Leveraging the UNT Digital Library Infrastructure to Support our Extended Communities

Mark Phillips
UNT Libraries
March 12, 2015
The UNT Libraries Digital Collections
1,043,589 unique items
1,033,011 of those available in to anyone in the world
Three Distinct Digital Library Interfaces
The Portal to Texas History
280+ Partner Institutions from across Texas
Letters, Photos, Maps, Books, Documents, Newspapers and much much more
Over 2.8 million pages of Texas Newspapers
Over 82,000 Maps
Service hub in Texas for the Digital Public Library of America (DPLA)
By far the most visible digital library “collection” we operate
The Gateway to Oklahoma History
About the Gateway

The Gateway is an online repository of Oklahoma history. You may browse through hundreds of thousands of newspaper pages dating from the 1840s to the 1920s.

The Gateway provides free access to 182,874 issues and 916,190 pages of historical newspaper content.

Other Resources

The Oklahoma Historical Society has numerous resources for Oklahoma history. Visit the Research Center section to find out more about our books, manuscripts, maps, photographs, audio, video and other newspapers.

Recent Additions

The Gateway is continually adding new digital newspapers. View some of the latest materials here.
Operated for the Oklahoma Historical Society
Currently two major collections:

Oklahoma Newspaper Collection 1 Million pages
Oklahoma Publishing Company Photos 235,000
Expected to include over a million photos in a few years...
The UNT Digital Library
More UNT focused content

UNT Created
UNT Collected
UNT Purchased
(Mostly non-Texas related content)
UNT Scholarly Works
What's Inside this Collection.

The Scholarly Works Collection is home to materials from the UNT community's research, creative, and scholarly activities. It serves as UNT’s Open Access Repository. This collection brings together articles, papers, artwork, music, research data, reports, presentations, and other scholarly and creative products representing the expertise in our university community.

UNT Scholarly Works aims to:

- Provide easy access to valuable scholarly and creative materials from the UNT community
- Promote discovery through effective search and navigation tools
- Secure long-term access through stewardship and preservation
- Ensure sustainability through continuing system improvements
- Showcase UNT’s research and creative achievements to a worldwide audience

To learn more about UNT Scholarly Works, see our web page or contact us at untrepository@unt.edu.
UNT Data Repository
What's Inside this Collection.

The UNT Data Repository is a central archive to provide long-term preservation and access to the research data of the UNT scholarly community. The Data Repository works in conjunction with the UNT Scholarly Works repository to ensure long-term access to the full range of research outputs from UNT.

| Texas Digital Newspaper Program Issue Dataset for IFLA/Roosttech Analysis Added: 02/14/2014 |
| UNT Libraries Collection Development Dataset, 2002-2003 Added: 01/21/2014 |
| UNT Libraries Collection Development Dataset, 2012-2013 Added: 01/21/2014 |
UNT Theses and Dissertations

Electronic Theses and Dissertations (ETD)
Digitized Theses and Dissertations
Theses and dissertations represent a wealth of scholarly and artistic content created by masters and doctoral students in the degree-seeking process. In 1999, the University of North Texas was one of the first American universities to begin requiring electronic theses or dissertations (ETDs) for graduation. Some ETDs in this collection are restricted to use by the UNT community.
Collection from across Campus
Digitized collections from the UNT Libraries
Born Digital Collections
Content that supports our students and faculty
And the extended community
Some examples include:
Documenting Plate Waste in Middle School Cafeterias Using Digital Still Photography

What’s Inside this Collection.

These images document the disappearance of food and beverage items by photographing trays before and after schoolchildren ate lunch. The images were taken in North Texas during lunchtime in urban middle schools that participated in the National School Lunch and School Breakfast Programs.

About the Project

These photographs come from USDA-ERS Project #10.253, "Testing a Food Choice Innovation for Middle School Cafeterias," conducted from October 2010 to May 2011. Researchers were P. Connors, C. Bednar, B. Davenport, and L. Kennon.

Equipment and Procedure

Three Canon PowerShot 1400 cameras with 8 GB memory cards were used. An apparatus forming a T-aerial was constructed using ½ inch PVC pipe with fittings and a GorillaGripper to suspend each camera at a height of two feet above a black FOAMCORE board cut to fit the top shelf of a food trolley. To position trays directly below the camera an 8x15
Description: Images taken at a North Texas middle school documenting the food on a lunch tray and the remains on the same tray after the meal was consumed. These images are part of a study to document what food students are eating.

