Launching the Texas Data Repository

How to Implement TDR at Your Institution

Kristi Park, Director, Texas Digital Library
Santi Thompson, University of Houston

Texas Digital Library

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

Texas Digital Library

Part 1: Background

Texas Digital Library

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.

Texas Digital Library

Goals for TDR:

Provide infrastructure (and other) support that will help ***researchers***:

- **comply** with funding mandates
- receive greater **recognition** for for research data and researchers (DOIs, discoverability)
- produce **better research** (more reproducible, more efficient)

Texas Digital Library

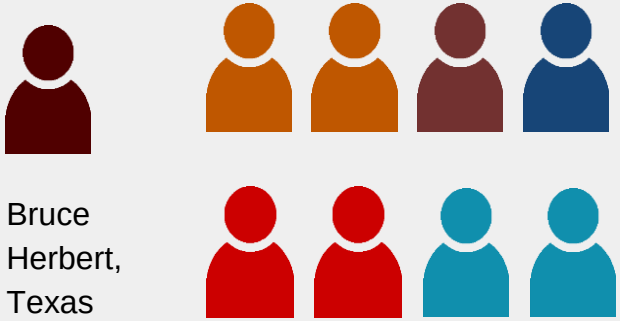
Goals for TDR (cont.):

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

Texas Digital Library

TDL Data Management Working Group



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)

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

Texas Digital Library

TDL Dataverse Implementation Working Group

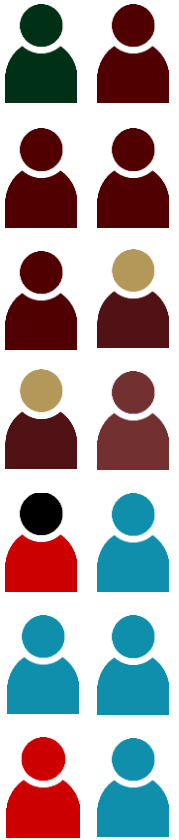
Charge: Pilot test, assess, and launch a consortial repository for research data archiving and management.

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)

