Developing an Open Source Digital Scholarly Research Ecosystem
Local and Global Possibilities

Ray Uzwyshyn, Ph.D. MLIS
Director, Collections and Digital Services
Texas State University Libraries
What is a Digital Scholarly Research Ecosystem?

Ecosystem of Several Software Components to Enable Faculty and Student Research

Ecosystem Metaphor Look at Relationships in the Digital Environment
Specifically Focuses Upon the Discrete Component Relationships with the Networked Digital Environment
Six Main Software Components

- Digital Collections Repository (Dspace)
- Research Data Repository (Dataverse)
- Identity Management System (ORCID)
- ETD Management System (VIREO)
- User Interface Software (OMEKA)
- Open Journal Software (OJS3)

Hardware: Digitization Lab
Developed and Implemented
Texas State University Libraries, 2014-2019
General Characteristics
Digital Scholarly Research Ecosystem

- Open Source Software
- Customizable Components
- Active Developer Communities
Collocating Digital Components in Networked Research Ecosystem Enables Connections and/or Larger Network Effects
Network Effects: Metcalfe’s Law
Early Telecommunications Law for Ethernet (1993)

The Systemic Value of Compatibly communicating components grows as the square of their number increases.

Component Networks may be Internal and/or External
Texas State University Libraries
Digital Scholarly Research Ecosystem
Primary Components
Institutional Digital Collections Repository (Dspace)

Organizes, centralizes and makes accessible research and knowledge generated by the institution’s research community:

- Pre-prints
- Faculty Publications
- White Papers
- Conference Presentations
- Graduate Student Theses and Dissertations
- Historical Legacy Application
Primary Use Case Value
Application of Structured Metadata Schema for Search Engine Optimization
Enabling Accessibility and Multiple Points of Access
Efficacy of Structured Metadata Schema Application for Search Engine Optimization Accessibility and Multiple Points of Access
Digital Collection Repositories Gives Insight and another window into Faculty/Student Research (Statistics)
Research Data Repository

Texas State University Dataverse
A platform for publishing and archiving Texas State University's research data.

CAPTURE
Project Data from Experiments, Surveys Researchers and Scientists

CATALOG
Assign Metadata Schema, Specialized and Disciplinary Taxonomies, DOI, UNF

MANAGE
Administrative Online Research Data Archives

FIND/VIEW
Retrieve, Download Relevant Data Sets Instantaneously

Synthesize Research
Verification, Insight, Discovery Visualization, Harvesting and Linked Data
Texas State University Dataverse/TDL Dataverse
Larger Consortial Online Research Data Repository

Dataverse Architecture
(Consortial)
Digital Scholarly Research System
Secondary Components
Vireo, Omeka and OJS  
(Dependent on Primary Content Repositories)

Electronic Thesis and Dissertation Management System  
Addresses Intermediary steps in the ETD Process  
Bridges Student Thesis/Dissertation Submission with  
Graduate School Review, Online Publication and ETD Preservation

Open Source User Interface Software  
allows an elegant portal or gateway entrance to digital collections  
data repositories, large research projects - linking text, image  
media and datasets

Open Access Academic Journal Software for the academic refereed journal workflow and online publishing
Researcher Identity Management System

ORCID

- Allows publications from a researcher to be found, linked and aggregated across multiple information systems
- Gives Researchers Unique Number (ORCID ID)
  Connecting and Disambiguate Scholars names
  Maria Hernandez, Biochemist
  Maria Hernandez, M.D. or Astrophysicist
- Can also act as a Network Hub

ORCID is a hub connecting the research landscape
Digital Scholarly Research System
Tertiary Components
The Digitization Lab Hardware & Specialized Software
The Digitization Lab
Expands Possibilities for Faculty Research Projects

Digitization possibilities on media levels range from OCR to image, book, manuscript & journal digitization, 3D objects, posters, audiovisual material maps, GIS and visualization technologies (IIIF etc)
Combining Components
System Synergies
Digital Scholarly Research Ecosystem
Combining These Research Ecosystem Components
Opens Possibilities For Digital Scholarship & Partnership Opportunities

- **Cognitive Cartography/Multimedia Archives**
  - (Video, Text, GIS, Images, Field Notes)
  - **Dick Reavis: National Tour of Texas**

- **Multimedia, Digital Archives/Retrospective ETD Projects**
  - (Digital video, online exhibit images, text, digital archives)
  - **Severo Perez: And the Earth Did not Swallow Them**

