

Preserving eScholarship and Digitized Special Collections

Distributed Digital Preservation

Bill Donovan
donovawf@bc.edu

Summary

As stewards of eScholarship and digitized special collections, **we are responsible** for saving these and other treasures **effectively and economically**.

One approach for digital preservation is being spearheaded by the **MetaArchive Cooperative**; collections are replicated by peer institutions to guard against loss. The MetaArchive approach is **one model** for cultural memory organizations to consider adopting/adapting for their own use.

Rationale for this talk

- Not recruiting for MetaArchive Cooperative
- DDP = a work in progress
- Just one approach, but promising
 - Adaptable for other “CMO” consortia?
 - Cultural memory organizations (CMOs)
- Perspective of just one member
- Ulterior motive: convince management

eScholarship@BC

eScholarship@BC

Boston College Libraries
Log On

[Digital Collections](#) [Search](#) [Results](#) [Previous Searches](#) [Settings](#) [Basket](#) [Help](#)

Guest

[Simple Search](#) [Advanced Search](#)

Select collection: eScholarship@BC

GO

A word or phrase:

☒ Contains ☐ Exact ☐ Starts With

Browse eScholarship@BC

[Graduate Theses and Dissertations](#) (204)

[Carroll School of Management](#) , [Connell School of Nursing](#) , [Graduate School of Arts & Sciences](#) ...

[Undergraduate Honors Theses](#) (263)

[College Honors Program](#) , [Economics Honors Program](#) , [History Honors Program](#) ...

[Research Unit, Center, or Department](#) (924)

[Center for Retirement Research](#) , [The Church in the 21st Century](#) , [Economics Department](#) ...

[Faculty Publications](#) (1552)

[Arts & Sciences](#) , [Carroll School of Management](#) , [Connell School of Nursing](#) ...

[Dept. of Business Law, Carroll School of Management](#) (94)

[History of Boston College](#) (25)

[GSAS](#) , [Occasional papers](#) , [Women](#)

© 2007-2009 Boston College Libraries

Special Collections

Bobbie Hanvey photographic archives

Digital Collections | Boston College Libraries

THE COLLECTION SERIES LISTING BOBBIE HANVEY USING THE COLLECTION ABOUT THE PROJECT



About the Bobbie Hanvey Photographic Archives

Bobbie Hanvey is an award winning photographer living and working in Northern Ireland. The John J. Burns Library acquired the first part of the Bobbie Hanvey Photographic Archives in 2001, with major additions to the collection made in 2008 and 2009, bringing the total number of negatives in the Hanvey Archives to more than 50,000.

Hanvey's photographs comprise a comprehensive documentation of people and life in the North of Ireland beginning in the 1970's through circa 2007. Some photographs feature subjects photographed in various other locations in Ireland. The collection contains portraits, candid images (including weddings and other social events) as well as journalistic images covering public, paramilitary and political activity.


Boston College Libraries, in order to create access for scholars and any interested parties, has initiated a project to digitize, describe, and host the photographs.

The collection is arranged in 13 series, four of which have been digitized thus far. Descriptions and links to each series in the Libraries' online digital repository are found under [Series Listing](#).

Selections from the Collection

Images shown are selections from the [four series](#) (out of 13) digitized thus far and available via the web.

To view the slideshow full screen, click the expand icon 

To view selections from a particular series, click the gallery icon 

[Browse a selection of the John J. Burns Library's photos on Flickr](#)