14
Working Group members

7
TDL member institutions

1
Santi Thompson (Chair)

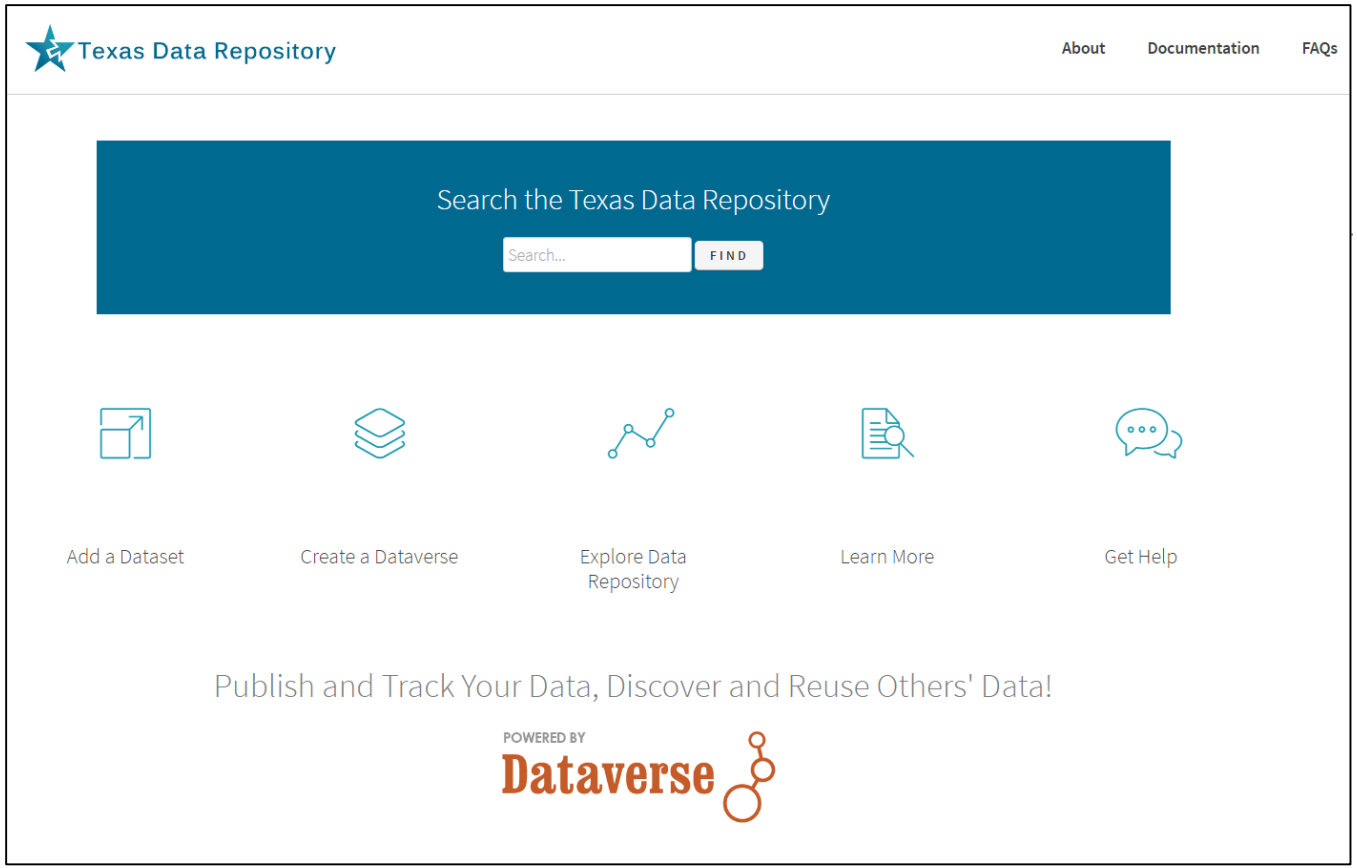


Texas Digital Library

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 Digital Library



Texas Data Repository

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

Texas Digital Library

Part 2: Capabilities of Dataverse

Texas Digital Library

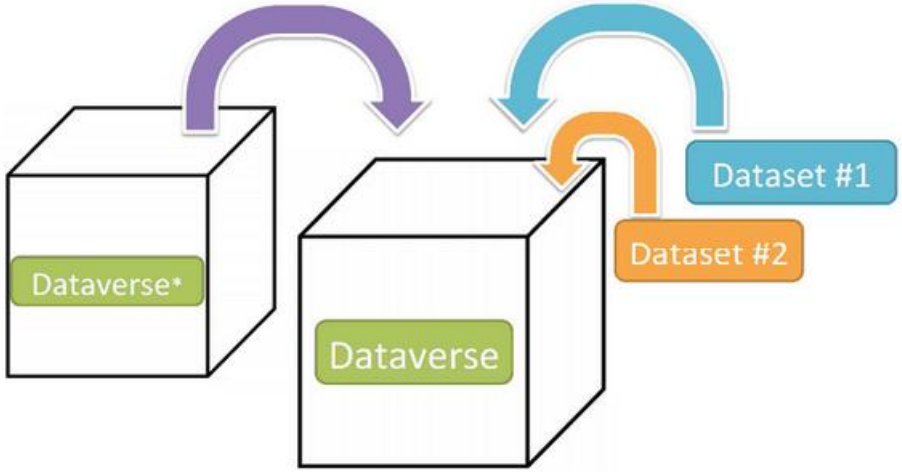
Benefits Overview

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

Texas Digital Library

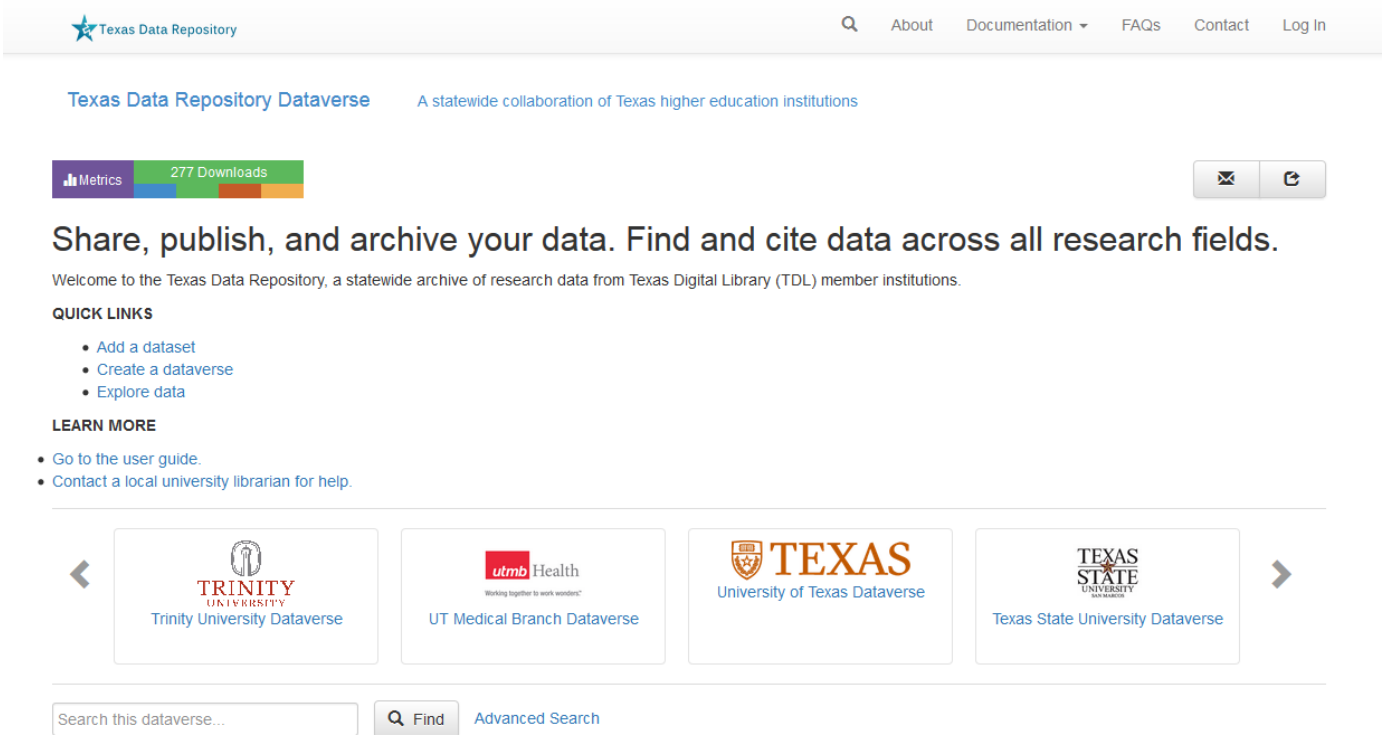
Dataverse and Datasets

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

John's Dataverse

Texas Data Repository Dataverse > John's Dataverse > **John Doe Dataset**

Metrics 0 Downloads

Publish Edit

John Doe Dataset Draft

Doe, John, 2016, "John Doe Dataset", http://dx.doi.org/10.18738/T8/CI0DDX, Texas Data Repository Dataverse, DRAFT VERSION

Download Citation

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

View Differences

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

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, Texas Research Data Repository Dataverse, V3 [UNF:6:pPsX4Snp7wO7DoipwKaz8w==]

Download Citation

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

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

Files | Metadata | Terms | Versions

Terms of Use

Waiver

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.

CC0 - "Public Domain Dedication"

Institute of Classical Archaeology Dataverse (University of Texas at Austin, Institute of Classical Archaeology)