Creator(s): Connors, Priscilla

Location(s): United States - Texas

Creation Date: March 30, 2011

Partner(s): UNT College of Merchandising, Hospitality and Tourism

Collection(s): Documenting Plate Waste in Middle School Cafeterias Using Digital Still Photography

Usage: Total Uses: 41
Past 30 days: 3
Yesterday: 0
Student Lunch Tray: 01_20110330_01B5916
Statistics for Documenting Plate Waste in Middle School Cafeterias Using Digital Still Photography

8,396 Total Uses / 1,414 Total Items (2,830 files) / 1,414 Visible / 0 Hidden

Usage by Month/Year

<table>
<thead>
<tr>
<th>Year</th>
<th>January</th>
<th>February</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>July</th>
<th>August</th>
<th>September</th>
<th>October</th>
<th>November</th>
<th>December</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>612</td>
<td>488</td>
<td>155</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1,255</td>
</tr>
<tr>
<td>2013</td>
<td>43</td>
<td>53</td>
<td>16</td>
<td>20</td>
<td>17</td>
<td>44</td>
<td>480</td>
<td>424</td>
<td>366</td>
<td>353</td>
<td>510</td>
<td>548</td>
<td>2,874</td>
</tr>
</tbody>
</table>
APIS for Documenting Plate Waste in Middle School Cafeterias Using Digital Still Photography

Introduction

The UNT Digital Library provides public access to number of application programming interfaces (APIs) to the collections within the system. Below are examples of APIs available for Documenting Plate Waste in Middle School Cafeterias Using Digital Still Photography that can be used openly by those interested in programatically accessing data from this system. You do not need to apply for a special key to use these APIs.

For additional information about these APIs or if you have general questions about machine interaction with the UNT Digital Library please contact Mark Phillips.

Note that all example URLs below use the same protocol and server name, http://digital.library.unt.edu/explore/collections/DPWMSC/. We only show the URL paths and parameters below to save space.

The API

OAI-PMH

The Open Archives Initiative's Protocol for Metadata Harvesting (OAI-PMH) allows programmatic access to this collection's metadata. Two metadata formats are currently supported, the standard oai_dc and the UNT Libraries native metadata format until.

Below are example URLs which demonstrate some of the standard views of this OAI-PMH repository:

- oai/
  base URL for OAI-PMH repository
- oai/?verb=Identify
  Display information about this repository
- oai/?verb=ListMetadataFormats
  List available metadata formats
- oai/?verb=ListSets
  List available sets
Texas Fashion Collection

What’s Inside this Collection.

The UNT Texas Fashion Collection is dedicated to the preservation and documentation of historically significant fashion, and serves as an educational and inspirational resource for students, researchers, and the general public. This vital resource is part of the UNT College of Visual Arts + Design.

About the Physical Collection

The collection of top designers’ works began in 1938 through the efforts of Stanley and Edward Marcus. The Dallas Fashion Group saw that it came to UNT in 1972 to support a growing fashion design program. Today, the collection includes over 18,000 items. Between February and August of 2013, the entire collection was packed and moved to a new facility on the UNT Campus.

At this time, the Texas Fashion Collection is undergoing extensive inventory and documentation in preparation for providing online access to all objects that have been photographed.

About the Digital Collection

RAG Ladies visit the
Texas Fashion Collection, November 21, 2013
Added: 03/08/2014

Festival Attire - Akha
Hill Tribe
Added: 03/07/2014

Belly Dancer’s
Costume - Banjara
Peoples of India
Added: 03/07/2014
College of Music Recordings

What's Inside this Collection.

About This Collection

The College of Music Recordings include doctoral, ensemble, faculty, guest, and senior recitals from the UNT College of Music, beginning with spring 2006. Access to these recordings is restricted to the UNT community.

