Cooperative Curation

From single IRs to statewide/consortial repositories and beyond

Tyler Walters
Dean, University Libraries
Virginia Tech

Symposium on Cooperative Curation: Inter-institutional Approaches to Supporting Scholarly Communication, Georgia Tech, Atlanta, GA
August 8, 2012
Cooperative Curation: Overview

• Focus on institutional repositories
  – Quick history, purpose, experience to date

• Cooperative strategies and developments
  – Statewide & consortial IRs
  – Digital preservation networks and data repositories

• Brief overview of GALILEO Knowledge Repository
A Quick History of IRs from...

...single to consortial IRs and cooperative curation

• 2002 – IRs in the U.S. begin...
  “through institutional repositories (IRs), universities hold the potential to “permanently change the landscape of scholarly communication””


• SPARC’s position paper advocating IR development
  (Raym Crow, SPARC position paper, Aug. 2002)

• ARL & ACRL support IRs as reform effort in scholarly communication, achieve OA to publicly funded research
The Promise of IRs

• By 2009, 229 IRs were operational in the U.S.

• By 2012, Open DOAR lists 274 IRs!

• IRs can deliver on increasing visibility and sustainability of digital intellectual output
  – Promote collaboration
  – Provide valuable resources to public, enhance teaching, learning and research
  – Create an outlet for / preserve digital scholarship
The Challenge

• Much of a college or university’s (C&U’s) intellectual output never appears in a permanent, traditionally published form

• Exists as disorganized pockets of born digital objects scattered on hard drives, departmental servers, and flash drives

• While C&U’s digital scholarly output reflects substantial investment of resources and effort, they:
  – Lack curatorial stewardship;
  – May be inaccessible;
  – Exist on unsustainable hardware, software, or lack individual support;
  – Need future-proofing migration strategies
Disappointment...

Have IRs realized the far-reaching changes to scholarly communication that Lynch, Crow, and others envisioned?

- ~4% of U.S. C&Us host an IR, lack resources, expertise

- IRs track with library funding: 78% are hosted by universities with ARL membership
  - Yet ARL institutions represent only 3% of public post-secondary schools and 10% of four-year institutions

- Only 2 public HBCUs have an IR, and potential of repository services for 2-year colleges is virtually unexplored
The Key is Collaboration

Cooperative approaches are key to addressing financial and technical barriers with IR adoption

• Institutions have banded together, generating efficiencies to:
  – Achieve consortial pricing
  – Preserve digital information
  – Digitize library holdings
  – Catalog materials
  – Host library systems
  – Share information resources
Collaborative Network Creators and Participants

• Curating institutions are forming collaborative entities that reach beyond grant-project timeframes

• Fostering long-term sustainable practices to support common needs, including technology development, repository management, and preservation

• Coordinating intra- and inter-institutional alliances, maintain clear documentation about roles and responsibilities for each in these multipronged efforts

*Learning to forge healthy collaborative relationships through credibility, reliability, transparency, and self-awareness*
Collaborative Strategies for Digital Curation/Preservation

• Wherever possible, collaborative or community-based approaches to digital curation are likely to be more effective and sustainable for a variety of reasons.

• Effective use of these strategies and community-managed resources requires local investment and capacity building.

• Yet... there are only a few groups who have worked to build collaborative IR programs.
The Founding of Statewide IRs

eScholarship  http://escholarship.org/

• 2002: UC 10-campus system
  47,000 papers, 15M downloads
  Provides suite of OA, publishing, research tools that enable universities and scholars to have control over their scholarship

Texas Digital Library  http://www.tdl.org/

• 2005: Partnership of Texas A&M University, Texas Tech University, the University of Houston, and the University of Texas at Austin
• Combine member resources to provide cost-effective collaborative solutions to digital storage, publication, and preservation of research, scholarship, teaching materials
• 15 institutions (large and small)
Statewide/Consortial IRs

OhioLink Digital Resource Commons  http://drc.ohiolink.edu/
• 2007: Statewide platform for saving, discovering and sharing creative materials produced by the University System of Ohio and Ohio’s private colleges
• 24 institutions currently, 6 more coming

• 2008: consortial digital repository service offered by Colorado Alliance of Research Libraries to member institutions and affiliates. ADR builds on open source platforms and establishes a preservation-oriented system
• 7 institutional repositories
Statewide/Consortial IRs

NITLE  [http://www.longsight.com/blog/nitle-dspace](http://www.longsight.com/blog/nitle-dspace)

- Mid-2000s: paid service to member institutions for inclusion in a single, multi-institution repository. Hosted by the Longsight Group.
- 20 institutional repositories

GALILEO Knowledge Repository [http://www.library.gatech.edu/gkr/](http://www.library.gatech.edu/gkr/)

- 2010: 10 institutions (4 hosted/6 harvested), Univ. System of Georgia
- More information to follow at this symposium and workshop!

