

**MIT Libraries' DSpace Business Plan Project
Final Report to the Andrew W. Mellon Foundation**

July 2002

Submitted by Mary R. Barton and Julie Harford Walker

An award from the Andrew W. Mellon Foundation granted to the Massachusetts Institute of Technology in April 2000 has enabled MIT Libraries to develop a business plan to transform its DSpace research project into a sustainable technology platform and service administered by MIT Libraries and adopted by the Institute's producers and consumers of digital scholarly materials.

Table of Contents

Executive Summary	3
Introduction	5
DSpace Overview	5
Methodology	6
DSpace Business Plan Overview	8
Further Research	20
Lessons Learned	22
References	24
Appendix A: MIT Faculty Survey Results	25
Appendix B: Total DSpace Costs at MIT	33

Executive Summary

Project Description

An award from the Andrew W. Mellon Foundation granted to the Massachusetts Institute of Technology in April 2000 has enabled MIT Libraries to develop a business plan to transform its DSpace research project into a sustainable technology platform and service administered by MIT Libraries and adopted by the Institute's producers and consumers of digital scholarly materials. DSpace is MIT Libraries' innovative institutional digital repository designed to manage, host, preserve, and enable distribution of the scholarly output of MIT's faculty. Developed as a joint research project of the MIT Libraries and Hewlett-Packard (HP) through invent@MIT, the HP-MIT Alliance, it reflects MIT's mission to "generate, disseminate and preserve knowledge" and will provide MIT faculty with a stable long-term storage and content management system to house their digitally formatted work.

DSpace is currently in pre-production release to four Early Adopter Communities at MIT and will be offered to the entire MIT campus in September 2002. The next phase of the project will include a collaborative federation plan for DSpace implementation and further development at other pre-eminent research universities around the world. DSpace code will be made freely available to all through an Open Source license.

Service Offering

DSpace Core Services are comprised of two distinct but interconnected service elements, Interactive Services and Operations Services. DSpace Interactive Services offer a fully functional system that allows DSpace Community members to accomplish all tasks necessary to submit and access items in DSpace. Additionally, MIT Libraries provides Operations Services to host and preserve faculty materials, establish and deliver ongoing support for DSpace Communities, respond to customer inquiries, and supply system monitoring, back-up, and recovery. Services beyond those in the DSpace Core Services will be offered through DSpace Premium Services and may be offered on a fee-for-service basis. Potential Premium Services include E-Conversion Services, Metadata Services, Custom Repository Services, and User Reporting Services.

Governance

The DSpace team will receive guidance from three advisory groups, the DSpace Faculty Advisory Board, the DSpace Policy Committee, and the DSpace Operational Advisory Committee. The DSpace Faculty Advisory Board was formed to provide guidance from the perspective of MIT target users of DSpace, both as contributors of content and as end-users of the system. The DSpace Policy Committee will make policy decisions related to DSpace services, standards, and functionality within the Libraries context. The DSpace Operational Advisory Committee will provide two-way communications between the DSpace team, the library units, and other users regarding needed bug fixes, system enhancements, and future system development.

Management, Staffing, and Training

The Libraries recognize that the ability to foster and respond to rapid early growth effectively will be critical to the acceptance of DSpace by MIT Communities. Although the skills

required to run DSpace exist among current Libraries staff, it was determined that relying exclusively on existing staff would provide a service far too fragmented for success and, thus, the DSpace team will be staffed primarily with dedicated resources. Two new positions have been approved to provide centralized management of the DSpace service offering, a DSpace User Support Manager and a DSpace Systems Manager. The two new positions will join the Libraries' staff in support of the launch of DSpace in the Fall of 2002 and will report to the Libraries' Associate Director for Technology, who will have overall responsibility for DSpace. In advance of the DSpace launch, the User Support Manager and Systems Manager will be conducting general training sessions for all members of MIT Libraries and in-depth training for designated units of the Libraries.

Communications Plan

DSpace has a communications plan designed to reach all potential target audiences including Libraries staff, faculty and researchers, consumers of DSpace content, MIT alumni, potential university federation partners, and the general public. With the support of the Libraries' Communications Manager, HP Public Relations, and the MIT News Office, regular updates and information will be disseminated through a variety of channels such as Institute publications, faculty meetings, and on-line publications. An event has been planned for early November 2002 to announce the formal launch of the DSpace service within MIT.

Cost Model

The cost model captures all of the expenses that MIT Libraries will incur for ongoing DSpace staffing and operations for Core Services. Costs are reported in three categories: incremental, principal, and comprehensive expenses. Incremental costs represent the expenses associated with dedicated staff and equipment. Principal costs include existing staff with a significant portion of their activities redirected to DSpace. Comprehensive costs include expenses that enhance, but are not vitally necessary to providing, the most basic DSpace service. Expenses associated with incremental staff constitute the majority of DSpace costs.

Operations Funding Model

MIT Libraries plan to transition DSpace from its reliance on outside funding to a more sustainable funding model. Consistent with the Libraries' mission, Core Services will be offered free of charge to all registered members of the MIT community. In keeping with MIT's mission, content will be offered as freely as possible via the Web to the public. This service strategy precludes seeking user or subscription fees for means to support the ongoing operations of DSpace. The proposed funding model will rely upon a number of potential resources including, but not necessarily limited to, support from the Institute, revenue from Premium Services, and support from corporate and federation partners. Support may take the form of financial support or in-kind assistance. Collectively, these contributions will cover the operational costs of DSpace, as well as some future development needs.

Introduction

DSpace is MIT Libraries' innovative institutional digital repository designed to manage, host, preserve, and enable distribution of the scholarly output of MIT's faculty. Developed as a joint research project of the MIT Libraries and Hewlett-Packard (HP) through invent@MIT, the HP-MIT Alliance, it reflects MIT's mission to "generate, disseminate and preserve knowledge" and will provide MIT faculty with a stable long-term storage and content management system to house their digitally formatted work.