Metrics 3 Downloads

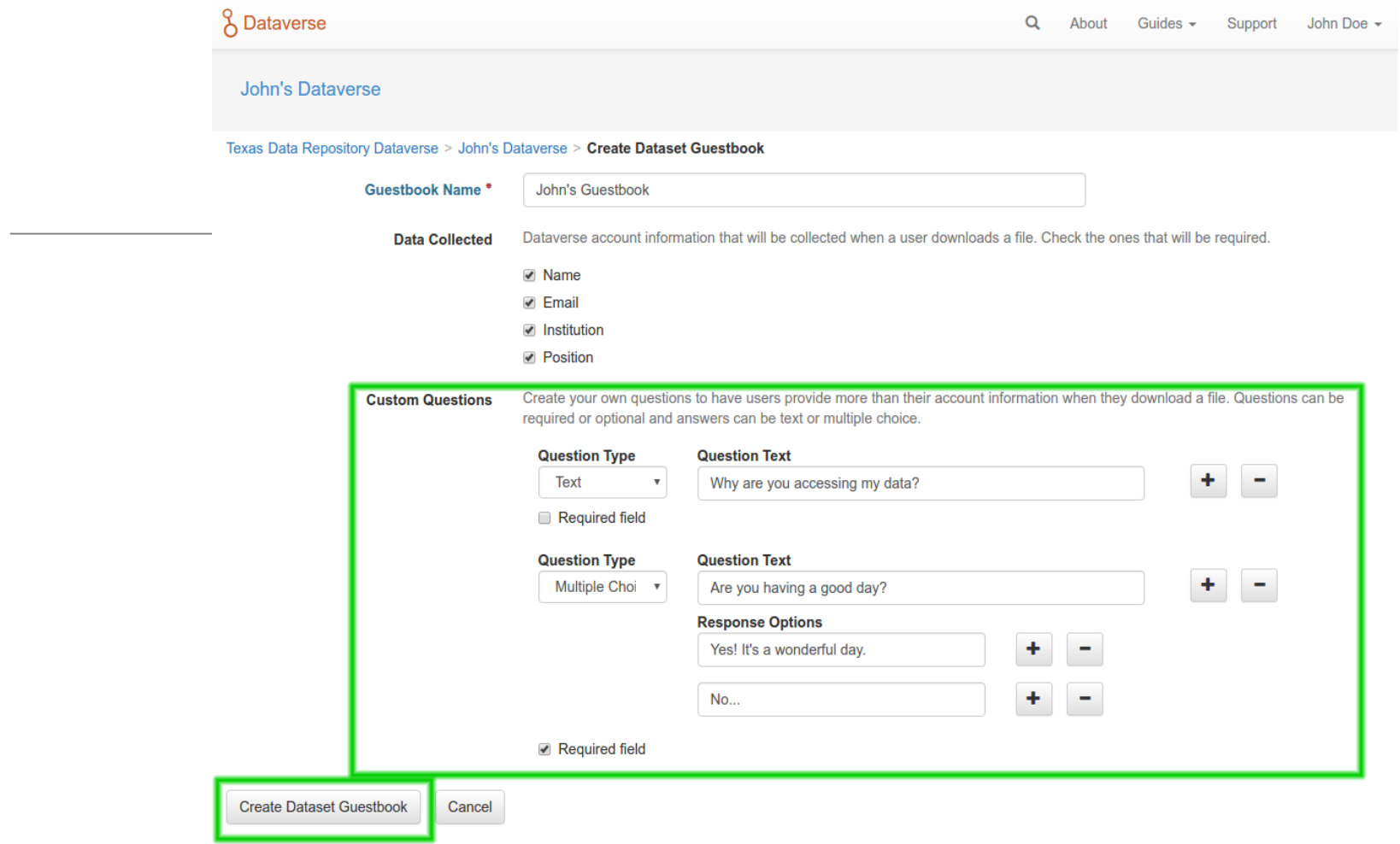
Spatial extents of ICA projects

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

Cite Data

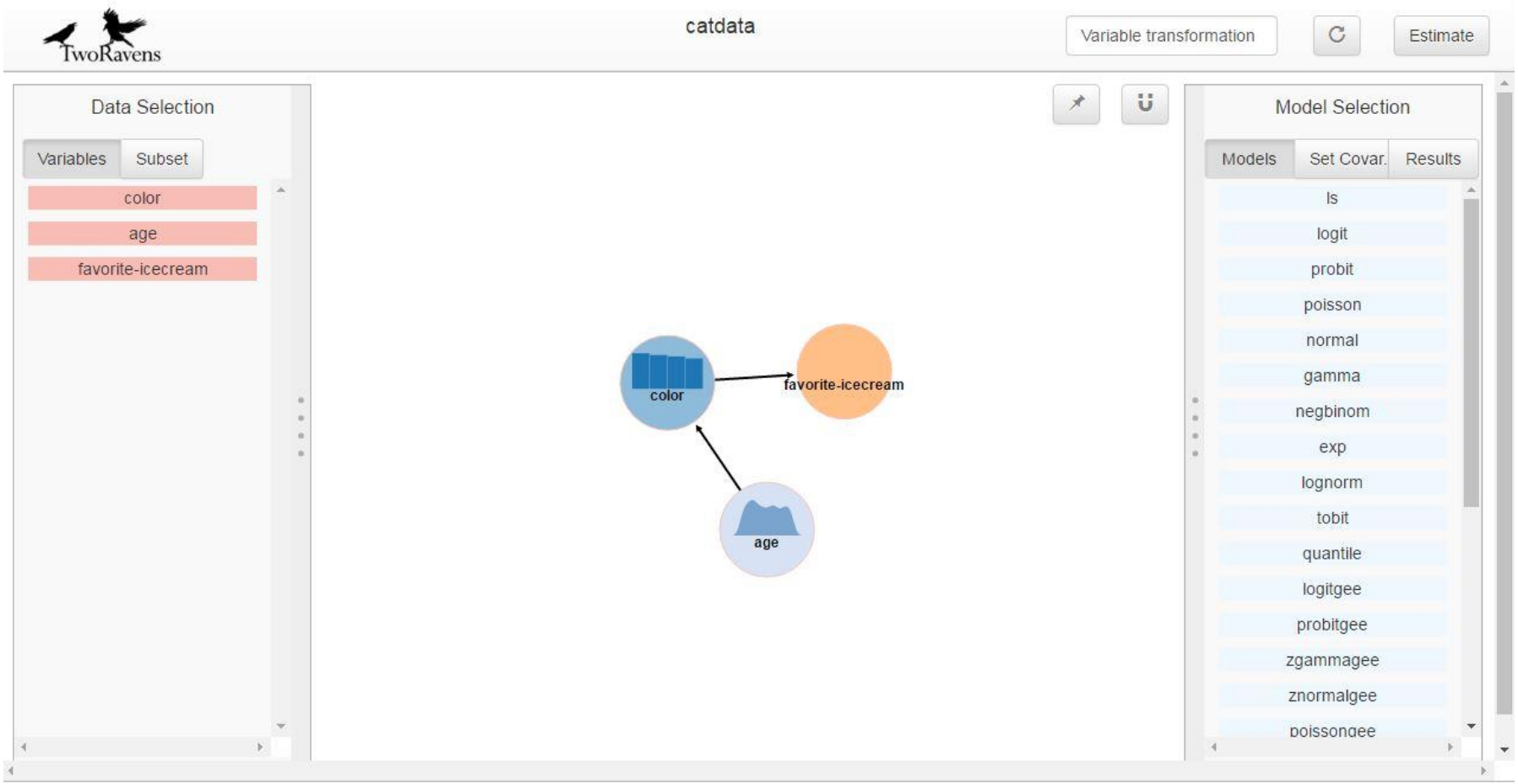
If you use these data, please add this citation to your scholarly resources. Learn about Data Citation Standards.

