily extended.

8.4. Networked Digital Library of Theses and Dissertations (USA)

NDLTD, www.ndltd.org, is an international digital library initiative supporting education [NDLTD1, NDLTD2]. It represents a rather distinctive approach to DLE development, which fits well with UNESCO interests, and has had strong UNESCO support in recent years. A number of developing nations have become involved, since there are immediate benefits, and since there is a clear and simple sustainability model. NDLTD has benefited from support of the latest in digital library technologies, and so also is interesting since it is based upon open source software.

8.4.1. Education focus

NDLTD focuses on graduate education and research. It deals with a key phase of graduate studies, namely the preparation of a thesis or dissertation (TD), but also can cover undergraduate theses or graduate project reports. Though historically TDs were submitted in paper form, NDLTD aims to move the community of scholars rapidly toward an emerging genre of electronic theses and dissertations (ETDs).

Given the emerging collection of ETDs, students as well as other researchers and even educators derive significant benefit. By and large, universities make ETDs freely available, and NDLTD services support centralized discovery. There is intense interest in this content from around the world, since it serves many educational purposes.

First, students preparing their own thesis or dissertation can learn from the works of others. Of course, they can learn specific content, related to their research interest. But also they can learn about style and organization, presentation, and even about use of multimedia and hypermedia technologies. Since TDs generally have long bibliographies, as well as extensive literature reviews, students may be aided in carrying out their own literature review, and may have an easier time getting oriented to the related work in their field. Indeed, the world's collection of TDs constitutes one of the largest and most comprehensive collections of overviews of research in varied specialty areas. Further, since TDs provide details on the methods used by researchers, they also are a rich mine regarding research methods and particular methodologies. By mining the TDs of the world, students should also be able to avoid wasting time to reinvent the work already carried out by others, when that related work is not easily discovered.

Second, since ETDs are electronic documents, it is possible to work with small portions that fit in with a specialized need. If the ETD is a traditional document, formed simply as "electronic paper" (e.g. as a PDF version of a Word document), then even if the entire work is in a single file, tools will allow a single page or small span of pages to be isolated. If the ETD is broken into files, such as for each chapter, then these smaller segments can be readily accessed alone. Further, if there is multimedia content, a single image, video, audio segment, data file, spreadsheet, or other component can be utilized. Teachers can show these components in class, can assign them to students in connection with an online assignment, or can encourage their study in connection with project work.

Further, ETDs may help students identify topics that they can explore in conjunction with their own research. Many theses have sections that summarize open problems. Others suggest additional work that can be done to extend the recently completed study. According to one report, a student who prepared an ETD in South Africa was contacted by a student at the University of California, Berkeley, USA, who wanted to extend the first student's efforts.

Finally, working with the NDLTD collections and services, students gain familiarity with modern digital library systems. Since 1998, NDLTD has provided federated search access to a number of sites maintained by members, and more sophisticated technology can now help in that regard [MARIAN1]. Today, the MARIAN (Multiple Access Retrieval of Library Information with Annotations) system, a modern research digital library system that is being enhanced with multilingual support, can help enhance access to ETDs [MARIAN3]. Indeed, current research on generating digital libraries based on formal specifications is allowing MARIAN to support content from a variety of sites, through a variety of search and harvesting mechanisms, with views governing how the internal semantic network representation is managed to support user information needs [MARIAN2].

8.4.2. Learning by doing

Further educational benefit derives with regard to NDLTD when students learn by doing. Many students learn best this way, rather than by simply reading or listening to lectures. Since a thesis or dissertation may take months or years to develop, and often is an object of pride to a young scholar, learning while preparing an ETD may have particularly lasting effect. Part of the learning process deals with exposure to the latest document processing and electronic publishing technologies.

Thus, as time moves forward, it is more and more common for students to prepare their thesis or dissertation with a computer. Then it is a small matter, if proper prior planning was undertaken at the university, for the student to prepare a version suitable for long-term preservation (e.g. using PDF or XML, as well as suitable standards for multimedia, like JPEG and MPEG). Finally, it is quite simple for students to fill in a form (that thus captures needed metadata), and to upload files, thus getting familiar with electronic submission processes.

Such exposure to electronic publishing and author-submission procedures is particularly important for modern scholars. Many funding agencies, including the US NSF, require proposals to be prepared and submitted electronically. Many conferences and journals require papers to be managed in a similar fashion. Those reviewing papers or proposals often must access others' works from a DL-like system, and must submit their reviews in similar fashion. In short, the activities of preparing and submitting an ETD are good preparation for a number of commonly required scholarly tasks.

Clearly, however, technologies related to electronic publishing, as well as to use of digital libraries, rapidly change. Fortunately, there is a good tracking of students learning through work on their ETDs with young researchers engaging in common activities. In other words, if universities always ensure that their students who work on ETDs learn about state-of-the-art methods in electronic publishing and in digital library access, then all will be prepared to work as scholars after graduation.

This learning process actually has an even broader effect. First, students who will make their works widely available have greater need to learn about matters of digital preservation, patent, copyright, and intellectual property than they might if only a paper document were prepared. Thus, they may be exposed to these important issues, which unfortunately are often are poorly understood by many senior members of academia. Second, since students often turn to their mentors when preparing their theses or dissertations, there is increased exposure of the faculty to many of the above-mentioned issues and skills. Thus, ETD activities often have a stimulating effect on campus scholarly publishing and library budget discussions, and may facilitate institutional progress on these matters.

8.4.3. A cost-effective packaging

While many DLE-type efforts discussed in this survey require considerable investment, and have not yet led to sustainable programmes, work with ETDs clearly illustrates how sensible packaging of efforts can lead to beneficial and cost-effective practices. The following illustrates how this may work. Though there are many variants, we focus here on the simplest approach, which was adopted at Virginia Tech, USA, after 1996, and has since been adapted in a variety of other contexts.

Since the beginning of 1997, students preparing a TD at Virginia Tech have conformed to the ETD requirement then put in place (after pilot efforts dating back to 1987). In particular, as part of their graduate activities, all who prepare a thesis or dissertation must do so electronically, and must submit their ETD for review by the graduate school. Once the work is approved, the automated system that handles submission and subsequent processing then allows library handling, including cataloguing. Works are made accessible in accordance with the approval form that has been signed by the student author as well as the student's graduate or supervising committee. The university library accepts no paper submission, so the entire workflow is managed electronically. This leads to so much saving in terms of paper handling and transport, avoidance of binding and shelving, and even student copying, that the overall programme costs less than the earlier paper-based approach.

These savings result from the connection of training of graduate students with the skills they need as scholars, and then placing responsibility on their shoulders, through a requirement. It is facilitated by the use of automation software that is made available as open source by Virginia Tech. That software even includes allowing a university to become a data provider as part of the Open Archives Initiative, facilitating development of a global union catalog of ETD metadata.

Benefits are considerable, even beyond the above-mentioned savings and learning. First, student research becomes much more available. Whereas paper theses and dissertations historically have been read by very few beyond the set of faculty involved in advising and examining the student, ETDs are typically read by hundreds or thousands. Given the investment of time and effort by the student, and the institutional support that they receive over several years of study, it is clearly sensible to leverage that investment by making student work accessible. Indeed, based on experience at a number of NDLTD sites, a university may have hundreds of thousands of accesses per year to its ETD collection, thus expanding its visibility in the research community. This system may be of particular value in developing countries, where research is now rarely disseminated on the international scene; in the world of ETDs, every participating university can benefit from a truly level playing field.

On the global scale, the benefits are even clearer. If there is full participation in NDLTD worldwide, there should be more than 200,000 ETDs made available each year. If production follows the awarding of graduate degrees, this will level the playing field regarding research access. Furthermore, whereas there are efforts under way to make journals more widely available globally, due to economic barriers it is likely that it will take a number of years for smaller universities, especially in developing countries, to gain access to 200,000 journal articles per year.

8.4.4. Current status

There are national ETD efforts in Australia, Brazil, and Germany in addition to expanding efforts in India, Korea, and USA, as well as support by national institutions like the British Library and the national libraries of Canada and Portugal. All told, there are over 150 members of NDLTD, with roughly half in the USA, but with fastest growth outside the US. Some members, like OhioLink, which represents the 79 colleges and universities in Ohio, demonstrate that the number of universities engaged in ETD production is well over the 150 count of members mentioned. Now that NDLTD, which was launched in 1996, is maturing as a federation, more and more of its members are shifting to require ETDs, which will ensure a rapid increase in ETD production.

Many NDLTD members are leading universities, and many are known for their innovative practices. California Institute of Technology, Johns Hopkins University, Louisiana State University, Michigan Tech, Massachusetts Institute of Technology, Pennsylvania State University, University of Texas at Austin, University of Virginia, Virginia Tech, and West Virginia University are among the more than 60 institutions that have joined in the USA. International members come from many countries, including Australia, Brazil, Canada, China, Finland, France, Germany, Hong Kong, India, Korea, Mexico, Norway, Portugal, Russian Federation, South Africa, Singapore, Spain, Sweden, and United Kingdom. In addition, NDLTD benefits from members that represent disciplinary interests such as for physics (PhysDiss) and mathematics (MathDiss).

UNESCO support is focusing on increasing ETD efforts in South America, the Caribbean, Eastern Europe, Africa, and other regions. UNESCO funded preparation of online and CD-ROM versions of an ETD Guide (http://etdguide.org), with over 400 pages, encouraging and supporting new initiatives with guidance for students, faculty, administrators, and trainers. The guide is available in English, French, and Spanish. UNESCO offices in Uruguay and France (headquarters) are coordinating efforts, working in concert with NDLTD. Some of the support will allow interested parties from key areas to attend a special training programme undertaken in conjunction with ETD 2003, the latest in a series of annual international conferences, scheduled for Berlin in May 2003.

8.4.5. Advanced technology

NDLTD has been actively engaged in the development and deployment of the latest electronic publishing and digital library technologies. In 1988, a document type definition for electronic dissertations was developed to facilitate authoring using SGML, and in subsequent years various tools were tested to help in such preparation. In recent years this has led to document type definitions as well as schemas that relate to authoring using XML. In the early 1990s, as soon as PDF was available in test form, ETDs were prepared using that format. And many ETDs have made use of standards like JPEG and MPEG as soon as they were established.

In 1998, access to the distributed collections of ETDs was facilitated by a federated search system set up by NDLTD at its collection site, www.theses.org. In 1999, NDLTD was represented in the Santa Fe, New Mexico, USA, meeting that led to the Open Archives Initiative, and NDLTD has been one of the lead groups involved in the OAI steering and technical committees ever since. As a result of over two years of international meetings and discussion, the ETD metadata standard, ETD-MS, was devised, and along with MARC 21, is the recommended representation format, which, along with Dublin Core, all NDLTD OAI data providers are encouraged to support. OCLC is working to make available more than four million TD records through OAI, and Virginia Tech maintains a union catalogue for ETDs that was among the first union collections harvestable as part of OAI.

Digital library support for NDLTD continues to be state-of-the-art. In the mid-1990s, studies were undertaken with the test versions of the IBM digital library software, and full-text content searches were demonstrated. The above-mentioned Open Digital Library framework was tested early on for applicability to NDLTD, and an experimental service using that component-based approach has been available for over a year. The Virtua system, one of the most comprehensive commercial library automation packages, now extended to be a full digital library system, has been deployed by VTLS Inc. in connection with its support of the union catalogue, with interfaces supporting 14 languages and content available already in six languages.

MARIAN, the research digital library system developed at Virginia Tech, has worked with NDLTD data since it first became available. MARIAN has demonstrated a variety of types of support for distributed collections, including through OAI, Harvest, and Z39.50 [MARIAN1]. MARIAN also supports a modular approach to DL development, but in a more principled way than most, building as it does upon a theoretical framework for digital libraries (5S, which builds upon five key constructs: societies, scenarios, spaces, structures, and streams). NDLTD support is now available as driven by a description of NDLTD collections and services in the 5S specification language, 5SL [MARIAN2]. MARIAN supports Unicode and searching in multiple languages, incorporates a view mechanism similar in concept to that in database systems, supports varied ontologies, and is implemented in Java in versions available as open source software, scalable to run on laptops or as a distributed system running on a network of computers [MARIAN3].

Because of its educational focus, its international scope, its simple sustainability model, and its incorporation of the latest DL technology, NDLTD is particularly amenable to integration with other UNESCO efforts, such as the move toward DLEs.

