A Model for Publishing and Linking Transportation Datasets and Reports

TRB Webinar

Lisa Zilinski
Carnegie Mellon University
@l_zilinski

David Scherer
Purdue University
@davidascherer

Kendra Levine
University of California Berkeley
@tranlib

March 24, 2015
Today’s Agenda

- Introduction to the workflows
  - Institutional Repository
  - Data Repository
- Convergent points in the workflows
- Outcomes of convergence
- Live demonstration & best practices
  - Citability
  - Altmetrics
- Final comments & questions
Part 1: Service Model and Workflows
The Purdue Service Model

Purdue e-Pubs and the Purdue University Research Repository (PURR)

Publications

- **e-Pubs**
  - Purdue e-Pubs is a service of the Purdue University Libraries, providing online publishing support for original publications as well as hosting for Purdue affiliated articles, reports, conference proceedings, student scholarship, and more. Contact the Libraries to discover opportunities to bring additional Purdue-affiliated scholarship online.

  - **At a Glance**
    - Top 10 Downloads
    - Recent Additions
    - Paper of the Day

- **Browse**
  - Departments and Centers at Purdue

- **Recent Additions**
  - 50 most recent additions

- **Paper of the Day**
  - Absolute measurement of hadronic branching fractions of the ηC(→ππną) meson

  - View Larger

Data

- **PURR**
  - The Purdue University Research Repository (PURR) provides an online, collaborative working space and data-sharing platform to support the data management needs of Purdue researchers and their collaborators.

  - **Start Your Research Project**
    - Create a Data Management Plan
    - Upload Research Data to Your Project
    - Publish your Dataset

  - **Featured Dataset**
    - Linking Pressure and Separation through Interstellar Arsenic in Polycyclic Aromatic Molecules

  - **Do you have a question?**
    - Ask a Librarian

  - **Your Question Message**
    - You will be prompted and guided toward answering your own question.
IR Technical Report Workflow

a) PI drafts proposal

b) Data Management Plan Created

c) PI submits proposal to funding agency

d) Proposal accepted by funding agency

e) PI submits draft report

f) SAC review of report

g) PI revises draft report

h) PA/PI & SAC confer

i) Final report production process

q) Post Production:
- Data/Tech Report Link
- Tech Summary
- DOIs
- Persistent URL
- Indexing
- Archiving
- Print on Demand

r) Measurements of Impact:
- Altmetrics
- Citations
- Downloads
- Access

PI = principal investigator
SAC = study advisory committee
PA = project administrator
Data Repository Workflow

1. PI drafts proposal
2. Data Management Plan Created
3. PI creates project in PURR
4. Project group collaborates in PURR
5. PI submits data set
6. SSL verifies data set
7. Data set published / archived with DOI

- PI = principal investigator
- PURR = Purdue University Research Repository
- SSL = Subject Specialist Liaison

Measurements of Impact:
- Citations
- Downloads
- Access
Part 2: Points of Convergence
Combined Workflow & Points of Convergence

1. DMP
2. DOI
3. Metadata

**Notes:**
- PI = principal investigator
- SAC = study advisory committee
- PA = project administrator
- PURRE = Purdue University Research Repository
- SSL = Subject Specialist Liaison

**Post Production:**
- Data/Tech Report Link
- Technical Summary
- DOIs
- Persistent URL
- Indexing
- Archiving
- Print on Demand

**Measurements of Impact:**
- Altmetrics
- Citations
- Downloads
- Access
Part 3: Outcomes of Convergence
Points of Convergence

Datasets (Purdue University Research Repository)


Report (Purdue e-Pubs Repository)

Linking Data & Publications
Linking Data & Publications
Part 4: Live Demonstration
Demonstration Examples

Report


Dataset

Part 5: Things to Consider
Citability

DOIs

- CrossRef DOIs for the report publications
- DataCite DOIs for the dataset publications

Recommended Citations

- Recommended citations directly provided within the metadata records
- Access to Google Scholar citation reporting and tracking
Discoverability

Internal

- Repository-based indexing and searchability
- Linkage to research programs web presence
- Connection of all affiliated citations (Reports to Datasets)

External

- Indexing through Google Scholar
- Inclusion in institutional library catalog
- Additional indexing in broader federated library cataloging systems or content indexes
Accessibility

Open Access and Freremium Versions

- PDF and e-Book versions of reports are Open Access in Purdue e-Pubs
- Datasets published with Creative Commons Licenses (CC-0, CC-BY)
- Print-on-Demand (POD) available for printed versions

User Experience

- External searching (e.g., Google Scholar)
- Multiple versions of reports available to meet users needs/experience
- Embedded Streaming Preview in PURR to sample Datasets prior to download
Measurements of Impact

Downloads/Views

- Monthly automated author and administrator-level reports detailing monthly and life-time total downloads
- Author Dashboards
- Access to Google Analytics reporting
Measurements of Impact

Altmetrics

- Embedded access to altmetric data, including social media dissemination, readership, and citations
Part 6: Conclusion/Best Practices
Conclusions/Best Practices

- **Linked Workflows**
  - Coordinate Resources
  - Anticipate Needs

- **Early Interaction and Anticipation**
  - Employ good data management principles and practices

- **Consider publication attributes**
  - Citation management
  - Increase visibility/discoverability

- **Monitor Usage and Access**
  - Employ both quantitative and qualitative measurements and metric points
  - Communicate impact
Thank you!

Lisa Zilinski  
Carnegie Mellon University  
ldz@cmu.edu

David Scherer  
Purdue University  
dscherer@purdue.edu

Kendra Levine  
UC Berkeley  
kklevine@berkeley.edu

March 24, 2015