Digital Preservation: Context and Content
Binghamton University (NY, USA)

- Binghamton is the “premier public university in the northeast” *Fiske Guide to Colleges* (2010)
- Undergraduates: 11,787
- Graduate students: 3,108
- Average SAT score range: 1180-1340
- ~ 2.4 million volumes in the libraries
- Started as Triple Cities College in 1946
- Became a “University Center” in 1965
- Part of the State University of New York system
Digital preservation: A definition (1 of 2)

- A “series of managed activities necessary to ensure continued access to digital materials for as long as necessary”
- “Actions required to maintain access to digital materials beyond the limits of media failure or technological change.”
- “materials may be records created during the day-to-day business of an organisation; "born-digital ... or the products of digitisation projects.”
JISC Beginner’s Guide to Digital Preservation elaborates:

- **Managed**: Digital preservation is a Management problem.
- **Activities**: The policy needs to filter down to a list of processes: tasks that can take place at specified times and in specified ways.
- **Necessary**: What needs to be done. How long do you want to preserve the objects for? Discussions about the activities needed to achieve a level of preservation are necessary.
- **Continued Access**: Access is the key here. Most objects in the public sphere are preserved to support access and retrieval.
- **Digital Materials**: Digital materials, digital objects, call them what you will. This is the stuff you are preserving. Different objects require different processes.
Local Content as the Future of Academic Libraries?

- At least in regards to Physical Collections?
- Google Books, HathiTrust
- To a large degree journal collections and government documents already being “managed” outside of libraries
Why digital preservation? (2 of 2)

- Furthers the role of University Archives to the digital world
- Not a new idea, a new format
- Majority of new material is published in digital format
  - Scholarly Articles
  - Campus newsletters
  - Course catalogs
  - Web sites

Photo by Klara Kim. © CC BY-NC-SA
A University is a collection of niche markets (especially when it comes to faculty/researchers)
Low-hanging fruit

Course Catalogs

Newsletters

Faculty Research

Special Collections/University Archives

Also need to be opportunistic (blood serum collection)

Images cc-by-nc-2.0: http://www.flickr.com/photos/bycp/
At-risk materials
“Digital asset management (DAM) consists of management tasks and decisions surrounding the ingestion, annotation, cataloguing, storage, retrieval and distribution of digital assets.” (Wikipedia, 2011)

Preservation is not mentioned

“Often, those creating digital materials or designing digital content management systems do not take great interest in their long-term preservation.”

(Trusted Digital Repositories: Attributes and Responsibilities An RLG-OCLC Report, 2002 (p. 18))
Digital Preservation: Binghamton’s Solution

- Experimented with various “Digital Content” systems including Content Pro, CONTENTdm, DSpace, EPrints
- None of these has preservation “built-in”
- Building our own was not practical
  - Staffing levels
  - Lack of expertise
- Rosetta by Ex Libris
Rosetta and discovery

- No preservation systems is useful if there is no access (especially for a University Library)
- Rosetta does not have a public discovery layer
- Rosetta’s Digital Publishing System is flexible so there are options
- Primo for discovery
  - First University to use Primo with Rosetta
  - Works well with other library systems such as Aleph and Primo Central
  - One stop shopping
Implementation timeline

- Early December 2010: Initial training
- January 2011: Installation and functional training
- February 2011: 2 Metadata librarians begin working at Binghamton
- End of March 2011: Went live with first collection
- May 2011: Present at ELAG!
- May 2011: Go live beta with Primo for Aleph, etc.
- More collections coming soon to Rosetta!!
Overall staffing

- Systems: 1 person (~0.5 FTE)
  - Project Management
  - Systems/Technical
- Metadata/Cataloging: 3 people (~1.0 FTE)
- User Interface: Part of Web Services Librarian's Time
- Special Collections: Not directly involved with implementation, but relied on heavily for collection level expertise
  - Could really use a Digital Curation Librarian
O/S and hardware

- RHEL 5.5
- IBM x3550 M2 (2x – one production, one test)
  - 24 GB RAM
  - (2) 2.93 GHz Intel x5570 Xeon Processors
  - (2) 146 GB hard drives
- HP StorageWorks MSA2300i Controller
  - 24 TB (pre-mirroring)
- Tape backup
What is put into Rosetta?

- Each file (or group of files) deposited is an Intellectual Entity or IE
- With an IE are:
  - Metadata forms (in DC format)
  - Submission formats (how a file is to be uploaded)
  - Copyright statement to be used
  - Access right forms (who can view the content)
Rosetta metadata decisions

- Dublin Core fields
- Non-standard fields
- Create a master list
- Dublin Core is so flexible that it is inflexible
Staffing for metadata

- Metadata Librarians create the metadata forms and set up the rules to use them
- Other staff, student help from within the Libraries will create metadata
- Projects ingested from outside the library will get our help but use their own staff/students
Metadata Form

Ex Libris Rosetta Administration

Home > Ongoing Configuration > Material Flow Infrastructure > Metadata Forms (+ Lookup Lists) > Update Form

<table>
<thead>
<tr>
<th>Type</th>
<th>Generic</th>
<th>Created on</th>
<th>Created by</th>
</tr>
</thead>
<tbody>
<tr>
<td>Updated on</td>
<td>16/05/2009</td>
<td>16/05/2009</td>
<td>Ex Libris</td>
</tr>
<tr>
<td>Updated by</td>
<td></td>
<td>Ex Libris</td>
<td></td>
</tr>
</tbody>
</table>

- **Design Information**
- **Fields Information**

<table>
<thead>
<tr>
<th>Name</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
</table>

| Form Layout   | 1 Column      |

- **Save Form**
- **Preview**
- **< Back**
Create metadata form
Choose your properties for the field

Properties for: Title

- **Label**: Title
- **DC Tag**: title
- **Validator**: No Validator
- **Default Value**: Title
- **Tooltip**: Title of item, if there is no given title; assign one, try to be unique.