This report will share what we learned through research into economic models in scholarly communication, how we defined a sustainable business model for MIT's DSpace system, the DSpace business plan itself, and our plans for the next phase of the project, which include a collaborative federation plan for DSpace implementation and further development at other pre-eminent research universities around the world. The evolving nature of DSpace necessitates that this plan is a snapshot of our current business strategy, one that is evolving as more users adopt the service. It reflects the organizational and technical infrastructure and resources of MIT. Federators will be able to scale DSpace to match the needs and resources of their own institutions, creating a repository with the scope of digital formats and the scale of content that address their constituent's needs. DSpace was developed under the guidelines of the Open Source Initiative which will enable system customization and collaborative development among federation partners.

DSpace Overview

DSpace was developed in response to expressed faculty needs for an easy-to-use, dependable service that could manage, host, preserve, and distribute faculty materials in digital formats. It offers faculty the advantages and convenience of web-based submission and dissemination. DSpace can accommodate a variety of genres (e.g. documents, datasets, images) and formats (e.g. text, audio, video, images).

DSpace differs from other digital library initiatives in that it employs a distributed submission process and seeks to capture newly created digital research materials in a broad range of formats. DSpace is also well suited to housing digitized historic collections to enhance the contextual reference for newly submitted works (e.g. an historic collection of working papers to augment newly submitted works in the series, providing researchers with a sense of "completeness" for the collection). Self-defined subsets of the MIT academic and research community (such as schools, labs, centers, and departments) will determine for themselves what research materials may be submitted. These self-defined, depositing Communities will also determine who may have access to archived works, with options ranging from a worldwide audience to a select few. There is no charge for submitting to or viewing digital material in DSpace, although the MIT Libraries reserve the right to introduce future fees as needed through Premium Services offerings to satisfy the exceptional resource requirements of a particular depositing Community.

MIT Libraries released a test version of the DSpace service in February 2002 to a set of four early adopter Communities within MIT who will demonstrate the viability of the submission process. Their evaluation and feedback will be incorporated into the Fall 2002 release that will be made available to Communities throughout MIT. Worldwide access to the contents also will be provided at that time, with the exception of any items that have been specifically restricted from public access.

Methodology

Literature Review and Primary Research

We began this project by thoroughly investigating digital library initiatives and the economics of information and quickly learned that libraries are faced with a tremendous number of new challenges in the digital realm.

Through our library colleagues we learned that faculty are producing and capturing more and more research digitally in numerous formats. However, much of this research is in imminent danger of loss because the storage media often are degrading on professors' shelves or the hardware or software used to access data is obsolete. "The timeframe for decisions and actions to prevent loss is a matter of years, not decades." (Hedstrom)

Universities have traditionally relied upon institutional libraries and archives to preserve the scholarly record. Yet preservation processes and techniques for print materials that were perfected over hundreds of years must be recreated for digital assets. Digital materials present additional legal and regulatory complexity beyond those traditionally associated with print materials.

The ubiquity of the Internet leads scholars increasingly to demand that their research needs be met on-line. A Digital Library Federation study found that faculty members are increasingly using the Internet to access library resources from the comfort of their offices. (Greenstein, 2002) Further they increasingly avail themselves of the Internet to share their research with colleagues worldwide, many through their own or discipline specific web sites, yet this provides a disparate, unreliable, and unpredictable source of information for researchers and colleagues.

Scholarly Communications Stakeholder Review

Our review of digital library and related solutions, both for-profit and non-profit, revealed a highly fragmented market with individual initiatives selecting different aspects of the problem to address, and different means to achieve solutions. We included independent scholarly communication initiatives (e.g. Scholarly Publishing & Academic Resources Coalition (SPARC)), university library scholarly communication initiatives (e.g. California Digital Library's eScholarship), commercial digital asset management vendors (e.g. CONTENTdm), digital preservation initiatives (e.g. Preserving Access to Digital Information (PADI)), digital

library organizations (e.g. Digital Library Federation (DLF)), and standards initiatives (e.g. Open Archive Initiative (OAI)) in our assessment.

We were able to develop an understanding of how institutionally based digital repositories like DSpace can best respond to the needs of institutional constituents. We confirmed that DSpace is relatively unique in that there are no other institutionally based repositories that invite the breadth of file formats, have digital preservation as an explicit objective, and provide a flexible, decentralized community-based submissions process. Initiatives that bear the greatest similarities to DSpace include the California Institute of Technology's EPrints project and the University of California's eScholarship Repository, but neither encourages the breadth of file formats that can be found in DSpace.

We also learned that the need for institutional repositories is widespread based on the multitude of new projects and the inquiries we received from other universities. For example, Ohio State University, in their recently completed proposal for the development of "Knowledge Bank", identified DSpace as a potential component of their broad-based institutional repository initiative.

Further, our review revealed the various business models in use. Independent digital library and scholarly communication initiatives typically must recover the cost of operations by charging participants. SPARC finances its efforts through coalition member fees that support operating expenses and help build a capital fund to provide start-up money for SPARC projects. A full member of SPARC pays annual membership dues and makes a purchase commitment to support the SPARC-endorsed journals that fit their collection development needs. JSTOR charges subscribers both a one-time, up-front fee and a yearly maintenance fee.

A common model among university-based scholarly communication initiatives is sponsorship by the university through its library. The California Digital Library sponsors the University of California's eScholarship program. California Institute of Technology's EPrints project and xxx.arXiv.org also rely on institutional funding.

Operational Planning

We developed the DSpace business plan collaboratively with MIT Libraries staff, the DSpace development team and MIT's Information Systems organization. Our participation in weekly DSpace development team meetings, the DSpace Transition Planning Group, and the Early Adopter Librarian team was invaluable in our research into the operational implementation of an institutional digital repository.