*These programs are exceptional, however. Most states lack collaborative strategies for supporting and promoting open access to scholarly information*
Consortial Repository Developments: Digital Preservation Networks

MetaArchive cooperative
synergies
DURACLoud
PeDALS
Persistent Digital Archives & Library System
COPPUL
Lukii
US Gov Docs
Data-PASS
DATA PRESERVATION ALLIANCE for the SOCIAL SCIENCES
DPN
LOCKSS
LOTS OF COPIES KEEP STUFF SAFE
The Alabama Digital Preservation Network
Preserving Alabama's Digital Resources
Academic Preservation Trust
CLOCKSS
Consortial Repository Developments: Digital Data Repositories

There are many more!
Consortial repositories in the active workflow

- Repositories are being woven into "virtual ecosystems", they are holistic and support communities of practice
  - Early stages/deposit: raw/early phase data, notes, etc.
    - Annotating, sharing within research groups, commenting, etc.
    - Research proposal writing, project planning, etc.
  - Production Tools:
    - Discovery, analysis, visualization, and text/data/image mining are being used in concert with repositories
  - Virtual communities and their communication tools:
    - e.g., social media and community networking capabilities
Repositories in the larger research ecosystem/cyberinfrastructure

Protein Crystallography Research Data and Metadata Workflow

From Capture to Publication

Future?: Virtual lab system
How do IRs and “papers” fit in?

• Services over an active content layer that is backed by/harvested into a federated archive infrastructure based on institutional resources

(slides from SEAD)
IR Disparities and Need for Cooperative Curation

- The IR disparity holds true for the thirty-five institutions in the University System of Georgia (USG), where 4 had a repo/ETDs:
  - Georgia Tech
  - Georgia State University
  - University of Georgia
  - Valdosta State University

- Lack of resources, rather than interest, is cause of disparity

- In Nov. 2007, USG reps in statewide stakeholder meeting, discussed prospect of repositories at their campuses.
  - Indicated high interest in a systemwide IR service, with 100% rating central hosting and meta-searching services as highly important
GKR: Grant Project Basics

• Approved by Regents Advisory Committee on Libraries, August 2004

• 3-yr project to finalize several components toward establishing a comprehensive statewide repository, 2009-12

• $1,664,650 project: $857,005 (IMLS) & $833,946 (insts’ share)

• Leads: Georgia Institute of Technology & University of Georgia

• Participants (10):
  • Georgia State Univ., Georgia Southern Univ., Kennesaw State Univ., Georgia Health Sciences Univ., Valdosta State Univ., Albany State Univ., College of Coastal Georgia, Georgia Gwinnett College
GKR Project Deliverables

- Build a repository of standardized metadata harvested from IRs within the USG (GT and GKR Technical Committee)

- Hosting service for IRs at three USG institutions, GGC added (GT)

- Establish IR-related services:
  - Copyright research (GT)
  - Digitization (UGA)
  - Content Submission (GT/VSU)
  - Preservation (GT/MetaArchive)

- Partner with NGCSU to conduct an assessment survey of the USG faculty’s usage and perceptions of IRs

- Document and make GKR organizational model available to others interested in establishing statewide IRs
GKR Project & Deliverables

• Distinction between GKR and existing repository programs:
  – Promoting consortial IR work in other states
  – Building, testing, and launching a replicable, collaborative IR model – project staff dedicate a year to disseminating the results of the GKR initiative nationally

• Host a national symposium on statewide/consortial repositories
  – Create instructional materials
  – Conduct consortial IR training

• GKR Initiative’s desired outcome:
  – Advance the state of scholarly communication nationally by positioning consortia for success in implementing their network of institutional repositories
  – Meeting the needs of GKR’s in-state stakeholders
Thank You!

Tyler Walters
Dean, University Libraries
Virginia Tech

540.231.5595
tyler.walters@vt.edu
tywalters1 = Google/twitter/FB, etc.