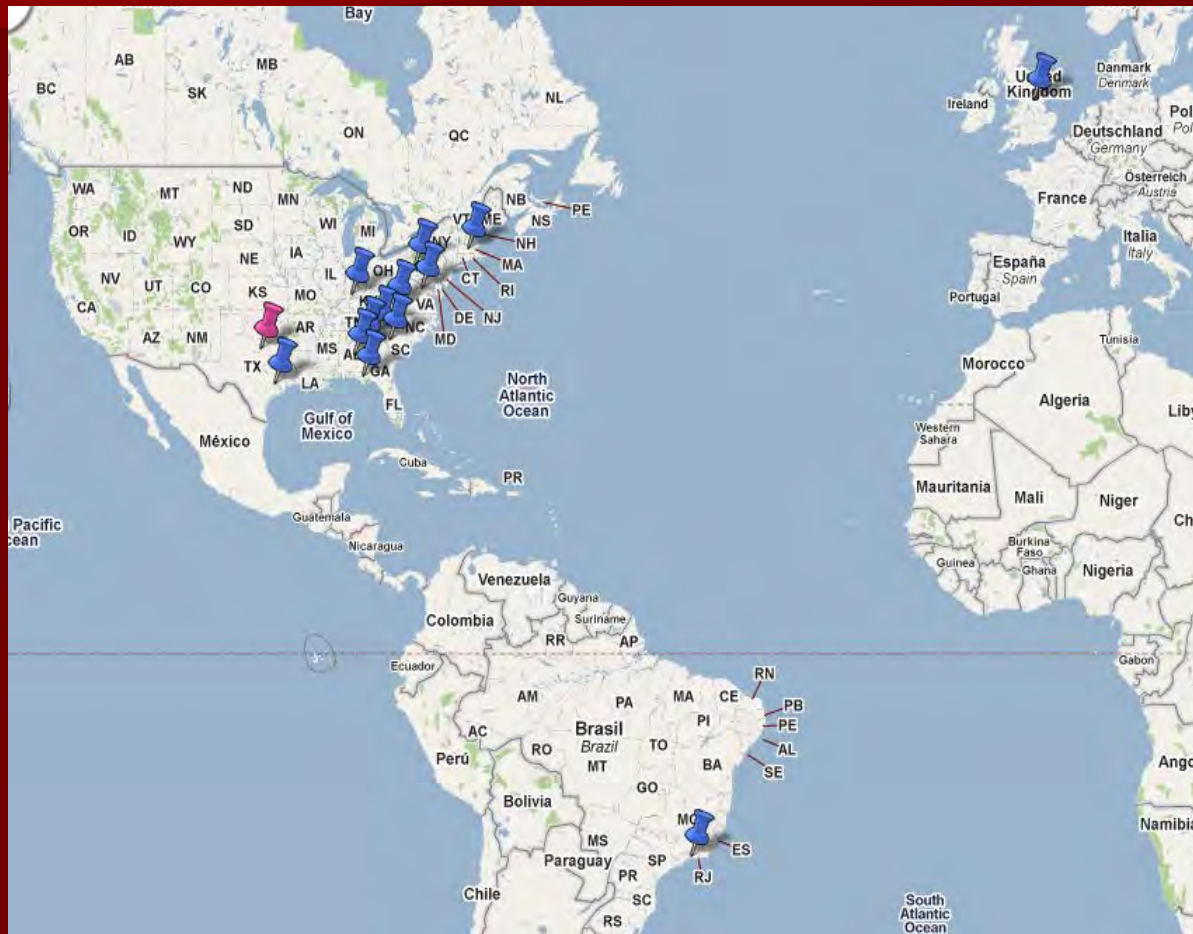


"Digital Preservation" defined

- "Digital preservation" combines **policies**, **strategies** and **actions** that ensure access to digital content over time.
- <http://www.ala.org/ala/mgrps/divs/alcts/resources/preserv/defdigpres0408.cfm>

Distributed Digital Preservation (DDP)

geographically dispersed sites



Participating Institutions:

Auburn University
Boston College
Clemson University
Emory University
Florida State University
Folger Shakespeare Library
Georgia Tech
Library of Congress
Penn State University
Pontificia Universidade Católica
Rice University
University of Hull
University of Louisville
University of North Texas
University of South Carolina
Virginia Tech

"MetaArchive Cooperative"?

- low-cost, high-impact DDP for "CMOs"
 - e.g. libraries, research centers, and museums
- founded in 2004; funding from:
 - NDIIPP (Library of Congress)
 - NHPRC (National Archives)
- **Not** vendor-based; enable CMOs to own and control the process of digital preservation for themselves.

MetaArchives's networks



1. **The Southern Digital Culture Network** was established in coordination with the Library of Congress in 2004. In this network, we collaborate to preserve a wide range of digital content, including music, video, images, photographs, artwork, text documents, websites, e-journals, and other digital files, all of which relate to the history and culture of the American South.
2. **The Electronic Theses and Dissertations Network**, a collaborative endeavor of the Networked Digital Library of Theses and Dissertations and the MetaArchive Cooperative, preserves the electronic theses and dissertations of its member institutions.
3. **The Transatlantic Slave Trade Network** preserves an international collection of maps, ship logs, manifests, images and other materials relating to slave trade voyages across the Atlantic Ocean.

MetaArchive's ETD network





Policies & Strategy --- 1

- Flat, Trim, Tight-Knit organization
 - P2P: no supermember, no host institution
 - Minimal overhead, bureaucracy
 - Emphasis on communication & collaboration
 - Committees: steering, technical, content, preservation
- Self-sufficiency
 - avoid outsourcing; retain control
 - cost containment, understand & refine process
 - sustainable sources of funding

