Digital Asset Management System (DAMS) Infrastructure: A Collaborative Metadata Pilot

Digital Libraries Federation
Spring Forum 2004

Yong-Mi Kim
Judy Ahronheim
University of Michigan

Presentation Overview

- Background
- DAMS at the University of Michigan
- DAMS architecture
- DAMS metadata issues
- Future directions

Vision

- Create a robust infrastructure to ingest, manage, store and publish digital rich-media assets and their associated metadata.
- Streamline the “workflow” required to create new works with digital rich-media assets.
- Build an environment where assets are easily searched, shared, edited and repurposed in the academic model.
- Provide a campus-wide platform for future application of rights declaration techniques (or other IP tools) to existing assets.

Background

2000 – Summer – UMTV looking for video archive solution

2000 – Fall – UM team visit to CNN. Eureka! – the “indoor plumbing” of rich media asset management.

2001 – Spring – Digital Asset Management defined as a possible university infrastructure (PIRC report influence).

2001 – July – Limited proof of concept authorized.

2001 – December - Proof of concept complete.

2001 – Winter – Informix purchased by IBM. Ascential fate uncertain.

2002 – Spring-summer: Cross-campus team develops RFP.

2002 – July – RFP issued for DAMS Living Lab.

2002 – Fall – Five finalists narrowed to three vendor teams: Bearing-Point (KPMG Consulting)+Documentum; Sun+Artesia; IBM+Ancept. Site demos and Web tests continue to March, 2003.

2003 – March – IBM+Ancept selected for final negotiations by 8-unit team.

2003 – Spring – Candidates for Affiliate partnerships sought.

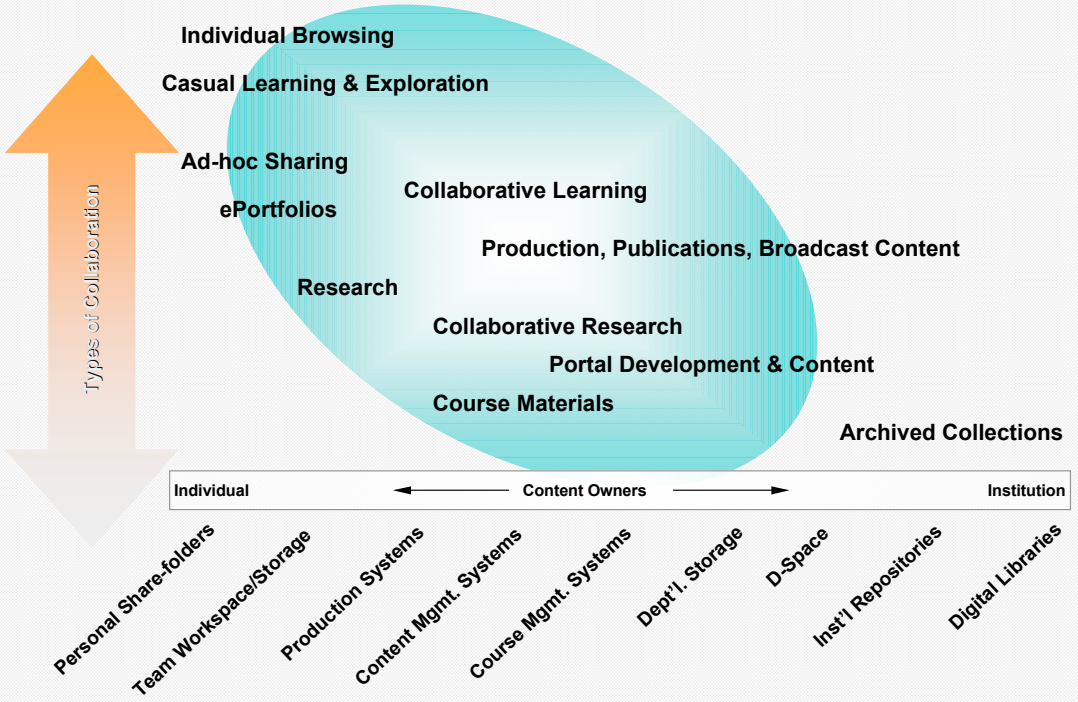
2003 – August – Hardware and software install begins in Living Lab.

