Developing a Digital Scholarship Research Ecosystem

Ray Uzwyshyn, Ph.D. MBA MLIS
http://rayuzwyshyn.net
Director, Collections and Digital Services
Texas State University Libraries

Presented for ICEIT2020
St. Anne’s College, Oxford, UK
February 2020
What is a Digital Scholarly Research Ecosystem?

Network of Several Software Components to Enable Faculty and Student Research and Raise Research Profiles
General Characteristics of This Digital Scholarly Research Ecosystem

- Open Source Software
- Customizable Components
- Active Developer Communities
Digital Ecosystem Consists of 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
Collocating Open Source Digital Components in a Networked Research Ecosystem Enables Larger Connections and/or Network Effects

Ecosystem Metaphor Looks at Component Relationships

Simple Larger Idea
Network Effects: Metcalfe’s Law

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

Component Networks may be Internal and/or External
Together, These Digital Ecosystem Components Enable Various Parts of the Academic Research Cycle

**Pragmatic Levels**
- Think & Plan
- Discover
- Gather & Analyse
- Write & Publish
- Share/Impact

**Abstract Levels**
- i. Identification of knowledge e.g. undertaking literature reviews 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

**Research Cycle**

Social media: A guide for researchers (2011), p15
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 (Faculty and Graduate Students):

- Pre-prints
- Faculty Publications
- White Papers
- Conference Presentations
- Graduate Student Theses and Dissertations
Primary Use Case Value
Application of Structured Metadata Schema for Search Engine Optimization
Athermal annealing of low-energy boron implants in silicon

Donnelly, David W., Southwest Texas State University; Dept. of Physics;
Covington, B. C., Southwest Texas State University;
Grun, J., Naval Research Laboratory, Washington, DC;
Fischer, R.P., Naval Research Laboratory;
Peckerar, M., Naval Research Laboratory;
Felix, C. L., United Industries Inc.

Comments:

Recommended Citation:

Download

This item appears in the following Collection(s)
- Faculty Publications-Physics

https://digital.library.txstate.edu/
Application of Structured Metadata Schema for Search Engine Optimization Leads to Accessibility and Multiple Points of Access
Research Data Repository

https://dataverse.tdl.org/dataverse/txstate
The Dataverse can be configured as a single Instance or as a Consortial Model

https://dataverse.tdl.org/

(Texas Aggregates 22 Individual Instances, through the Texas Digital Library)
Digital Scholarly Research Ecosystem

Secondary Components
(Dependent on Primary Digital Repositories for Content and Data)
Vireo

Electronic Thesis and Dissertation Management System

• Addresses Intermediary steps in the ETD Process
• Connects with Both the Collections Repository And Data Repository so students can publish and link their theses dissertations and data
• Bridges Student Thesis/Dissertation Submission with Graduate School Review, Online Publication and ETD Preservation
Researcher Identity Management System

- Gives Researchers Unique Number (ORCID ID) Connecting and Disambiguate Scholars names
  - Maria Hernandez, Biochemist
  - Maria Hernandez, M.D. or Astrophysicist

- Allows publications from a researcher to be found, linked and aggregated across multiple information Systems.

- Allows Papers in the collections repository and datasets in data repository can be associated with ORCID ID’s for aggregation of research profiles.

ORCID is a hub connecting the research landscape

Orcid can act as a Network Hub aggregating from several sources and connecting to other internal and external networks.
Omeka and OJS

**Open Source User Interface Software**
Allows an elegant portal/gateway entrance for larger scholarly research projects, digital collections and data repositories - linking text, image media and datasets and acting as a front end for connecting components and component networks.

