DLESE: A Case Study in Sustainability Planning

Mary Marlino, Tamara Sumner, Karon Kelly, and Michael Wright
Digital Learning Sciences
University Corporation for Atmospheric Research and University of Colorado at Boulder

Abstract: Sustaining open access educational digital libraries presents unique challenges and opportunities. This paper describes these challenges and opportunities, and presents the processes and strategies that were developed to address them at the Digital Library for Earth System Education (DLESE).

Introduction

For the past seven years we have been operating the Digital Library for Earth System Education (DLESE – www.dlese.org), with generous funding from the Geoscience Directorate of the National Science Foundation (NSF). Like all good things, grants from the NSF end at some point; in DLESE's case, in Fall 2007. We were tasked with developing and implementing a sustainability plan that would ensure DLESE users continued access to the educational resources and collections in the library for the "foreseeable future."

DLESE is a large, geoscience education community undertaking involving scientists, educators, and library builders from many institutions across the nation. The goal of this grassroots, community-led project is to provide searchable access to high-quality, online educational resources for K-12 and undergraduate Earth system science education (Marlino et al., 2001). These resources include simulations, maps, lesson plans, lab exercises, data sets, virtual field trips, and interactive demonstrations. The primary capabilities and content embodied in the library include:

(1) Access to approximately 13,500 digital educational resources, organized into 41 thematic collections contributed by 25 different institutions. Resources in the library were created by a wide variety of individual faculty members, agencies, and institutions and are held (stored) on local servers.
(2) Tools to support collection development and curation. The DLESE Collection System (DCS) enables collection developers to catalog educational resources (lesson plans, modules, data, imagery, etc.), news and opportunities announcements, and annotations about resources.
(3) A sophisticated discovery service, supporting both searching and browsing, based on the Lucene engine (Weatherley, 2004). Users can search DLESE collections by keyword, grade-level, educational resource type, and educational standard.
(4) A variety of mechanisms for user contributions and community building, including a "Suggest a Resource" feature, a system enabling both teachers and learners to submit teaching tips and other informal comments, a News and Opportunities service, hosting of listservs for geoscience education groups, and a community newsletter.
(5) A technical infrastructure supporting all of these capabilities based on open platforms and open standards, e.g., Lucene, OAI-PMH, java, javascript.
(6) An active and significant user base, exceeding over a million library sessions annually.

Sustainability Planning

We were asked by NSF to develop a sustainability plan that would:

- Continue to make all DLESE resources widely available for the purpose of education
- Acknowledge NSF support for the initial development of DLESE in all future publications
- Honor the DLESE Intellectual Property Policy established July 2, 2002, the essential element being that the IP for metadata or technologies created by community members for the library would remain the property of their home institution.

In addition, NSF asked us to convene an advisory board to provide guidance on sustainability planning, criteria for decision-making and selection of new business models or host environments, and recommendations for new hosts. This board was composed of recognized experts in geoscience, library operations and strategic planning, and business.

To develop this plan, we went through a structured analysis process consisting of the following major steps:

- Defining core library components and determining what should be sustained
- Developing a taxonomy characterizing different operational levels
- Developing cost estimates for different operational levels
- Developing criteria for selecting new business models or host environments
- Developing a range of models characterizing different hosting configurations
- Conducting an IP audit

Defining core library components and determining what should be sustained

A key challenge in sustainability planning is disaggregating components of the library and determining what should be sustained. This is a complex issue with no single correct answer. Digital libraries have many different, yet interdependent, components such as content, technology infrastructure, and end-user services. In this process, we defined the components of library operations to include system administration, application support, content processes, workflows, maintenance, and use metrics. "Content" refers to the development and curation of collections, educational resources, and the library portal website. We defined library services as those providing support for library developers and end-users. We developed a taxonomy characterizing four levels of service reflecting different levels of demands on the host environment.

Level 1 Service, the minimum required by NSF, was characterized by offering access to library collections as static HTML pages available on the Web, i.e., each item-level metadata record in the library would be rendered as a web page and users would access the DLESE collections directly through their web browsers or via commercial search engines.

Level 2 Service focused solely on sustaining library content and was characterized by providing users with access to DLESE collections through a third party site, such as a university or public library. In this level, metadata records would be ingested into the collection management systems already in place in the host environment.

Level 3 Service focused on the continuation of library operations and selected end-user services, i.e., users would access curated collections and services through the DLESE.org interface. This required a host environment to provide technical hosting for both hardware and software systems, collection curation, and support for selected library services.

Level 4 Service was based on modifying library systems to achieve significant integration with the NSDL technical infrastructure, specifically the NSDL Data Repository released in early 2007. This Fedora-based infrastructure is operated by NSDL on behalf of its community and it currently provides reliable technical operations at no additional cost to NSDL member libraries.

Developing cost estimates for different operational levels

The primary challenge that we faced in developing reasonable cost estimates for future operations is that any reliable estimate is completely dependent upon the new host's technical and human resource infrastructure, as well as the level of service to be provided. Given these uncertainties, we parsed out the major functions that would have to be undertaken to guarantee Level 3 Service, and based our estimates on what this level of

sustainability would approximately cost with our existing institutional infrastructure and talent base intact. We made a considerable effort over the past year to streamline the operational costs associated with DLESE, including discontinuing support for many community and library developer services, and automating workflows and maintenance procedures around continued end-user and collections services as much as possible.