Participants

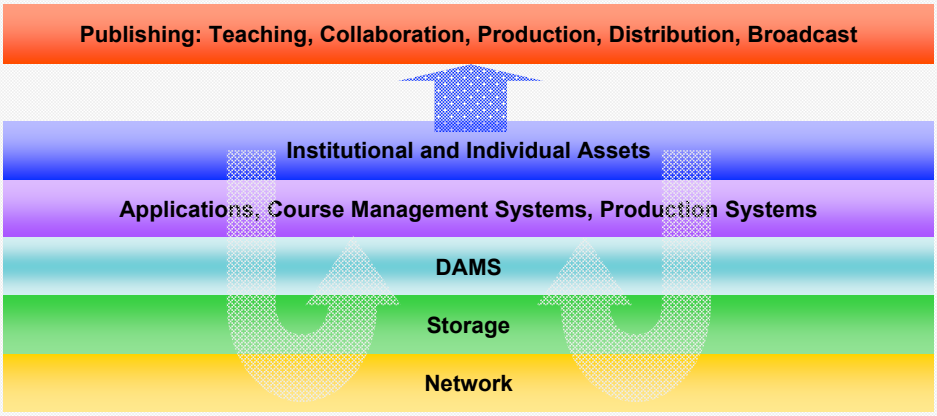
School of Education
School of Dentistry
College of Literature, Science and the Arts
School of Nursing
College of Pharmacy
School of Social Work
Information Technology Central Services
University Libraries

Business School
School of Information
Media Union

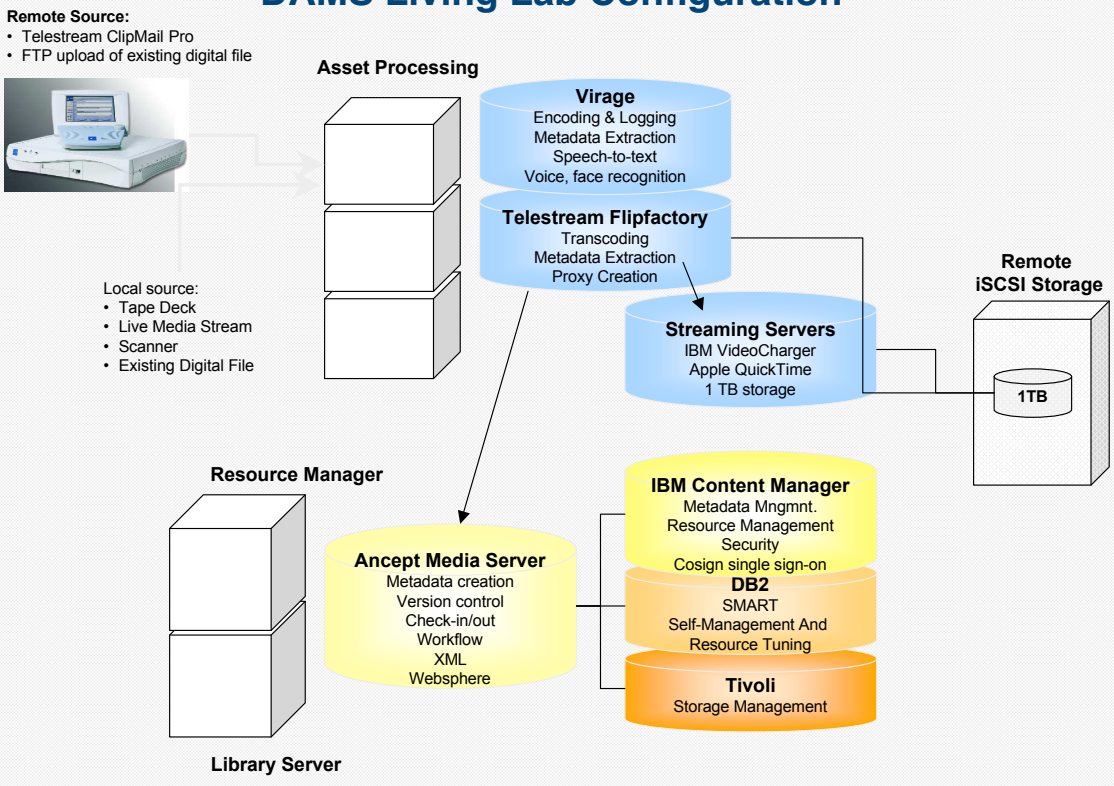
What space does DAMS occupy?



What is the place of DAMS in the campus infrastructure?