Operational implementation of DSpace at MIT Libraries began under the guidance of the DSpace Faculty Liaison and continued with the DSpace Transition Planning Group. Formed as a working group charged with outlining operational, marketing, management and staffing plans, the group comprised the Associate Director of Collection Services; Associate Director of Technology; the Heads of Dewey Library for Management and Social Sciences, Document Services, Bibliographic Access Services, the Libraries' Systems Office, and the Institute

Archives and Special Collections; DSpace's Faculty Liaison and System Curator; a Senior Strategist from MIT's Academic Computing/IS organization; and the DSpace business plan team. The work of this group is represented heavily throughout our business plan.

An additional group, the Early Adopter Librarians task force, staffed with Subject Specialists from each of the Libraries, provided additional insight into DSpace's potential faculty and researcher user base. Subject Specialists, who work directly with faculty members and other library users on their research questions, contributed significantly to the team's understanding of these interactions. This group also was key in understanding the organizational impact that the implementation of DSpace would have on Libraries staff.

Faculty Survey

We administered a survey to the tenure track faculty of MIT, in order to learn about their perceptions and anticipated use of DSpace. Respondents were roughly representative of the overall Institute's department and tenure mix. Our findings served as a means to validate aspects of the business plan as well as the Transition Planning Group's plans for marketing DSpace. A summary of the survey results can be found in Appendix A.

In order to validate our results (based on ~10% response rate), we are in the process of administering the survey to additional faculty members randomly selected from among the non-respondents. We will analyze the results from this random sample to confirm that our responses are representative of the entire faculty.

Cost Model

We developed a cost model to capture the full economic cost of operating DSpace including staff impact, space, hardware and other Libraries resources, only some of which will result in differential cash flows. The remainder of the costs are important to capture for MIT Libraries' planning purposes. We gathered data from the Transition Planning Group staffing model, HP, MIT Libraries' records and MIT central accounting. Our model is designed to reflect the costs of operating DSpace; therefore we specifically ignore system development costs which included \$1.8 million for development as well as 3 FTE HP staff and approximately \$400,000 in system equipment.

DSpace Business Plan Overview

The DSpace business plan addresses five key areas:

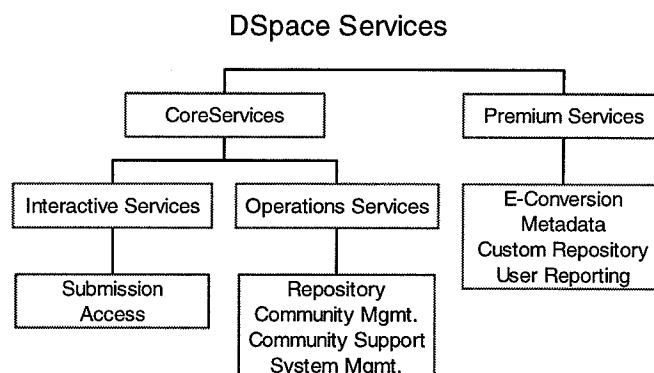
- Distribution License
- Service Description
- Service Delivery
- Building Awareness and Driving Adoption
- Financial Model

The Distribution License section explains MIT Libraries' rights in distributing DSpace content. The Service Description provides an overview of the DSpace services that MIT Libraries will or may offer now and in the future to support user needs. Service Delivery describes the organizational infrastructure that has been approved by the Libraries' management Steering Committee for managing and growing the service. Building Awareness and Driving Adoption outlines publicity channels and events to promote usage of the system. And the Financial Model describes in detail the cost model for ongoing operations and potential sources of funding.

Distribution License

MIT Libraries will not own the content of DSpace but rather will hold a non-exclusive license to distribute and preserve items. It is the aim of DSpace to make all of the content available freely, but there will be certain instances where access to digital objects will necessarily be restricted, and those requests will be honored. However, MIT Libraries reserve the right to limit the amount of restricted content that individual Communities may deposit in DSpace.

Service Description



The DSpace Service is divided into two main areas: Core Services and Premium Services. Core Services are freely available to Community members and consumers of DSpace content, as applicable. Premium Services are specialized services designed to meet the extraordinary needs of Community members and may be offered on a fee for service basis.

DSpace Core Services

DSpace Core Services are comprised of two distinct but interconnected service elements, Interactive Services and Operations Services. DSpace Interactive Services offer a fully functional system that allows DSpace Community members to accomplish all tasks necessary to submit and access items in DSpace. Additionally, MIT Libraries provide Operations Services to host and preserve faculty materials, establish and deliver ongoing support for DSpace Communities, respond to customer inquiries, and supply system monitoring, back up, and recovery.

DSpace Interactive Services

Submission Services

DSpace offers faculty a flexible, easy-to-use submission process. A depositor simply completes a brief submission form and grants permission to distribute and preserve the work. At the discretion of the individual Community, new submissions to a Community's Collections may then be subjected to a review and approval process, customized to a given research community's needs. The workflows of the submission process as well as any reviewer, metadata editor, and administrator roles are established during a Collection's initial set-up on DSpace. Communities requiring assistance with bulk importing of items into DSpace can request assistance from the Libraries through Premium Services.

Access Services

DSpace's Access Services help users browse and search the repository to discover items of interest, request one or more items from the system, and take delivery of the requested items. In the future, DSpace may offer e-mail notification Premium Services when items of interest to individual users are added to DSpace.

DSpace Operations Services

Repository Services

MIT Libraries provide storage and preservation management services to ensure the longevity of all deposited materials. For each submission, DSpace provides persistent storage, including appropriate back-up and recovery procedures; assigns a unique persistent identifier that is appropriate to cite in other works; supports a range of service levels for preservation and notifies users of DSpace's preservation constraints; stores provenance information; and maintains an auditable history and record of changes to the archive. Initial storage limits will be established for all Communities based on the size of their Community. Communities with greater storage needs may be served through Premium Services.