About the UNT College of Music

One of the nation's largest music schools, the UNT College of Music provides a dynamic learning environment for both future professionals and the broader university community. The College of Music offers instruction in the areas of composition studies; conducting and ensembles; instrumental studies; jazz studies; keyboard studies; music education; music history, theory, and ethnomusicology; and vocal studies. With more than fifty performing ensemble groups, a full schedule of student recitals, and frequent visits by guest artists, the college brings music lovers nearly a thousand concert events each year.

Browse the holdings of this collection
4,500 audio recordings of recitals and concerts
600 HD videos of recitals and concerts
Joseph C. Britton, Jr. was born on October 14, 1942 in Fort Worth, Texas. He earned a Bachelor’s of Science and a Master’s Degree in Biology from Texas Christian University (1963 and 1965 respectively). He earned a Ph.D. from George Washington University in 1970. His career began as the Assistant Director of Exhibits at the Smithsonian Institution National Museum of Natural History in Washington, D.C. where his early studies were of the marine bivalve family Lucinidae. Dr. Britton returned to his hometown of Fort Worth, Texas in the early 1970s and accepted a faculty position in the Biology Department at Texas Christian University. His research interests then transitioned to freshwater bivalves. He surveyed Texas waters for native mussels and studied the ecology and distribution of the invasive freshwater mussel Megalonalas nervosa.
Megalonaias nervosa, Specimen #1655

Description:
One preserved mussel specimen including both the left and right valves. The specimen exhibits a rhombooidal shape; moderately thick shell; double looped beak sculpturing; white internal coloring; tan external coloring; external sculpturing in the form of ridges. Collected in the Guadalupe basin. The specimen measures between 0 - 60 mm in length and was assessed to be recently dead when collected.

Creator(s):
- Hammontree, Sarah
- Pulliam, Lauren
- Perry, Heather
- Mabe, Jeffery

Location(s):
United States - Texas - DeWitt County

Creation Date:
June 11, 2011

Partner(s):
Elm Fork Natural Heritage Museum
Environmental Policy Collection

What’s Inside this Collection.

About This Collection

The Environmental Policy Collection contains a variety of open access resources that provide a balanced view on environmental issues and their potential consequences. The current focus of the collection is climate change, and the documents come from such agencies and individuals as:

- United States Climate Change Science Program
- National Oceanic and Atmospheric Administration
- International Geosphere-Biosphere Programme
- Climate Change and Agriculture and Food Security
- Earth System Science Partnership
- Intergovernmental Panel on Climate Change
- United Nations Environment Programme
- World Meteorological Organization
- University of North Texas (UNT) graduate students

About This Project

The UNT Libraries are working to identify, collect, organize, and
What's Inside this Collection.

This selection of materials from the Technical Report Archive and Image Library (TRAIL) includes hard-to-find reports published by various government agencies. The technical publications contain reports, images, and technical descriptions of research performed for U.S. government agencies prior to 1975. Topics range from mining, desalination, and radiation to broader physics, biology, and chemistry studies. Some reports include maps, foldouts, blueprints, and other oversize materials.

Browse the holdings of this collection
All content is loaded into our digital preservation system
A system for managing BagIt bags
Coda
Dashboard

The Coda system acts as a digital archive for items in the UNT Libraries' Digital Collections. This dashboard presents a non-technical overview.

1,107,681 Bags
242.9 TB Disk Space Used
119,877,725 Files
4,437,107 PREMIS Events
0 Queue Entries
Validation Entries