9. Current Digital Libraries for Education: European projects

9.1. The Distributed National Electronic Resource and the hybrid library (UK)

DNER [DNERLP, DNERAV] is a generic term used to describe the wide range of information-related activities and services supported by the Joint Information Systems Committee (JISC). The DNER coordinates informational activity in three broad areas:

1. The creation of a strategic national resource of educational and learning materials: this consists of collections of materials made available on a national level via networked systems designed to deliver them to the end-user.
2. The creation of a framework for community resources: institutions and individuals within institutions also manage or create resources that may be articulated as part of a wider "community" resource.
3. The creation of a resource discovery framework for a global resource collection.

The collections themselves consist of a wide range of resource types: scholarly journals, monographs, textbooks, abstracts, manuscripts, maps, music scores, still images, geospatial images, vector and numeric data, moving picture collections, and sound collections.

For content and service delivery the predominant approach is plain HTTP, but this limits how the services can be used because of the limited semantic structures in the protocol. The aspiration here is to make more services available through more structured protocols such as Z39.50, LDAP, and the Open Archives Initiative. In that way, their data is more directly accessible and manipulable for others to use. This aspiration is broadly in line with the vision of the "semantic Web" espoused by the World Wide Web Consortium. The most important difference between the DNER and hybrid library projects is the most obvious one. Whereas the hybrid library projects are based in individual institutions, the DNER is a national service.

9.2. Scholnet and Cyclades: Extending the role of digital libraries (EU)

Scholnet (IST-1000-20664) and Cyclades (IST-2000-25456) are two digital library projects funded by the European Union 5th Framework Programme and coordinated scientifically by the IEI – CNR [SCCA01]. Both projects aim at extending the role of a digital library by providing services that support remote communication and collaboration among scholars. In particular, the goal of Scholnet is to develop a digital library providing an enhanced set of specialized services, while Cyclades is focused on the need to develop a service environment on top of large heterogeneous and multidisciplinary interoperable archives.

Scholnet (http://www.ercim.org/scholnet) aims at enabling the immediate dissemination and accessibility of technical documentation within a globally distributed multilingual community. Scholnet also aims at developing a digital library infrastructure to support communication and the collaboration within networked scholarly communities.

In order to achieve this objective, Scholnet will provide:

- *Traditional digital library services on multimedia documents:* These services enable scholars to communicate through the publication of and access to not only textual documentation, such as technical reports, project deliverables, or workshop proceedings, but also videos of tutorials or seminars (possibly synchronized with corresponding textual slides), training sessions, project presentations, demos, etc.
- *Handling of document annotations*: Annotations can be textual notes, ratings, links, etc., associated with either the entire document or its parts. Annotations can be authored by different people and will have public or group restricted access privileges.
- *Monolingual and multilingual search and retrieval services:* Monolingual search is provided in all of the project languages; if a user specifies the search language, the system searches only those documents that contain information in that language. In addition, a cross-language search facility allows users to query in their own language and retrieve documents matching the query in other languages.
- *Automatic personalized information dissemination service:* When a new document arrives in the digital library, a proactive facility sends messages to those users who, on the basis of their system-maintained profiles, are potentially interested in its contents.

From the technical point of view the Scholnet infrastructure will be built by extending, and partially rethinking, the basic services provided by the ERCIM Technical Reference Digital Library (ETRDL) (http://www.iei.pi.cnr.it/DELOS/ETRDL).

Cyclades (http://www.iei.pi.cnr.it/cyclades) will develop an open, collaborative virtual archive service environment supporting both single scholars and scholarly communities in carrying out their work. In particular, it will provide functionality to access large, heterogeneous, multidisciplinary archives distributed over the Web and to support remote collaboration among the members of communities of interest.

Cyclades will run in the data environment composed by the archives that adhere to the Open Archives Initiative's harvesting protocol specifications (http://www.openarchives.org). From the technical point of view, Cyclades will consist of the following federation of independent but interoperable services:

- *Access:* supports harvest-based information gathering, plus indexing and storage of gathered information in a local database.
- *Query and browse:* supports the users in formulating queries and develops plans for their evaluation. In particular, it provides an advanced multilevel browse facility, completely integrated with the search facility, that allows one to browse at the levels of schema, attributes, and documents.
- *Collection*: provides mechanisms for dynamically structuring the overall information space into meaningful (from some community's perspective) collections.
- *Personalization:* supports information personalization on the basis of individual user profiles, and profiles of the working communities the user belongs to. User and community profiles are automatically inferred by monitoring the user behavior.
- *Recommendation:* provides recommendations about new published articles within a working community. The choice about what recommendations to send, and to whom, is based on the profiles of both the user and the working community.
- *Collaborative work:* supports collaboration between members of remotely distributed working groups by providing functionality for creating shared working spaces referencing users' own documents, collections, recommendations, related links, textual annotations, ratings, etc.

10. Advanced frameworks and methodologies related to DLEs

10.1. Instructional course development in the presence of learning module repositories

A gap exists between current DLEs and their use in course development. DLE evolution in this direction can be envisaged as applying the successful experiences of instructional course development in the presence of teaching module repositories. Impressive experience in this respect has been accumulated in the Cooperative Program for Operational Meteorology, Education and Training (COMET) that was formed in 1989. The COMET programme [COMET] was originally envisioned as a broad effort to affect meteorology education and training in the United States. However, the programme has recently been involved in activities to enhance meteorology education in universities and meteorological services throughout the world. The COMET mission is to serve as a premier resource to support, enhance, convey, and stimulate scientific knowledge about the weather for the benefit of weather-information providers, educators, and users.

A typical COMET multimedia-based learning (MBL) module contains one to 12 hours of highly interactive instruction and incorporates case studies, graphics, animations, and video to provide an effective educational experience. Concepts are introduced via both computer text and spoken dialogue and are reinforced by displays of such graphic materials as time-sequenced satellite data, radar data, or videos demonstrating laboratory experiments. At various points in each module, students have the opportunity to test their understanding of the concepts presented. If the student answers incorrectly or would like more detailed information, she/he can access additional material that is typically presented by an expert on the topic.

COMET modules provide interactive, multimedia training on satellite meteorology, marine meteorology, hydrology, fire weather, and much more for the meteorologist, educator, and weather enthusiast. The COMET programme has instructional designers, author-programmers, staff meteorologists, graphic designers, and animators who produce computer-based training modules and related learning materials.

Currently COMET does not espouse the concept of a digital library. At the same time, to make the graphic images created in COMET projects available to the sponsors and the public, to enhance media reuse by the sponsors, and to enhance educational materials available to the meteorological community at large, the COMET programme has developed a Hydrometeorological Multimedia Database. Besides the images, instructional components can be found by text search applying keywords in a query. These examples show how a digital library of instructional materials can be organized to make materials customizable and reusable in new courses. Finally, information on various weather cases (e.g. on severe weather, hurricanes, winter weather, fire weather, floods, cyclones) is arranged into a simple classification system. They include radar and satellite data, model data, text data, upper-air and profiler data, and other contributed data sets. These case studies are being used in National Weather Service offices and university classrooms across the United States.

Some of the instructional modules contain concept maps in their definitions showing interrelationships of concepts used (Figure 10). For example, Mesoscale Convective Systems (MCSs) appear in many forms, ranging from a relatively disorganized mass of convective cells, to highly organized convective lines. At times the systems appear like small synoptic-type cyclones with bands of convective cells spiraling around the cyclone centre. The instructional modules and exercises provide the opportunity to apply many of the concepts presented in both the Conceptual Models and the Physical Processes portions of the Web site (http://www.comet.ucar.edu/modules/MCSMatrix.htm).

Such concept maps may serve as hints to define an ontology and to structure the information relevant to the instructional course on MCSs. This would provide for relating to those definitions of DL resources that contain any relevant information. Such a DL mediation interface is considered to be an advanced DL feature.

10.2. Learning objects for reuse

Learning objects are elements of a new type of computer-based instruction grounded in the object-oriented paradigm of computer science. Instructional designers can build small (relative to the size of an entire course) instructional components that can be reused a number of times in different learning contexts. Learning objects include multimedia content, instructional content, learning objectives, instructional software and software tools, and persons, organizations, or events referenced during technology-supported learning [LOMN00].

The Learning Objects Metadata Working Group (a working group of the LTSC of the IEEE) describes the purpose of learning objects as being: "to enable computer agents to automatically and dynamically compose personalized lessons for an individual learner". Regretfully, this intention has not succeeded in any instantiation.

In fact, diverse definitions have arisen for learning objects and similar terms used with the same meaning [CLOIDP]. People have coined the term "instructional design theory", which is in its infancy. The online book [IULOOB] that contains a collection of philosophical papers (e.g. [CLOIDP, NOIOBJ, RBLPSS]) discusses the nature of learning objects and instructional design theory as if no information systems design with reuse existed before.

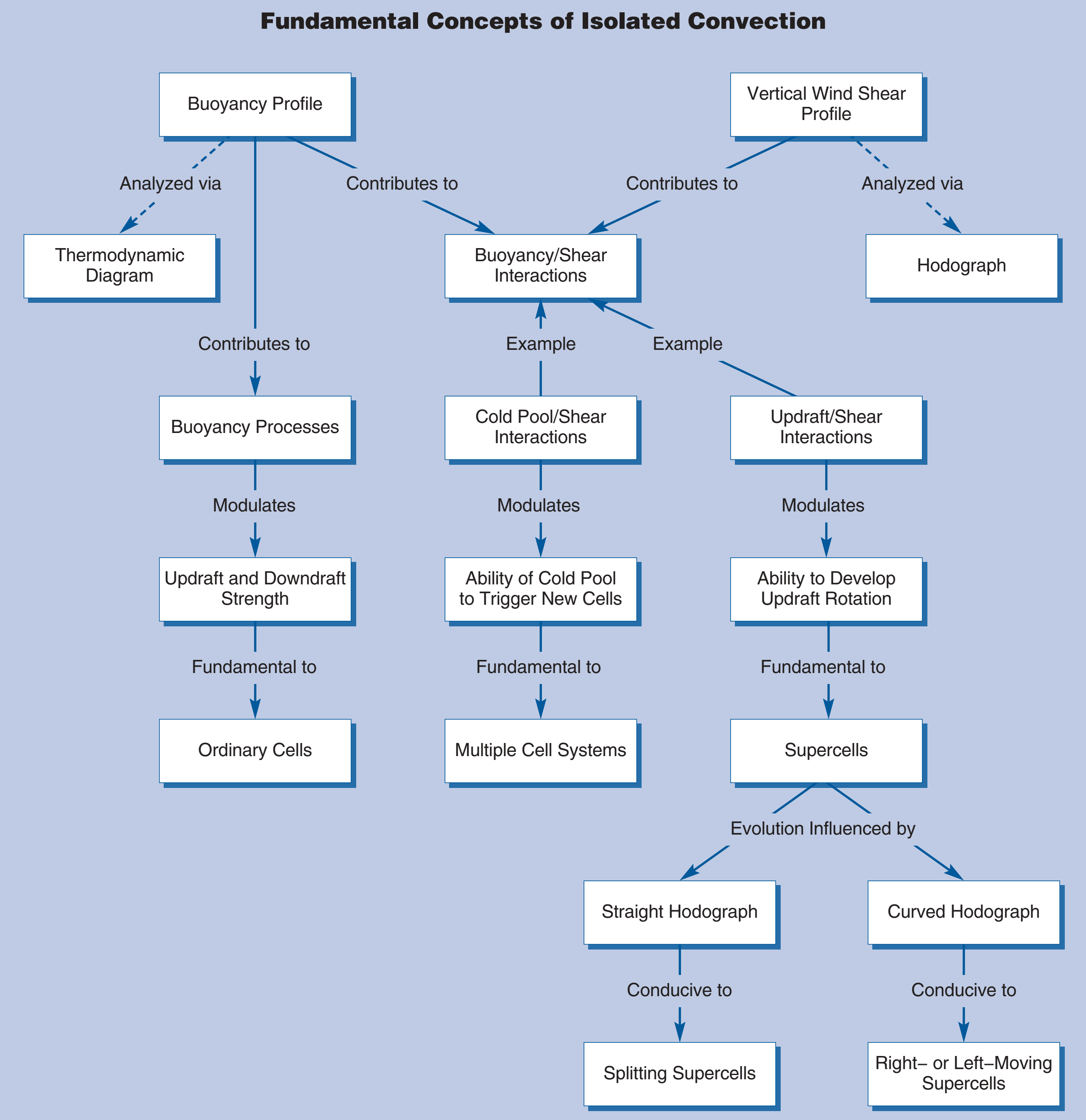

Figure 10. Concept map from COMET Web module "Mesoscale convective systems: Squall lines and bow echoes"

At the same time, a note on the COMET experience [COMIDP] shows that an engineering approach does exist for at least the top-down process of instructional development. The process includes the following phases: Initial Requirements Definition, Analysis and Planning, Design, Content Development, Component Development, Product Development, Publication, Implementation, Evaluation. Deep modifications are required to make the process a mixture of top-down and bottom-up (reuse of preexisting learning objects) approaches. The component definitions (learning objects) should appear alongside the requirements definitions. Both kinds of definitions should be provided with proper modeling allowing users to accomplish three things: (1) Users should be able to discover learning objects (stored in DLEs) and fragments of learning objects that are potentially relevant to the requirements definitions. "Relevant" means that learning objects are ontologically (contextually) suitable and fragments of requirements definitions can be substituted for learning object definitions so that users will not notice the substitution (frequently, it might be necessary to adapt learning objects to match requirements definitions). It should also be possible to specify generic learning objects that could be further customized. (2) Users should be able to compose adapted learning objects into larger definitions that would refine the requirement definition. (3) Finally, users should be able to justify that the resulting construct can serve as an implementation of the requirements definition.