Community Management Services

DSpace's Community Management Services are a set of consultative and applied services designed to meet the needs of individual Communities. The MIT Libraries staff will provide guidance in developing a Community or Collection's workflow process, metadata requirements, and distribution policies as well as launching a custom DSpace Community homepage. A standard set of on-line reports will also be made available to the Community to assist in the self-management of their Community and Collections. Additional custom reports will be made available through Premium Services.

Community Support Services

The MIT Libraries will provide both web-based and telephone support to all registered DSpace community members. Support will be provided Monday through Friday, 9:00am to 5:00pm. All inquiries will receive a return response via e-mail or telephone within one business day.

System Management Services

System Management Services are the back-office support for DSpace, working to ensure optimal system performance and high availability. The MIT Libraries, with support from MIT's Information Systems team, will provide a comprehensive set of technical services for DSpace system operations, including system monitoring, testing and debugging; system administration; monitoring and upgrading DSpace utility programs and middleware; and developing system enhancements.

DSpace Premium Services

The MIT Libraries plan to introduce a set of Premium Services to meet the exceptional resource requirements of DSpace Communities. DSpace was designed as a distributed system that allows Communities to manage their own Collections, and the DSpace Core Service includes all necessary functionality and support to achieve that objective. We anticipate that some DSpace Communities may nonetheless choose to outsource certain activities to the Libraries or may put extraordinary demands on the system that require additional library resources for support. For those Communities, MIT Libraries will offer Premium Services as a means of managing the impact of those user demands on Libraries staff and DSpace resources and will reserve the right to introduce fees as needed to aid in cost recovery. The primary aim of Premium Services is therefore not to generate funds but rather to allow MIT Libraries to respond to user needs. We expect MIT Libraries' Premium Services will be competitive with off-campus offerings and will offer DSpace Communities a reliable source of expertise.

As with any new service, it is difficult to determine how users will employ specific features of the system and the resulting services that they will demand. The broad Premium Services areas identified thus far have been divided into the following categories: E-Conversion Services, Metadata Services, Custom Repository Services, and User Reporting Services. Fee based service will be introduced only when user demand warrants, the system functionality is developed, and, where necessary, resources are obtained to support them within the Libraries. The Libraries plan to offer these services to ensure that DSpace offers a full set of resources to meet faculty and researcher's needs.

The following are brief summaries of our current plans for each Premium Service area.

E-Conversion Services

E-Conversion Services will assist individual submitters or Communities in two main areas: the creation of digital content from non-digital materials (e.g. scanning print materials to create a PDF) and custom, on-demand transformation of materials from one format to another (e.g. from Microsoft Word into PDF). Already offered through MIT Libraries' Document Services department, E-Conversion Services will be extended and scaled to fit the needs of the DSpace Communities. For particularly large or specialized jobs, Document Services staff will attempt to identify and recommend reliable outside vendors.

Metadata Services

Metadata, "vocabulary for describing resources that enable more intelligent information discovery systems" (Dublin Core web site: <http://dublincore.org/about/>), is an emerging field that has grown increasingly complex with the proliferation of descriptive metadata schemas. Each item in the DSpace system will have its own metadata record that is entered by end-users when content is submitted. Currently, Dublin Core is the metadata standard used by the DSpace system, but we anticipate that Communities will demand discipline specific metadata schemas. Identifying and establishing the most advantageous metadata schemas is an area where Communities can draw on the experience and knowledge of the Libraries.

The DSpace Metadata Services will likely focus on descriptive metadata in the initial service offering, however, it is our intention that Metadata Services encompass administrative, structural, technical, rights, and preservation metadata as these schemas further develop and evolve. Over the next two years, we anticipate the introduction of descriptive metadata services in the following three areas:

- *Metadata consulting* – includes needs assessment, survey and feasibility study of available metadata schemas, and recommendation
- *Metadata research services* – includes a variety of services such as metadata crosswalks, advice on appropriate taxonomies, consultation on tools for metadata creation, capture and management, authority control advisory service, and preservation metadata
- *Metadata creation and support services* – includes encoding and Dublin Core metadata entry and review of metadata submissions for data consistency

Custom Repository Services

To manage the flow of items into DSpace and ensure that sufficient and equitable allocation of storage space is available to all DSpace Communities, the Libraries will be monitoring the amount of materials that are submitted to DSpace by each Community through a growth monitoring system. Normal growth expectations, set according to the

number of submitters in a community, will establish norms and upward limits for both the number of items that can be submitted each year and the size of an individual item. Those individuals or Communities that have storage needs that exceed the normal limits will be asked to subscribe to our Premium Customer Repository Services, through which the standard DSpace storage space will be expanded to meet that Community or individual's requirements.

User Reporting Services

User Reporting Services will be designed to provide individuals and Communities with customized information about DSpace's contents:

- *Research Alert Services* - email notifications sent to individual end-users when items of interest are added to DSpace based on a profile or search query filled out by the end user
- *Targeted Notification Services* - email notifications sent out to a listserv or electronic discussion list that is selected by a community as a suitable forum for displaying its work
- *Hot Topic Citations* - Compilations of items organized around high-demand or "hot" topics provided to end-users periodically in print or online
- *Custom Reporting Services* - Electronic or printed reports for faculty or communities giving information such as size of communities and collections; number of viewings; number of content downloads by item, collection or community; and number of items added by community, collection, or author

Service Delivery

Management and Staffing

The Transition Planning Group developed a management and staffing model for DSpace that anticipates the need for dedicated attention to build the Community base and support the technology infrastructure in the first few years of operation. The Libraries recognize that the ability to foster and respond to rapid early growth effectively will be critical to the acceptance of DSpace by MIT Communities.