**Open Access Academic Journal Software** for refereed journal online publishing, workflow and connections with background research and datasets etc. through Dataverse/Dspace connections.
The Digitization Lab

- Expands Possibilities for Faculty Research Projects

- Possibilities range from OCR, image, book, manuscript & journal digitization to 3D objects, audiovisual material, GIS and visualization technologies (IIIF etc)

(i.e. Digitization Lab’s IIIF Framework can create internal or globally distributed Image Libraries.)
Combining Research Ecosystem Components
Opens Amazing Possibilities For Digital Scholarship & Collaboration 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
Human Resources and Implementation Paths

- **System Administrator/Programmer**
  server infrastructure set-up/maintenance/basic customization

- **Digital Collections Librarian/Specialist**
  Software Administration, Marketing, User Support

(Further Expansion Possibilities: Web Developer, Programmer Project Manager, Metadata Librarian, Digitization Specialist, GIS Specialist, Data Visualization Specialist, AI Specialist/Post-Doc)

**Timelines, 1-5 Years (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.
# Ecosystem Assessment and Results

## Quantitative and Qualitative Measures

### Annual Usage Growth

<table>
<thead>
<tr>
<th>System</th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Downloads</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSpace</td>
<td>318,742</td>
<td>385,163</td>
<td>341,224</td>
<td>972,359</td>
<td>1,010,349</td>
</tr>
<tr>
<td>ETDs</td>
<td>158,240</td>
<td>200,373</td>
<td>328,420</td>
<td>470,437</td>
<td>505,658</td>
</tr>
<tr>
<td>Dataverse</td>
<td>n/a</td>
<td>n/a</td>
<td>455</td>
<td>3,451</td>
<td>2,043</td>
</tr>
</tbody>
</table>

| **Items Added** |       |       |       |       |       |
| DSpace        | 1,437  | 1,546  | 1,660  | 2,135  | 2,720  |
| ETDs          | 1,174  | 1,326  | 1,581  | 1,789  | 2,218  |
| Dataverse     | n/a    | n/a    | 28    | 33     | 53     |

| **ORCID IDs** |       |       |       |       |       |
| ORCID         | 190    | 316    | 438    | 545    | 669    |

| **Hosted Journals** |       |       |       |       |
| OJS             | 1      | 2      | 2      | 3      | 4      |

### LibQual Biannual Survey

2013-2019, Faculty and Student System Perceptions, Comments

Making electronic resources accessible from my home or office:

- Perceived 2013
- Perceived 2015
- Perceived 2017
- Perceived 2019
Summary

Reflections

Placing Digital Scholarship Components within an Ecosystem Paradigm Enables:

1) Better Guidelines and Roadmaps for Developing Digital Scholarly Components

2) Pathways Forward and Evolutionary Possibilities for Future System Development

3) New Possibilities For Researchers within the academic research cycle
Questions, Comments

Ray Uzwyshyn, Ph.D.  MBA MLIS  
Director, Collections and Digital Services  
Texas State University Libraries  
ruzwyshyn@txstate.edu, 512-245-5687  
http://rayuzwyshyn.net
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
Future Pathways
Networked Global Scholarly Research Environment
Research Universities and Digital Research Ecosystems

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

• ~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.

• Enable Top 2-3% Research Institutions Globally, 1000 Institutions beyond the US and Canada.
  (This represents the other 90% of Research Libraries Globally)
One Server Per Research Institution
2020-2025

- Empower 1000 Research University Institutions/Research Libraries Globally

- Gift each Research University One Configured Server Ecosystem with 6 Open Source Scholarly Research Software Components, < $1000.00 US/Server or set up Fractional Server Space with Mirror Sites Globally (SAAS)

- Set Up brief weeklong training over five continents

- Connect Networks

- Measure the Effects
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?
Brainstorming 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?
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
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.
- **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
Larger Digital Scholarly Research Projects Can Act as Qualitative/Quantitative Benchmarks

1) Project Completion (Milestones)
2) Usage Statistics

Complexity

Cognitive Cartography/Multimedia Archives
(Videó, 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)
Pedagogos University Yearbooks

Projects, Prototypes Grant Partnerships

Faculty Digitization Proposals/Partnerships
Combining Components
System Synergies
Digital Scholarly Research 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)
**Assessment and Results**