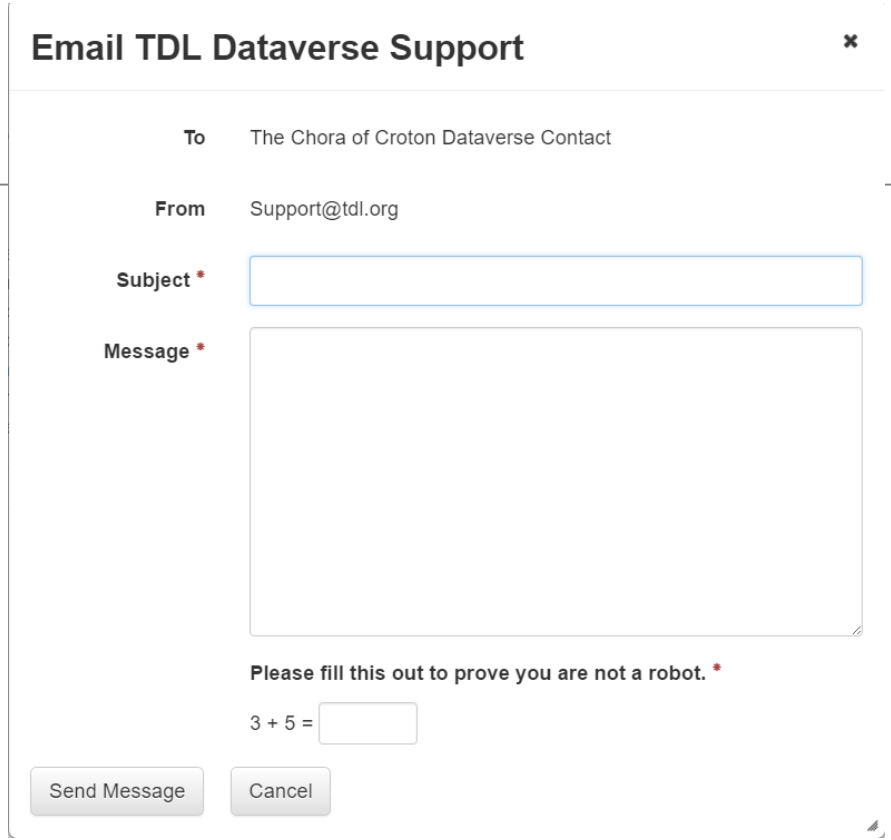
Description	This dataset represents the rough location and spatial extents of ICA's field research projects carried out between 1974 and 2013 in the agricultural territories of Metaponto and Croton in southern Italy and Chersonesos in Crimea.
Subject	Arts and Humanities
Keyword	archaeology, archaeological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


Dataverse
About
Guides
Support
John Doe
John's Dataverse
Texas Data Repository Dataverse > John's Dataverse > Create Dataset Guestbook
Guestbook Name *
John's Guestbook
Data Collected
Dataverse account information that will be collected when a user downloads a file. Check the ones that will be required.
Name
Email
Institution
Position
Custom Questions
Create your own questions to have users provide more than their account information when they download a file. Questions can be required or optional and answers can be text or multiple choice.
Question Type
Text
Question Text
Why are you accessing my data?
Required field
Question Type
Multiple Choi
Question Text
Are you having a good day?
Response Options
Yes! It's a wonderful day.
No...
Required field
Create Dataset Guestbook
Cancel