The Transition Planning Group identified two new positions to provide centralized management of the DSpace service offering. Although the skills required to run DSpace exist among current Libraries staff, the Transition Planning Group determined that relying exclusively on existing staff would be far too fragmented and, thus, has chosen to staff the DSpace team with dedicated resources. Additionally, the plan anticipates the eventual need for a dedicated Product Manager, a role that in the interim will be assigned to the Associate Director for Technology.

The two new positions, a DSpace User Support Manager and a DSpace Systems Manager, will join the Libraries' staff in support of the launch of DSpace in the Fall of 2002. These two

positions will report to the Libraries' Associate Director for Technology, who will have overall responsibility for DSpace. The User Support Manager will manage documentation, training, and the communication with and set-up of DSpace Communities. The User Support Manager also will serve as the initial contact and coordination point for all Premium Services. The Systems Manager will have primary responsibility for the technical management of DSpace, including system monitoring, backup and recovery, and will serve as the liaison with MIT's Information Systems (IS) organization. Through a Service Level Agreement, IS will provide Libraries with optimally conditioned space for the DSpace server rack, redundant power and network connections, and offsite removable media shuttling.

Governance

The DSpace team will receive guidance from three advisory groups, the DSpace Faculty Advisory Board, the DSpace Policy Committee, and the DSpace Operational Advisory Committee. The DSpace Faculty Advisory Board was formed to provide guidance from the perspective of MIT target users of DSpace, both as contributors of content and as end-users of the system. As DSpace is broadly adopted at MIT, and more is learned about user practices, Community needs and system functionality in practice, we anticipate a continued need for this advisory role. The Board will be convened on an annual basis to provide this vital user feedback to the future development of DSpace functionality or associated Library services. Members of this group include faculty and administrators from all five Institute schools, the Director of Libraries, the Associate Director for Technology, and a divisional librarian.

The DSpace Policy Committee will make policy decisions related to DSpace services, standards, and functionality within the Libraries context. Committee decisions deemed to have significant financial, service or public relations impact will be presented to the Libraries' Steering Committee for their approval. The Policy Committee will be chaired by the Associate Director for Technology and its membership will be comprised of the heads of several MIT Libraries service units, the Information Technology Librarian for Public Services as well as the DSpace Systems Manager and DSpace User Support Manager.

The DSpace Operational Advisory Committee will provide two-way communications between the DSpace team, the library units, and other users regarding needed bug fixes, system enhancements, and future system development. The Committee will be chaired by the User Support Manager and will include among its membership the DSpace Systems Manager, a representative subject specialist/reference librarian from each divisional library, a staff member with metadata expertise, a Libraries' Systems Office representative, and an Institute Archives representative. Recommendations from the team will be provided in an advisory role to the Associate Director for Technology. Designated representatives from DSpace Communities may also participate on the Operational Advisory Committee until such time as there is a sufficient number of Communities established in DSpace to warrant the formation of a DSpace User Group.

Training

We aim to make DSpace an integral part of the daily business of the MIT Libraries. As with other academic information resources that the Libraries provide, librarians are trained and prepared to instruct members of the MIT community in the use of that resource. In advance of the DSpace launch, the User Support Manager and Systems Manager will be conducting general training sessions for all members of MIT Libraries and in-depth training for designated units of the Libraries.

Building Awareness and Driving Adoption

DSpace has already engendered substantial interest throughout MIT and beyond. Demand for participation in the Early Adopter Program exceeded capacity, with more queuing for early participation in the full roll out. Several individual faculty members have expressed interest in storing a terabyte of data or more, far exceeding our current ability to accommodate data in those amounts. These innovators and early adopters, who make up the early market of technology adopters, are “visionaries” and risk-takers who are eager to try out new innovations that may lead to dramatic improvements and changes (Rogers, 1995). The challenge for DSpace is to span the chasm, “where adoption may falter or even plummet”, between the early market and the mainstream market. (Moore, 1995)

A useful model for crossing the chasm is the *innovation-decision process* (Rogers, 1995, p. 163), “through which an individual (or other decision-making unit) passes from first knowledge of an innovation, to forming an attitude toward the innovation, to a decision to adopt or reject, to implementation of the new idea, and to confirmation of this decision.” The communications plan that has been developed by the User Support Manager, the Libraries’ Communications group, the MIT News Office, and HP’s Public Relations employs communications programs for each step of the decision process targeted at the mainstream market of faculty members, who characteristically will “wait and see” before trying new technology applications and are looking for “compelling value.” (Geoghegan 1994)

The first step is to inform faculty about the capabilities and benefits of DSpace in order to foster a high level of interest among potential Communities. To persuade new Communities to become established in DSpace, the DSpace User Support Manager regularly meets with individual faculty members and research communities, has delivered training and information sessions, and is planning a presentation at the new faculty orientation. Faculty members from the Early Adopter Communities will also be important contributors at this stage as they share information about DSpace with other professors. The official DSpace launch event will target MIT Faculty and include case studies about how the Early Adopter faculty members are using the DSpace system.

We believe that DSpace has a “compelling value” proposition for faculty and Communities that will drive the decision to adopt. Our survey indicates that faculty highly value the

diversity of file formats, preservation, search and retrieval, and the Libraries' management of this type of system on their behalf. Establishing a consistent message about key attributes using a compelling, pragmatic value proposition is critical to engaging the mainstream market. The value proposition for DSpace will be a common component in all communication messages.

Success in the last two stages of the innovation-decision process, implementation and confirmation, rely heavily on a positive user experience with ample opportunity for feedback and support. The DSpace Operations Services described earlier in this report are designed to provide a smooth Community implementation. We plan to develop User Reporting Services that report on how often a Community's submitted content is being accessed as a confirmation of the value of the system. Further, the DSpace Operational Advisory Committee and the User Group that we hope to establish in the future will be a forum for Community members to provide feedback on the service and learn from how others are using the system.