For software components, an approach similar to the above has been experienced in a specific tool [CBISDT]. For instructional development, however, this remains an open issue. The state of current research only underscores this observation [DLREUSE, COMPLO].

10.3. Community organization around vast specialized educational resources and data

Development of specialized educational resources and data requires adequate organizational efforts in the relevant community. MeteoForum [METFOR], a pilot programme developed jointly by the COMET and Unidata programmes of the University Corporation for Atmospheric Research in Boulder, Colorado, USA, may serve as an example of such efforts. Under the MeteoForum concept, the international network of Regional Meteorological Training Centres (RMTCs) of the World Meteorological Organization (WMO) will work collaboratively with universities to enhance their roles of training and education through information technologies and multilingual collections of resources. MeteoForum will build a stronger sense of community among the RMTCs through Internet-based interactions and through the process of sharing educational concepts, educational materials, and hydrometeorological data with one another and with universities.

COMET has established a multilingual MeteoForum Web site to foster interaction and communication among MeteoForum members. The site, http://www.meteoforum.ucar.edu, will be coordinated with the WMO "Virtual Library".

The COMET programme brings to MeteoForum its recognized excellence in (1) distance learning (lectures, student interaction, and a large collection of educational graphics that are accessible via the Web or on CD-ROMs); (2) an outreach programme which involves partnerships between the academic research community and forecasters; (3) residence courses in advanced meteorology and hydrology that include lectures and hands-on laboratory activities; and (4) case studies, a collection of meteorological data sets for specific events (available online). Some of its vast resources have been translated into Spanish. COMET has the expertise to help MeteoForum create new materials or translate existing materials (e.g. case studies that focus on regional phenomena, such as Amazon squall lines or El Niño related floods).

Unidata [UNIDAT] activities that are of potential use to MeteoForum include: (1) providing tools to visualize, analyse, organize, receive, and share data; (2) facilitating data access to a broad spectrum of observations and forecasts (most in real time); (3) supporting faculty who use Unidata systems at colleges and universities; and (4) building a community where data, tools, and best practices in education and research are shared.

10.4. Information content for science and research more broadly

Important resources for research and education include not only archives of data but also streams of real time data, as well as software services. Such resources should become part of DLE content. We now consider DLEs from this perspective, taking the Unidata Program Center (UPC) [UNIDATI, UNIDATP] as an example.

UPC offers software and services that enable universities to acquire and use atmospheric and related data on their own computers, often in real time or "near-real time" — that is, the data are sent to participants almost as soon as the observations are made. The UPC's software and services are available to any college or university at no cost. Member institutions provide their own computers, network connections, human resources, and other requirements for participation, including access fees for certain data. Located in Boulder, Colorado, the UPC serves more than 150 universities. Through computer networking, Unidata participants are members of a mutually supportive "virtual community" — a nationwide group of electronically linked individuals who hold common academic interests in the atmospheric and related sciences and who share similar needs for data and software.

The nationwide Unidata community has established a cooperative network of computers running the Local Data Manager package and exchanging data in real time via the Internet. Known as the Internet Data Distribution (IDD) system, this network is linked to data sources at various locations and provides universities with various types of data at low cost. Participants in the IDD system can access numerous data streams, including information from surface and upper-air observing stations worldwide, radar systems nationwide, forecasting models run by the National Weather Service (NWS), national lightning detection networks, wind profilers, and satellite-borne images. In addition, Unidata does provide mechanisms for accessing some archived data sets and case studies, and some Unidata sites do archive data streams in raw, encoded form.

Unidata was founded in the atmospheric science domain; however many universities employ Unidata systems or data to support education and research that falls outside that domain. This reflects a national trend toward interdisciplinary education and research that Unidata will support by continuing to emphasize generality and broad utility in its systems, especially the data management components.

Descriptive information (metadata) about core data stream contents and about other data accessible through Unidata will be made accessible through special Unidata Web pages. Part of this objective will be to advertise and link to Web pages created by the NWS and others, such as those describing the new Geostationary Operational Environmental Satellite (GOES) sensors and the quantities they measure.

Just as the Web and digital library technologies have simplified the process of publishing and accessing multimedia documents, the Thematic Realtime Environmental Distributed Data Services [THREDDS] will provide needed infrastructure for publishing and accessing scientific data in a similarly convenient fashion. These services are intended for students, educators, and researchers to publish, contribute, find, and interact with data relating to the Earth system in a convenient, effective, and integrated fashion.

THREDDS will establish both an organizational infrastructure and a software infrastructure. A team of data providers, software tool developers, and metadata experts will work together to develop a software framework that allows users to publish, find, analyse, and display data residing on remote servers. The heart of THREDDS, however, is metadata contained in the publishable inventories and catalogues (PICats). Based on XML, PICats can be created in many different ways. Educators will incorporate PICats of illustrative data sets into educational modules that also include the tools for data analysis and visualization. Indeed students will eventually be able to use PICats to point to data sets related to their research projects, just as they now use URLs to point to relevant documents. Since they are text-based, PICats can be "harvested" and indexed in digital libraries using specialized tools that make use of the internal structure and semantic content, as well as tools similar to those used by current document search engines.

Data collections are a cornerstone of environmental research and education. New levels of accessing and using data are now achievable because of evolving technologies, even as the amount and variety of Earth system data are increasing daily. Recent parallel progress in the worlds of scientific data management and education-oriented digital libraries is highlighting common needs to discover widely distributed data sets and to use unfamiliar data meaningfully with a comprehensive set of analysis tools for:

- visualizing complex, multidimensional data;
- integrating and overlaying data from multiple sources;
- gracefully handling coordinate systems, measurable quantities, units of measure, and sampling variations.

The THREDDS system has four main components (Figure 11 — compare with Figure 5). First, data-access protocols, such as DODS, HTTP, and FTP, provide Internet access to scientific data sets, using URLs to name the data sets. These existing protocols are already in wide use by data servers in the scientific community. Second, the proposed PICats provide lists of available data sets and a framework for specifying the semantics of data sets, sometimes called "use metadata". The PICats will contain data set inventory and data set description components. Third, the proposed PICat servers are distributed processes that monitor a set of PICats and provide integrated discovery services. Fourth, existing

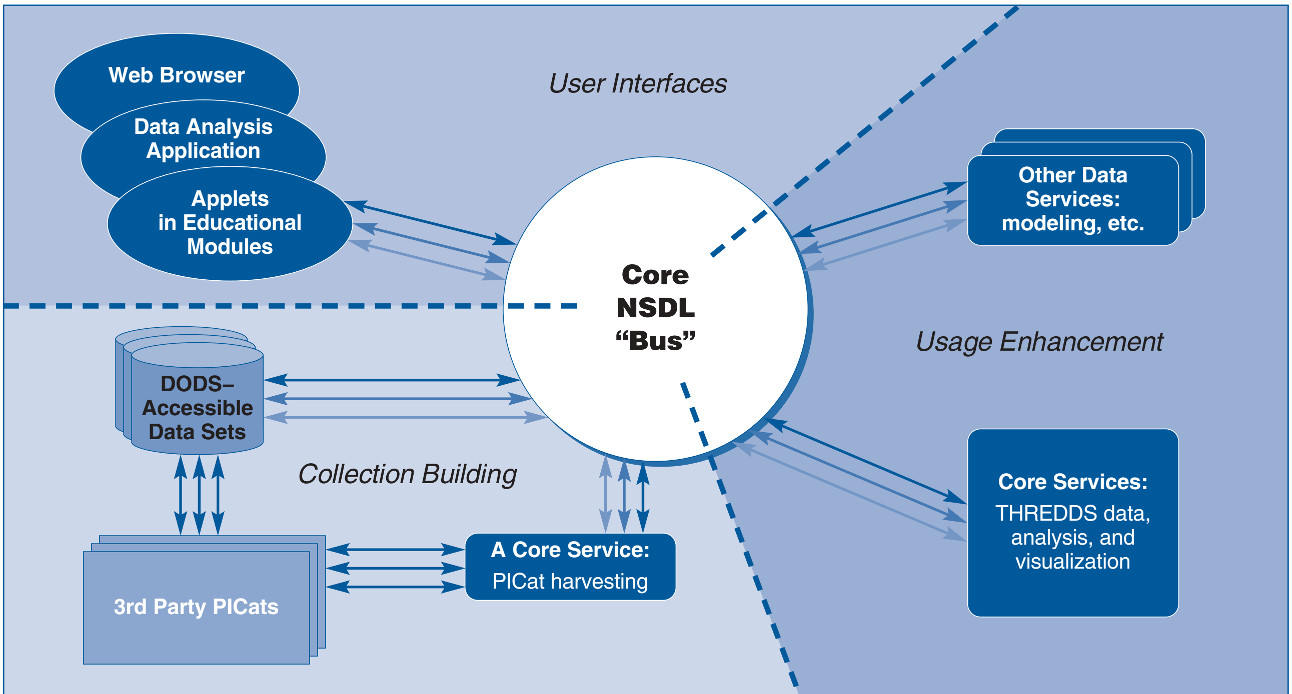

Figure 11. THREDDS PICat elements within NSDL architecture

visualization and analysis clients will be extended to connect to the PICats, PICat servers, and the data servers themselves.

The PICat server is a digital library service that enables the discovery of scientific data-set inventories, and allows searching by category, standard quantity, and possibly space and time regions. An online digital library can include a user interface (or portal) to the THREDDS PICat Server, allowing users to search, discover, and download scientific data sets from a browser.

10.5. Cyberinfrastructure

The term "cyberinfrastructure" [CYBERI] was recently coined by the US NSF management to connote not only advanced scientific computing but also a more comprehensive infrastructure for research and education based upon distributed but federated networks of computers, information resources, online instruments, and human interfaces. It provides a convenient way to talk about infrastructure based on information technology (IT), in contrast to more traditional science infrastructure.

In combination, IT capabilities will in combination allow access to complex services as well as raw computing resources through the network, enabling both collaboration and sharing over distance and time.

This picture is consistent with the vision of the Grid,[8] the modern Internet, distributed computing, and collaboratories. It is the basis for what some are calling *e-science*. A schematic of such cyberinfrastructure services is shown in Figure 12.

An environment is envisioned in which raw data and recent results are easily shared, not just within a research group or institution but also between scientific disciplines and locations. There is an exciting opportunity to share insights, software, and knowledge, and thus to reduce wasteful re-creation and repetition. Key applications and software that are used to analyse and simulate phenomena in one field can be utilized broadly in other fields. However, this will only take place if the standards and underlying technical infrastructure are accessible by all.

[8] More recently the term Grid has evolved to mean a more comprehensive structure linking people, information, and tools/facilities as indicated in Figure 12. The term xGrid, where x can be a discipline or a place, is now coming into use, e.g. the BioGrid or the MGrid (Michigan Grid). In this sense the terms "Grid" and collaboratory represent similar ideas.