search text within a bag
1,107,681 Bags
<table>
<thead>
<tr>
<th>Ark ID</th>
<th>Bagged Date</th>
<th>URLs</th>
<th>ATOM</th>
<th>Size</th>
<th># Files</th>
</tr>
</thead>
<tbody>
<tr>
<td>ark:/67531/codanpvd</td>
<td>March 9, 2015</td>
<td>urls</td>
<td>ATOM</td>
<td>122.5 MB</td>
<td>3,732</td>
</tr>
<tr>
<td>ark:/67531/codanpve</td>
<td>March 9, 2015</td>
<td>urls</td>
<td>ATOM</td>
<td>138.6 MB</td>
<td>134</td>
</tr>
<tr>
<td>ark:/67531/codanpvg</td>
<td>March 9, 2015</td>
<td>urls</td>
<td>ATOM</td>
<td>467.9 MB</td>
<td>4</td>
</tr>
<tr>
<td>ark:/67531/codanpvh</td>
<td>March 9, 2015</td>
<td>urls</td>
<td>ATOM</td>
<td>907.7 MB</td>
<td>1,830</td>
</tr>
<tr>
<td>ark:/67531/codanpvi</td>
<td>March 9, 2015</td>
<td>urls</td>
<td>ATOM</td>
<td>2.4 GB</td>
<td>1,302</td>
</tr>
<tr>
<td>ark:/67531/codanpzn</td>
<td>March 9, 2015</td>
<td>urls</td>
<td>ATOM</td>
<td>43.3 MB</td>
<td>1,694</td>
</tr>
<tr>
<td>ark:/67531/codanpzw</td>
<td>March 9, 2015</td>
<td>urls</td>
<td>ATOM</td>
<td>19.0 MB</td>
<td>934</td>
</tr>
<tr>
<td>ark:/67531/codanpzr</td>
<td>March 9, 2015</td>
<td>urls</td>
<td>ATOM</td>
<td>19.6 MB</td>
<td>982</td>
</tr>
<tr>
<td>ark:/67531/codanpzs</td>
<td>March 9, 2015</td>
<td>urls</td>
<td>ATOM</td>
<td>14.5 MB</td>
<td>702</td>
</tr>
<tr>
<td>ark:/67531/codanpzt</td>
<td>March 9, 2015</td>
<td>urls</td>
<td>ATOM</td>
<td>37.5 MB</td>
<td>1,974</td>
</tr>
<tr>
<td>ark:/67531/codanpzu</td>
<td>March 9, 2015</td>
<td>urls</td>
<td>ATOM</td>
<td>28.3 MB</td>
<td>1,006</td>
</tr>
<tr>
<td>ark:/67531/codanpzw</td>
<td>March 9, 2015</td>
<td>urls</td>
<td>ATOM</td>
<td>14.7 MB</td>
<td>654</td>
</tr>
<tr>
<td>ark:/67531/codanpzw</td>
<td>March 9, 2015</td>
<td>urls</td>
<td>ATOM</td>
<td>7.3 MB</td>
<td>390</td>
</tr>
<tr>
<td>ark:/67531/codanpzw</td>
<td>March 9, 2015</td>
<td>urls</td>
<td>ATOM</td>
<td>7.8 MB</td>
<td>334</td>
</tr>
<tr>
<td>ark:/67531/codanpzy</td>
<td>March 9, 2015</td>
<td>urls</td>
<td>ATOM</td>
<td>42.2 MB</td>
<td>1,638</td>
</tr>
<tr>
<td>ark:/67531/codanpzy</td>
<td>March 9, 2015</td>
<td>urls</td>
<td>ATOM</td>
<td>31.3 MB</td>
<td>1,606</td>
</tr>
</tbody>
</table>
119,877,725 files
249 TB of primary data
Capacity of 400 TB
<table>
<thead>
<tr>
<th>Name</th>
<th>URL</th>
<th>Capacity</th>
<th>Used</th>
<th>Available</th>
<th>Full</th>
<th>Last Updated</th>
</tr>
</thead>
<tbody>
<tr>
<td>coda-001</td>
<td><a href="http://libdigitarch10.library.unt.edu/coda-001/">http://libdigitarch10.library.unt.edu/coda-001/</a></td>
<td>23.0 TB</td>
<td>22.8 TB</td>
<td>169.0 GB</td>
<td>99.28%</td>
<td>Nov. 1, 2014, 10:06 p.m.</td>
</tr>
<tr>
<td>coda-003</td>
<td><a href="http://libdigitarch10.library.unt.edu/coda-003/">http://libdigitarch10.library.unt.edu/coda-003/</a></td>
<td>23.0 TB</td>
<td>22.9 TB</td>
<td>134.7 GB</td>
<td>99.43%</td>
<td>Oct. 31, 2014, 10:56 p.m.</td>
</tr>
<tr>
<td>coda-004</td>
<td><a href="http://libdigitarch10.library.unt.edu/coda-004/">http://libdigitarch10.library.unt.edu/coda-004/</a></td>
<td>23.4 TB</td>
<td>23.2 TB</td>
<td>147.9 GB</td>