Financial Model

Cost Model

Our cost model was developed to assist MIT Libraries in staff planning, understanding the budget necessary to sustain DSpace and the factors that may significantly affect our projections. In order to estimate costs we consider the changes that the implementation of DSpace has made in staff, support services, space and hardware requirements thus far and how we expect their impact on total costs will change with growth of the system.

We report costs in several categories according to their allocation to the DSpace service. Incremental costs reflect the only costs that will create new expense categories and therefore include only the expenses associated with newly created dedicated staff positions, their direct expenses and system equipment. All Incremental costs represent actual cash flows. Principal costs include the expenses associated with existing staff who will be significantly redirected to ensure the successful implementation of DSpace. Comprehensive costs include expenses for public service staff, and all other expenses such as professional travel that enhance, but are not vitally necessary to providing, the most basic DSpace service. The sum of Incremental, Principal and Comprehensive costs is reported as Total Expense. We report the allocation of Incremental, Principal and Comprehensive costs as applicable in each major category, Staff, Direct Expense, and System Equipment. A breakdown of costs by category is included as Appendix B.

Staff

The Transition Planning Group members tasked with the staffing plan conducted interviews with a broad range of staff members including public service, administrative, development, communications, cataloging, archiving, and Director's office staff. They categorized the impact on individual staff members from minimal to high. We used the results of the

Transition Planning Group's staffing plan as input data for staffing costs, translating the impact levels to staff allocation percentages.

In our model we assume that the two new positions will act as dedicated staff. We also incorporate the expectation that at some point, a dedicated Product Manager will be hired. Costs for these positions, including staff salaries and benefits, office supplies and equipment are included in Incremental costs.

Principal costs include those attributable to the Associate Director for Technology, who will at the outset serve as the Product Manager for DSpace, and the Systems Office staff. These individuals will experience an immediate and substantial change in the day-to-day course of their business because of DSpace. Staff in this category are included in the current MIT Libraries budget but will allocate at least 20% of their time to DSpace. We allocate a portion of the cost of salary and benefits plus any DSpace dedicated travel for these two positions. Our model incorporates the expectation that when the dedicated Product Manager is hired, the allocation of time of the Associate Director for Technology will be decreased.

Comprehensive costs include an allocation for staff whose jobs will be changed by the adoption of DSpace, minimally at the outset but growing over time. Although these individuals must be trained to access this resource, DSpace merely represents another in a suite of electronic resources that the library already provides.

In the interest of understanding the full impact of DSpace on Libraries planning, in our model we recognize the contribution of all affected staff but we ignore costs for those staff whose job functions will not be changed due to the implementation of DSpace. We also recognize that there will be increased demand for Document Services' E-Conversion Services. Those staffing costs are excluded because Document Services operates as a fully cost recovered center, and is expected to continue to do so in the future. Further, the DSpace service begins with the assumption that materials to be deposited are already in digital form.

Direct Expenses

Incremental Direct Expenses include offices supplies and expenses for the dedicated staff, the Information Systems Service Level Agreement fees. Principal Direct Expenses include staff travel, conference fees and other meeting expenses.

System Equipment

We outlined growth scenarios to our development partners at HP and worked with an HP hardware expert and a representative from MIT's IS organization who helped us develop alternatives for scaling the DSpace system. The cost estimates were based on HP hardware costs and are included in Incremental Expenses. Other vendors may have higher or lower costs associated with hardware to meet these same specifications.

MIT Libraries is well positioned to delay any major hardware purchases until the actual growth of the system is better understood. As we gather data on the initial usage of DSpace through summer 2002, the DSpace User Support Manager will be studying the demands on the DSpace system in terms of cumulative content. Based on our current best estimates of typical demand, the hardware donated by HP is more than sufficient to handle the anticipated transaction volume, is well configured for disaster recovery, and can scale to store approximately 3TB of data with the purchase of additional disks. DSpace was architected with scalability in mind. The asset store is decoupled from the application services such that the two can be scaled independently.

Federation Note

Our model reflects costs specific to MIT, where DSpace will be implemented as a full-scale digital repository in an organization with limited excess resources. However, DSpace is a fully scalable system, operable on an individual personal computer and DSpace code will be offered at no charge through an Open Source license. The costs that a federating institution may incur in the implementation and operation of DSpace are a factor of the intended use of the system, current staff availability, and information systems resources. For example DSpace implemented as a pre-print server in a small university or department may add little or no new expense.

Funding

Since the DSpace project began over two years ago, MIT Libraries has been fortunate to have the support of the Andrew W. Mellon Foundation and HP. Grants from these two organizations have provided the initial funds for the Libraries to develop a robust digital repository and plan for its enduring future. The Libraries will make the transition from reliance on external funding sources to a more self-sufficient set of funding sources to support the operational service.

Sustaining the DSpace service will require funding to cover staffing, direct expenses, and equipment. As future demand warrants, MIT Libraries may undergo significant one-time development projects to expand DSpace's services to meet Institute specific needs, necessitating one-time funding.

A primary goal of DSpace is to encourage widespread adoption of the system among faculty and researchers who submit their digital objects. In support of this, DSpace will provide its Core Services free of charge to all registered Community members. MIT Libraries is committed to the support of MIT's mission to "generate, disseminate and preserve knowledge", therefore DSpace content, with the exception of restricted items, will be provided free of charge to the public. This necessitates that we find sources other than user or subscription fees to support the ongoing operations of the DSpace service.

We propose a funding model for MIT Libraries that is supported by a number of resources in the form of financial support or in-kind assistance. During the first year, the operational plan

focuses on stabilizing the system and building the Community base and content. It is the intention of MIT Libraries to seek support from MIT in the Libraries' annual operating budget as an extension of the Libraries' traditional role of capturing the scholarly record of the Institute and an alternative to the current disparate and sometimes inadequate means to manage its digital assets.