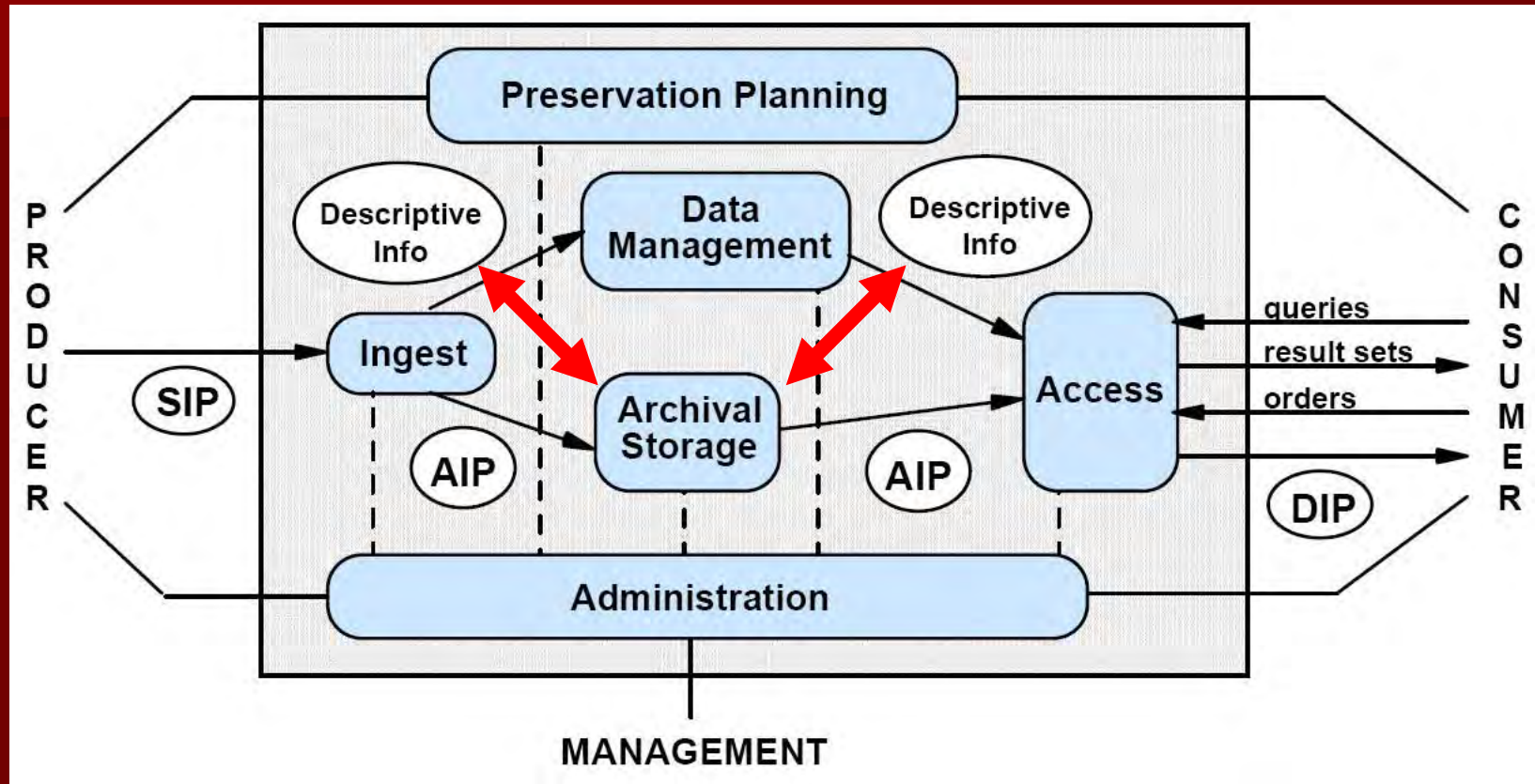
Policies & Strategy --- 2

- Caches (dark archives)
 - 6 replications
 - Access only via contributing member
- Active monitoring of the integrity of stored digital content --- NOT just back-ups
- For ETDs, discovery via Networked Digital Library of Theses & Dissertations, [NDLTD](#)

Local actions/responsibilities

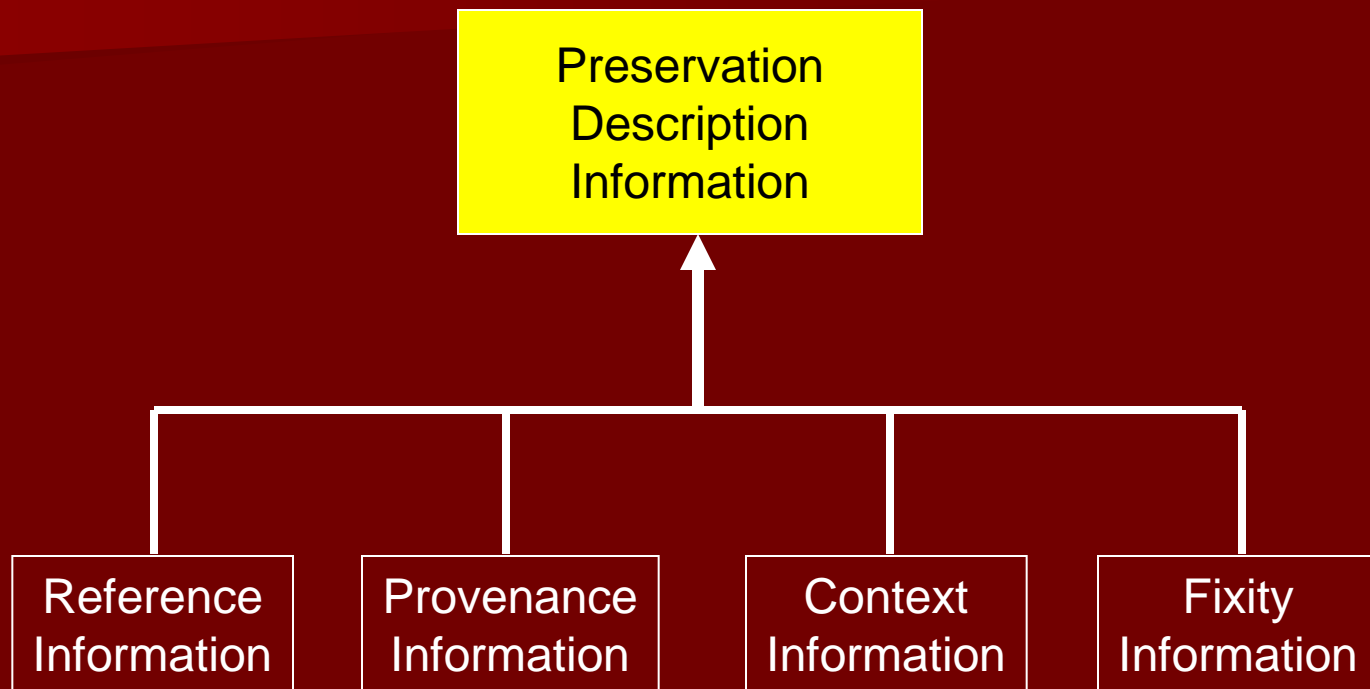
- Skills & infrastructure
- Copyright responsibility
- Data wrangling
 - Format choices
 - Proprietary versus open formats
 - Bit preservation versus migration
 - Filenaming & directories
- Preservation information (OAIS)