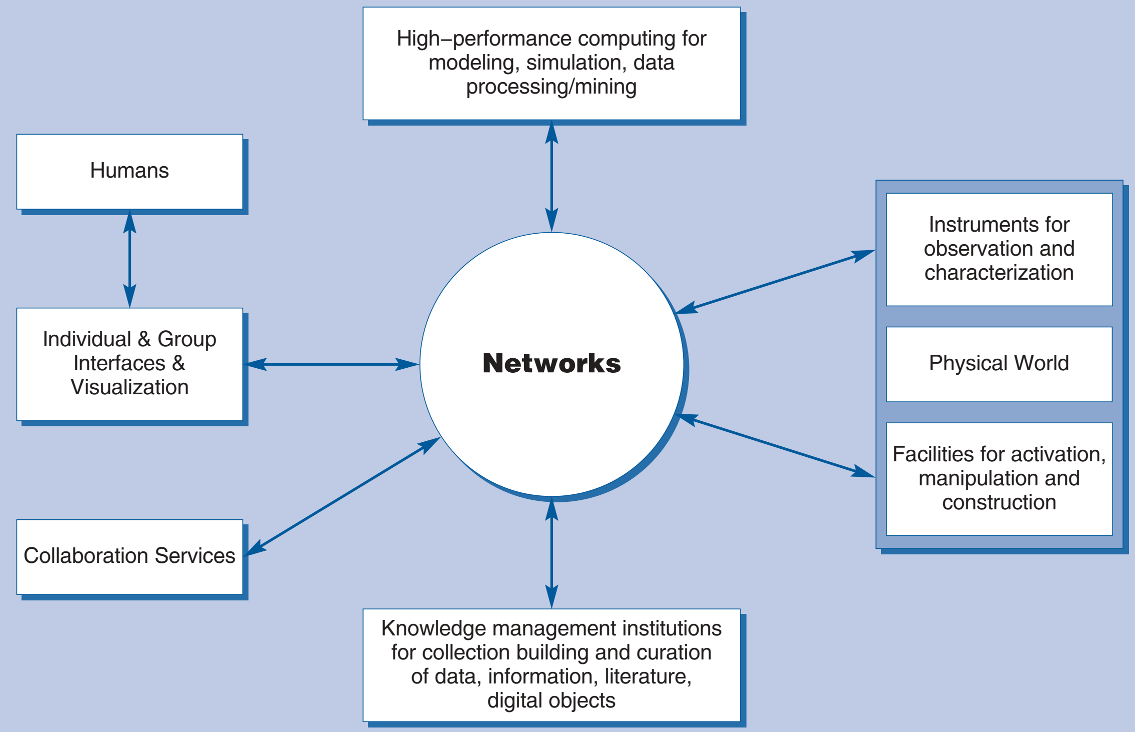

Figure 12. Schematic of services for cyberinfrastructure to support research

The Grid, built around the Internet and Web, is an infrastructure designed to provide scalable, secure, high-performance mechanisms for discovering and negotiating access to remote resources. Ultimately, it should allow scientific collaborators to share resources on an unprecedented scale, and allow geographically distributed groups to work together very effectively.

The new infrastructure is to be applied to educate the next generations of scientists using the best techniques, and to ensure broader participation without respect to field boundaries, institutional wealth, personal origin or bodily ability. It will maximize international collaboration and resource sharing through standardization and networking.

There are also significant risks and costs:
- Researchers in different fields and at different sites will adopt different formats and representations of key information, which will make it forever difficult or impossible to combine or reconcile their information.
- If no decision is made to curate and store raw and intermediate research results indefinitely along with polished and final publications, irreplaceable data will be lost.
- Effective use of cyberinfrastructure can break down artificial field boundaries, while differing tools and structures can isolate scientific communities for years.

Collaboration among disciplines is growing at an unprecedented pace and now includes, in some cases, hundreds of scientists working on a single project across the globe. Cyberinfrastructure must support this type of collaboration in a reliable, flexible, and cost-effective manner.

A significant need exists in many disciplines for long-term, distributed, and stable repositories for data and metadata that institutionalize public-domain data holdings. These repositories must provide tutorials and documentation on data format, quality control, interchange formatting, and translation, as well as tools for data preparation, fusion, mining and knowledge discovery, and visualization. A key element associated with filling this need is the development of middleware and related data storage strategies.

More and more disciplines are expressing a compelling need for nearly instantaneous access to selected databases (both local and distributed) and related services, particularly because such access often drives the actual collection of the data.

Users are expressing the need for nearly instantaneous access to real time data streams from observing platforms or computations, where such information feeds prediction models and decision support tools used in time-critical decision-making. Users also require an ability to control complex instruments remotely with exceptional network quality of service. For example, the National Virtual Observatory (http://www.us-vo.org/nvo-proj.html) links astronomy with cyberinfrastructure in the forms of Grid computing and federated access to massive data collections. A similar Grid-based approach enables neuroscientists to remotely control high-energy electron microscopes that are located a continent away: cf. the Telescience project at http://www.npaci.edu/envision/v16.2/telescience.html.

10.6. Data grids

One of the aims of the Grid is to promote the open publication of scientific data. If this is realized then it is expected that many of the Grid advances will come from applications that can combine different information about a single entity to gain a more complete picture of that entity, and to aggregate similar information about different entities. Achieving this will require support for integrating data from multiple data sources.

A good example of a data grid involves the Storage Resource Broker (SRB) [SRBDLE] that has been developed at the San Diego Supercomputer Center (SDSC) to provide "uniform access to distributed storage" across a range of storage devices, via a well-defined API. The SRB supports file replication, and this can occur either off-line or on-the-fly. Interaction with the SRB is via one of several APIs: the end-user can use a client-side GUI[9] or a standard Web browser.[10] SRB servers are managed by an administrator, with authority to create user groups. A key feature of the SRB is that it supports metadata associated with a distributed file system, such as location, size, and creation date information. It also supports the notion of application-level (or domain-dependent) metadata, specific to the content and not generalizable across all data sets. In contrast with traditional network file systems, SRB is attractive for Grid applications in that it can handle very large volumes of data, which can transcend individual storage devices, because of its metadata awareness and capabilities, and many other features, such as file replication, authentication, etc.

Features of SRB include:

- A metadata server (MCAT) that holds information on the data, users, and resources managed by the SRB: it can also be used to hold application-specific metadata. However, a limitation is that there appears to be no general mechanism for connecting MCATs into a hierarchy (for example, to allow the scalable federation of servers).
- A logical naming scheme for data sets: the mapping from logical name to the physical file is done automatically when a data set is accessed.
- A federation facility that allows a set of SRB servers to offer a single interface to clients.
- Authentication can be through the Grid security infrastructure.

As well as files, the SRB can manage data stored in relational databases. SRB looks to be a possible candidate to support THREDDS functionality [THREDDS].

10.7. Curriculum–based interfaces for DLEs

10.7.1. Project 2061: Atlas of science literacy

Project 2061 is a long-term effort of the American Association for the Advancement of Science (AAAS) to reform primary and secondary education so that all citizens attain science literacy — that is a basic understanding of the natural and social sciences, mathematics, technology, and their interactions [ATLAS]. The project is creating a coordinated set of tools and services — books, CD-ROMs, online resources, and professional development workshops — that educators, parents and families, and community leaders can use to make meaningful and lasting improvements in teaching and learning for all students. Thus the project supports education standards.

Science literacy should be approached not as a collection of isolated abilities and bits of information, but as a rich fabric of mutually supporting ideas and skills that must develop over time. From primary school to high school, what students learn should build on what they learned before, make sense in terms of what else they are learning, and prepare them for what they will learn next. To help students achieve science literacy, educators need to see how the ideas and skills that students learn in different grades and topics — and even disciplines — depend on and support one another. Atlas

[9] http://www.npaci.edu/dice/srb/inQ/inQ.html

[10] http://www.npaci.edu/dice/srb/mySRB/mySRB.html

[ATLAS] depicts this pattern of connections in a set of "strand maps" that provide a graphic representation of students' growth of understanding. Each map displays the ideas, skills, and connections among them that are part of achieving literacy on a particular topic, showing where each step along the way comes from and where it leads. "Strand" denotes identifiable concepts or stories that are developed in groups of benchmarks across different grade levels.

The maps in Atlas do not prescribe a particular curriculum or instructional strategy. Instead, they present a framework meant to inspire a variety of different ways to design and organize learning experiences suited to local circumstances. The maps in Atlas are built from benchmarks – the learning goals presented in the AAAS's Benchmarks for Science Literacy.

The horizontal lines in each map (see Figure 13 for an example) delineate the US grade ranges (where K-12 covers US primary and secondary education levels) in which most students should be able to achieve particular benchmarks. In each map, a few principal strands are pointed out to help the reader find things in the map and get a sense of its content. These strands are labeled along the bottom of the map (Figure 13). A line between boxes implies that understanding one benchmark contributes to understanding the other.

This notation is emphasized in Atlas to show that the structuring and definition of any discipline can be standardized. Each benchmark may serve as a template denoting resources in digital collections relevant to the benchmark's educational content.

10.7.2. Science literacy benchmarks as possible navigation and search interface for NSDL

A formal coding of the benchmarks as semi-structured conceptual graphs (SCGs) is proposed in [CURINT] to be used in NSDL. Each registered collection and digital object within a collection is mapped back to such SCGs. The system admits a set of mapping predicates, such as *example_of, discovery_of, description_of, explanation_of, anecdote_about*, etc. When a library holding $H is placed in the collection, the author may fill in a set of predicates, e.g. *example_of* ($H,