Institutional economies are served by developing a robust managed repository in the Libraries with their well-developed expertise in collection management, cataloging and digital asset access. The DSpace service will provide a less expensive and higher quality alternative to multiple systems of servers maintained in departments and research centers across campus.

MIT Libraries have made every effort to control costs during the implementation of DSpace. The system is completely scalable, allowing it to grow as demand warrants. The total annual cost of operating DSpace, with plans to scale content to 3 TB in three years, currently represents less than 2% of the Libraries annual budget.

In-kind contribution in the form of new development for DSpace is expected to come from several sources. HP, along with MIT Libraries, the MIT Lab for Computer Science and the World Wide Web Consortium, developed a three-year plan for further research using the DSpace platform and content. Funding for this project would come from invent@MIT, the HP-MIT Alliance, and applicable research results will be migrated into the operational DSpace system. MIT Libraries also intends to develop a federation of institutions such as university libraries, cultural heritage institutions, and government agencies who use DSpace. Collaborative development among the DSpace federation participants will be an important source of enhancements for the system. User identified needs addressed by developments made by federation members (e.g. in the areas of digital collection management, preservation, cross-institute system interoperability and scholarly communications) will become part of the basic DSpace system allowing the system to advance far more rapidly than would be possible in a single instantiation. With funding from the Cambridge-MIT Institute, a cooperative venture between Cambridge University and MIT supported by the British government, the first federated implementation of DSpace will begin in Fall 2002 at Cambridge University. The project will focus on developing interoperability, supporting digital course materials, and using the digital content of DSpace as a test bed for digital preservation techniques.

Alternative Funding Model

The main purpose of Premium Services is not the generation of funds but rather to allow MIT Libraries to offer as complete and valuable a service as possible to users while controlling the impact on the Libraries' scarce resources. However, fees for services may be required to augment other funding sources and provide for the Core Service offering. Those services for which MIT Libraries are the sole source provider, such as Custom Repository Services, or clearly can establish added value over other competitive services offer the best opportunity for system cost recovery by the Libraries. Further definition of the Premium Services and market validation of the demand will be conducted starting in the Fall 2002, when DSpace becomes fully operational campus-wide.

Further Research

MIT Business Plan

Usage

User activities will provide MIT Libraries with an opportunity to learn more about actual use of DSpace including data about Community set-up time, rate and growth of submissions, and types of file formats submitted. The development team built several tools for monitoring system use including DSpace-specific web session tracking and system troubleshooting. The DSpace team will study reports from these tools, along with utilization analyses and standard web server and database logs, for internal planning purposes and will also share the statistics with Communities upon request. We designed the cost model for DSpace as a flexible management tool so that information can be incorporated to maintain accurate and useful projections.

Premium Services

The Premium Services are currently in the conceptual stage, born from feedback from Early Adopters and potential DSpace Communities, and require market validation to determine their extensibility to other DSpace Communities. During the remainder of the Early Adopter program and into the initial launch, The Associate Director for Technology, in her role as DSpace Product Manager, will monitor user demand and introduce new services as needed. As DSpace is federated there may be opportunities for the MIT based Premium Services to reach a larger audience and or expand in scope. Just as some MIT Communities may choose to outsource activities, Communities that develop at other universities may do the same.

Further policy decisions

DSpace policies are necessarily still evolving and will continue throughout the life of the service. A sub-set of the DSpace Transition Planning group served as an interim policy advisory group and then proposed the more permanent DSpace Policy Committee to make decisions related to the DSpace services, standards and functionality. Future policy decisions may have business implications that will need to be reflected in the financial model.

Digital Preservation

File Formats

The scale and scope of the digital items in DSpace will provide a rich test bed for conducting research on digital preservation. MIT Libraries' policy is currently to accept digital objects in all formats, including individualized formats. The policy further defines preservation service levels, categorizing file formats according to whether they are known and supported, known and unsupported, or unknown and unsupported. For "unsupported" file formats, DSpace merely guarantees to preserve and return the bits, while for "supported" formats DSpace makes

a further promise to maintain the usability of the content in its original context. Going forward, the DSpace team proposes to maintain a registry, collaboratively developed with the Digital Library Federation, documenting format specifications to support preservation and future digital archeology. The aim will be to develop digital preservation best practices that will increasingly enable DSpace to maintain the content in an immediately accessible form.

Preservation metadata

As important as preserving the digital item itself is the recording of historical information about the technology and strategies used to preserve it, information known as preservation metadata. In the same research vein as that planned for file formats, MIT Libraries intends to use the content of the DSpace system to explore standards for preservation metadata schema.

Publishing

DSpace's distributed Community structure provides an ideal platform for "e-communities" or faculty-initiated journals. MIT Libraries are discussing a relationship with the Electronic Publishing Initiative at Columbia (EPIC) University to develop publishing tools that would be offered through the DSpace service in support of new publishing ventures.

Federation

Interoperability

A federated DSpace system will allow research into the benefits of network effects, virtual journals with editors at multiple universities, distributed costs of format migration, cross-institutional searches, etc.

Policy and Standards Development

A collective group of institutions concerned with digital preservation and dissemination will likely have a stronger impact on policy and standards development and adoption than individual institutions. The DSpace federation will draw on the expertise of its membership to examine such areas as copyright policy, intellectual property rights management, and interoperability with Learning Management Systems.

Governance

As part of the formation of the DSpace Federation, MIT Libraries will be exploring possible governance structures. The governance structure will provide guidance on such areas as integration of new system functionality, policy and standards development, and cross-institute collection management and services.

Lessons Learned

DSpace represents the cornerstone project in the MIT Libraries' ongoing applied research agenda and the successes from this project will serve as best practices for future endeavors. Among the many learning opportunities that the business plan project provided, perhaps the most important were those related to organizational change.

