From library repository to university-wide service

Stanford Digital Repository as a case study

Hannah Frost
SDR Services Manager
Stanford University Libraries

Open Repositories ❖ Helsinki ❖ June 12, 2014
The Stanford Digital Repository (SDR) is...

- a Fedora-based preservation repository,
- with Hydra-based deposit, management and access services,
- for library collections, domain-specific scholarly resources, and content of enduring value generated by the Stanford community.

http://library.stanford.edu/sdr
Current Profile of the SDR

Staffing

- Operations & Development
- Service Management
- Parts of many more...

Volume

- 530K objects
- 192M files
- 549 collections
- 147 TB
IR content is modest ...

- 3300 ETDs
- 1300 objects from Stanford depositors
- All else are library collections

But growing rapidly:

- Objects: ▲ 40% in 6 months
- Depositing student programs & courses: ▲ 275% in 1 year
SDR IR Services Development Timeline

1997

- Need identified

2006

- SDR 1.0 in prod
- Faculty Advisory Board convened

2007

- SDR 1.0 in prod
- Hydra Project born

2008

- ‘08
- ETDs launched

2009

- ‘09

2010

- ‘10

2011

- ‘11
- Online Deposit launched
- Data Management Services formalized

2012

- ‘12

2013

- ‘13

2014

- ‘14
Stanford Digital Repository – Online Deposit

The SDR is a service supporting long-term management of scholarly information resources at Stanford. Faculty, students, and researchers use the SDR to promote and protect the products of their work. The benefits of this service distinguish the SDR from other content storage or management options on campus: deposited scholarly content is preserved in a robust, reliable, and secure environment and is available from persistent URLs (PURLs) with optional access controls.

The growing body of content deposited in the SDR includes:

- scientific research data like this data set
- digital humanities research data like this corpus
- honors theses, like this undergraduate work
- images, audio and video like this podcast
- software and computer games like DOOM
- student projects, like this final team report
- technical reports like this publication
- archival collections like Project MKULTRA

Sign in via WebAuth to access your deposited content

Contact the SDR to become a depositor

Find SDR Deposits using SearchWorks

» Data Management Plans and SDR
» SDR Services information web site
Code and data supplement for "Perceived depth in natural images reflects encoding of low-level luminance statistics"

Author/Creator: Cooper, Emily A. (Author)  Norcia, Anthony M. (Author)
Date created: 2013–2014
Type of resource: Software, multimedia
Genre: Dataset
Abstract: The online Supplemental Material contains high-resolution versions of all images used in the perceptual experiments. Also included is MATLAB code for performing the analyses reported in the article. This includes the natural scene statistics analysis, the image manipulation, and the perceptual experiment analysis (raw response data from both experiments is provided).
http://purl.stanford.edu/yg499ys5636
Contact: emilyacooper@gmail.com
Collection: Stanford Research Data
Subject: natural scene statistics
depth perception
image manipulation

File (Download) Description
CooperNorciaSupplementalMaterial.zip (300 MB)

Use and Reproduction: User agrees that, where applicable, content will not be used to identify or to otherwise infringe the privacy or confidentiality rights of individuals. Content distributed via the Stanford Digital Repository may be subject to additional license and use restrictions applied by the depositor.

License: This work is licensed under a Creative Commons Attribution 3.0 Unported License
Self-deposited content

Jan 2013 - May 2014
1300 objects total

Archival documents 36%
Student reports 14%
Technical reports 12%
Honors theses 11%
Publications 10%
Data sets 8%
Audio/video 6%
Software (2%)
Conference papers (1%)
Other (1%)
More on the SDR today

• Agile software and service development
• Focus on research data, student works
• Favor timely deposit with low barrier tools
• Ever mindful of TDR
Foundations for Repository Services

Foundations for Repository Services

Currently for the SDR

**Content**
- No shortage!
- Increasingly diverse

**Context**
- Low barrier metadata requirements facilitates adoption, but at risk of reduced quality for discovery
- Bolstering metadata processes, support

**Access**
- Major focus
- Revamping discovery environment to improve access to IR content
Key Issues for Growing SDR’s IR Services

1. Thinking like a business
2. Supporting the service
3. Recognizing the preservation challenges
1. Thinking like a business

Repository services require business acumen

• Know the competition and complementary services
  – Figshare, Dryad, ICPSR, etc.

• Policy and legal issues
  – Deposit agreement framework

• Financing and cost modeling
  – “Under construction”
Marketing and Communications

Know what the depositor values
• Saving time!
• Library services?
• Institutional affiliation?

Pay attention to language
• Long-term preservation ➔ Persistent access

Emphasize the benefits
Stanford Digital Repository

“Making deposits into the repository has been very convenient and helpful to our research and teaching here at Stanford.”

David Chen
Graduate Student, Electrical Engineering

The Stanford Digital Repository supports management of scholarly information resources of enduring value to Stanford University. Faculty, students, and researchers use SDR services to promote and protect the products of their work. Scholars around the world use content in the SDR in their research. The benefits of this service distinguish the SDR from other content storage or management options on campus: deposited content is preserved in a robust, reliable, and secure environment for access by scholars today and for generations to come.
Other Marketing Techniques

- Blog posts
- Newsletters
- Ads
- Faculty meetings
- Events on key issues (e.g., US OSTP mandate for science data)
- “Deposit Parties”
- Focus on grad students, young faculty
- Build on familiarity with ETD system
2. Supporting the Service
The SDR Service Team

Since 2013

Matt
Nicole
Regina
Hannah
Rosy
Stella
Amy
Service Team Function and Roles

• Power user and local expert
• Work directly with campus depositors, encourage other librarians to use SDR
• Develop user documentation
• Advise on software requirements and development priorities
Selecting Research and Other Works for Deposit to the Stanford Digital Repository

When it is time to preserve the products of your research and creative output, you will need to carefully consider exactly which components of your work are most suitable for archiving and sharing. Different disciplines conduct research in different ways and produce content in different forms, so we have provided some general guidelines that are applicable across various fields, with some examples from the social sciences and musicology included online. For social science-specific questions, contact Ron Nakao, Data and Computational Social Science Librarian (ronbo@stanford.edu).

Example 6: Your work involves analyzing census data

- **Definitely deposit:**
  - study descriptions, codebooks, and summary statistics
  - results of your analysis

- **Do not deposit:**
  - original census data, as these are easily accessible elsewhere

- **Maybe deposit:**
  - intermediate versions of analyses or code if they are potentially useful to others or were used in publications or theses

- **Not necessary to deposit:**
  - incomplete, non-functional, or intermediate versions of code that would be of marginal utility to others
Benefits of the SDR Service Team

• Leverage relationships
• Develop digital skills of librarians
• Cross-pollinate ideas
• Reorients SDR as a library-wide service, not technology
• Contributes to sustainability

It’s how we scale.
3. Recognizing preservation challenges

Content is more diverse, less controlled:

• File formats
  – Robust file characterization processes

• File sizes
  – Impact on throughput and capacity planning
  – Especially media

• Challenges readiness for ISO 16363 audit
How Hydra Helps Overall

• Well-suited to building services incrementally

• Strong, growing, sharing community
  – Fostering a community of product managers and service managers
  – Bolsters overall sustainability of and trust in the SDR
Kiitos!

Hannah Frost
hfrost@stanford.edu