- **Type**: TextField
- **Mandatory**: Yes

[Save Field]
### Searching for “submarine” in Rosetta

#### Simple Search

**Object Type:** Intellectual Entity

**Find:** submarine

#### Results

<table>
<thead>
<tr>
<th>#</th>
<th>Title</th>
<th>Type</th>
<th>PID</th>
<th>SIP ID</th>
<th>Producer</th>
<th>Creation Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Submarine on display</td>
<td>Intellectual Entity</td>
<td>IE18479</td>
<td>7890</td>
<td>Demo Internal Producer</td>
<td>2011-03-25 07:50:00</td>
</tr>
<tr>
<td>2</td>
<td>Lifting boom on Sea Div.</td>
<td>Intellectual Entity</td>
<td>IE21715</td>
<td>8968</td>
<td>Demo Internal Producer</td>
<td>2011-03-26 21:05:07</td>
</tr>
<tr>
<td>3</td>
<td>Perry Cubmarine is prep...</td>
<td>Intellectual Entity</td>
<td>IE6809</td>
<td>3961</td>
<td>Demo Internal Producer</td>
<td>2011-03-25 03:03:06</td>
</tr>
<tr>
<td>4</td>
<td>Sea Diver II at work</td>
<td>Intellectual Entity</td>
<td>IE6812</td>
<td>3962</td>
<td>Demo Internal Producer</td>
<td>2011-03-25 03:03:10</td>
</tr>
<tr>
<td>5</td>
<td>Lifting boom on Sea Div.</td>
<td>Intellectual Entity</td>
<td>IE6815</td>
<td>3963</td>
<td>Demo Internal Producer</td>
<td>2011-03-25 03:03:15</td>
</tr>
<tr>
<td>6</td>
<td>Perry Cubmarine is rais...</td>
<td>Intellectual Entity</td>
<td>IE6818</td>
<td>3964</td>
<td>Demo Internal Producer</td>
<td>2011-03-25 03:03:21</td>
</tr>
<tr>
<td>7</td>
<td>Perry Cubmarine emerges...</td>
<td>Intellectual Entity</td>
<td>IE6821</td>
<td>3965</td>
<td>Demo Internal Producer</td>
<td>2011-03-25 03:03:25</td>
</tr>
<tr>
<td>8</td>
<td>Sea Diver II and submar...</td>
<td>Intellectual Entity</td>
<td>IE6572</td>
<td>3882</td>
<td>Demo Internal Producer</td>
<td>2011-03-25 02:56:45</td>
</tr>
<tr>
<td>9</td>
<td>Perry Cubmarine on Sea</td>
<td>Intellectual Entity</td>
<td>IE8564</td>
<td>4546</td>
<td>Demo Internal Producer</td>
<td>2011-03-25 03:47:00</td>
</tr>
<tr>
<td>10</td>
<td>Sea Diver &amp; Deep Diver</td>
<td>Intellectual Entity</td>
<td>IE14744</td>
<td>6643</td>
<td>Demo Internal Producer</td>
<td>2011-03-25 06:26:56</td>
</tr>
</tbody>
</table>
Perry Cubmarine on Sea Diver
Sea Diver (Boat)
Perry Cubmarine (Submarine)
Binghamton University Libraries
edwin a link

Available in the Library (37)

Collection
Rosetta (5,097)
Special Collections (32)

Resource Type
Images (5,097)
Books (30)
Audio Visual (5)
Journals (1)
Other (1)

5,134 Results for All Local Collections
Same search in Binghamton’s Primo

Search for "submarine" in the Digital Collections.

Results for Digital Collections:

- Title: Perry Cubmarine on Sea Diver
- Author: Sea Diver (Boat); Perry Cubmarine (Submarine)
- Is Part Of: Edwin A. Link Jr. Collection
- Description: Perry Cubmarine on Sea Diver.
- Publisher: Binghamton University Libraries
- Format: 1 photograph : b&w ; 20.5 X 25 cm. image/jpeg
- Language: English
- Source: Rosetta
- Subject (TGM): Boats
Can zoom in!
Comments and challenges

- Built for National Libraries
  - Comes with territory of being first University partner
    - Scales down well
  - Ex Libris extremely responsive and making improvements
  - Staffing
Future projects

- More Special Collections (digitization)
- Blood Serum
- University Photographer
  - 767 DVDs!
- Born Digital
  - Newsletters
  - Campus website
- Planning to be opportunistic
- Prioritization: Committee to set priorities
Digital Preservation: Context and Content: ELAG 2011

Edward M. Corrado
http://ecorrado.us/
ecorrado@ecorrado.us