OAIS = Open Archival Information System

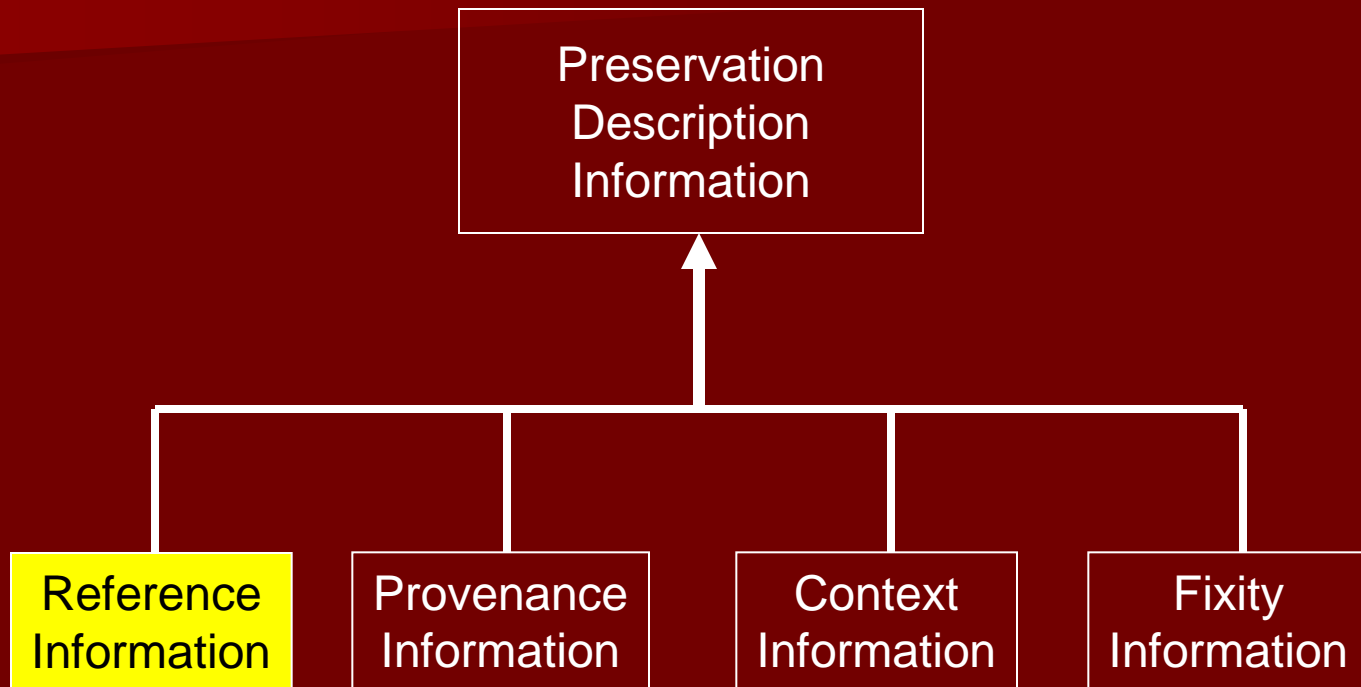


Adapted from: "Reference Model for an Open Archival Information System" CCSDS 650.0-B-1 (2002)