Developing Criteria for Hosts Selection

The criteria we developed for selecting a host were based on four factors: the mandates provided by NSF, the legal status, organizational capabilities, and financial stability of a potential host environment. The legal status of an organization refers to its ability to assume liability for the DLESE intellectual property, privacy and terms of use policies. Necessary organizational capabilities include significant experience with operating and maintaining server hardware and the software systems that underpin library operations, and experience in curating library collections, ideally digital collections where web-based resources are continually monitored for availability. The host organization must also have sufficient knowledge to answer support questions. For DLESE, these questions often require domain knowledge in both geoscience and education. It is also important that hosts demonstrate financial commitment and stability that could ensure library operations for at least three years. Ensuring stable library operations for a three year period seemed to be a minimally acceptable "return on investment" for a one year planning effort.

Developing a range of models characterizing different hosting configurations

As we considered the types of organizations or organizational configurations able to provide a suitable hosting environment, several models emerged:

Sponsorship Model. In this model, DLESE sustainability would be undertaken by an allied Earth science professional society, a public or government agency, or sponsorship by a private foundation. This model does not necessarily assume that the sponsoring agency would be the actual host institution for operational services.

Hybrid Model. This model is a variation of the above, but blends public and private sector support. Again, this model does not assume that the sponsoring agency would be the actual host institution for operational services.

Adoption Model. In this model, an institution (collegiate, private, government, etc.) would subsume DLESE operating costs into an existing budget as part of their institutional remit. That is, the institution would consider the mission of DLESE and its operating requirements to be so closely aligned with its core mission that DLESE would become an additional service that the institution would provide to its stakeholders and core constituencies.

Partnership Model. In this model, multiple organizations or organizational entities in one institution would assume responsibility for different components of library development under a collaborative agreement. An example of this would be a partnership between the National Center for Atmospheric Research (NCAR) Library for curation services, the NCAR Computational and Information Systems Laboratory (CISL) for hosting services, and the University Corporation for Atmospheric Research (UCAR) Office of Programs for system support and upgrades.

Our considerations of these various models indicated that the partnership model was the most promising one for DLESE. In our discussions with potential hosts, we found that very few organizations have the combination of mission alignment and all three capabilities – technical, library, and domain knowledge and skills – in house.

Conducting an IP audit
A key issue related to the transition of DLESE to another host or operator concerns intellectual property rights and ownership of collections and technical infrastructure. UCAR (DLESE's host institution under NSF funding) owns 45% of the metadata in DLESE, 15% is in the public domain (e.g., NASA data), and the remainder are owned by 22 other institutions. The core infrastructure and technologies for DLESE developed at UCAR will continue to be available on SourceForge under a GPL open source license. One of the lessons learned was that obtaining licenses or permissions related to transfer of intellectual property rights between institutions can be an extended process.

Reflections and considerations
We are pleased to report that as a result of our sustainability planning, we have successfully negotiated agreements with UCAR and NCAR to collaboratively provide the capabilities to continue Level 3 Service; i.e., continue to make DLESE collections and core end-user services available through DLESE.org. It is a common refrain amongst those discussing sustainability planning to remark that it should be taken into account from the beginning of a project. We found it extraordinarily useful to have an advisory board dedicated to sustainability planning. In retrospect, the DLESE effort would have benefited from establishing a separate advisory board focused solely on this challenge early in the project. In addition to bringing in invaluable forms of expertise, this separate board could fulfill a vital function in providing advice about how to strike a balance between providing community services and controlling costs.

Second, as presaged above, developing a disciplined model for cost control is a critical element for long-term project sustainability. There is a difficult and delicate tension between building the library and building the community. Both must go hand-in-hand, but in our experience, they have very different cost structures. It is incumbent upon the broader digital library and scientific communities to develop a more detailed and thorough understanding of the cost structures and benefits for different architectures around collaboration and distributed construction, particularly in this era of eScience.

Finally, we recognize that sustaining the library's community of developers and users is perhaps the most important, albeit most difficult, aspect of library sustainability. One of the most enduring artifacts of the DLESE experience is a community with an enhanced level of digital library expertise, sharing resources for the common good. A frequently cited definition of "sustainability" is the one created by the Brundtland Commission (United Nations, 1987), which defined sustainable development as development that "meets the needs of the present without compromising the ability of future generations to meet their own needs." This philosophy has guided our planning over the past 12 months as we considered our sustainability options to ensure that our library and our library community remain vibrant and relevant in years to come.

References
Marlino, M., T. Sumner, D. Fulker, C. Manduca, and D. Mogk. 2001. The digital library for earth system education (DLESE): Building community, building the library. Communications of the Association for Computing Machinery (ACM); 44, no.5, 80-81.

Weatherley, John. 2004. DLESE search service documentation. http://www.dlese.org/dds/services/ddsws1-1/service_specification.jsp

United Nations. 1987. Report of the World Commission on Environment and Development. General Assembly Resolution 42/187, December 11.