DSPACE: DURABLE DIGITAL DOCUMENTS

Daniel Chudnov

DSpace, a joint MIT/Hewlett-Packard software development, will provide a long-term digital storage and preservation facility. The development is community based, indexed using qualified Dublin Core elements and its architecture uses the OAIS Reference Model. The possibility of federated services and dynamic links to other projects are also being investigated.

*Daniel Chudnov, Systems Curator, The DSpace Project, Massachusetts Institute of Technology Libraries, Building 10-500, 77 Massachusetts Avenue, Cambridge, MA 02139-4307 USA
E-mail: dchud@mit.edu*

DSpace is a joint development project of the Massachusetts Institute of Technology Libraries and Hewlett-Packard Company. Its mission is to establish a library service to capture, distribute, and preserve the digital, intellectual output of the MIT community. We are developing a software platform for long-term digital content storage and preservation, and implementing this platform as a service of the Libraries.

The entire MIT community, from individuals to departments and schools, will be able to use DSpace to distribute and preserve digital forms of research, in fixed versions ready to share publicly ("fixed", rather than many continuously updated drafts during creation, which will remain the researchers' responsibility). Such versions might include a paper shared as a pre-print or post-publication, a milestone software release, or a video of a public artistic performance. To accommodate the diversity of research at MIT, DSpace will store and preserve many file types, including text-based, software code, datasets, and binary formats such as images and multimedia audio/video. Among these we will support many standard file types with appropriate value-added services (such as, data refreshing or migration), where specifications for those types are stable and broadly accessible.

DSpace content will be organized into communities and collections, with different groups empowered to manage user privileges, item ingest workflows, copyright, controlled vocabularies, and access controls within the collections of their community. A given community may also choose to organize collections in various ways, such as specific themes of research, groups of researchers, content types, or historical periods. Independent of how specific communities choose to organize collections, all items stored in DSpace will receive a unique and persistent URL, enabling researchers to trust that research stored

in DSpace will be continuously accessible from a standard, stable location. DSpace will also support use restrictions such as MIT-only, dated availability (an interval after print publication, for instance), or freely available to the world.

In addition to preserving type-specific metadata formats such as MARC for books and technical reports, or FGDC for spatial datasets, all items stored in DSpace will be indexed using qualified Dublin Core elements. This will provide standardized description across content types and genres for searching, browsing, and metadata harvesting from DSpace to other systems. We will publish our Dublin Core qualifiers, which are influenced by crosswalks published by the U.S. Library of Congress and the developing Library Application Profile of the Dublin Core Libraries group.

The DSpace software architecture is deeply informed by the Open Archival Information System (OAIS) Reference Model. The development team, consisting of HP technical staff, consultants, and MIT librarians, is designing DSpace to address the long-term preservation needs specified by the OAIS. These needs include well-defined models for information packages, access services and ongoing administration.

DSpace is guided by a Steering Committee and Technical and Faculty Advisory Boards, composed of senior research, technical, and administrative staff from HP and MIT. In addition to the HP-funded development project, a concurrent project is funded by the Andrew Mellon Foundation to research and prepare a DSpace Business Plan. This effort is led by marketing and finance staff of the MIT Libraries, hired using Mellon grant funds, who coordinate closely with DSpace development efforts. We are all working together to plan the many roles and tasks needed for the Libraries' long-term management of the DSpace service, including the many necessary community interactions.

The DSpace Business Plan will also investigate possibilities for federating services or content between many institutions, using locally managed instances of the DSpace platform. MIT Libraries' staff, funded under a separate planning grant from Mellon to investigate a dynamic e-journal archiving project, are also working with DSpace to consider potential technical and service relationships between the projects.

The DSpace development project is funded by HP for two years, with an early adopter development release for select MIT programs beginning in fall 2001, MIT-wide public beta in February 2002, and full implementation by summer 2002. DSpace will be publicly released using a BSD-style open source software license in late 2001 or early 2002.

For more information please visit: dspace.org.

———