My background

https://matienzo.org/
Stanford Digital Libraries as infrastructure and services
Stanford Libraries by the numbers

- 12 million volumes
- 75,000 current serials
- 500,000 cartographic resources (maps, images)
- 59 million manuscript pages across Special Collections
- 120 countries from which we acquire
- 400 languages represented
- 110,000 new volumes acquired each year
Stanford University Libraries (SUL)

- - -

- 16 Libraries
- ~450 staff
- 68 subject / information specialists
- 8th largest academic unit at Stanford (by budget)
- Plus...
  - Stanford University Press
  - LOCKSS
- Art & Architecture (Bowes)
- Archive of Recorded Sound
- David Rumsey Map Center
- Earth Sciences (Branner)
- East Asia
- Education (Cubberley)
- Engineering (Terman)
- Green
- Lathrop
- Marine Biology (Miller)
- Music
- Stanford Auxiliary Library 1&2 (SAL)
- SAL 3
- Science (Li and Ma)
- Special Collections
- Stanford Digital Repository
What Does a Library Do?

- Select resources to add to its collections
- Acquire those resources
- Describe the resources to enable discovery
- Store & manage items until such time as they are needed
- Provide access to the items to patrons on request
- Preserve the items so they are available to future scholars
- Support teaching, scholarship, and research needs
Digital Library Systems and Services
The Digital Library: Content, Services & Infrastructure

- Ebooks
- Google Books
- Digitized mss, texts, images, media
- Born-digital materials (data sets, theses, articles, new media, etc.)

- Services
  - Discovery
  - Delivery
  - Use
    - Analysis
    - Annotation
    - Citation
  - Collaboration
  - Publication

- Content

- Infrastructure
  - Digital preservation infrastructure
  - Identifiers, Resolvers
  - Server, storage and data center facilities
SDR: Manage, Preserve, Access

SDR at a Glance

- **2006**: 1st year of operation
- **1.6 million** Digital Objects
- **227 million** Files
- **1,800+** Collections
- **467 TB** Unique Data
- **930+** Depositors from 40+ departments
SDR’s mission:

The Stanford Digital Repository (SDR) provides services to make scholarly resources available over the long term by helping ensure their integrity, authenticity, and reusability. To fulfill its mission, the SDR must be secure, sustainable and trustworthy.
Funder Mandates & Other Use Cases

1. Supplemental data files for published articles
2. Reproducible Research: sharing data & code
3. Complying with data management plans
4. Archiving pre-prints (with embargo support)
5. Publication archives
6. Open access articles
7. Dark archiving patent information
8. Enriched data access (e.g., large file access, video streaming, archived websites)
Digitization

Stanford has digitized millions of items in the last 20 years:

- Books
- Images
- Maps
- Manuscripts
- Audio
- Video
- Film
- Legacy digital files
- 3D objects

To support

- Collection development
- Digital access
- Research projects
- New forms of research (AI, ML, etc.)
- Preservation
Digitization

Stanford has digitized millions of items in the last 20 years:

- Books
- Images
- Maps
- Manuscripts
- Audio
- Video
- Film
- Legacy digital files
- 3D objects

To support

- Collection development
- Digital access
- Research projects
- New forms of research (AI, ML, etc.)
- Preservation
Preserving legacy research

Hopkins Marine Station CalCOFI hydrobiological survey of Monterey Bay, CA: 1951 - 1974

Project Goals

1. Digitize everything;
2. Rehouse materials in archival boxes;
3. Make content available via Stanford Digital Repository (SDR);
4. Convert select data from PDF to actionable format (tabular text).
Capturing new research

Chris Gerdes
Director, Revs Program
Simplifying publication management

Stanford Digital Repository

Seismic Collapse Risk Assessment of Buildings: Effects of Intensity Measure Selection and Computational Approach

COLLECTION

John A. Blume Earthquake Engineering Center Technical Report Series

Blume Earthquake Engineering Center

Home  About  News and Events  Publications  Research  Affiliates  Links

Blume Center Technical Reports

<table>
<thead>
<tr>
<th>No.</th>
<th>Title</th>
<th>Authors</th>
<th>Date Published</th>
</tr>
</thead>
<tbody>
<tr>
<td>191</td>
<td>Ground motion simulations: validation and application for civil engineering problems</td>
<td>L. Burks and J. Baker</td>
<td>Oct 2015</td>
</tr>
<tr>
<td>190</td>
<td>A framework for engineering sustainable composites based on time-dependent material properties and environmental impact assessments: an application to bio-based composite design</td>
<td>S.A. Miller, M.D. Lepech and S.L. Billington</td>
<td>Sept 2015</td>
</tr>
<tr>
<td>189</td>
<td>Design of Composite RCS Special</td>
<td>D. Kathuria, H. Yoshikawa, S.</td>
<td>Sept 2015</td>
</tr>
</tbody>
</table>
Honoring Student Achievement
Replaying lectures, performances, oral histories
Feeding social science research
Harvesting the web
Supporting Access, Discovery, and Delivery

- SearchWorks
- Spotlight
- EarthWorks
- The International Image Interoperability Framework (IIIF)
- Requests and Course Reserves
SearchWorks

- Discovery environment ("library catalog") for Stanford University
- Allows for searching and browsing of materials from SDR and Stanford’s libraries (inc. some outside of SUL), and articles
- Integrated access to digital resources, request management for delivery of materials, and citation generation

https://searchworks.stanford.edu/
Spotlight

Full-featured website for showcasing digital collections:

- Visually attractive, responsive, and accessible website
- Self-service exhibit creation
- Powerful searching and browsing, multiple search result views, include gallery, slideshow, and map views
- Integrated with SDR and IIIF

Uses: rich digital libraries, digital exhibits, student portfolios and teaching tool

EarthWorks

- Discovery of geospatial assets
  - Purchased
  - Licensed
  - Stanford-produced
  - Multi-institutional

- Direct access to geospatial web-services
  - Preview
  - Download
  - Sideload to online services

https://earthworks.stanford.edu/
International Image Interoperability Framework

- Open international community building specifications for rich access to image and other resources
- Used to support deep zoom, presentation, annotation, and comparison of images
- Expanding to provide support for audiovisual materials and 3D resources

https://iiif.io/
How digital library projects happen: content and projects
## Digital Object Life Cycle

<table>
<thead>
<tr>
<th>Original</th>
<th>Capture</th>
<th>Accession: PURL/SDR</th>
<th>Index: SearchWorks</th>
<th>Showcase: Spotlight</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.jpg" alt="Original Image" /></td>
<td><img src="image2.jpg" alt="Capture Image" /></td>
<td><img src="image3.jpg" alt="Accession Image" /></td>
<td><img src="image4.jpg" alt="Index Image" /></td>
<td><img src="image5.jpg" alt="Showcase Image" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Accessible:</th>
<th>Preserved:</th>
<th>Discoverable:</th>
<th>Curated:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limited</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>
Stakeholder Roles and Responsibilities

- Content owner
  - Project champion
  - Selection & decision making
  - Preparation
  - Description
  - Funding

- DLSS
  - Project management
  - Technical resources
  - Content management
  - User experience

- Conservation
  - Material review
  - Treatment
  - Stabilization
  - Housing
  - Handling

- Metadata
  - Planning
  - Metadata mapping
  - Best practices for access & discovery
Digitization/Acquisition

Content
- Digitize or Acquire
- Accession in SDR
- Publish in SW
- Showcase in SL

Metadata
- Plan & Map
- Create & Enhance
- Load to SDR
- Review & Remediate
Digitization/Acquisition is a Team Effort

- Content Owner
  - Project request
  - Content selection
  - Descriptive metadata
  - Rights and display behavior
  - Funding (if available)

- DLSS
  - Cost/time estimate
  - Work plan
  - Digitization/Data targets
  - Quality control
  - Progress reports

- Conservation Services
  - Assessment
  - Stabilization
  - Repair

- Metadata Department
  - Guidance on standards & best practices
  - Planning & mapping
  - Transformation
What always takes longer than you think

- Determining rights and access
- Creating descriptive metadata
- Competing priorities across projects
- Digitization or data wrangling complications
Accessioning

Content
- Digitize or Acquire
- Accession in SDR
- Publish in SW
- Showcase in SL

Metadata
- Plan & Map
- Create & Enhance
- Load to SDR
- Review & Remediate
Creating Digital Objects is a Team Effort

**Content Owner**
- Descriptive metadata
- Delivery/display specifications
- Remediation

**DLSS**
- Assembly in Digital Object Registry (DOR)
- Well-formed objects
- PURL
- Long-term preservation

**Metadata Department**
- Metadata review
- Enhancement
- Metadata loading assistance
Accessioning process
Assembly
What always takes longer than you think

- Modeling the digital content
- Incomplete metadata
- Technical errors
Publishing in SearchWorks
Publishing is a team effort

- Content Owner
  - Quality control & remediation on:
    - Content
    - Metadata
    - Display
    - Rights

- DLSS
  - Develop, operate & maintain discovery environments
  - Remediation of technical issues

- Metadata Department
  - Remediation guidance & support
Minimum Metadata for SearchWorks

- Title
- Creator
- Language
- Type of Resource
- Date
- & preferably more!

Metadata supporting discovery
What always takes longer than you think

- Remediation of descriptive metadata
- For archival materials in the SDR, adding PURL links to finding aids
Building Exhibitions in Spotlight

---

Content

- Digitize or Acquire
- Accession in SDR
- Publish in SW
- Showcase in SL

Metadata

- Plan & Map
- Create & Enhance
- Load to SDR
- Review & Remediate
Building an Exhibit is a Team Effort

- Content Owner
  - Create exhibit concept
  - Liaise with service team
  - Coordinate text creation & layout

- DLSS
  - Develop, operate & maintain Spotlight
  - Service team: document, train & consult

- Metadata Department
  - Metadata display guidance & support
Why Spotlight?
Exhibit Creator Perspective

Easy-to-use, form-based exhibit building process:

- Site identity (title, subtitle, contacts, masthead image)
- User and role management
- Metadata fields to display on page-by-page basis
- Easy to create browse categories
- Import and export exhibit content and configuration settings
- Widget-based page configuration
What always takes longer than you think

- Exhibit planning – purpose, context & narrative
- Supplemental content identification
- Refinement with key stakeholders & service team
Ongoing Maintenance

Error reported via SW-Feedback → Fixed by Metadata Dept → Content updated → Changes are: Automatically published to PURL and SW

Manually pulled into Spotlight
Summary
Thank You!

Questions?

Mark A. Matienzo / @anarchivist
Stanford Libraries
matienzo@stanford.edu