<td>99.38%</td>
<td>Nov. 8, 2014, 3:56 p.m.</td>
</tr>
<tr>
<td>coda-005</td>
<td><a href="http://libdigitarch10.library.unt.edu/coda-005/">http://libdigitarch10.library.unt.edu/coda-005/</a></td>
<td>23.4 TB</td>
<td>23.3 TB</td>
<td>118.3 GB</td>
<td>99.51%</td>
<td>Nov. 6, 2014, 6:41 a.m.</td>
</tr>
<tr>
<td>coda-007</td>
<td><a href="http://libdigitarch10.library.unt.edu/coda-007/">http://libdigitarch10.library.unt.edu/coda-007/</a></td>
<td>22.6 TB</td>
<td>22.4 TB</td>
<td>146.1 GB</td>
<td>99.37%</td>
<td>June 6, 2014, 3:53 p.m.</td>
</tr>
<tr>
<td>coda-008</td>
<td><a href="http://libdigitarch10.library.unt.edu/coda-008/">http://libdigitarch10.library.unt.edu/coda-008/</a></td>
<td>24.0 TB</td>
<td>23.8 TB</td>
<td>168.3 GB</td>
<td>99.32%</td>
<td>Nov. 11, 2014, 9:10 a.m.</td>
</tr>
<tr>
<td>coda-009</td>
<td><a href="http://libdigitarch10.library.unt.edu/coda-009/">http://libdigitarch10.library.unt.edu/coda-009/</a></td>
<td>24.0 TB</td>
<td>23.9 TB</td>
<td>149.7 GB</td>
<td>99.39%</td>
<td>Feb. 18, 2015, 11:44 a.m.</td>
</tr>
<tr>
<td>coda-010</td>
<td><a href="http://libdigitarch10.library.unt.edu/coda-010/">http://libdigitarch10.library.unt.edu/coda-010/</a></td>
<td>24.0 TB</td>
<td>23.9 TB</td>
<td>139.1 GB</td>
<td>99.43%</td>
<td>Jan. 31, 2015, 10:10 a.m.</td>
</tr>
<tr>
<td>coda-011</td>
<td><a href="http://libdigitarch10.library.unt.edu/coda-011/">http://libdigitarch10.library.unt.edu/coda-011/</a></td>
<td>24.0 TB</td>
<td>10.1 TB</td>
<td>13.9 GB</td>
<td>41.89%</td>
<td>March 10, 2015, 1:20 p.m.</td>
</tr>
<tr>
<td>coda-012</td>
<td><a href="http://libdigitarch10.library.unt.edu/coda-012/">http://libdigitarch10.library.unt.edu/coda-012/</a></td>
<td>24.0 TB</td>
<td>945.9 GB</td>
<td>23.1 TB</td>
<td>3.85%</td>
<td>Feb. 25, 2015, 10:48 a.m.</td>
</tr>
<tr>
<td>coda-013</td>
<td><a href="http://libdigitarch10.library.unt.edu/coda-013/">http://libdigitarch10.library.unt.edu/coda-013/</a></td>
<td>24.0 TB</td>
<td>32.8 MB</td>
<td>24.0 TB</td>
<td>0.0%</td>
<td>Jan. 21, 2015, 3:03 p.m.</td>
</tr>
<tr>
<td>coda-014</td>
<td><a href="http://libdigitarch10.library.unt.edu/coda-014/">http://libdigitarch10.library.unt.edu/coda-014/</a></td>
<td>22.6 TB</td>
<td>32.8 MB</td>
<td>22.6 TB</td>
<td>0.0%</td>
<td>Jan. 21, 2015, 3:03 p.m.</td>
</tr>
<tr>
<td>coda-015</td>
<td><a href="http://libdigitarch10.library.unt.edu/coda-015/">http://libdigitarch10.library.unt.edu/coda-015/</a></td>
<td>24.0 TB</td>
<td>32.8 MB</td>
<td>24.0 TB</td>
<td>0.0%</td>
<td>Jan. 21, 2015, 3:03 p.m.</td>
</tr>
<tr>
<td>coda-016</td>
<td><a href="http://libdigitarch10.library.unt.edu/coda-016/">http://libdigitarch10.library.unt.edu/coda-016/</a></td>
<td>24.0 TB</td>
<td>32.8 MB</td>
<td>24.0 TB</td>
<td>0.0%</td>
<td>Jan. 21, 2015, 3:04 p.m.</td>
</tr>
<tr>
<td>coda-017</td>
<td><a href="http://libdigitarch10.library.unt.edu/coda-017/">http://libdigitarch10.library.unt.edu/coda-017/</a></td>
<td>22.6 TB</td>
<td>32.8 MB</td>
<td>22.6 TB</td>
<td>0.0%</td>
<td>Jan. 21, 2015, 3:04 p.m.</td>
</tr>
</tbody>
</table>