The importance of dedicated business and operational planning conducted in parallel with the research and development process is paramount. Through the work of the Transition Planning group, the Early Adopter Librarians, and our business planning team, we were able to anticipate organizational issues, align development plans with resource realities and communicate findings to the staff as they developed.

A second finding was that in a dynamic organization like the MIT Libraries, it was important to communicate the development of the DSpace project and vision with the staff. Early and accurate communication served to generate awareness and acceptance within not only the Library staff but also in faculty and researchers for whom the public service staff represent the point of contact with the Libraries.

Cross-functional teams, which included members from across the Libraries and Information Systems organizations, were involved in the business planning process. Members of the MIT Libraries staff from all departments participated in working groups such as the Transition Planning group and the Early Adopter Librarians group. Their knowledge and experience proved invaluable in analyzing staffing implications, creating a representative management structure, developing a communication plan, and envisioning a future set of value-added DSpace services. Their participation will continue to play a key role in integrating DSpace into the Libraries' daily operations.

Not surprisingly, strong leadership and expertise from the Director of Libraries, the Associate Director for Technology, the Transition Planning group, and the DSpace Project Leader were also key components of not only the business planning project, but also the DSpace development. Pivotal decisions about system policies and guidance on the strategic direction of the service were required throughout the development of an innovative new system such as DSpace. They provided a source of communication to address staff concerns about the impact of DSpace on Libraries' resources and provided a forum for receiving feedback on the proposed staffing plan. Their advocacy also has contributed significantly to the enthusiasm and interest in DSpace generated throughout the MIT campus and beyond.

Through the implementation of a system that alters traditional library processes we also learned a tremendous amount about managing project uncertainty. De Meyer et al. (2002) assert "the challenge in managing uncertainty, to whatever degree, is to find the balance between planning and learning. Planning provides discipline and a concrete set of activities and contingencies that can be codified, communicated and monitored. (Eppinger, 1997) Learning permits adapting to unforeseen or chaotic events." A primary example of this came from input we received from faculty, Libraries staff, and other universities regarding DSpace's applicability

to a wide variety of issues in the academic and scholarly communication environments. We therefore gave careful consideration to how we manage these opportunities and how we define DSpace's function in the digital realm, planning which helped us prioritize DSpace's foremost function as an operational service for faculty. All other aspects of the DSpace domain will be in support of this goal, including DSpace's role as a research platform for digital library issues; its extensibility as an educational, publishing or administrative platform; and its ability to encourage new ways of intra- and inter-institutional cooperation and collaboration.

Fortunately, the uncertainty that we encountered was a positive aspect of the project because it was a reflection of our growing understanding of DSpace's users and the exciting problems that DSpace can help address. The Transition Planning Group, the Early Adopter Librarians, and other experts in the Libraries helped us anticipate and plan for future opportunities. We also incorporated the management of project uncertainty into the DSpace business plan by designing a flexible cost model adaptable to the feedback obtained through the use of the system by the entire campus.

A valuable by-product of the integration of research projects in the Libraries' agenda is professional development for Libraries' staff, which can have a broad impact on traditional library services. For example, the lessons in the management of uncertainty were shared with the Transition Planning Group and the Libraries' Document Services department. The development of new metadata schemas and tools in the Bibliographic Access Services group will be applicable in print collections as well. Demand for new metadata development with the adoption of DSpace in new disciplines will continue to drive new knowledge in the department.

DSpace is a significant extension of the Libraries' traditional role. By addressing challenges libraries face in the digital realm, MIT has the opportunity to influence the future of digital preservation and scholarly communication both of which have the potential to have far-reaching effects on the academic environment. MIT Libraries' applied research approach to library services has further established the Libraries as an active participant in faculty members' research processes.

References

De Meyer, Arnoud, Cristoph H. Loch, and Michael T. Pich. "Managing Project Uncertainty: From Variation to Chaos." *MIT Sloan Management Review* 43 (2002): 60-67.

Eppinger, S.D. and R.P. Smith, "A Predictive Model of Sequential Iteration in Engineering Design," *Management Science* 43 (1997): 1, 104-1, 120; and J. Mihm, C.H. Loch and A. Huchzermeier, "Modelling the Problem Solving Dynamics in Complex Engineering Projects," working paper 2001/48-TN, INSEAD, Fountainbleau, France, 2001

Fleishauer, Carol, Mary Barton, Margret Branschofsky, Dan Chudnov, Catherine Friedman, Keith Glavash, Carl Jones, Jeff Merriman, Sarah Mitchell, MacKenzie Smith, Megan Sniffin-Marinoff, Julie Harford Walker. "Report of the DSpace Transition Planning Group to the MIT Libraries Steering Committee." Cambridge, MA., 19 April 2002.

Geoghegan, William H. 1994. *What ever happened to instructional technology?* [online]. International Business Schools Computing Association. Available from World Wide Web: (<http://franklin.scale.uiuc.edu/scale/links/library/geoghegan/wpi.html>)

Greenstein, Daniel and Leigh Watson Healy. "National Survey Documents Effects of Internet Use on Libraries." *CLIR Issues* 27 (2002). Available from World Wide Web: (<http://www.clir.org/pubs/issues/issues27.html - national>)

Hedstrom, Margaret. *Digital preservation: a time bomb for Digital Libraries*. Available from World Wide Web: (<http://www.uky.edu/~kiernan/DL/hedstrom.html>)

Moore, G.A. 1995. *Inside the Tornado*. New York: HarperCollins Publishers, Inc.

Rogers, E.M. 1995. *Diffusion of Innovations*. (4th ed.). New York: Free Press.

"Dublin Core Metadata Initiative Overview." Available from World Wide Web: <http://dublincore.org/about/>