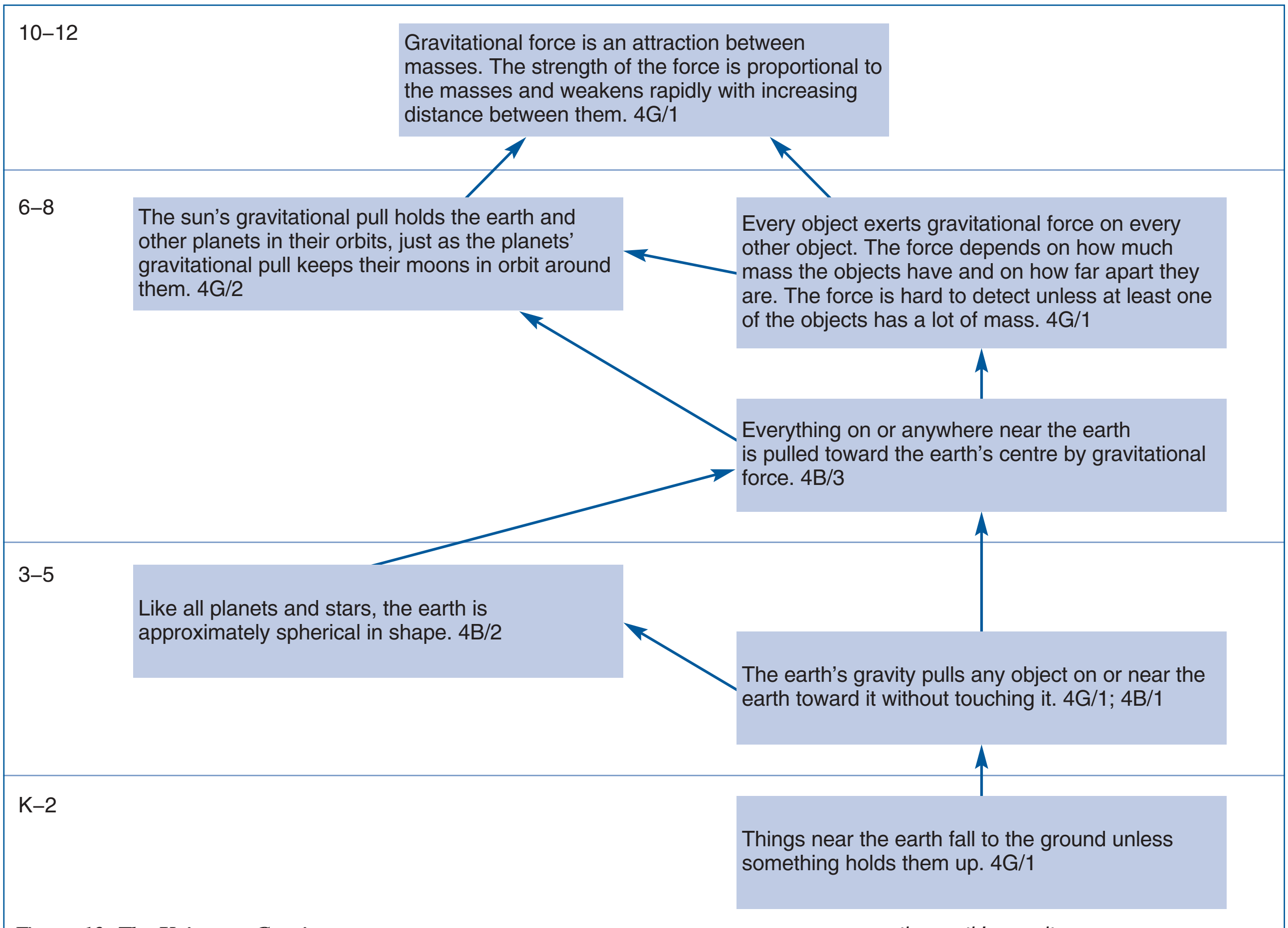

Figure 13. The Universe: Gravity map

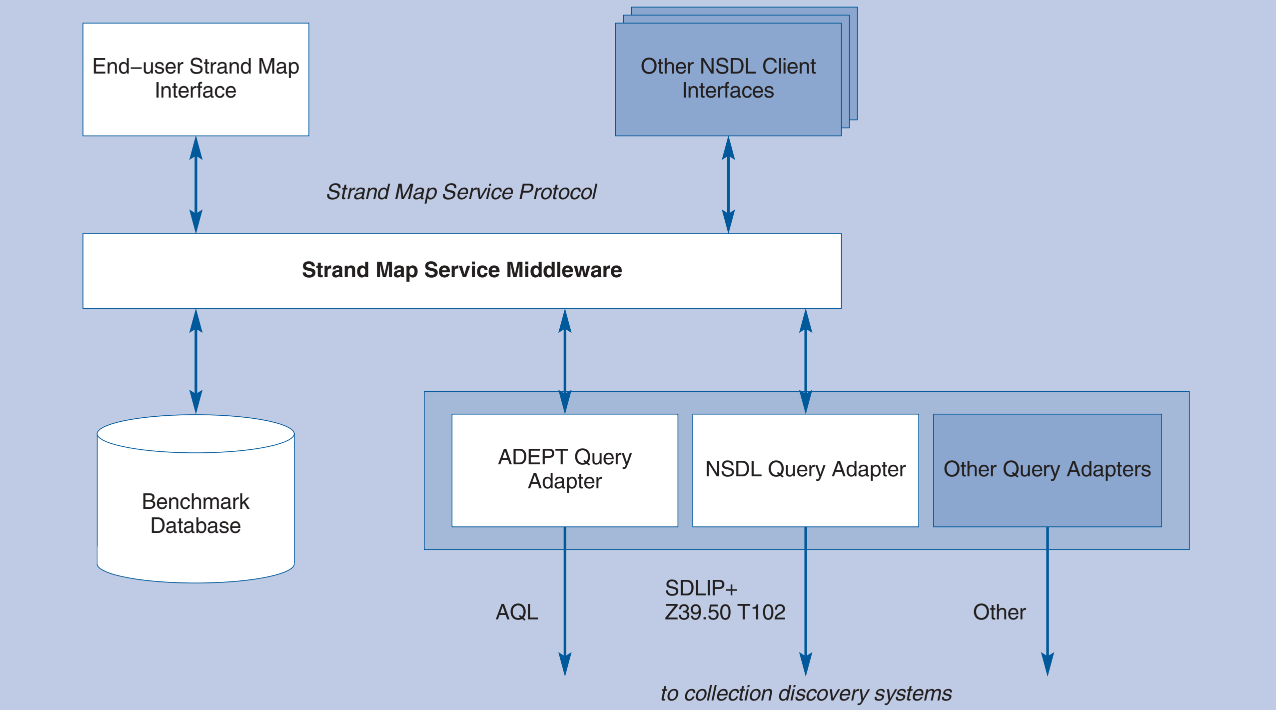

Figure 14. Architecture of the strand map service

atom), *description_of* ($H, hydrogen_atom) where the first term is a reference to the holding and the second term denotes a concept that the digital object "is about". When such SCGs are applied, it becomes possible to navigate and search in a digital library based on concept and discipline-oriented structuring of information. Note the difference between such "curriculum"-oriented structuring of a library and the conventional bibliography-oriented structuring.

Advanced applications based on SCGs are indicated in [CURINT]. Suppose an educator wants to evaluate a textbook, represented as a linear graph of chapters, where each unit is a graph of concepts. The system employs graph-matching techniques to determine whether the textbook contains certain concepts by first searching for the units touched by each chapter, and then by computing the difference between the concept set within the unit and the sets covered by the chapter. Similarly, the system can also determine whether the order of concept coverage in the textbook preserves the same dependency structure as the SCGs.

Another project Strand Maps as an Interactive Interface to NSDL Resources [STMINT], is also oriented toward discovery of educational resources that support the learning goals, or benchmarks, articulated in the strand maps. During the two-year performance period, investigators from the University of Colorado, the American Association for the Advancement of Science, the Digital Library for Earth System Education Program Center, and the University of California at Santa Barbara will work together to develop the service middleware and protocol that enable the strand map service to be integrated into library discovery systems and other NSDL services. The project includes development of other related services. One goal of this project is to provide educators with a tool to locate resources that are aligned with benchmarks through well-designed interactive interfaces and educationally tailored query mechanisms. The strand map service will be analogous to the UCSB place name gazetteer. The architecture of the service is shown in Figure 14. The basic function of the middleware (along with the other components) would be to find resources (of type X) "aligned" to a particular benchmark.

10.8. Concept–based organization of learning materials and courses

10.8.1. Concept space specifying a domain of interest

Issues of proper definition and structuring of learning materials in learning environments are analysed in [ADECOA]. The idea of access to digital collections by content is based on the premise that "scientific concepts" and "relationships between concepts" provide a powerful level of granularity with which to support effective access and use for learning.

The sets of concepts and interrelationships developed over many centuries by scientists and mathematicians constitute nothing less than the fundamental building blocks from which useful representations of reality are created, applied, evaluated, and modified.

The structure of the educational experience should reflect the structure of the concept space specifying the domain of interest. Understanding the knowledge in some domain of science requires students to understand how sets of concepts and their interrelationships are developed and applied in representing the phenomena of the domain. It looks both feasible and valuable to organize, access, and use scientific knowledge explicitly in terms of the sets of concepts that underlie some domain of scientific knowledge rather than in terms of the information objects contained in the domain.

Within the framework of existing digital library technology, it is proposed to develop:

- a concept model for representing concepts and their interrelationships;
- domain-specific knowledge bases of such representations;
- associated DL collections of "illustrative materials" concerning different aspects and attributes of the concepts;
- services supporting the creation, modification, viewing, and use of concepts for various purposes in learning contexts.

It is assumed that such an explicitly concept-based support for learning leads students to a deeper understanding of: (1) the nature, structure, and classes of the concepts that, together with the interrelationships between the concepts, provide a basis for scientific development in specific domains of knowledge; (2) the scientific roles of various classes of concepts across the spectrum of scientific activities; and (3) the global structure of some domain of scientific knowledge in terms of the underlying framework of concepts. Advantages that may accrue to instructors include the efficient reuse and repurposing of the knowledge bases of concepts and the associated collections and services in creating instructional support materials.

There is a growing consensus that science education should be a meaningful learning activity in which students learn to think like scientists rather than solely remembering information. In [ADECOA] concepts are considered to belong to one of three classes:

- *Abstract concepts:* have operational semantics defined in terms of syntactic (or computational) manipulations of symbolic representations. Three possible subclasses are syntactic (linguistic) concepts, logical concepts, and mathematical concepts.
- *Methodological concepts:* have a semantics defined in terms of the various classes of scientifically well-defined operations that may be carried out in relation to them. Possible subclasses of this class include concepts relating to procedures for scientific observation, scientific information seeking, scientific data gathering and interpretation, scientific hypothesis construction, theory construction, and scientific communication.
- *Concrete concepts:* have a semantics defined in terms of scientifically well-defined operations that provide the concept with an interpretation of defined phenomena in the domain of scientific knowledge. The concept of river discharge, for example, has a characterizing set of operations that defines the concept in terms of various sets of measurement procedures that may be carried out in real-world contexts. An example of such a procedure is one to determine the amount of water passing across some cross section of a river during a given interval of time. These concepts are, by and large, the class of concepts used in model and theory construction.

10.8.2. ADEPT concept–based architecture

The Alexandria Digital Earth Prototype (ADEPT) is being extended with a knowledge representation system that includes a knowledge base of concepts, a collection of associated materials, and a set of services for use in supporting an introductory course in physical geography at UCSB in fall 2002. The ADEPT concept-based architecture [ADEVLE] is shown in Figure 15.

The instantiation of this architecture involves:

- *Constructing a knowledge base of concepts:* Given the existence of a set of concepts for some domain, a knowledge base of these concepts is constructed by taking the abstract concept model as the basis for an XML schema from this model, and creating an XML record for each concept.
- *Constructing a collection of information* illustrating specific facets of the concepts: creating a DL collection of information illustrating the facets of some concept (such as representations of the measurement operations for some measurable concept) involves cataloguing the item as part of an ADEPT collection. This involves using the ADN metadata standard for learning objects, with an extension that involves a field for the concept name(s) and facet(s).

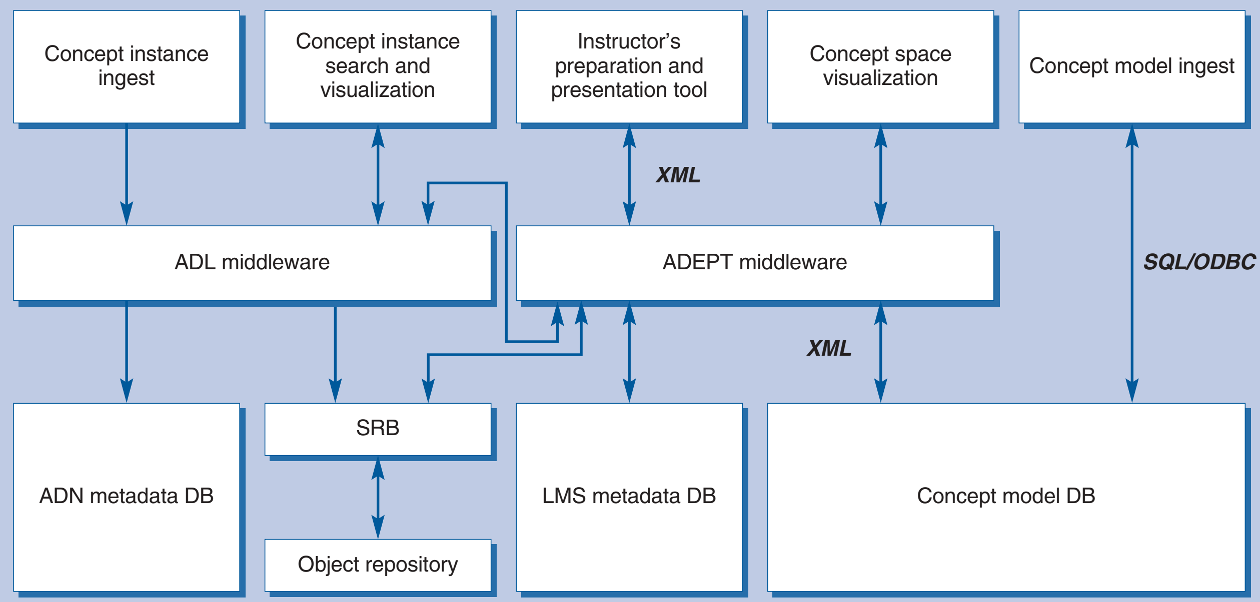

Figure 15. The ADEPT concept-based architecture

- *Creating services* that provide access to the knowledge bases and collections: Three sets of services over the concept knowledge bases and associated collections are planned: services for manipulation of entries in the concept knowledge base and the collection of illustrative items; services for searching the concept knowledge base and the collection of illustrative items; services providing both textual and visual views of the concept knowledge base and collections.
- *Developing a client that provides access to the knowledge base, collections, and services* from the perspective of the general user: This client includes functionality supporting the following operations on both the knowledge base and the collection: (1) creation and modification; (2) access; and (3) the creation of views, and the use of the material for various learning purposes.

