OJS and Dataverse: Open Source Data Deposit

ACRL 2015

Alex Garnett
Public Knowledge Project
Simon Fraser University
@axfelix
The story to date

2 year Sloan Foundation grant (2012-2014):

**Public Knowledge Project (PKP)**
- Simon Fraser University
- Stanford University (John Willinsky)

**Dataverse Network Project**
- Harvard University’s Institute for Quantitative Social Science (IQSS) (Gary King & Merce Crosas)
- Micah Altman – Director of Research at MIT
**Project Proposal**

**Who?:** Address the needs of **journals** in addition to researchers and data managers.

**What?:** Enable journals to seamlessly manage the submission, review, and publication of data associated with published articles.

**How?:** Build the needed **technology** and create **awareness / incentives** regarding the importance of data sharing and preservation.

**Why?:** Help increase the **replicability** and **reusability** of research data. **Sustainable scholarship** (low overhead) meets **sustainable preservation practices**.
Integrating Open Source Systems

1. Open Journal Systems (OJS) [Willinsky 2005]

2. Dataverse Network [King 2007; Crosas 2011]
A repository for research data that takes care of long term preservation and good archival practices, while researchers can share, keep control of and get recognition for their data.
Open source journal management and publishing system created by PKP to expand & improve access to research.
Author submission with supplementary files

Jen Author

Abstract

Author submits article with one or more supplementary files selected for deposit in journal's Dataverse.

Author edits metadata. Study cataloguing information is also updated.

Editor edits metadata. Study cataloguing information updated again.

Full Text:

PDF
Which Workflow?

1. Submit paper
   - Approve paper + ask for data
   - Publish paper in OJS + release data in Dataverse

2. Submit paper + data
   - Approve paper + review data
   - Publish paper in OJS + release data in Dataverse

3. Submit paper + data
   - Reject paper + review data
   - Release data in Dataverse

+ 4, 5, 6, ....

Slide acknowledgement: Merce Crosas
Advantages to Integration

→ Reference implementation so others can reuse and extend (extension of SWORD).
→ Blind review of supplemental data
→ Streamlining authors’ article and data deposit process.
→ Permanent 2-way linking of the published article with its archived data.
→ Increase visibility/access, and encourage data citation, replication and re-use.

Photo: Flickr Commons