Citation Metadata

Dataset Persistent ID	doi:10.18738/T8/Q2IT68
Publication Date	2016-05-03
Title	Data from excavations of the Late Roman farmhouse at San Biatio
Author	Institute of Classical Archaeology (The University of Texas at Austin)
Contact	Use email button above to contact. Trelogan, Jessica (UT Libraries)
Description	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. (1980)
Subject	Arts and Humanities
Keyword	Late Roman archaeology Magna Grecia
Production Date	1980
Production Place	Metaponto, Italy
Depositor	Trelogan, Jessica
Deposit Date	2016-05-03
Kind of Data	field observations


TwoRavens
catdata
Variable transformation
Estimate
Data Selection
Variables
Subset
color
age
favorite-icecream
color
favorite-icecream
age
Model Selection
Models
Set Covar.
Results
ls
logit
probit
poisson
normal
gamma
negbinom
exp
lognorm
tobit
quantile
logitgee
probitgee
zgammagee
znormalgee
poissongee



Texas Digital Library

Other Benefits

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

Texas Digital Library

Secure infrastructure

- Amazon Web Services
- Controlled Access
- Data back ups

Texas Digital Library

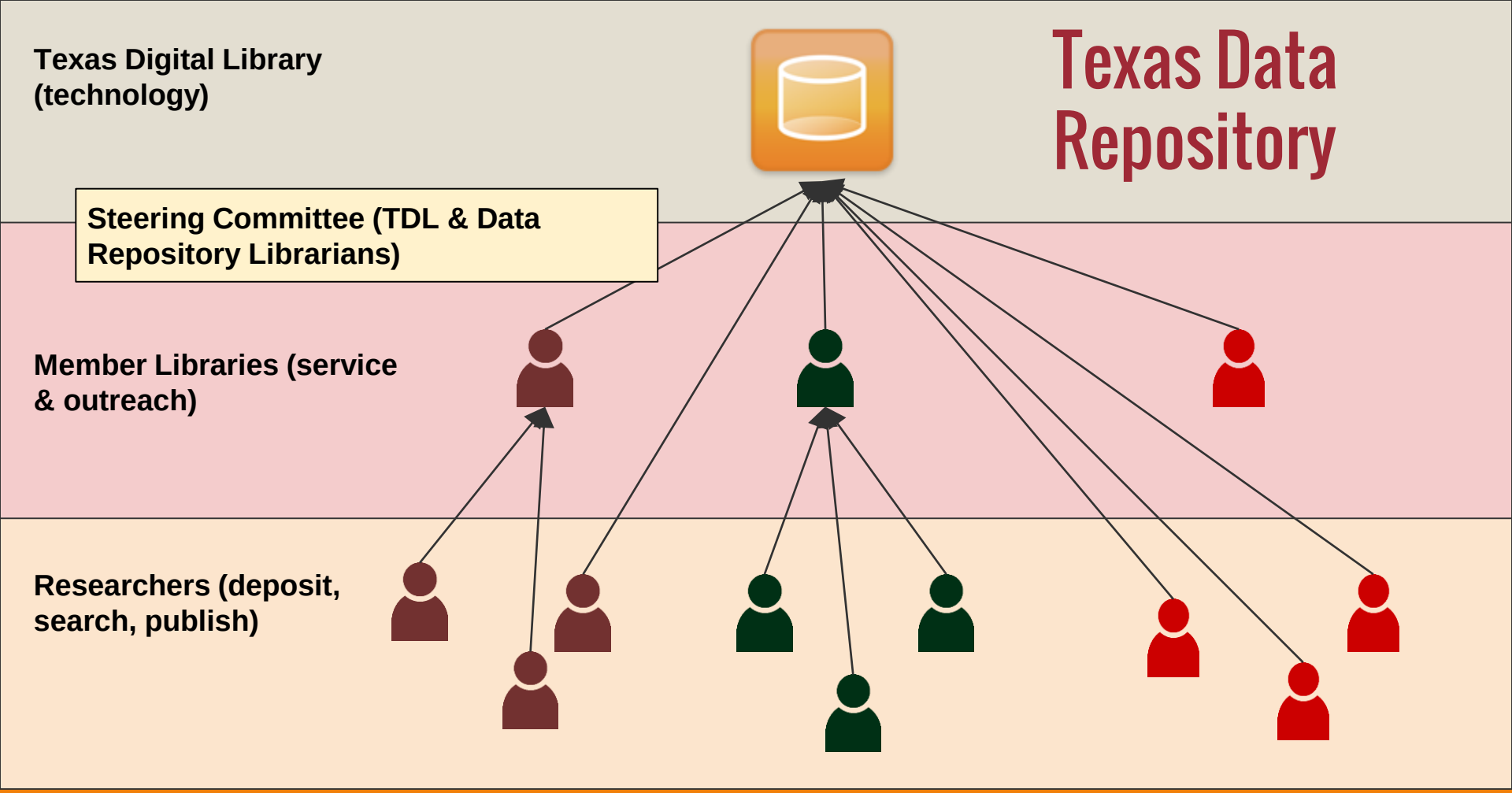
Part 3: How the Service Works

Texas Digital Library

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


Texas Digital Library (technology)
Texas Data Repository
Steering Committee (TDL & Data Repository Librarians)
Member Libraries (service & outreach)
Researchers (deposit, search, publish)

