Launching the Texas Data Repository
How to Implement TDR at Your Institution
Kristi Park, Director, Texas Digital Library
Santi Thompson, University of Houston
Today’s agenda

1. Background of the Texas Data Repository
2. Capabilities of Dataverse
3. How the Service Works
4. Local Implementation
5. Wrap Up & Questions
Part 1: Background
The Texas Digital Library is a consortium of Texas higher education institutions that builds capacity for preserving, managing, and providing access to unique digital collections of enduring value.
Goals for TDR:
Provide infrastructure (and other) support that will help researchers:
• comply with funding mandates
• receive greater recognition for research data and researchers (DOIs, discoverability)
• produce better research (more reproducible, more efficient)
Goals for TDR (cont.):

Provide infrastructure (and other) support that will help libraries:

- Become an integral part of the Research Data Lifecycle.
- Provide services at whatever level they are able
- Benefit from collaborative infrastructure/management and mutual support.
- Showcase institutional collections of research data
TDL Data Management Working Group

**Charge:** Help the TDL determine what kinds of data management services it could provide at a consortial level.

- Develop criteria
- Evaluate proposed projects
- Investigate issues
- Document findings
- Make recommendations for services

Bruce Herbert, Texas A&M (Chair)

Maria Esteva (TACC); Colleen Lyon (UT Austin); Jeremy Donald (Trinity); Martha Buckbee (UT Southwestern); Christie Peters, Santi Thompson (UH); Kristi Park, Ryan Steans (TDL)
TDL Dataverse Implementation Working Group

**Charge:** Pilot test, assess, and launch a consortial repository for research data archiving and management.

**Members:** Denyse Rodgers (Baylor); Bruce Herbert, Sean Buckner, Wendi Kaspar, Cecilia Smith (TAMU); Ray Uzwyshyn, Todd Peters (Texas State); Christopher Starcher (Texas Tech); Jeremy Donald (Trinity); Kristi Park, Ryan Steans, Nick Lauland, Laura Waugh (TDL)
Group timeline

- September 2015: Working Group formed
- April-May 2016: Repository Pilot Project
- September 2016: Soft Launch
- November 15-16: Data Symposium/TDR training
- December 2016: Formal Launch
Texas Data Repository

- Centrally hosted, collectively managed
- Built in open-source Dataverse
- De-identified data only

Search the Texas Data Repository

Add a Dataset  Create a Dataverse  Explore Data Repository  Learn More  Get Help

Publish and Track Your Data, Discover and Reuse Others' Data!

POWERED BY Dataverse

Texas Digital Library
Part 2: Capabilities of Dataverse
Benefits Overview

- Data storage and sharing
- Data versioning
- Permissions and Terms of Use
- Measuring impact and tracking use
- Discipline-based metadata schema
- Data visualization
Dataverse and Datasets

Schematic Diagram of a Dataverse in Dataverse 4.0

Container for your Datasets and/or Dataverses*

* Dataverses can now contain other Dataverses (this replaces Collections & Subnetworks)
Add Data
Organize, publish, and archive

Share Data
Share data with trusted group

Version Data
Maintain multiple versions of data

Share, publish, and archive your data. Find and cite data across all research fields.

Welcome to the Texas Data Repository, a statewide archive of research data from Texas Digital Library (TOL) member institutions.

QUICK LINKS
- Add a dataset
- Create a dataverse
- Explore data

LEARN MORE
- Go to the user guide.
- Contact a local university librarian for help.

Search this dataverse... Find Advanced Search
Texas Data Repository Dataverse > John’s Dataverse > John Doe Dataset

John Doe Dataset [Draft]

Doe, John, 2016. "John Doe Dataset", http://dx.doi.org/10.18738/T8/C10DDX, Texas Data Repository Dataverse, DRAFT VERSION

If you use these data, please add this citation to your scholarly resources. Learn about Data Citation Standards.

Description
This is John Doe's dataset. This is an example of editing the metadata.

Subject
Arts and Humanities

Files | Metadata | Terms | Versions

| DRAFT | Citation Metadata: Description (1 Changed); View Details | John Doe | August 16, 2016 |
| 2.0   | Files (Added: 1); Terms of Use/Access Changes; View Details | John Doe | August 16, 2016 |
| 1.0   | This is the first published version. | John Doe | August 15, 2016 |

View Differences
**Test GIS Dataset: Bikeways of College Station, TX**

Smith, Cecilia, 2015, "Test GIS Dataset: Bikeways of College Station, TX", [http://dx.doi.org/10.5072/FK2/PQIEBO](http://dx.doi.org/10.5072/FK2/PQIEBO), Texas Research Data Repository Dataverse, V3 [UNF.6.pPsX4Snp7w07D0jpwKaz2w==]

If you use these data, please add this citation to your scholarly resources. Learn about Data Citation Standards.