**TOTALS:** 398.9 TB 243.2 TB 155.7 TB 60.98%
There are 3 premis events associated with ark:/67531/codanpvd:
PREMIS Events logged for various repository activities
There are **3 premis events** associated with ark:/67531/codanpvd:

<table>
<thead>
<tr>
<th>Event ID</th>
<th>Event Date</th>
<th>Event Status</th>
<th>Linked Object(s)</th>
<th>Classified Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>7d4d2de426184f23bda1097ded35360a</td>
<td>2015-03-09 02:04:05</td>
<td>Success</td>
<td>ark:/67531/codanpvd</td>
<td><a href="http://purl.org/net/untl/vocabularies/preservationEvents/#ingestion">http://purl.org/net/untl/vocabularies/preservationEvents/#ingestion</a></td>
</tr>
<tr>
<td>942bedce334f484c80cc6261a6376804</td>
<td>2015-03-09 14:44:33</td>
<td>Success</td>
<td>ark:/67531/codanpvd</td>
<td><a href="http://purl.org/net/untl/vocabularies/preservationEvents/#fixityCheck">http://purl.org/net/untl/vocabularies/preservationEvents/#fixityCheck</a></td>
</tr>
<tr>
<td>0355c70f2461483c9a497a004ee5e39e</td>
<td>2015-03-09 14:44:33</td>
<td>Success</td>
<td>ark:/67531/codanpvd</td>
<td><a href="http://purl.org/net/untl/vocabularies/preservationEvents/#replication">http://purl.org/net/untl/vocabularies/preservationEvents/#replication</a></td>
</tr>
</tbody>
</table>
Ingest
Replication
Fixity Checks
Migration
There are 4,437,107 total events. Here are the 10 most recent events.