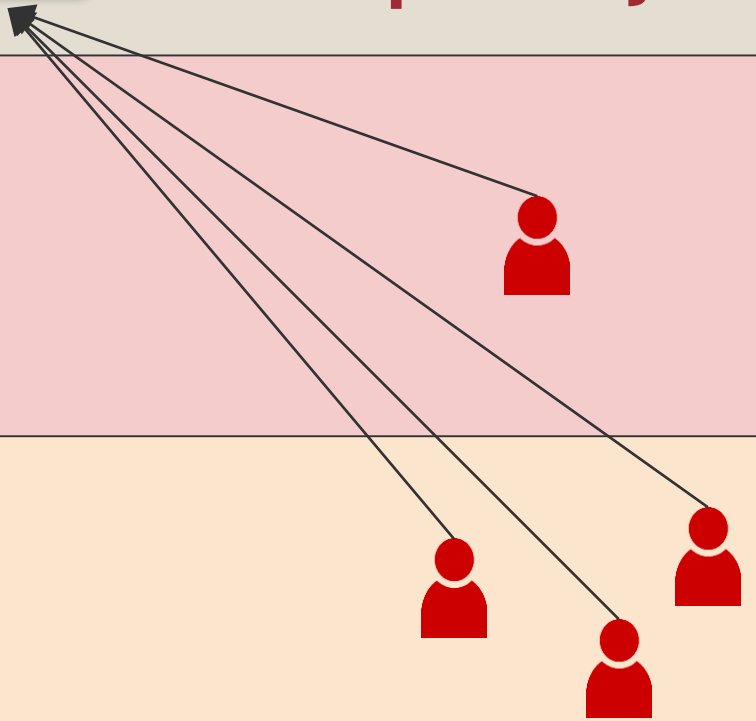
Texas Digital Library (technology)

Texas Data Repository

Steering Committee (TDL & Data Repository Librarians)

Member Libraries (service & outreach)

Researchers (deposit, search, publish)



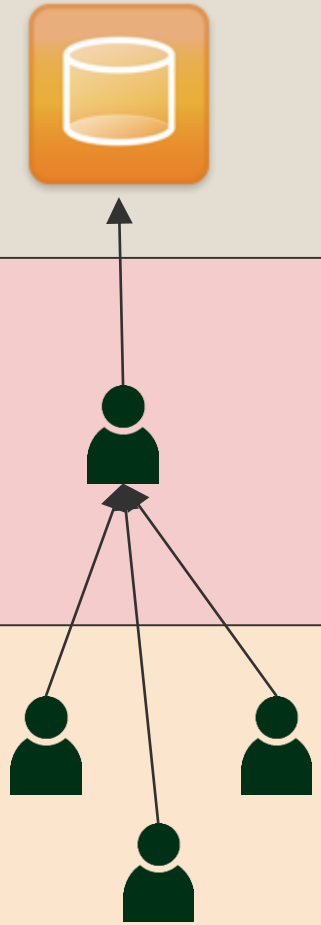
Texas Data Repository

Texas Digital Library (technology)

Steering Committee (TDL & Data Repository Librarians)

Member Libraries (service & outreach)

Researchers (deposit, search, publish)



Texas Data Repository Dataverse A statewide collaboration of Texas higher education institutions

Metrics | 172 Downloads

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.