10.8.3. Physical geography course in a concept–based architecture

The course on physical geography being developed may be viewed as a set of subtrajectories relating to each of the major topics covered in the course and corresponding to the chapter organization of a typical textbook. All of the highest-priority concepts are being modeled with the use of an XML schema. The collections of associated materials that illustrate the various facets of each concept model are being created as regular ADEPT (multimedia) collections, accessible by the operational search middleware services of ADEPT. The metadata created for each item follows the ADN metadata content standard for instructional materials augmented with an additional field for the relevant concept(s) and their facets.

The services associated with the knowledge base and collection include services for creating a trajectory through concept space that may, for example, form the backbone of a session constructed by the instructor in relation to the following high-level lecture plan:

- The instructor motivates the session with collection items containing (typically) implicit representations of the phenomena encompassed by the concepts of hydraulic geometry. These may include, for example, aerial photographs of rivers taken over one significant portion of their length and representing (implicitly) aspects of the concept of downstream hydraulic geometry. This concept refers to the changes in the discharge, width, average depth, velocity, etc., of a river as it flows downstream. Items may also include videos taken at a cross-section of a river that show the changes in the discharge, width, average depth, velocity, etc., of the river over time in response to a rainfall event.
- The instructor begins a discussion of the concepts that might be useful in constructing empirical and theoretical representations of hydraulic geometry. All such concepts may be found by constructing from the knowledge base a view of all the concepts that relate to the topic of hydraulic geometry, including the concepts of width (w), average depth (d), average velocity (v), and total discharge (Q) at a cross-section of a river channel. The students might be

led to develop specific models of concepts not previously specified in any detail. Using items from the collection, for example, the instructor may lead the students to a representation of a river as a sequence of cross-sections. The instructor may also encourage development of explicit representations for cross-sections in terms of the product of width and average depth.

- The instructor may then interact with the students in establishing hypothetical relationships between concepts (i.e. hypotheses) that may be considered not only to be important characterizations of hydraulic geometry but, if verified, to provide (partial or full) objective representations of the concept of hydraulic geometry. The instructor, may, for example, provide the students with data sets from the collection that can be used for constructing and testing hypotheses about the mathematical relationships of concepts.

A prototype interface is being implemented as a three-screen (three-window) system, including the lecture window containing a trajectory through the concept space, the knowledge window reflecting the concept base, and the collection window of selected resources that illustrate presented concepts.

11. Summary

1. Digital libraries have become a core ingredient, a collective memory of the educational environments of today and of the future. Hybrid libraries have already become widely used components of many universities around the world.

2. In several countries (USA, UK, Germany) a national digital library for education in science, engineering, and technology is being developed as an important ingredient of the national educational infrastructure. While the development of a DLE is a continuous process of collecting, classifying, conceptualizing, and using information, the process is paralleled by rapid technological advancements. Together, these two developments lead to the evolution of DLE frameworks and methodologies for their application.

DLEs in each country are dependent on the language used by it's educational community, the national culture, and the national traditions in education. This should be taken into account for the globalization of DLEs.

3. The development of national DLEs requires the involvement of various groups in society. Besides educators and learners, the community includes members of professional societies, information providers, researchers, and representatives of industries. Interrelationship of the interests of the community members should be addressed, in order to develop a reasonable strategy leading to DLE sustainable development and gradual evolution. This process leads to formation of a wide community around the DLE, providing for development, governance, collecting of information, and use in education. Large DLEs should function as nonprofit organizations. Experience from NSDL as a national project is expected to show how such challenges can be met.

Building sustainable and scalable communities where education-oriented best practices and information are shared (national and global communities, domain- or discourse-oriented communities) requires specific attention and organizational measures during DLE development and evolution.

If DLEs are to succeed, there must be serious social change, at least, in two important aspects: community building and development of a sharing culture. In particular, such change should result in getting educators working with researchers and with professional societies, educators who are willing to share education resources as well as their time (through such activities as development of reusable instruction modules, mentoring, reviewing, and advising). Educators must feel empowered to reuse resources, with community support, if required. New schemes for rewards and policies for promotion and tenure must be developed. New outlets, like the ACM *Journal of Educational Resources in Computing*, which is one of various related efforts to motivate people to share high-quality resources, must be created.

4. Organizational frameworks and interdependencies of components of the DLE infrastructure are quite complicated. These interdependencies impose serious challenges on the decisions taken during DL development. For example, users of existing specialized course modules may rely on software products developed by industries or military organizations. If such courses are to be included in the national DLE, the issues of availability of such products for the academic environment must be resolved. MeteoForum (see sections 10.3, 10.4) is an example of an organizational structure required to solve such problems.

A reasonable compromise must be found between the approach of selecting only high-quality education-oriented objects for DLEs (as in DLESE) and the approach of collecting a broad class of materials including millions of objects (the NSDL strategy).

5. DLEs in science include not only textual and multimedia information, but also real time data — the results of measurements provided by specialized organizations on the national level or globally. DLEs also provide access to expensive scientific instruments (e.g. electronic microscopes, telescopes) and to specific services. Such possibilities create efficient conditions to involve learners in research in the early stages of their education.

6. DLEs should provide various services, such as cataloguing, archiving, selective dissemination of courseware and other instructional materials developed internationally, annotation, evaluation, cross-lingual search and retrieval, personalization, recommendation, instructor support, and copyright management.

7. A gap exists between the technological and informational possibilities of existing DLEs and the availability of systematic reports by teachers and learners about actual experience using them. During the work on this survey, we found some success stories that have been collected at the University of Michigan, USA, showing how "collective memories" can be successfully used in a classroom. Analysis and evaluation of the real practice of DLE use require specific attention from DLE governance bodies and from the respective funding organizations. Systematic evaluations are still required to analyse the effectiveness of the new models of learning environments and new pedagogical methods built on DLEs (e.g. being student-centric, enabling group work on real world problems, customization of educational programmes for different students, interaction of learners with the required subject domain, curriculum-based access to learning objects, course design with reuse of existing educational modules, etc.).

8. Alternative approaches (visions) for DLE frameworks.

Existing DLE architectures, technologies, and methodologies for using them have not yet become mature. DLE frameworks start simply, being based in metadata and organized according to the metaphor of a conventional library. With time and experience, these frameworks are expected to evolve into more knowledge-based systems using conceptual definition of subjects, curricula, and ontologies introduced (in parallel with the metadata layer) for the subject definitions in various areas of teaching (e.g. see sections 10.7, 10.8).

According to the framework based on the metaphor of a conventional library, a DLE as a collective memory can be considered a container extending the conventional library (cataloguing) practice. In this case the granularity of the memory is at the level of "information entities" — electronic versions of books, journal articles, images, and videos. Metadata schemas support retrieval focused on information entities (as in the conventional tradition of library bibliographic cards), not on subject structuring and the respective granularity of retrieved items. Such an approach looks reasonable at least because of the large heritage of traditional information entities and the significant difficulty in getting access to proper information items by content.

On the other hand, analysis of DLEs and specifically of DLEs for different branches of science shows that information in such libraries (acting as collective memories) should be structured differently. Textbooks and courses are not good information entities any longer. "Bibliographic cards" are not suitable for information discovery. Educational domains in different branches of science should be properly structured. More suitable entities would be concept spaces, theories, models, hypotheses, experimental results and measurements, curricula, and educational modules. Scientists have spent centuries to reach well-defined structures, concepts and theories in the various branches of science. These definitions cannot be used following the conventional library metaphor, but are more suitable as a guiding principle for information structuring and search in DLEs (see sections 6.3, 10.7, 10.8).

For this reason, the gradual evolution of DLEs from the current framework based on the metaphor of a conventional library to more knowledge-based organization is expected. With time and experience, these frameworks will be upgraded with conceptual definitions (ontologies) of subject domains and curricula along with the conventional metadata so that information resources can be registered in accordance with the proper subject definition and granularity. This trend will also lead to a higher level of coherency of the information collected in a specific subject domain, by contrast with metadata use, where collected materials are more diverse though less relevant to the subject.

There is a discussion that DLEs need to be constructed to facilitate access by both humans and computer, i.e. must be semantic and support cross-resource and cross-domain searching and harvesting. To allow for best dissemination practices, the reusability of resources, and harvesting at semantic levels beyond OAI, current research, such as the work related to the semantic Web (e.g. DAML/OIL and OWL), should not be overlooked.

Several directions of research and development surveyed in this report indicate that such evolution has already started.

9. Current DL technologies constitute a combination of well known database, information retrieval, and web techniques. Some promoters of DL technologies often prefer not to notice that they really reuse well known techniques for DL purposes. Just a few examples:

- Metadata registries are often just databases with schemas defining structures of the respective metadata.
- OAI is similar to some well known data-warehousing techniques.
- "Subject mediator" technologies aim at capturing community agreements on data structures, thesauri, and ontologies for specific subject domains in a meta-information database. Such definitions are in a sense analogous to the registries of metadata elements, though they are much more advanced. Subject mediators are designed to

convert and reconcile a diversity of structures, terminology, and ontology of multiple data sources relevant to the subject domain with those of the mediator, as well as to provide a uniform query interface to these data sources registered at the mediator.

According to this approach, the subject domain model is to be defined by the experts in the field independently of relevant information sources. This model may include specifications of data structures, terminologies (thesauri), concepts (ontologies), methods applicable to data, and processes (workflows) characteristic of the domain. These definitions constitute the specification of a subject mediator. The process of developing a specification for a subject domain in the respective community is called mediator consolidation. After the definitions are consolidated, an operational phase of the mediator starts. During the operational phase, information providers can disseminate their information for integration into the subject domain independently of each other and at any time. To disseminate, they should register their information sources at the mediator. The registration assumes contextualization of sources in the mediator (reconciling of terminological and ontological differences) and defining source structures and behaviors in terms of the mediator schema.

If we think about what DC is or what application profiles are or what Z39.50 profiles are or what the Alexandria metadata schema is (many examples of various metadata schemas can be given), we easily see that they specify "schemas" of sets of attributes characterizing information resources in respective domains. These schemas were the result of consolidation in special working groups, conferences, and meetings held in the respective communities. In the process of registration of a source at the mediator, the differences in mediator and source terminology, ontology, and structuring are to be reconciled. Converting, harmonizing, and maintaining element sets of different metadata registries constitute a simplified instance of this process.

This observation is important: mediators are a more general technique than metadata registries. Mediated schemas are more general than schemas of metadata registries. Mediator technology can support a range of new DL technologies that are emerging (see above); metadata registry technologies cannot.

- The search bucket mechanism [ADEPTA] is just a specialized mediator designed for a concrete schema defining "collection-neutral search buckets". More can be reached in the framework of a Local-as-View (LAV) mediator technique. In LAV a mediator's schema is a schema of a subject domain. The subject domain schema is agreed to by the domain community independently of the existing collections. Each relevant collection can be registered at the LAV mediator as a materialized view. Query rewriting algorithms exist to transform mediator queries into query plans above the registered collections. Some benefits a general mediator architecture can bring, when compared to buckets, are:
  a) More flexibility in developing geo-referenced mediators (instead of rigid bucket schema, any geo-referencing mediator can be defined and implemented).
  b) The possibility of introducing data mediators (not only a search mediator) over various geo-referenced collections.
  c) Buckets in their current form probably should be treated as personalization over some mediators.
  d) General collection registration tools can be reused. An object-oriented data model can be introduced into the ADEPT architecture. Query rewriting algorithms of the mediator are more powerful and sound than those of the current bucket search facilities. Extensions for semantic mediation are playing an increasingly important role in federations of scientific data sets, and it can be expected that they might do so in DLs as well. Essentially, semantic data mediation can be seen as an extension of database mediators and query rewriting, supported by ontologies and other knowledge representation techniques.

The observations considered in item 9 lead to a conclusion that tighter collaboration with specialists in knowledge systems and mediation systems is required for developing DLEs and related technologies.