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Date</th>
<th>Event Type</th>
<th>Linked Object(s)</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>d14b893f5a3649748be4b7c32b93edcb</td>
<td>March 10, 2015, 9:40 p.m.</td>
<td>⬪ <a href="http://purl.org/net/untl/vocabularies/preservationEvents/#replication">http://purl.org/net/untl/vocabularies/preservationEvents/#replication</a></td>
<td>ark:/67531/codanqfg</td>
<td>success</td>
</tr>
<tr>
<td>9f931d3aeef224a9d8c53d648fc95f767</td>
<td>March 10, 2015, 9:40 p.m.</td>
<td>⬪ <a href="http://purl.org/net/untl/vocabularies/preservationEvents/#fixityCheck">http://purl.org/net/untl/vocabularies/preservationEvents/#fixityCheck</a></td>
<td>ark:/67531/codanqfg</td>
<td>success</td>
</tr>
<tr>
<td>832ac11f29b34a2bb0b24906ce91af39</td>
<td>March 10, 2015, 9:40 p.m.</td>
<td>⬪ <a href="http://purl.org/net/untl/vocabularies/preservationEvents/#replication">http://purl.org/net/untl/vocabularies/preservationEvents/#replication</a></td>
<td>ark:/67531/codanqfg</td>
<td>success</td>
</tr>
<tr>
<td>bf914f19f174ae5a2af6e142a8999c</td>
<td>March 10, 2015, 9:40 p.m.</td>
<td>⬪ <a href="http://purl.org/net/untl/vocabularies/preservationEvents/#fixityCheck">http://purl.org/net/untl/vocabularies/preservationEvents/#fixityCheck</a></td>
<td>ark:/67531/codanqfg</td>
<td>success</td>
</tr>
<tr>
<td>6474474d36af4590bcd0925b1169a0</td>
<td>March 10, 2015, 9:39 p.m.</td>
<td>⬪ <a href="http://purl.org/net/untl/vocabularies/preservationEvents/#replication">http://purl.org/net/untl/vocabularies/preservationEvents/#replication</a></td>
<td>ark:/67531/codanq3d</td>
<td>success</td>
</tr>
<tr>
<td>755d3a5e3af946fbb3b1e71d093d8a39</td>
<td>March 10, 2015, 9:39 p.m.</td>
<td>⬪ <a href="http://purl.org/net/untl/vocabularies/preservationEvents/#fixityCheck">http://purl.org/net/untl/vocabularies/preservationEvents/#fixityCheck</a></td>
<td>ark:/67531/codanq3d</td>
<td>success</td>
</tr>
<tr>
<td>8b4a6baa5b334b8caa0bc8a132b0efb1</td>
<td>March 10, 2015, 9:39 p.m.</td>
<td>⬪ <a href="http://purl.org/net/untl/vocabularies/preservationEvents/#replication">http://purl.org/net/untl/vocabularies/preservationEvents/#replication</a></td>
<td>ark:/67531/codanqfv</td>
<td>success</td>
</tr>
<tr>
<td>2c51154538ee438fa64f6db9858f9303</td>
<td>March 10, 2015, 9:39 p.m.</td>
<td>⬪ <a href="http://purl.org/net/untl/vocabularies/preservationEvents/#fixityCheck">http://purl.org/net/untl/vocabularies/preservationEvents/#fixityCheck</a></td>
<td>ark:/67531/codanqfv</td>
<td>success</td>
</tr>
<tr>
<td>f1e497a1cbfd409dab5f7434162ba4c</td>
<td>March 10, 2015, 9:37 p.m.</td>
<td>⬪ <a href="http://purl.org/net/untl/vocabularies/preservationEvents/#replication">http://purl.org/net/untl/vocabularies/preservationEvents/#replication</a></td>
<td>ark:/67531/codanq49</td>
<td>success</td>
</tr>
<tr>
<td>f2d9afa0dd20436fbe4deda256b3b361</td>
<td>March 10, 2015, 9:37 p.m.</td>
<td>⬪ <a href="http://purl.org/net/untl/vocabularies/preservationEvents/#fixityCheck">http://purl.org/net/untl/vocabularies/preservationEvents/#fixityCheck</a></td>
<td>ark:/67531/codanq49</td>
<td>success</td>
</tr>
<tr>
<td>Event Identifier</td>
<td>d14b893f5a3649748be4b7c32b93edcb</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td>----------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identifier Type</td>
<td><a href="http://purl.org/net/untl/vocabularies/identifier-qualifiers/#UUID">http://purl.org/net/untl/vocabularies/identifier-qualifiers/#UUID</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Type</td>
<td><a href="http://purl.org/net/untl/vocabularies/preservationEvents/#replication">http://purl.org/net/untl/vocabularies/preservationEvents/#replication</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Date</td>
<td>March 10, 2015, 9:40 p.m.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Details</td>
<td>Replication of object ark:/67531/codanqfg Using Harvester, revision 8718:8838</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date Added</td>
<td>March 10, 2015, 9:40 p.m.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linking Agent Identifier Value</td>
<td><a href="http://coda.library.unt.edu/agent/codaReplicationVerification">http://coda.library.unt.edu/agent/codaReplicationVerification</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linking Agent Identifier Type</td>
<td><a href="http://purl.org/net/untl/vocabularies/identifier-qualifiers/#URL">http://purl.org/net/untl/vocabularies/identifier-qualifiers/#URL</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linking Agent Role</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outcome</td>
<td><a href="http://purl.org/net/untl/vocabularies/eventOutcomes/#success">http://purl.org/net/untl/vocabularies/eventOutcomes/#success</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outcome Detail</td>
<td>Replication time for object ark:/67531/codanqfg: 0:02:21.445838 Object oxum was 4633589186.1910 Total data transfer time was 0:02:15.268115. Average throughput was 33452.0 KB/s</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linked Objects</td>
<td>ark:/67531/codanqfg</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ATOMPUB
We've operated this system for six years now.
Through one hardware cycle
1,107,681 bags