OAIS preservation information

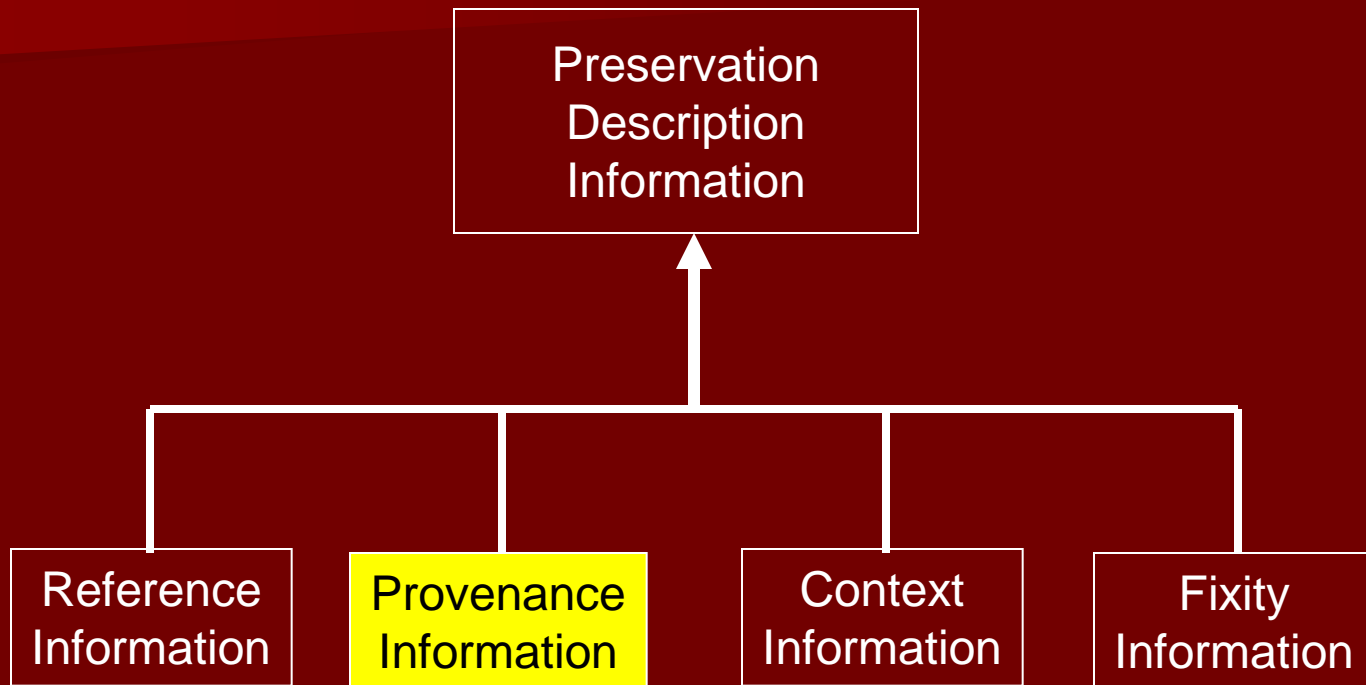


OAIS preservation information



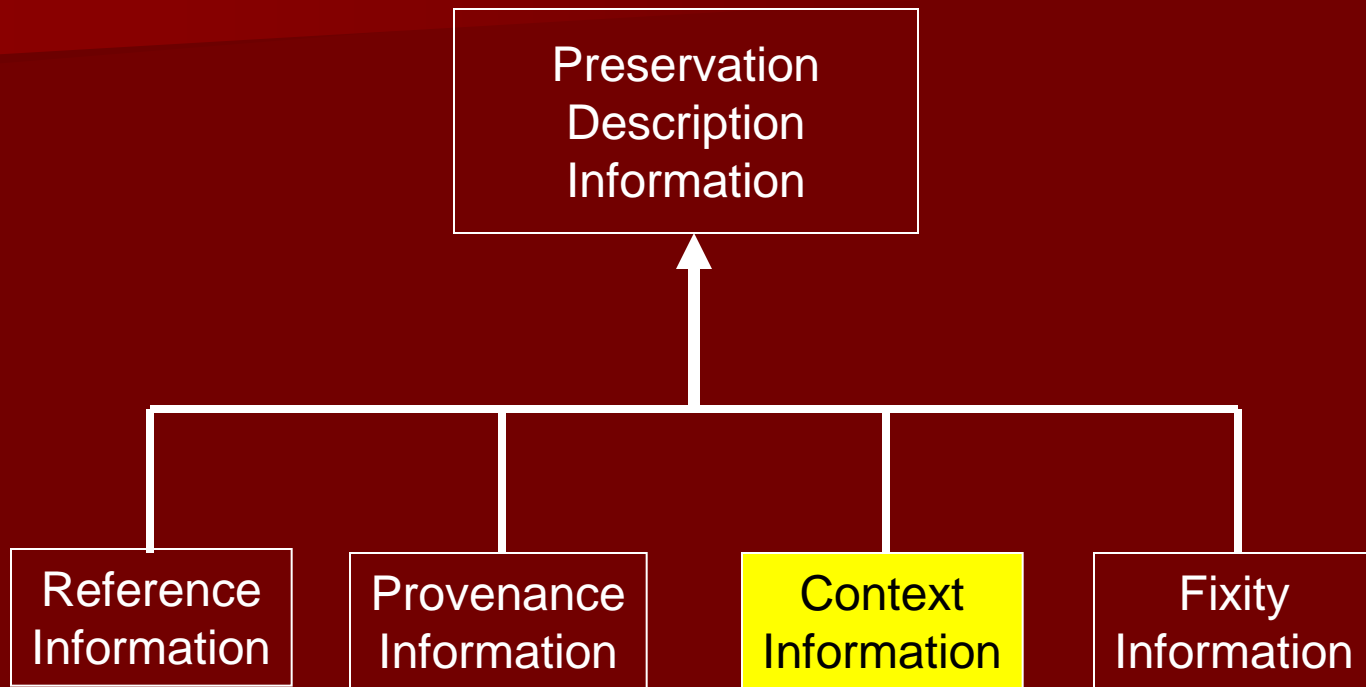
... identifies, and if necessary describes, one or more mechanisms used to provide assigned identifiers for the Content Information. It also provides identifiers that allow outside systems to refer, unambiguously, to a particular Content Information. An example of Reference Information is an ISBN.