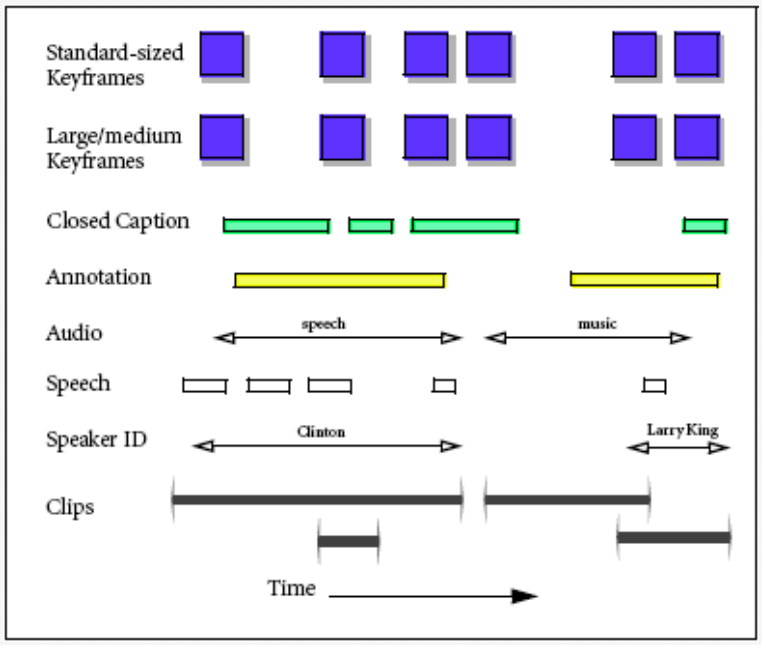
DAMS Living Lab Configuration



Workflow

- Images
  - Ancept Media System (AMS)
- Video
  - FlipFactory
    create files in various compression formats
  - VideoLogger
    automatic metadata generation from files
  - AMS
    descriptive metadata entered

VideoLogger metadata



VideoLogger

- Speech recognition
  - Generates text from speech
    time-stamped stream of text from dictionary words
  - Uses base dictionary
  - User-defined speech recognition dictionary
    -> issue - dictionary construction
    select dictionary terms
    record dictionary terms read aloud

Legacy metadata

- Metadata needs of four partnership units surveyed:
  - LS&A - Language Resource Center
  - Education - Mathematics Teaching and Learning to Teach project
  - Pharmacy
  - Dentistry

Legacy metadata status

- LS&A - legacy metadata available online; can be exported with some modification
- Education - legacy metadata exists but is scattered; structure does not allow easy search and retrieval
- Pharmacy - metadata needs to be extracted or generated from objects
- Dentistry - some legacy metadata available

DAMS metadata issues

- **Heterogeneous** collections with domain-specific description needs, vocabularies
- Metadata entry will be handled by non-specialists
- Granularity of description
  - Enable search across subject areas
  - Provenance

Metadata in DAMS

- Descriptive metadata
  - Dublin Core
    - General, simple metadata standard intended for use by non-specialists
  - IMS
    - Metadata standard focusing on educational uses
- System metadata
  - File size, file name, file format, etc.

UM_Core

- Minimal level of description for DAMS assets
  - Elements from Dublin Core
  - Local elements

Search across collections done through UM_Core elements

UM_Core elements

- **DC_Title**
- **DC_Creator**
- DC_Subject
- UM_SecondarySubject
- DC_Description
- **DC_Publisher**
- DC_Contributor
- **DC_Date**
- DC_Type
- **DC_Format**
- **DC_Identifier**
- DC_Source
- DC_Language
- DC_Relation
- DC_Coverage
- DC_Rights
- UM_Publisher
- UM_AlternatePublisher

Dentistry

- Assets for pilot
  - 15 most commonly requested videos of dental procedures
  - 300 diagnostic quality oral histology images
- High-level browse categories specific for dentistry defined
- Metadata schema for videos defined
- Issues
  - Controlled vocabularies (e.g., SNODENT)
  - Incorporating other schema elements (e.g., DICOM)
  - Permissions and privacy (HIPAA)

Education

- Assets for pilot
  - Videos of classroom interaction
  - Associated transcripts
  - Presentations and papers based on above
- Issues
  - Explore using IMS, GEM
  - Educating asset producers (faculty)
  - Permissions and privacy

Long term issues

- Integration with other academic tools, portals
- Relationship to Library, Institutional Repository, federated catalog searching
- Intellectual property, copyright, use and misuse policy
- Privacy, access control

Additional information

http://sitemaker.umich.edu/dams