242.9 TB

119,877,725 files

Monthly Bags Added  Running Bags Total  Monthly Files Added  Running Files Total  Monthly Size Total  Running Size Total


0  30,000,000  60,000,000  90,000,000  120,000,000  150,000,000  180,000,000  210,000,000  240,000,000  270,000,000
Understanding Repository Growth at the University of North Texas: A Case Study

Mark E. Phillips  
University of North Texas  
Denton, Texas 76205  
mark.phillips@unt.edu

Lauren Ko  
University of North Texas  
Denton, Texas 76205  
lauren.ko@unt.edu

ABSTRACT
Over the past decade the University of North Texas Libraries (UNT) has developed a stable digital library infrastructure for use in carrying out its core mission to support teaching, learning, research, scholarship, and community engagement. However, the cost and complexity of the repository, itself, are often not understood. The advantages of the repository include: the ability to automate the function of publishing and sharing information, and the ability to perform complex queries and analyses that are not possible in other repositories. The disadvantages of the repository include: the ability to store large amounts of data, and the ability to perform complex queries and analyses that are not possible in other repositories. The advantages of the repository include: the ability to automate the function of publishing and sharing information, and the ability to perform complex queries and analyses that are not possible in other repositories.

1. INTRODUCTION
Understanding the characteristics and growth patterns in digital library, information retrieval, and user behavior is complex. In many situations, digital library programs advertise the size of their repository in relation to the number of items that are publicly available or the number of pages that provide the repository, such as those given by Google Scholar [1]. These metrics are often aggregated in databases such as the OpenDOAR system [2], which allows users to obtain an understanding of the size of a repository. These metrics, however, are not always easy to obtain and organizations such as OpenDOAR [3] have had to be a non-trivial amount of work to provide accurate statistics for record counts. The number of digital objects available in a repository is only a number of metrics that are important in understanding the growth of a digital repository over time. Additional metrics that may be of interest include the number of files that comprise a digital object, the size of the digital object, and the time that objects were added to the repository.

Another important factor to consider is that often the public facing side of a digital repository does not fully reflect the preservation information used to store the digital object files. In many organizations there are systems that maintain all files related to an object in one place, while other systems separate these systems and manage them differently. The UNT Libraries operates a number of public facing digital libraries, including the Portal to Texas History [4], the Digital Library [5], and The Gateway to Oklahoma History [6] operated by the Oklahoma Historical Society. Each of these public facing libraries makes available digital resources and associated metadata for consumption by users around the world. Digital objects are generally presented in Web deliverable formats, such as HTML, PDF, and PEP, and so on. These libraries have a number of public facing digital libraries that are not as active as the other public-facing libraries, but they are not as active as the other public-facing libraries.

When setting up this paper, there were 6,884,068 total digital objects in the public facing digital collections system with 642,133 currently accessible by users. In the digital archival repository, there are 68,998 digital objects occupying 56 TB of storage and accounting for 9,685,321 files. The analysis in this paper is based on the digital objects held in the UNT repository.

2. RESEARCH QUESTIONS
The following research questions drive this analysis:

- What are the growth trends in digital content acquisition for the UNT Libraries Digital Collections?
- What are the average characteristics in relation to size, scale, and frequency of ingest of digital objects?
- How has the percentage change been driven by the UNT Libraries Digital Collections?

3. BUILDING THE DATASET
The UNT Libraries stores its primary digital objects in a purpose built digital archival system developed by the UNT Libraries called Codex. This system is responsible for ingest, monitoring, and dissemination of the rich digital content files that make up the UNT Libraries Digital Collections. The Codex system has been in full operation at UNT since September 2009 when
Current Project
Working with University of Florida to pilot a peer-reviewed option for the TRAC process
Each institution prepares a full TRAC self audit
And then we each review/critique these audits.
The goal is to have something that balances between a self audit and a full external certification.
Goals going into the project was transparency with all documentation being released at the end of the process.
Started in Sept 2014
Scheduled to complete in Sept 2015
Huge amount of work from both institutions
In closing
Leveraging digital library infrastructure
Creating a sustainable program
The “how” becomes a known answer
Like digital stacks
Allowing us to focus on the content
And the connections people make with it.
Thank you

mark.phillips@unt.edu

http://digital.library.unt.edu

http://texashistory.unt.edu/