<table>
<thead>
<tr>
<th>Description</th>
<th>GIS shapefile of existing and proposed bike lanes and paths for the city of College Station, TX. The shapefile is composed of 7 files.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>Social Sciences; Other</td>
</tr>
<tr>
<td>Keyword</td>
<td>GIS, Bikeways</td>
</tr>
<tr>
<td>Notes</td>
<td>This is a test GIS dataset. Originally produced by the city of College Station, it is now in the public domain and freely distributable</td>
</tr>
</tbody>
</table>

**Terms**

<table>
<thead>
<tr>
<th>Files</th>
<th>Metadata</th>
<th>Terms</th>
<th>Versions</th>
</tr>
</thead>
</table>

**Terms of Use**

Our Community Norms as well as good scientific practices expect that proper credit is given via citation. Please use the data citation above, generated by the Dataverse.

**Waiver**

CC0 - "Public Domain Dedication"
Institute of Classical Archaeology Dataverse (University of Texas at Austin, Institute of Classical Archaeology)

Spatial extents of ICA projects

Institute of Classical Archaeology, 2016, "Spatial extents of ICA projects", doi:10.18738/T8/ZGS0J3, Texas Data Repository Dataverse, V1

Description
This dataset represents the rough location and spatial extents of ICA's field research projects carried out between 1974 and 2013 in the agricultural territories of Metaponto and Croton in southern Italy and Chersonesos in Crimea.

Subject
Arts and Humanities

Keyword
archaeology, archaeological survey, archaeological excavation, Classical Archaeology

Notes
This data layer was compiled from a variety of sources (at a wide range of scales and accuracies) over the course of 20 years. These data are only for reference and illustrative purposes, and should not be used for analysis.
<table>
<thead>
<tr>
<th>Dataset Persistent ID</th>
<th>doi:10.18738/T8/Q2iT68</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publication Date</td>
<td>2016-05-03</td>
</tr>
<tr>
<td>Title</td>
<td>Data from excavations of the Late Roman farmhouse at San Biaio</td>
</tr>
<tr>
<td>Author</td>
<td>Institute of Classical Archaeology (The University of Texas at Austin)</td>
</tr>
<tr>
<td>Contact</td>
<td>Use email button above to contact. Telogian, Jessica (UT Libraries)</td>
</tr>
<tr>
<td>Description</td>
<td>This dataset includes raw field observation data from excavations at the farmhouse at San Biagio conducted by ICA in 1980 in addition to the results of study done in preparation for publication in The Chora of Metaponto 4: The Late Roman Farmhouse at San Biagio, published by UT Press in 2012. (1960)</td>
</tr>
<tr>
<td>Subject</td>
<td>Arts and Humanities</td>
</tr>
<tr>
<td>Keyword</td>
<td>Late Roman archaeology, Magna Grecia</td>
</tr>
<tr>
<td>Production Date</td>
<td>1980</td>
</tr>
<tr>
<td>Production Place</td>
<td>Metaponto, Italy</td>
</tr>
<tr>
<td>Depositor</td>
<td>Telogian, Jessica</td>
</tr>
<tr>
<td>Deposit Date</td>
<td>2016-05-03</td>
</tr>
<tr>
<td>Kind of Data</td>
<td>field observations</td>
</tr>
</tbody>
</table>
Email TDL Dataverse Support

To: The Chora of Croton Dataverse Contact

From: Support@tdl.org

Subject *

Message *

Please fill this out to prove you are not a robot. *

3 + 5 = 

Send Message  Cancel
Other Benefits

- Integration with OJS
- DOI and Citation Metadata Aggregated by SHARE
- APIs
Secure infrastructure

• Amazon Web Services
• Controlled Access
• Data back ups
Part 3: How the Service Works
Scope & Collecting policies

- Research data and products
- Any discipline; any file type
- Midsized-data set size
- Free of confidential or sensitive information
Texas Digital Library (technology)

Steering Committee (TDL & Data Repository Librarians)

Member Libraries (service & outreach)

Researchers (deposit, search, publish)
Share, publish, and archive your data. Find and cite data across all research fields.