10. The subject of Digital Libraries for Education is too broad to cover exhaustively in one survey. Through consideration of several carefully chosen projects, this survey report attempts to concentrate on advanced topics, to analyse the current state of the technology, and to foresee the probable directions of its forthcoming evolution. The report could not consider every educational discipline, mostly concentrating on education in the natural sciences and engineering. Thus the specificity of DLEs for many other disciplines needs to be investigated further. Geographically, the report is based on information produced mostly in the USA and Europe. For completeness, the DLE development programmes in the rest of the world need additional serious analysis. Even collecting information about the state-of-the-art in different countries is difficult, because of the insufficient level of information available and the diversity of the presentation languages.

Several important issues were not sufficiently analysed in this survey and will require separate discussion:

- Sustainability and economic issues are crucial for DLE development. What are the visions here? What are the approaches? What may be funded by governments, and what by private companies or organizations? What will result from volunteer efforts undertaken in connection with educational institutions, or in connection with professional societies? How can the numeric dominance of students be leveraged to have their efforts lead to benefits (through sharing of theses, dissertations, technical reports, portfolios, etc.)?
- Preserving the national language as well as a cultural and historical identity in the education and globalization of DLEs. For example, is a move toward an international DLE reasonable? Or is it better to have national DLEs that somehow interoperate? If so, who would help guide and ensure interoperability? The argument that each nation needs a DLE is not well justified. Can every nation afford one? Why not have regional DLEs? Why not have DLEs for language groups? Are differences in language really so important that we must have different DLEs when content may be in different languages? What about countries like India where there are many languages? What of the fact that mathematics education is less tied to natural language because the language of mathematics is used?

11. Digital libraries are becoming a core ingredient, a collective memory, of the educational environments (global, national, university or domain-oriented) of today and of the future. A main conclusion of the survey is that to provide a competitive education, different countries should establish their own DLEs (e.g. as a national DLE, collaboratively with other DLEs, or as a regional DLE). They cannot passively wait until suitable global digital educational content is formed. The digital content of DLEs remains dependent on the language (or language groups) used by the educational community in each country, as well as the culture and national traditions in education. A significant amount of time is required to form the national community around DLEs, collect the DLE content, and educate specialists to develop, maintain, and govern DLE.

On the other hand, a DLE is distinguished from other ICTs applicable to education (e.g. multimedia, distance learning) by several important features:

- To establish a DLE (after technology is installed as software components), serious efforts are required to collect (harvest, integrate, gather, register) the digital resources, and to maintain and continuously extend them. If the digital content is not completely borrowed from another DLE, this process requires specific organizational efforts and investments. It is not a task that can be done by a separate individual (the way an educator can individually establish and use multimedia technology preparing courses). Governance, maintenance, and a community must be arranged around the DLE to make it sustainable.
- To make a DLE useful, additional efforts are required to provide for preserving the proper quality of the digital content. This is also not an individual effort. Various organizations in society must be involved in the process of creating digital content of the required quality.
- To make information in digital form widely available requires supporting rights of access and use, including copyright, preservation of the integrity of the document, licensing, and payment for use.
- In DLEs with digital content a wide set of interrelated services require administration and development.

DLEs are complex technologies that will require specific considerations and planning to provide for their deployment (compared to other, more compact technologies).

Acknowledgements

This analytical survey is devoted to recent research and development in the area of Digital Libraries in Education. Because this is very active area of research, it was not possible to reflect all the related works in one survey. The cited works are listed in the References. We would like to thank the community of authors of these publications and apologize to them in the likely event that we have misinterpreted some of their ideas.

The authors express their deepest gratitude to the experts in the field who read the survey draft and provided detailed remarks and suggestions that helped to improve the survey: Terry Smith (UC Santa Barbara, USA), Lillian N. Cassel (Department of Computing Sciences, Villanova University, USA), Bernd Wegner (TU Berlin, Germany), David Casal (University of East Anglia, UK), and Efim Kudashev (SRI RAS, Russian Federation).

Specific gratitude the analytical survey Working Group expresses to Mary Marlino for the arrangement of the technical editing of the text of the analytical survey in Boulder. We appreciate the contributions of the National Science Digital Library Core Integration Office, the Digital Library for Earth System Education, and especially the services of Marie Boyko of the University of Colorado in the preparation of this final manuscript. The editing contributed a lot to the quality of the survey text.

References

[ACRA98] President's Information Technology Advisory Committee Interim Report to the President, August 1998, National Coordination Office for Computing, Information, and Communications 4201 Wilson Boulevard, Suite 690, Arlington, VA 22230.

[ADECOA] Structured models of scientific concepts as a basis for organizing, accessing, and using learning materials. Terence R. Smith, Marcia L. Zeng, ADEPT Knowledge Organization Team. JCDL, 2002.

[ADEDLE] ADEPT/DLESE/NASA. ADN joint metadata content model. http://www.dlese.org/Metadata/, 2001.

[ADEPRO] The Alexandria Digital Earth Modeling System (ADEPT). NSF Proposal. http://www.alexandria.ucsb.edu/adept/proposal.pdf, 1999.

[ADEPTA] Greg Janée, James Frew. The ADEPT Digital Library Architecture. JCDL, 2002.

[ADEPTS] T.R.Smith, G.Janée, J.Frew, and A.Coleman. The Alexandria digital earth prototype system. In Proceedings of the Joint Conference on Digital Libraries. June 2001.

[ADEVLE] Michael Freeston. The ADEPT Virtual Learning Environment. Talk at the SDSC Seminar on DLE: *Bringing Scientific Databases to the Classroom*. http://kbi.sdsc.edu/events/DLE-06-02

[ALEXBI] Alexandria digital library: Publications, research papers, current bibliography. http://www.alexandria.ucsb.edu

[ATLAS] Atlas of Science Literacy. Project 2061. AAAS. Washington, DC, 2001.

[CASSEL] L. Cassel and E. Fox. Introducing the ACM Journal of Educational Resources in Computing (JERIC), editor-in-chiefs' introduction. 1(1), March 2001, http://doi.acm.org/10.1145/376697.382399

[CBISDT] Dmitry O. Briukhov, Leonid A. Kalinichenko. Component-based information systems development tool supporting the SYNTHESIS design method. Proceedings of the conference on Advances in Databases and Information Systems (ADBIS), LNCS, Springer Verlag, 1998.

[CLOIDP] David A. Wiley. Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. The Instructional Use of Learning Objects, online book, http://www.reusability.org/read/

[COMET] Cooperative Program for Operational Meteorology, Education and Training, http://www.comet.ucar.edu/

[COMIDP] COMET Instructional Development Process Outline, http://www.comet.ucar.edu/

[COMPLO] Jean R. Laleuf, Anne Morgan Spalter. A Component Repository for Learning Objects. A Progress Report JCDL '01, June 24-28, 2001.

[CONV93] Laurillard, D. Rethinking University Teaching - a framework for the effective use of educational technology, London, Routledge, 1993.

[COVB99] Borgman C.I. What are digital libraries? Competing visions. Information processing & management, 35(3), 1999, 227 – 243.

[CRIM] Edward A. Fox, Rachelle S. Heller, Anna Long, and David Watkins. *CRIM: Curricular Resources in Interactive Multimedia*, in Proceedings ACM Multimedia '99, Orlando, FL, Nov. 1-5, 1999.

[CURINT] Ontology Services for Curriculum Development in NSDL. Amarnath Gupta, Bertram Ludäscher, Reagan Moore. Talk at the SDSC Seminar on DLE: *Bringing Scientific Databases to the Classroom* http://kbi.sdsc.edu/events/DLE-06-02, based on short paper at JCDL'2002 with the same title.

[CYBERI] Revolutionizing Science and Engineering through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure. January 2003, Daniel E. Atkins, Chair, http://www.cise.nsf.gov/evnt/reports/atkins annc 020303.htm

[DCQUAL] Dublin Core Qualifiers, 2000, http://purl.org/DC/documents/rec/dcmes-qualifiers-20000711.htm

[DCEDEX] Dublin Core Education Working Group, http://purl.org/DC/groups/education.htm

[DELI97] Matija Maroli, Marko Privosnik. A distance learning application based on a digital library of courseware 1997 IFIP, Chapman and Hall.

[DELO00] Digital Libraries for Distance Education (Editorial). D-Lib Magazine, October 2000, Volume 6, Number 10.

[DLCO00] Tefko Saracevic. Digital library evaluation: Towards an evolution of concepts. Library Trends, vol. 49, N 2, Fall 2000.

[DLESEC] Mary Marlino, Tamara Sumner. Research Issues for Educational Community-based Digital Libraries: The Digital Library for Earth System Education. Paper presented at the Annual Meeting of the American Educational Research Association New Orleans, LA, USA, April 1-5, 2002.

[DLESME] Michael Wright, Mary Marlino, Tamara Sumner. Meta-Design of a Community Digital Library. *D-Lib Magazine*, May 2002, Volume 8, Number 5.

[DLESNSF] The DLESE Program Center: Providing Infrastructure for a Distributed Community Library, NSF Proposal, http://www.dlese.org/documents/proposals/DPC_Infrastructure_dist2.pdf, 2002.

[DLFS99] Digital Library Policies, Organizations, and Practices. Digital Library Federation Survey, 1999.

[DLREUSE] Tamara Sumner and Melissa Dawe. Looking at Digital Library Usability from a Reuse Perspective. JCDL '01, June 24-28, 2001.

[DLTO01] Hans Roes. Digital Libraries and Education. Trends and Opportunities. *D-Lib Magazine*, July/August 2001, Volume 7, Number 7/8.

[DNERLP] DNER Learning and Teaching Programme, http://www.jisc.ac.uk/dner/development/programmes/learnteach.html

[DNERAV] Adding Value to the UK's Learning, Teaching and Research Resources: the Distributed National Electronic Resource (DNER), www.jisc.ac.uk/pub/index.html#dner

[EDLL00] Christine Borgmann, et al. Evaluating digital libraries for teaching and learning in undergraduate education: a case study of the Alexandria Digital Earth prototype. Library Trends, vol. 49, N 2, Fall 2000.

[EGEODL] G. H. Leazer, A.J. Gilliland-Swetland, and C.L. Borgman. Evaluating the Use of a Geographic Digital Library in Undergraduate Classrooms: ADEPT. ACM Conference on Digital Libraries, 2000.

[GEOFRA] JISC/NSF Proposal: Digital Libraries in Support of Innovative Approaches to Teaching and Learning in Geography. JISC/NSF Digital Libraries in the Classroom, JISC 07/01.

[GONCAL] Marcos André Gonçalves, Ming Luo, Rao Shen, Mir Farooq Ali, and Edward A. Fox. An XML Log Standard and Tool for Digital Library Logging Analysis. In Research and Advanced Technology for Digital Libraries, 6th European Conference, ECDL 2002, Rome, Italy, September 16-18, 2002, Proceedings, eds. Maristella Agosti and Constantino Thanos, pp. 129-143.

[ICADL02] E. Hoffman and E. Fox, *Building Policy, Building Community: An Example from the US National Science, Technology, Engineering, and Mathematics Education Library (NSDL)*, short paper for ICADL '2002, Singapore, Dec. 2002.

[IMEM01] *Digital Libraries in Education: State-of-the-Art Report.* Information materials prepared for the IITE expert meeting on DLE by Leonid Kalinichenko and Liang Zhang, May 2001.

[IULOOB] The Instructional Use of Learning Objects, online book, http://www.reusability.org/read/

[JISCMLE] Managed Learning Environments (MLEs) in Further Education: Progress Report, JISC Circular 7/00, July 2000. http://www.jisc.ac.uk/pub00/c07_00.html

[KERBER] J. G. Steiner, B. C. Neuman, and J. I. Shiller, *Kerberos: An Authentication Service for Open Network Systems,* presented at Winter 1998 Usenix Conference, 1998.

[LDAP05] W. Yeong, T. Howes, and S. Kille, *Lightweight Directory Access Protocol*, Internet Engineering Task Force, Request for Comments 1777, March 1995, http://www.ietf.org/rfc/rfc1777.txt

[LOMN00] Draft Standard for Learning Object Metadata, IEEE P1484.12/D5.0, 11 November 2000.

[LRBP11] IMS Meta-Data Best Practice and Implementation Guide Final Specification Version 1.1. Available at http://www.imsproject.org

[LRXB11] IMS Learning Resource XML Binding Specification Final Specification Version 1.1. Available at http://www.imsproject.org

[LTCOGN] Jonassen, David H. and Thos. C. Reeves. Learning with Technology: Using Computers as Cognitive Tools. In Handbook of Research for Educational Communications and Technology. New York: Macmillan Library Reference USA, 1996.

