DLA Annual Report 2011/12

Personnel Highlights
Successful national search for Digital Repository Librarian—Nathan Hall, and two Digital Technologies Development Librarians—Zhiwu Xie and Keith Gilbertson.

Partnering with the University Community

VTechWorks, the university’s first digital repository was born.

- Migrated 2400 records in time for arrival of the new Repositories Librarian in Aug.
  - Created spreadsheets of metadata for batch loading.
  - Wrote Perl scripts to batch load initial collections. Scripts pulled the files from various locations on DLA servers and created XML SIP packages and content lists.
  - Batch loaded images, ETDs, ejournals, and book collections
- Migrated 7,166 records from the April 16, 2007, Condolence Archives
  - Discontinued use of CONTENTdm, saving the library $6000/yr
  - Adopted Viewshare, hosted by Library of Congress, to display map and timeline
- Outreach led to 67 new users and ~300 new resources, such as
  - Nutrition, Obesity, and Health
  - Sinfonia Sacra, Opus 56
  - Engineers’ Forum
  - Blacksburg Topographical Map
  - Test Videos- Flexural Strength Experiments
- Implemented BioMedCentral deposits
- Client for computer science class – Multimedia, Hypertext, and Information Access, which developed:
  - VTechWorks interface proposal and prototypes
  - HTTP live streaming as a secure streaming method
- Hosting 9,842 information resources
  - The illustrations in this section are from an automatic thumbnail builder locally programmed.
  - 225,097 items accessed
  - 445,596 files downloaded

Data Management Plans
- Advised and assisted PIs and graduate assistants in writing plans
- Introduced over 100 faculty, students, staff in three FDI sessions
- Created Data Policy Watch wiki for ASERL/SURA
R & D

- Prepared DLA resources for web discovery tools:
  - Wrote metadata extraction tools (e.g., COinS) and crawler/indexers (e.g., python), scripts to export content (e.g., CSV file, OAI/ORE via atom/xml).
  - *Journal of Technology Education* successfully loaded into VTechWorks.
- Tested DSpace REST API, prototyped data and metadata APIs
- Finished migrating Fish Traits database from CNRE server to the IBM cloud
- Finished the first round of prototype to automatically discover geospatial data collections currently hosted at CGIT
- Coordinated with CGIT to store 14 GB of GARDEN project data

DLA's 'legacy collections' summary:
- Total number of files: 628,213
- Total disc usage: 2.8 TB
- 2,028 new files added, 8.1 GB of space.
- Total number of times collections accessed: 14,016,251
- Total queries conducted: 3,959,412

### Electronic Journals Online Access Data

<table>
<thead>
<tr>
<th>Year</th>
<th>Hits</th>
<th>Pages</th>
<th>HTML files</th>
<th>PDF files</th>
<th>Distinct pages</th>
<th>Unique visitors</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>4,120,946</td>
<td>2,673,618</td>
<td>2,127,328</td>
<td>512,467</td>
<td>16,500</td>
<td>942,418</td>
</tr>
<tr>
<td>2010</td>
<td>4,373,538</td>
<td>2,709,748</td>
<td>2,195,296</td>
<td>470,417</td>
<td>16,070</td>
<td>932,694</td>
</tr>
<tr>
<td>2009</td>
<td>3,745,749</td>
<td>1,891,890</td>
<td>1,506,445</td>
<td>362,258</td>
<td>14,573</td>
<td>732,032</td>
</tr>
<tr>
<td>2008*</td>
<td>3,145,258</td>
<td>1,574,919</td>
<td>1,322,831</td>
<td>234,347</td>
<td>11,812</td>
<td>593,128</td>
</tr>
</tbody>
</table>

eJournals
- Posted 774 new articles in 44 issues
- Reformatted 275 articles in 25 issues as HTML to provide full accessibility and comply with Sect. 508 of the Federal Accessibility Standards of the US Rehabilitation Act, and Virginia Tech Accessibility Standards.
- Added qualified DC metatags and COinS to 774 new articles, and these back issues:
  - journals w/ metatags and COinS: 1
  - issues w/ COinS: 143
  - issues w/ metatags and COinS: 71
  - index pages w/ metatags and COinS: 84
  - article pages w/metatags: 482
  - total pages w/ metatags: 582

Metadata offers several enhancements
- Facilitates access through Summon, Google Scholar, etc.
- It improves our ability to manage our collection.
- COinS is linked via LibX to our local OpenURL resolver
- COinS is read by bibliographic management software, such as Zotero
ETDs: Electronic Theses and Dissertations

- Edited Web pages to improve access and prepare for migration to VTechWorks
- Improved Resources for Developers of ETD Databases
- ~400 BTDs (bound theses and dissertations) records upgraded with committee members
- >300 ETD restrictions removed--permissions and over due release dates
- ETD-db v.2: Continued progress, e.g., Worked through Randy Crockett to access Banner, passed ED-

VT ImageBase

- Banded successfully so the public was never without access
- Metadata improvements made as requested, e.g., 268 records for Pakistani architect, Yasmeen Lari
- 392 patron requests for images, e.g.,
  - VT Visitor Center
  - History Museum of Western Virginia
  - Grove Cookbook
  - “Blacksburg Now and Then”
  - Globe and Mail, Toronto newspaper

Preservation

- 63 collections in 245 GB are archived with the MetaArchive Cooperative

Security Breaches

- Just one and we caught it before VT IT Security did: car insurance headers had been added to http://spec.lib.vt.edu/bicent/sites/yss.htm

Grants

Amazon Teaching Grant. $5000 worth of Amazon Web Services usage credit, 2012/13.
IBM Academic Skills Cloud. Free usage of up to 20 cloud instances on the IBM Smart Cloud Enterprise, 2012/14.
IMLS. ETD Lifecycle Management. 2011/13. $268,458. Led by University of North Texas Libraries with the NDLTD, Educopia Institute/MetaArchive Cooperative, and libraries at Virginia Tech, Rice, Boston College, Indiana State University, Pennsylvania State University, and the University of
Arizona. Develop and share a toolkit of guidelines, educational materials, and software tools for life-cycle data management and ETD preservation.

NEH. Chronicles in Preservation. 2011/13. $300,000. Led by Edupemia Institute with the libraries of University of North Texas, Boston College, Clemson, Georgia Tech, Penn State, San Diego Supercomputer Center, University of Utah, Virginia Tech. Study, document, and model the use of data preparation and DDP frameworks to preserve digital newspaper collections at MetaArchive-LOCKSS, Chronopolis-iRODS, and UNT-CODA.

Publications


Xie, Zhiwu; Liu, Jinyang; Van de Sompel, Herbert; van Reenen, Johann; Jordan, Ramiro. “Poor Man’s Social Network: Consistently Trade Freshness For Scalability.” USENIX WebApps ’12, Boston, MA.

Presentations

Library In Service Day
- DLA was invited to present a session about the department. Everyone in DLA participated.
- Created a DLA Pecha Kucha


McMillan, Gail. Copyright in Scholarship & Instruction and an Open Access Primer. VT FDI, Feb. 7.


- 14th International ETD Symposium: Cape Town, South Africa, Sept. 15.

Xie, Zhiwu. “Data Management @ Virginia Tech Libraries.” Data Management Day at University of Virginia, April 12.