Texas State University Dataverse

Texas A&M University Dataverse

Angelo State University Dataverse

Baylor University Dataverse

Part 4: Implementation

Texas Digital Library

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

Texas Digital Library

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

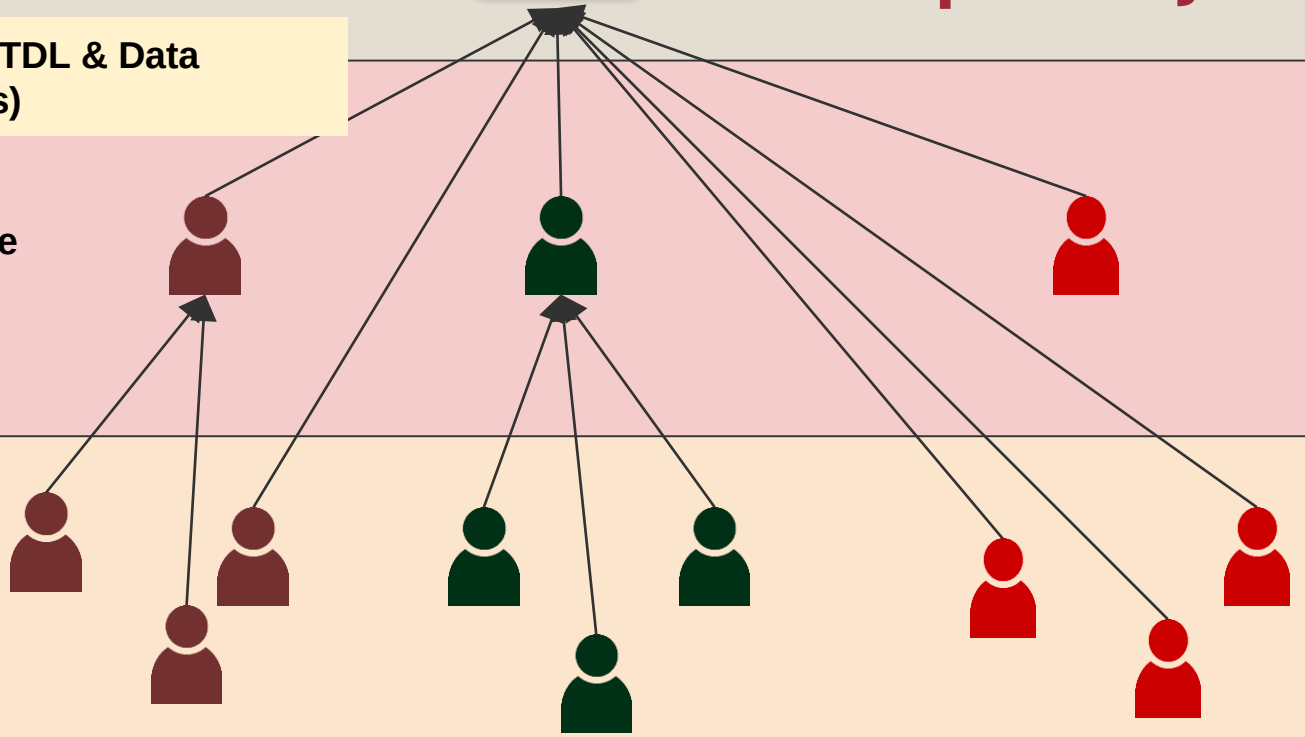
Texas Digital Library (technology)

Texas Data Repository

Steering Committee (TDL & Data Repository Librarians)

Member Libraries (service & outreach)

Researchers (deposit, search, publish)



TDR Roles & Responsibilities

TDL

- Technology implementation, hosting, backup
- Training & community development
- Coordination of steering committee
- Technical support (helpdesk)

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

Member Libraries

- Identification of “Data Repository Librarian”
- Local Campus Outreach
- Communication with TDL on institutional requirements (e.g. liaise with campus IT)
- Establishment of relevant institutional policies

Data Repository Librarian

- Serve as local liaison for repository users (registry of local experts)
- Maintain institutional collections of data (i.e. an “institutional dataverse”)
- Serve on repository Steering Committee
- Optionally, “value added” services to campus researchers assisting with data curation, data management planning, etc.

Researchers

- Deposit, share, and/or publish data
- Maintain self-deposited research data
- Adhere to TDR Terms of Use

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

Texas Digital Library

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.

Texas Digital Library

Repository Training

- TDL Data Symposium (November 15-16)
- December webinar
- User guide (http://data.tdl.org/user-guide/)

Texas Digital Library

Part 5: Wrapping Up

Texas Digital Library

More information

Texas Data Repository Website

Information Sheets

- Info Sheet For Researchers: http://data.tdl.org/wp-content/uploads/2016/09/Researcher-Information-Sheet.pdf
- Info Sheet For Librarians: http://data.tdl.org/wp-content/uploads/2016/09/Librarian-Information-Sheet.pdf
- Info Sheet For Administrators: http://data.tdl.org/wp-content/uploads/2016/09/Administrator-Information-Sheet.pdf

Other information

- User guide: http://data.tdl.org/user-guide/
- Metadata dictionary: http://data.tdl.org/wp-content/uploads/2016/09/TDR-Metadata-Dictionary.pdf
- Policies (external): http://data.tdl.org/policies/

TDL Dataverse Implementation Working Group Report: http://hdl.handle.net/2249.1/76364

Texas Digital Library

More information, cont.

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).

Texas Digital Library

Looking Ahead

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

Texas Digital Library

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.

Texas Digital Library

Questions and Discussion