[LTSA01] Draft Standard for Learning Technology — Learning Technology Systems Architecture. IEEE P1484.1/D8, 2001-04-06.

[MARIAN1] Marcos André Gonçalves, Robert K. France, and Edward A. Fox, MARIAN: Flexible Interoperability for Federated Digital Libraries. In Proceedings 5th European Conference on Research and Advanced Technology for Digital Libraries, ECDL'2001, September 4-8 2001, Darmstadt, Germany.

[MARIAN2] Marcos André Gonçalves and Edward A. Fox. 5SL – A Language for Declarative Specification and Generation of Digital Libraries. Long paper in Proc. JCDL'2002, Second Joint ACM / IEEE-CS Joint Conference on Digital Libraries, July 14-18, 2002, Portland, pp. 263-272.

[MARIAN3] Marcos André Gonçalves, Paul Mather, Jun Wang, Ye Zhou, Ming Luo, Ryan Richardson, Rao Shen, Liang Xu, and Edward A. Fox, Java MARIAN: From an OPAC to a Modern Digital Library System. In Proceedings of the 9th String Processing and Information Retrieval Symposium (SPIRE 2002).

[MCCKM] E. Fox and L. Kieffer. Multimedia Curricula, Courses and Knowledge Modules, ACM Computing Surveys, Dec. 1995, 27(4): 549-551.

[METAGE] E. D. Liddy and S. Sutton. Breaking the Metadata Generation Bottleneck, 2000, http://cnlp.org/research/project.asp?recid=6.

[METFOR] Arlene G. Laing, Linda Miller, Timothy C. Spangler, Dave Fulker. MeteoForum: A pilot project for advanced international meteorological training centers, 18th International Conference on IIPS, 2002, http://ams.confex.com/ams/annual2002/18IIPS/abstracts/28315.htm

[MLEB02] Jos Boys. Managed Learning Environments, joined up systems and the problems of organizational change. A JISC Report. March 2002.

[MLEJIS] Managed Learning Environments. http://www.jisc.ac.uk/mle/reps/infopack.html

[MLEVLE] MLEs and VLEs explained. JISC, http://www.jisc.ac.uk/mle/

[NDLTD1] E. Fox, M. Gonçalves, G. McMillan, J. Eaton, A. Atkins, and N. Kipp. The Networked Digital Library of Theses and Dissertations: Changes in the University Community. In special issue on *Information Technology and Educational Change* of The Journal of Computing in Higher Education, 13(2): 3-24, Spring 2002.

[NDLTD2] Hussein Suleman, Anthony Atkins, Marcos A. Gonçalves, Robert K. France, and Edward A. Fox, Virginia Tech; Vinod Chachra and Murray Crowder, VTLS, Inc.; and Jeff Young, OCLC. Networked Digital Library of Theses and Dissertations: Bridging the Gaps for Global Access - Part 1: Mission and Progress, Part 2: Services and Research. D-Lib Magazine, 7(9), Sept. 2001, http://www.dlib.org/dlib/september01/suleman/09suleman-pt1.html, http://www.dlib.org/dlib/september01/suleman/09suleman-pt2.html

[NOIOBJ] Andrew S. Gibbons, Jon Nelson, Robert Richards. The Nature and Origin of Instructional Objects. The Instructional Use of Learning Objects, online book, http://www.reusability.org/read/

[NSDINT] William Y. Arms, et al A Spectrum of Interoperability. The Site for Science Prototype for the NSDL. D-Lib Magazine, January 2002, Volume 8, Number 1.

[NSDLAR] David Fulker and Greg Janée. Components of an NSDL Architecture: Technical Scope and Functional Model, JCDL'2002.

[NSDLCO] Carl Lagoze (ed.). Core Services in the Architecture of the National Digital Library for Science Education (NSDL), JCDL, 2002.

[NSDLGOI] Interim NSDL Governance Model, Adopted 12/4/2001, http://doclib.comm.nsdlib.org/ InterimGovern-final-3.doc

[NSDGOS] Dave Fulker. A Short Term NSDL Governance Model. October 1, 2001, PPT file, http://www.smete.org/ nsdl/workgroups/governance/ NSDLgovernance32.ppt

[NSDLME] NSDL Standards Working Group, Metadata Resources Page, 2001, http://128.253.121.110/NSDLMetaWG/ IntroPage.html

[NSDLNSF] Collaborative Project: Core Integration – Leading NSDL toward Long-Term Success. Narrative, References, and Subcontractors' (Budgetary) Statements of Work for a Proposal to NSF, Submitted 6 May 2002.

[NSF996] International Digital Libraries Collaborative Research. Announcement Number NSF 99-6.

[OAI001] C. Lagoze and H. Van de Sompel, The Open Archives Initiative: Building a low-barrier interoperability framework, presented at Joint Conference on Digital Libraries, Roanoke, VA, 2001.

[ORGM81] Beer, S. (1981) The Brain of the Firm, 2nd Edn, Chichester: John Wiley, 1981.

[OSOSS] David A. Wiley, Erin K. Edwards. online self-organizing social systems: the decentralized future of online learning, 2002, http://wiley.ed.usu.edu/docs/ososs.pdf

[PATHWAY] Cathryn A. Manduca, et al. (Eds.) Pathways to Progress: Vision and Plans for Developing the NSDL. March 20, 2001, http://nsdl.comm.nsdlib.org/Comm.Center_Archive/workgroups/whitepaper.pdf

[PDLB97] Lesk M. Practical digital libraries, Books, bytes, and bucks. Morgan Kaufmann, 1997.

[RBLPSS] Michael J. Hannafin, Janette R. Hill, James E. McCarthy. Designing Resource-Based Learning and Performance Support Systems. The Instructional Use of Learning Objects, online book. http://www.reusability.org/read/

[SCCA01] Scholnet and Cyclades: Extending the Role of Digital Libraries. D-Lib Magazine, April 2001, Volume 7, Number 4.

[SDLIP] A. Paepcke, R. Brandriff, G. Janee, R. Larson, B. Ludäscher, S. Melnik, and S. Raghavan, *Search Middleware and the Simple Digital Library Interoperability Protocol*, D-Lib Magazine, 5 (3), 2000.

[SHIB01] Shibboleth, 2001, http://middleware.internet2.edu/shibboleth/

[SINHA00] A. K. Sinha, *Geoinformatics: A defining opportunity for earth science research. A white paper* http://www.geoinformaticsnetwork.org/whitepaper.html, 2000.

[SLRM01] Lee L. Zia. Growing a National Learning Environment and Resources Network for Science, Mathematics, Engineering, and Technology Education: Current Issues and Opportunities for the NSDL Program. D-Lib Magazine, March 2001, Volume 7, Number 3.

[SLWSJ98] REPORT OF THE SMETE LIBRARY WORKSHOP. Held at the National Science Foundation, July 21 to 23, 1998.

[SRBDLE] Arcot Rajasekar. Managing Digital Library Collections with the SDSC Storage Resource Broker. Talk at the SDSC Seminar on DLE: *Bringing Scientific Databases to the Classroom*, http://kbi.sdsc.edu/events/DLE-06-02

[STMINT] Strand Maps as an Interactive Interface to NSDL Resources. NSF Proposal Number DUE 0226286, 2002.

[SULEMA1] Hussein Suleman and Edward A. Fox. A Framework for Building Open Digital Libraries. D-Lib Magazine, 7(12), Dec. 2001, http://www.dlib.org/dlib/december01/suleman/12suleman.html

[SULEMA2] Hussein Suleman and Edward A. Fox. Designing Protocols in Support of Digital Library Componentization. In *Research and Advanced Technology for Digital Libraries, 6th European Conference, ECDL 2002, Rome, Italy, September 16-18, 2002, Proceedings*, eds. Maristella Agosti and Constantino Thanos, pp. 568-582.

[THREDDS] Ben Domenico. Thematic Realtime Environmental Data Distributed Services THREDDS. http://my.unidata.ucar.edu/content/software/thredds/index.html

[UNIDAT] Overview of the Unidata Program. http://www.unidata.ucar.edu/factsheets/overview.html

[UNIDATI] David Fulker, Sally Bates, Clifford Jacobs. Unidata: A Virtual Community Sharing Resources via Technological Infrastructure. Bulletin of the American Meteorological Society, 78(3).

[UNIDATP] Unidata 2003: Proposal to the National Science Foundation. Principal Investigator: David W. Fulker, http://unidata.ucar.edu/proposals/NSF2003/

[VLEEVA] Britain, Sandy and Liber, Oleg, *A Framework for Pedagogical Evaluation of Virtual Learning Environments*, http://www.jisc.ac.uk/jtap/htm/jtap-041.html

[VLEGRE] Archer, R. Virtual learning environments for distance learning in HEIs: a guide to resources. Unpublished report for IMRI, University of Nottingham, 2000.

[VLEPED] Britain, S.; Liber, O. A framework for pedagogical evaluation of virtual learning environments. JTAP, 1999, http://www.jisc.ac.uk/jtap/htm/jtap-041.html

[VLEREQ] Requirements for a Virtual Learning Environment. JISC, August 2000, http://www.jisc.ac.uk/vle

[VLETLS] Links to the VLE and MLE tools on the INSPIRAL Web site, http://inspiral.cdlr.strath.ac.uk/resources/toolssystems.html

[WTSM99] Raven Wallace. Teaching with the Web: Challenges in a Complex Information Space. Dissertation. March 1, 1999, http://www-personal.umich.edu/~ravenmw/disssumm.html

[Z203RLQ] P. Ryall, Type 203 Ranked List Query Spec, 1996, http://www.itl.nist.gov/iaui/894.02/projects/rlq/RLQ_spec_Feb96_ZIG.html

The Internet resources as at the end of 2002.

Appendix. Glossary of acronyms

AAAS: American Association for the Advancement of Science

ACM: Association for Computing Machinery

ADEPT: Alexandria Digital Earth Prototype

ADL: The US Department of Defense's Advanced Distributed Learning Network

ADN: ADEPT/DLESE/NASA Joint Metadata Framework

CITIDEL: Computer and Information Technology Interactive Digital Education Library

COMET: Cooperative Program for Operational Meteorology, Education and Training

CSTC: Computer Science Teaching Centre

DAML: DARPA Agent Markup Language

DL: digital library

DLE: Digital Library in Education

DLESE: Digital Library for Earth System Education

DLI: Digital Library Initiative

DNER: Distributed National Electronic Resource

DPC: the DLESE Program Center

EAD: Encoded Archival Description

ERCIM: European Research Consortium for Informatics and Mathematics

ESS: Earth System Science

ETD: Electronic Theses and Dissertations

ETRDL: ERCIM Technical Reference Digital Library

FGDC: Federal Geographic Data Committee standard for digital geospatial metadata

GILS: Global Information Locator Service

ICTs: information and communication technologies

IEEE: Institute for Electrical and Electronics Engineers

IITE: UNESCO Institute for Information Technologies in Education

IMS: instructional management systems

IT: information technologies

JCDL: Joint IEEE/ACM Conference on Digital Libraries

JDL: NASA's Joined Digital Library

JISC: Joint Information Systems Committee, UK

LOM: IEEE learning object metadata

LTSC: IEEE Learning Technology Standards Committee

MBL: multimedia-based learning

MLE: Managed Learning Environment

NASA: National Aeronautics and Space Administration, USA

NDLTD: Networked Digital Library of Theses and Dissertations

NSDL: National Science Digital Library

NSF: National Science Foundation, USA

NSTA: the National Science Teachers Association, USA

OAI: Open Archives Initiative

OIL: Ontology Inference Layer

OSOSS: Online Self-Organizing Social Systems

OWL: Ontology Web Language

PC: Policy Committee

PICats: Publishable Inventories and Catalogues

SCG: Semi-structured Conceptual Graphs

SCORM: Sharable Content Object Reference Model

SRB: Storage Resource Broker

STEM: Science, Technology, Engineering, and Mathematics

TDs: theses and dissertations

THREDDS: Thematic Realtime Environmental Distributed Data Services

VLE: Virtual Learning Environment

WSDL: Web Services Description Language

XML: eXtensible Markup Language