**Quantitative and Qualitative Measures**

### Annual Usage Growth

(Downloads, Number of Items, ORCID ID’s and Hosted Journals)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Downloads</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSpace</td>
<td>326,762</td>
<td>318,742</td>
<td>385,163</td>
<td>341,224</td>
<td>972,359</td>
<td>1,010,349</td>
</tr>
<tr>
<td>ETDs</td>
<td>136,985</td>
<td>158,240</td>
<td>200,373</td>
<td>328,420</td>
<td>470,437</td>
<td>505,658</td>
</tr>
<tr>
<td>Dataverse</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>455</td>
<td>3,451</td>
<td>2,043</td>
</tr>
</tbody>
</table>

### Number of Items

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DSpace</td>
<td>1,340</td>
<td>1,437</td>
<td>1,546</td>
<td>1,660</td>
<td>2,135</td>
<td>2,720</td>
</tr>
<tr>
<td>ETDs</td>
<td>967</td>
<td>1,174</td>
<td>1,326</td>
<td>1,581</td>
<td>1,789</td>
<td>2,218</td>
</tr>
<tr>
<td>Dataverse</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>28</td>
<td>33</td>
<td>53</td>
</tr>
</tbody>
</table>

### ORCID IDs

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ORCID</td>
<td>101</td>
<td>190</td>
<td>316</td>
<td>438</td>
<td>545</td>
<td>669</td>
</tr>
</tbody>
</table>

### Hosted Journals

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>OJS</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

**LibQual Biannual Survey**

2013-2019, Faculty and Student System Perceptions, Comments

Making electronic resources accessible from my home or office:

- Perceived 2019
- Perceived 2017
- Perceived 2015
- Perceived 2013

**Ecosystem Implemented in Stages, 2014-2019**

- 2014-2019
Ecosystem as System 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.

• Further Desired Connections can also guide developmental paths for both component software and the ecosystem.
Digital Collection Repositories Gives Insight and another window into Faculty/Student Research (Statistics)

<table>
<thead>
<tr>
<th>Item title</th>
<th>File downloads</th>
<th>Item views</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fear: A Psychophysiological Study of Horror Film Viewing</td>
<td>70,564</td>
<td>8,101</td>
<td>78,755</td>
</tr>
<tr>
<td>Study of Museum Lighting and Design</td>
<td>67,844</td>
<td>2,082</td>
<td>69,926</td>
</tr>
<tr>
<td>Female Figures of the Upper Paleolithic</td>
<td>62,848</td>
<td>2,103</td>
<td>64,951</td>
</tr>
<tr>
<td>Gender Differences in Parenting Styles and Effects on the Parent-Child Relationship</td>
<td>61,284</td>
<td>3,392</td>
<td>64,676</td>
</tr>
<tr>
<td>A Study of the Relationship Between Absenteeism and Job Satisfaction, Certain Personal Characteristics, and Situational Factors for Employees in a Public Agency</td>
<td>52,937</td>
<td>4,005</td>
<td>56,942</td>
</tr>
<tr>
<td>&quot;The Decoded Message of the Seven Seals,&quot; by David Koresh</td>
<td>48,721</td>
<td>23,917</td>
<td>72,638</td>
</tr>
<tr>
<td>Mobile Dating in the Digital Age: Computer-Mediated Communication and Relationship Building on Tinder</td>
<td>48,681</td>
<td>16,672</td>
<td>65,353</td>
</tr>
<tr>
<td>Bottled Water: Why is it so Big? Causes for the Rapid Growth of Bottled Water Industries</td>
<td>39,559</td>
<td>783</td>
<td>40,342</td>
</tr>
<tr>
<td>Introduction to Image Processing with Python and Jupyter Notebook</td>
<td>32,111</td>
<td>2,698</td>
<td>34,799</td>
</tr>
</tbody>
</table>
Network Effects Allow Opportunities Among Research Institutions
Digital Scholarly Ecosystem Timelines and Implementation Paths
Many Roads To Rome (1-5 Year Paths)

**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.

Ecosystem Implemented in Stages, 2014-2019
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**
Percent Increase in Article Citations by Discipline with Open Access Online Availability (Immediately Available Through Google)

- Physics
- Sociology
- Psychology
- Law
- Management
- Education
- Business
- Health Sci
- Political Sci
- Economics
- Biology

% Increase in Citations with Open Access

Range = 36%-250%
(Data: Stevan Harnad and Heather Joseph, 2014)