- **Online Exhibits/Digital Archives/ Online Academic Journals**
  - (Images and text, Omeka front end/Database back end, IIIF)
  - **Cabeza de Vaca La Relacion Digitization**
  - **Santiago Tafolla: Mexican Amer. Confederate Soldier**

- **Interactive Image Archives/Data & Research Projects**
  - (Image libraries, Interactive Commenting/Metadata)
  - **Texas State Flickr Commons**

- **Digital Libraries Archiving & Documentation Projects**
  - (Text, Metadata, OCR, Search, Zoom ability, Page Turning)
  - **Pedagogs University Yearbooks**

- **Faculty Digitization Proposals/Partnerships**

---

**Projects,Prototypes Grant Partnerships**

**Complexity**
Ecosystem Open Source Software Enables Core Research

- Articles, Theses, Dissertations in the collections repository can be associated with datasets in the data repository for reference, verification or reproducibility.

- Journal article citation lists can be associated with articles and datasets in the Collections and Data Repositories

- Papers in the collections repository and datasets in data repository can be associated with ORCID ID’s for aggregation of research profiles. Also, the University’s Faculty Profile Systems (Digital Measures)

- Further Desired Connections can also guide developmental paths for both component software and the ecosystem
Network Effects
Both In and Between Individual Components and In and Among Component Networks

1) ORCID Aggregates from Several Sources and Networks and Connects to Other Networks, Internal and External
2) OMEKA can act as a middleware front end connecting several components and component networks internally.
3) Digitization Lab’s IIIF Framework can create internal or globally distributed Image Libraries.
4) Dataverse can be configured as a single Instance or as a Consortial Model (Texas 22 Individual Instances, TDL)
Network Effects and Opportunities Among Research Institutions

Research Institution A
- The Dataverse Project
- Identity Management System
- DSPACE Online Institutional Digital Collections Repository
- OJS Online Journal System

Research Institution B
- The Dataverse Project
- Identity Management System
- DSPACE Online Institutional Digital Collections Repository
- OJS Online Journal System

Research Institution C
- The Dataverse Project
- Identity Management System
- Electronic Theses and Dissertations (ETD) Management System
- DSPACE Online Institutional Digital Collections Repository
- OJS Online Journal System
Assessment and Results
Quantitative and Qualitative Measures

<table>
<thead>
<tr>
<th>System</th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>Downloads</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSpace</td>
<td>330,668</td>
<td>396,650</td>
<td>656,778</td>
<td>1,015,314</td>
</tr>
<tr>
<td>ETDs</td>
<td>158,240</td>
<td>200,373</td>
<td>328,420</td>
<td>470,437</td>
</tr>
<tr>
<td>Dataverse</td>
<td>N/A</td>
<td>N/A</td>
<td>455</td>
<td>3,451</td>
</tr>
<tr>
<td>ORCID ID’s</td>
<td>190</td>
<td>316</td>
<td>438</td>
<td>545</td>
</tr>
<tr>
<td>OJS Journals</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Annual Usage Growth (Downloads)

LibQual Biannual Survey 2013-2019, Faculty and Student System Perceptions, Comments

Ecosystem Implemented in Stages, 2014-2019

Making electronic resources accessible from my home or office
Larger Digital Scholarly Research Projects Can Act as Qualitative/Quantitative Benchmarks

1) Project Completion (Milestones)
2) Usage Statistics

- **Cognitive Cartography/Multimedia Archives**
  - Video, Text, GIS, Images, Field Notes
  - *Dick Reavis: National Tour of Texas*

- **Multimedia, Digital Archives/Retrospective ETD Projects**
  - Digital video, online exhibit images, text, digital archives
  - *Severo Perez: And the Earth Did not Swallow Them*

- **Online Exhibits/Digital Archives/Online Academic Journals**
  - Images and text, Omeka front end/Database back end, IIIF
  - *Cabeza de Vaca La Relacion Digitization*
  - *Santiago Tafolla: Mexican Amer. Confederate Soldier*

- **Interactive Image Archives/Data & Research Projects**
  - Image libraries, Interactive Commenting/Metadata
  - *Texas State Flickr Commons*

- **Digital Libraries Archiving & Documentation Projects**
  - Text, Metadata, OCR, Search, Zoom ability, Page Turning
  - *Pedagogys University Yearbooks*

- **Faculty Digitization Proposals/Partnerships**
Summary Reflections

Placing Digital Scholarship Components within an Ecosystem Paradigm Usefully Enables:

1) Better Guidelines and Roadmaps for Developing Digital Scholarly Components
2) Pathways Forward and Evolutionary Possibilities for System Development
3) New Possibilities For Researchers working within the academic research cycle
Ecosystem Components Enable Various Parts of the Academic Research Cycle