OAIS preservation information



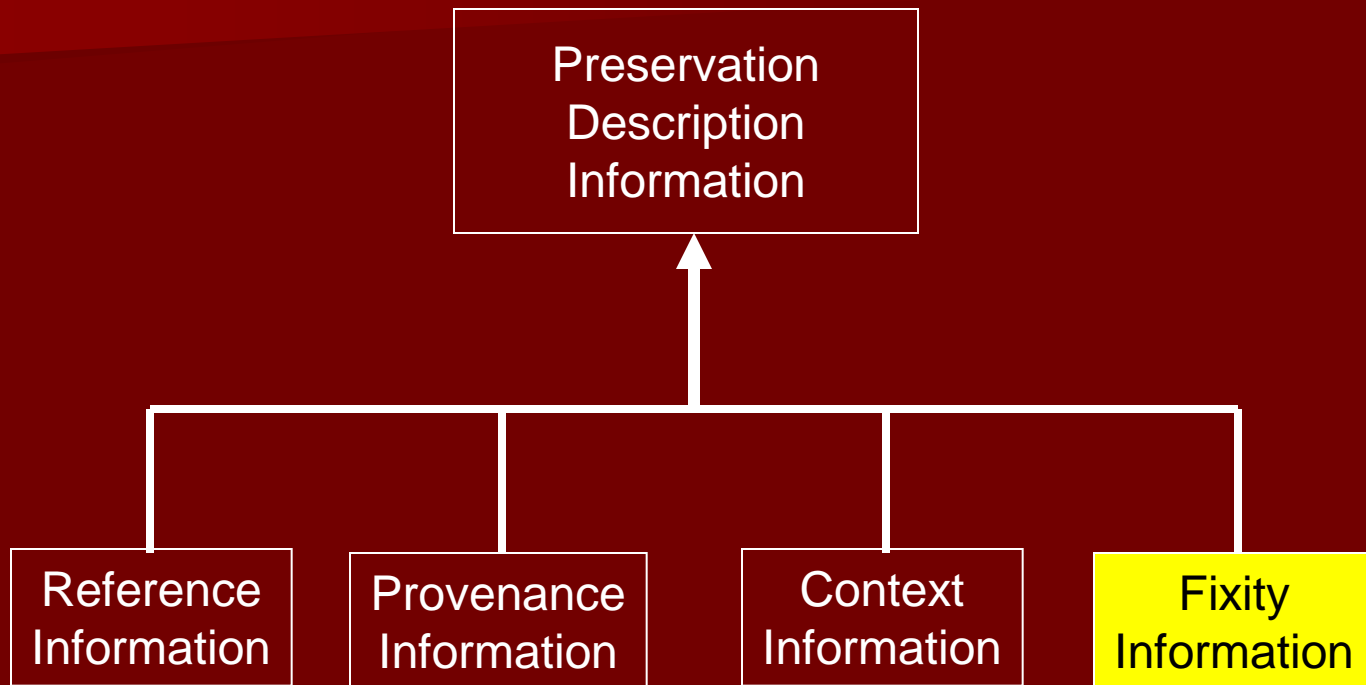
... documents the history of the Content Information. ... tells the origin or source of the Content Information, any changes that may have taken place since it was originated, and who has had custody of it since it was originated. Examples of Provenance Information are the principal investigator who recorded the data, and the information concerning its storage, handling, and migration.

OAIS preservation information



... documents the relationships of the Content Information to its environment. This includes why the Content Information was created and how it relates to other Content Information objects.

OAIS preservation information



... documents the authentication mechanisms and provides authentication keys to ensure that the Content Information object has not been altered in an undocumented manner. Example: Cyclical Redundancy Check code for a file.

MetaArchive hierarchy

- Archive (6⁺ caches per network)
 - Genre- or Format-based
- Collections (1⁺ per member)
 - Collection level metadata
- Archival unit (1⁺ per ingest)
 - e.g., all ETDs for each year

Lots of Copies Keep Stuff Safe

- LOCKSS open-source software/support to preserve web-published materials
- decentralized digital preservation infrastructure
- migrates content forward in time
- bits & bytes continually audited & repaired
- MetaArchive members also join LOCKSS

Private LOCKSS network (PLN)

- PLN is a LOCKSS network deployed by a set of like-minded institutions in order to preserve content in a closed preservation network.
- Not maintained by the Stanford University-based LOCKSS staff

Manifest page

Electronic Theses and Dissertations at Boston College

LOCKSS Manifest Page

Collection Info:

- Institution: Boston College
- Contact Info: etdadmin

Links for LOCKSS to start its crawl:

- [18640905-20081231](#) AU for 09/05/1864 to 12/31/2008



LOCKSS system has permission to collect, preserve, and serve this Archival Unit.

Archival unit

Electronic Theses and Dissertations at Boston College

Manifest list for AU 09/05/1864 to 12/31/2008

- [Benoit, Marie Saint Elphege 1967](#)
- [Benakis, Kristen 2004](#)
- [Koster, Katharine 2006](#)
- [Thompson, Matthew Charles Glenn 2005](#)
- [Gormley, Ann Marie 2005](#)
- [Goodhue, Laura 2005](#)
- [Aronzon, Stephanie 2005](#)
- [Kita, Caroline A. 2004](#)
- [Carr, Derek W 2004](#)
- [Voltolina, Vanessa 2006](#)
- [Caldwell, Robert 2003](#)
- [Kang, Bryan 2003](#)
- [Chapin, Daniel 2004](#)
- [Tress, Benjamin 2008](#)
- [Lawrence, Stephen Caleb 2007](#)
- [Zhao, Yu 2008](#)
- [Wieland, Laura Caroline 2008](#)
- [Sieber, Joshua Daniel 2008](#)
- [Rubio, Margarita 2008](#)
- [Rowan, Jamin Creed 2008](#)
- [Dougherty, Sara L. 2008](#)
- [Campbell, Leslie Ann 2008](#)
- [Burks, Heather Elizabeth 2008](#)
- [Fu, Shihe 2005](#)
- [Kauffman, Karen C 2008](#)



Done



zotero

Digital object and its metadata

Electronic Theses and Dissertations at Boston College

Manifest for Chang, Lauren Kawehionalani 2007

- [Metadata and Relationships](#)
- [ETD PDF](#)



