The Digital Archive @ MIT Project*

2000.03.03

Deliverables

(D1) Digital Repository
(D2) Submission & Retrieval
(D3) Searching & Sharing
(D4) Access Control
(D5) Rights Management
(D6) Community
(D7) Publishing

Research Areas

(R1) Metadata
(R2) Rights Management
(R3) Bulk Submission
(R4) Continuity

Planning (3/00-5/00)
"getting in position"
- survey relevant work
- user demand study
- gather the team
- basic architectural decisions
- identify areas for discovery

Phase 1 (6/00-12/00)
"get the blood pumping"
- develop repository
- submission from select few
- access from MIT/HP only
- open archives metadata
- address metadata standards
- identify widely useful formats
- develop hooks for searching
- develop access control
- develop continuity policy

Phase 2 (1/01-8/01)
"fill in core functions"
- develop submission "realms"
- available outside MIT/HP
- rights management framework
- publishing/printing capabilities
- metadata to RDF standards
- develop metadata "realms"
- make searching available

As more and more of our intellectual heritage finds its way into electronic form, libraries must take responsibility for capturing those documents that will form the foundation of tomorrow's scholarship. MIT hosts an astounding array of intellectual talent and the MIT Libraries wants to make sure that the contributions they make to their fields today is not lost tomorrow. Developing a Digital Archive of the electronic output of our talented faculty, students, and researchers is a critical extension of the MIT Libraries' role.

This project aims to develop a scalable Digital Archive with storage, submission, retrieval, searching, access control, rights management, and publishing capabilities. This Digital Archive will be designed in such a way that the underlying software, data models, and methods can be shared freely with other academic institutions and the resulting "distributed" Digital Archive could be searched and accessed as a single entity. The intention of this project is to develop this Digital Archive first at MIT, but then to roll out the project in such a way that it could be adopted by research universities worldwide. While this project focuses on the needs of the MIT community, we will strive early on to invite participants from beyond MIT so that we evolve a Digital Archive which can be replicated.

The Digital Archive @ MIT is a cooperative project of two organizations: the MIT Libraries and HP Labs. These two organizations will also depend on close working relationships with others such as the Worldwide Web Consortium (W3C) and the HP Labs Research Libraries. Other participating entities will include the MIT Lab for Computer Science, MIT Academic Computing, HP LaserJet Systems Group, HP Inkjet Solutions Group, and HP Labs University program.

The MIT Libraries seek a robust, maintainable, and above-all useful repository. The system must be scalable and provide a bona fide foundation for the collection of digital material from around the institute. Service rather than research is the primary focus for the involvement of the MIT Libraries in this project.

Even if the MIT Libraries and HP Labs do not realize the outcomes they desire, the project will still "succeed" if we commit sufficient effort to learn lessons from the process. We expect to be surprised by this project and aim to "learn by doing."

Eric Celeste, MIT Libraries, efc@mit.edu
William Wickes, HP Labs, william_wickes@hp.com

*looking for a real name, how about: Digital Document Repository, Haven, cPort, dSpace