Welcome to the Texas Data Repository, a statewide archive of research data from Texas Digital Library (TDL) member institutions.

QUICK LINKS

- Add a dataset
- Create a dataverse
- Explore data

LEARN MORE

- Go to the user guide.
- Contact a local university librarian for help.
Part 4: Implementation
Contracts and Cost

- Launch TDR under existing contracts with members, using an MOU to articulate the roles and responsibilities
- Member institutions must sign an MOU to use the repository in 2016-17
- No additional charge for use of the repository in the first year (2016-17)
- Limits on storage: 2GB per file and 10GB per project
Memorandum of Understanding

- **Purpose:** To clearly identify the roles and responsibilities of the member institution and Texas Digital Library (TDL) as they relate to the Texas Data Repository.
- **Definitions**
- **Roles & Responsibilities**
  - Member Institutions
  - Data Repository Librarians
  - Texas Digital Library
- **Ownership** (UT Austin as lead agency for TDL)
- **Signatures** (Library Administrator and Data Repository Librarian)
Texas Digital Library
(technology)

Steering Committee (TDL & Data Repository Librarians)

Member Libraries (service & outreach)

Researchers (deposit, search, publish)
<table>
<thead>
<tr>
<th>TDL</th>
<th>Technology implementation, hosting, backup</th>
<th>Coordination of steering committee</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training &amp; community development</td>
<td>Technical support (helpdesk)</td>
</tr>
</tbody>
</table>

**Steering Committee:** composed of Data Repository Librarians, provides advice and coordination of repository service

<table>
<thead>
<tr>
<th>Member Libraries</th>
<th>Data Repository Librarian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identification of “Data Repository Librarian”</td>
<td>Serve as local liaison for repository users (registry of local experts)</td>
</tr>
<tr>
<td>Local Campus Outreach</td>
<td>Maintain institutional collections of data (i.e. an “institutional dataverse”)</td>
</tr>
<tr>
<td>Communication with TDL on institutional requirements (e.g. liaise with campus IT)</td>
<td>Serve on repository Steering Committee</td>
</tr>
<tr>
<td>Establishment of relevant institutional policies</td>
<td>Optionally, “value added” services to campus researchers assisting with data curation, data management planning, etc.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Researchers</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Deposit, share, and/or publish data</td>
<td>Adhere to TDR Terms of Use</td>
</tr>
<tr>
<td>Maintain self-deposited research data</td>
<td></td>
</tr>
</tbody>
</table>
Authentication and User Accounts

Researchers/depositors will log in via local institutional accounts (via Shibboleth)

- Default permission to deposit, publish, create “dataverses” in:
  - Institutional dataverse
  - “Root” dataverse
  - Not other institution’s dataverses
Other types of user accounts

- **Data Repository Librarians** will log in via specially created accounts.

- Researchers can invite non-affiliated researchers to create accounts for sharing/collaboration. These user accounts do not have deposit permissions by default.
Repository Training

• TDL Data Symposium (November 15-16)
• December webinar
• User guide (http://data.tdl.org/user-guide/)
Part 5: Wrapping Up
More information

Texas Data Repository Website

Information Sheets


Other information

• User guide: http://data.tdl.org/user-guide/
• Policies (external): http://data.tdl.org/policies/

TDL Dataverse Implementation Working Group Report: http://hdl.handle.net/2249.1/76364
November 2016: TDL Data Symposium

On November 15-16, 2016, the TDL will hold its first ever Data Symposium at Baylor University Libraries. The two-day Symposium is designed to support the development of a core community of TDL member librarians who provide (or intend to provide) research data services on their campuses. Along with opportunities to discuss and learn about research data management generally, the Symposium will offer a half-day training workshop on Dataverse and the Texas Data Repository service.

December 2016: Formal launch webinar

After working throughout the fall of 2016 to onboard and train participating libraries, the TDL will formally launch the Texas Data Repository in December with a press release and a second webinar for TDL members (date TBD).
Looking Ahead

- Digital preservation
- Integrating with other TDL Applications
- Including data into local discovery systems
- Accommodating “big” data
- Re-architecting to scale
Get started

✓ Contact the TDL to begin implementation for your institution (support@tdl.org).
✓ Get TDL in contact with relevant IT personnel to implement Shibboleth.
✓ Review the documentation on the Texas Data Repository site at https://data.tdl.org.
✓ Review and sign the TDR MOU.
Questions and Discussion