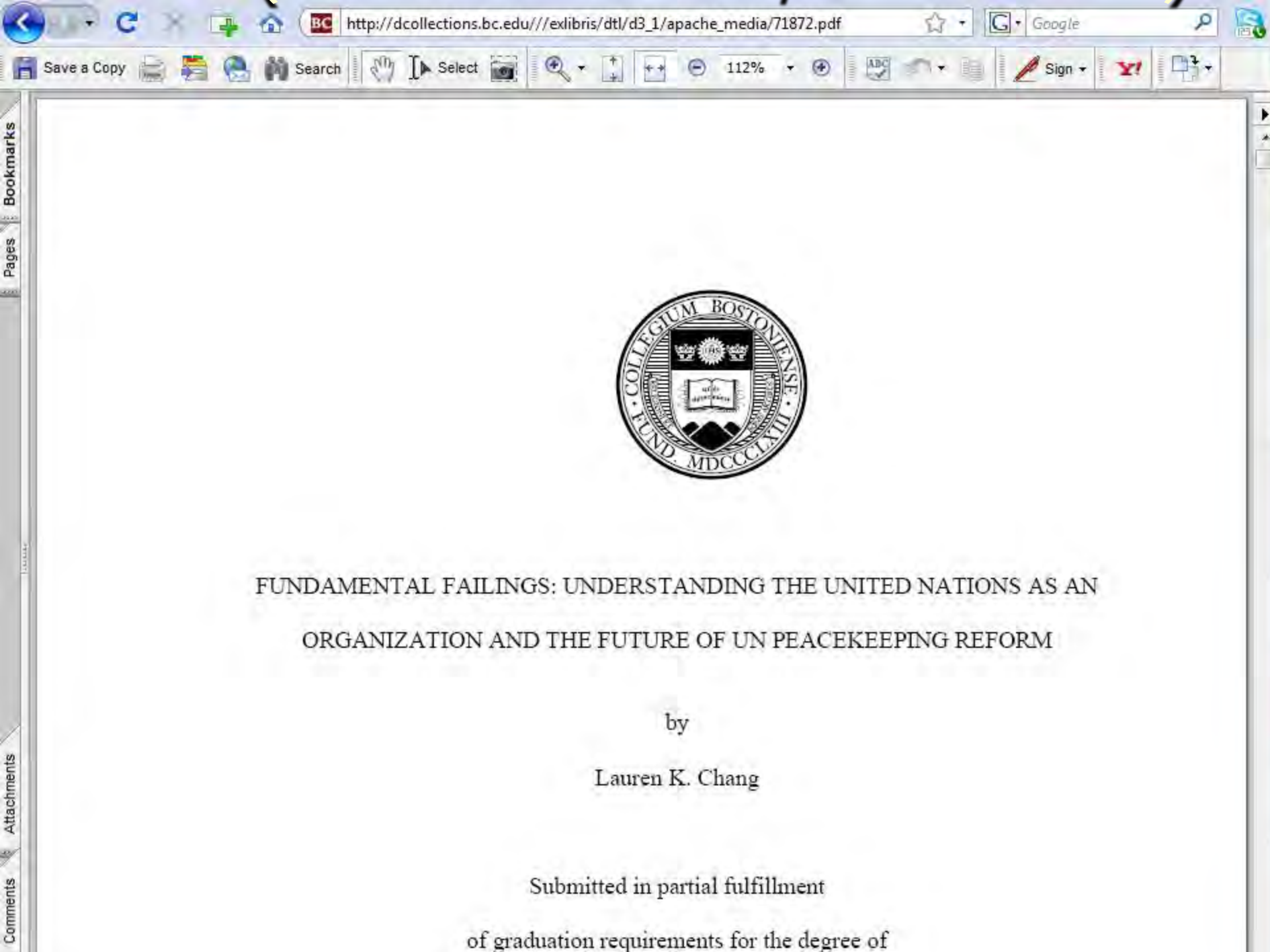
LOCKSS system has permission to collect, preserve, and serve this Archival Unit.

http://dcollections.bc.edu/webclient/DeliveryManager?metadata_request=true&GET_XML=1&pid=71872

