Coordinating Deposit and Compliance

Rebecca Kennison
Director, Center for Digital Research and Scholarship, Columbia University
ARL Fall Forum 2013
October 11, 2013
9:30-10:45 a.m.
Challenge #1: Unwieldy Workflows

Research Lifecycle at UCF

- Impact Measures & Prestige
- Research Concept
- Planning Cycle
- Experiment/Project
- Global Scholarly Community
- Data/Output
- Project Cycle
- Grant Management
- Compliance
- Fundraiser
- Draft Work
- Peer Review
- Publication Cycle
- Publication/Presentation
- Not yet supported
- Reviews/Revisions
- Institutional Repository
- High Performance Computing
- Research Data Management Infrastructure

21st Century Digital Scholarship Cycle

- Communication
- 21st Century Digital Scholarship Cycle
- Grant Conclusion
- Final Reports
- Preservation

Center for Digital Research and Scholarship
NIH Public Access Policy

Compliance requirements and workflow management for NIH Principal Investigators and authors acknowledging NIH grant support.

NIH

Wonderful guide with step-by-step screenshots:
CHORUS Conceptual Design
Challenge #2: Insufficient Infrastructure

• FundRef beta
  – Still really only prototype
  – Dirty data
  – No standards for grant info capture

• No interoperability between/among repositories

• Few reporting tools, none currently integrated
Phases of SHARE

PHASE FOUR
- Data Curation
- Linked Data
- Shared Distributed Resources in Digital Repositories

PHASE THREE
- Text and Data Mining
- Semantic Data
- APIs
- Bulk Harvesting
- ResourceSync
- Open Annotation

PHASE TWO
- Submission Workflow
- Usage Metrics
- Reporting
- Incorporation of OAI-ORE
- Certification
- Adoption of Best Practices

SHARE is operational for article deposit and access when Phase I is completed

PHASE ONE
- PI Identifier
- Award Number
- Publication ID
- Data Set ID
- Copyright Licensing
- Sponsoring/Funding Agency Name
- Reporting
- Core Usage Statistics
- Metadata Exposed to Search Engines
- SWORD
- OpenURL

REQUISITE CONDITIONS
- Principal Investigator (PI) Identifier
- Award Identification Number
- Copyright Licensing Assignment
- Digital Repository Designation ID Number
- Preservation Rights
Phases of SHARE

PHASE FOUR
- Data Curation
- Linked Data
- Shared Distributed Resources in Digital Repositories

PHASE THREE
- Text and Data Mining
- Semantic Data
- APIs
- Bulk Harvesting
- ResourceSync
- Open Annotation

PHASE TWO
- Submission Workflow
- Usage Metrics
- Reporting
- Incorporation of OAI-ORE
- Certification
- Adoption of Best Practices

SHARE is operational for article deposit and access when Phase I is completed

PHASE ONE
- PI Identifier
- Award Number
- Publication ID
- Data Set ID
- Copyright Licensing
- SWORD
- OpenURL

REQUISITE CONDITIONS
- Principal Investigator (PI) Identifier
- Award Identification Number
- Copyright Licensing Assignment
- Digital Repository Designation ID Number
- Preservation Rights
Challenge #3: Cash-Strapped Institutions

• No additional money from agencies
• Little or no additional money for storage
• No additional resources for sponsored projects/compliance officers
• No additional resources for libraries’ staff
Opportunity #1: Campus Communication

• Communications group from Libraries, Sponsored Projects, Research Compliance & Training to discuss resources, training, and messaging for our community
  – Monthly meetings
  – Shared folder filled with international grant mandates
  – Up-to-date coordinated information on Web sites

• Research Without Borders speaker series event in November will be devoted to discussion of OSTP memo
Opportunity #2: Repository Enhancements

• Have added RIOXX metadata fields to Academic Commons (our research repository) so that we can track publications and data by grant number, funder

• Are integrating SWORD protocol into our Fedora repository for interoperability with other repositories
  – NB: SWORD is big part of the SHARE proposal

• Are working fast and furious on implementing MODS/RDF for LOD interoperability, incl. spearheading discussions with larger community
Opportunity #3: Scalable Data Services

- Have raised our per-file size for Academic Commons from 2 GB to 10 GB (no cost) to accommodate large data sets and are allowing up to 100 GB files to be submitted (cost above 10 GB at $5/GB one-time)
- Have set up campus-wide EZID DOI service (http://www.cdlib.org/services/uc3/ezid/) for researchers to be able to obtain DOIs for their work and thus be able to contribute to data registries
Opportunity #4: Partnerships

• Proposal to National Endowment for the Humanities (along with Modern Language Association) to set up and run repository for NEH-funded researchers, already have sandbox clone of Academic Commons as proof-of-concept

• Discussions with other repositories about investigating/testing interoperability
Let’s talk!

Rebecca Kennison, Director
Center for Digital Research and Scholarship
Columbia University Libraries/Information Services
Tel.: 212-851-2812
Twitter: @rrkennison
E-mail: rkennison@columbia.edu
Web site: http://cdrs.columbia.edu