The academic research cycle

i. Identification of knowledge
   e.g. undertaking literature review using peer reviewed sources

ii. Creation of knowledge
    by professional researchers usually behind closed doors

iii. Quality assurance of knowledge
     e.g. peer review, filtering the best for publication

iv. Dissemination of knowledge
    e.g. publication, presentation at conference

Social media: A guide for researchers (2011), p1

Abstract Levels

Think & Plan

Pragmatic Levels

Discover

Gather & Analyse

Write & Publish

Share/Impact
Digital Scholarly Ecosystem Timelines and Implementation Paths

Many Roads To Rome

**Year 1** Digital Collection Repository and Digitization Lab

**Year 2** User Interface Software (OMeka), Identity Management System, ORCID

**Year 3** Data Repository

**Year 4** ETD Middleware (VIREO) and OJS Software

**Year 5** Complex Digitization Projects, IIIF Server, Faculty Grant Projects etc.
Human Resources

- **System Administrator/Programmer**: server infrastructure set-up/maintenance/basic customization
- **Digital Collections Librarian**: Administration, Marketing, User Support, Collections and Data Repository, OJS/ORCID
- **Metadata Librarian**: Dublin Core, Specialized Schema
- **Web Developer/Programmer**: OMEKA, System Integration
- **Project Manager/Department Head**: (PMP Certification)
- **Digitization Specialist**
- **GIS Specialist/Data Visualization Specialist**
- **AI Specialist/Post-Doc/CLIR Fellow**
Further References


Texas State University Libraries Website. https://www.library.txstate.edu/
Texas State Digital Collections Repository https://digital.library.txstate.edu/
Texas State Data Research Repository https://dataverse.tdl.org/dataverse/txstate
Texas State Online Research Identity Management System: https://guides.library.txstate.edu/researcherprofile/orcidTexas
Texas State Digital & Web Services: https://www.library.txstate.edu/services/faculty-staff/digital-web-services.html
Envisioning Future Possibilities
Networked Global Scholarly Research Environment

[Graph showing value for user versus number of users]
Can we Enable Scholarly Research Network Ecosystem Possibilities on Global Levels?

Is it Desirable or Time to Begin Thinking About Empowering a Global Research University Community?
Research Universities and Digital Research Ecosystems

• ~266-300 Research Institutions US & Canada, Carnegie R1 & R2, Very High or High Research Activity, 124 ARL Libraries

• ~1000-1250 Research Universities Worldwide
  QS Rankings and Times Higher Education Supplement. (40% Europe, 26.5% Asia Pacific, US/Canada 18%, Latin America 9% and Middle East/Africa.

• 26,000-40,000 Universities Globally. Research Universities 2.7% - 4.2% of all universities worldwide. Highest by Country: US 156, UK 76, Germany 45, Japan 44.

• Other Top 2-3% Research Institution Academic Libraries Globally, 1000 Institutions beyond the US and Canada. This represents the other 90% of Research Libraries Globally
Brainstorming & Antecedent Models

One Laptop Per Child
Dreamed up mid-late 90’s, Launched 2005

• Nicholas Negroponte, MIT Media Lab Founding Director
• Noble Initiative/Grand Ambitions
• Vision: Give each child in world access to a laptop with open source software for less than 100.00 $US/laptop
• Gage Effects For Education Globally
• Can We do the same thing for academic research globally?
One Server Per Research Institution 2020-2025

Simple Idea

• Empower 1000 Research University Institutions/Research Libraries Globally

• Give them One Configured Server Ecosystem with 6 Open Source Scholarly Research Software Components, < $1000.00 US/Server or set up Fractional Server Space Globally (SAAS)

• Set Up Brief Training

• Measure the Effects
Research Universities and Digital Research Ecosystems

- **124** ARL Research Libraries (US and Canada)
- **131** US Research Universities (Carnegie R1, Very High Research Activity)
- **135** Doctoral Universities (Carnegie R2, High Research Activity, US), ~266-300 Research Institutions US & Canada
- **1011** Research Universities Worldwide (40% Europe, 26.5% Asia Pacific, US/Canada 18%, Latin America 9% and Middle East/Africa.

**QS Rankings**

- **1250** Research Universities Worldwide, *Times Higher Education Supplement* (2.7% - 4.2% of all universities worldwide)
- By Country: **US 156**, UK 76, Germany 45, Japan 44
- Global Estimates of General University #’s **26,000-40,000**

Empower Other Top 2-3% Research Institution Libraries Globally, 1000 Institutions, the other 90% of Research Libraries Globally