<http://dcollections.bc.edu/webclient/DeliveryManager?pid=71872>

Metadata xml file

```
<mid>105836</mid>
<description xsi:nil="true"/>
<name>preservation</name>
<type>preservation_md</type>
- <value>
  <object xmlns="http://www.loc.gov/standards/premis"> <objectIdentifier> <objectIdentifierType>handle</objectIdentifierType>
    <objectIdentifierValue>2345/520</objectIdentifierValue> </objectIdentifier> </object>
  </value>
</md>
- <md shared="false">
  <mid>106647</mid>
  <description xsi:nil="true"/>
  <name>descriptive</name>
  <type>etd-ms</type>
  - <value>
    <?xml version="1.0" encoding="UTF-8"?><thesis> <title>Fundamental Failings: Understanding the United Nations as an Organization and the
    Future of UN Peacekeeping Reform</title> <creator>Chang, Lauren Kawehionalani</creator> <subject>Political Science, General</subject>
    <subject>United Nations</subject> <subject>Peacekeeping</subject> <subject>Organizational Theory</subject> <subject>Reform</subject>
    <description>This paper is an attempt to study the United Nations through the lens of organizational theory, and in particular, the theoretical
    framework as outlined by Allison and Zelikow in Essence of Decision, in order to understand the implementation patterns of the UN in regards to
    the Brahimi Report as reported and analyzed by the Henry L. Stimson Center. The findings of this report conclude that the UN is capable of
    change as demonstrated by its ability to comply with certain Brahimi Report recommendations, but is resistant to change, due to the structure of
    the organization. This does not mean, however, that it is fundamentally unable to do so. Attempts at reform must be able to circumvent these
    obstacles through targeted, direct action, for the Brahimi Report recommendations which received the highest implementation ratings were those
    incremental organizational reforms that targeted specific aspects of peacekeeping operations. Resistance to change within the UN, be it on behalf
    of individuals, departments, or Member States, is a huge obstacle to change, further compounding the obstacles to reform that the UN faces
    simply as an organization. Future reforms must thus be framed in a way that specifically grasps the attention of the groups/members involved in the
    reform, making the issue as pertinent and sensitive to them as it is for the success of UN peace operations in general. </description>
    <identifier>http://hdl.handle.net/2345/520</identifier> <contributor role="advisor">Hiroshi Nakazato</contributor> <date>2007</date>
    <publisher>Boston College</publisher> <rights>I hereby grant to Boston College the non-exclusive right to archive, reproduce and distribute my
    thesis, in whole or in part, and/or my abstract, in whole or in part, in and from an electronic format, subject to the release date subsequently
    stipulated in this submittal form and approved by my school. I represent and warrant that the thesis and its abstract are my original work, do not
    infringe or violate any rights of others, and that I make these grants as the sole owner of the rights of my thesis and its abstract. I warrant that I
    have obtained written permissions from the owner(s) of each third party copyrighted matter to be included in my thesis and will supply copies of
    such upon request by my school. I acknowledge that my school will not distribute my thesis or its abstract if, in their reasonable judgment, they
```

FUNDAMENTAL FAILINGS: UNDERSTANDING THE UNITED NATIONS AS AN
ORGANIZATION AND THE FUTURE OF UN PEACEKEEPING REFORM

by

Lauren K. Chang

Submitted in partial fulfillment

of graduation requirements for the degree of

Plug-in

An XML file that instructs the LOCKSS software how to ingest and preserve content.

Each cache on the network writes a plug-in for its collection, enabling other caches to replicate its content

Security

- Copies on different power grids
- All copies not accessible to one person
- Each cache secure and for DDP-only
- Security-enhanced Linux
- SSL-encrypted inter-cache communication
- IP address based Firewall exceptions

For more details...

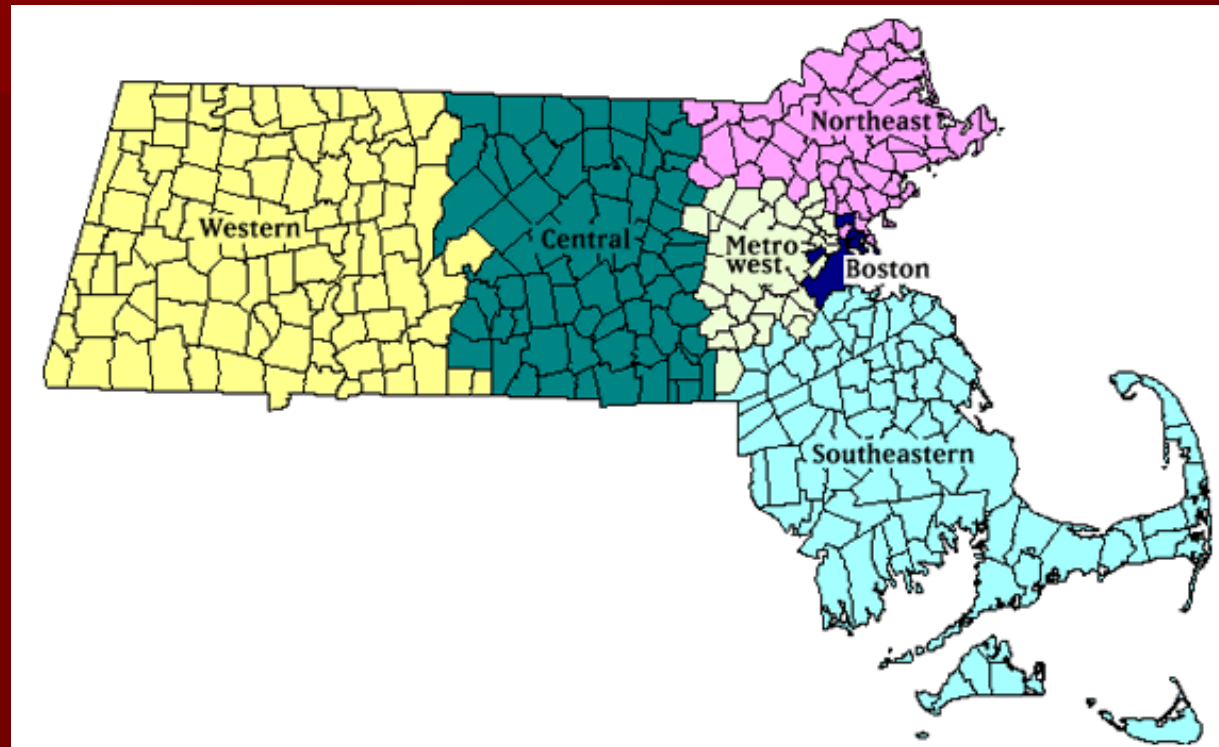
A Guide to Distributed Digital Preservation

Edited by Katherine Skinner
and Matt Schultz



<http://metaarchive.org/GDDP>

MA regional library systems



Massachusetts Networks:

CLAMS*

MBLN

SAILS*

NOBLE*

C/W MARS*

MVLC

Minuteman*

OCLN