Building a Bridge Between Journal Articles & Research Data

IDCC 2014 - Feb 26, 2014

Eleni Castro
Institute for Quantitative Social Science (IQSS)
Harvard University
@thedataorg

Alex Garnett
Public Knowledge Project
Simon Fraser University
@axfelix
Motivation

Photo: Jean Liu
“Any moves towards data sharing are dependent upon the cooperation of journals.”*  
– Sergiu Ghergina and Dr. Alexia Katsanidou

*from European Political Science 2013: Data Availability In Political Science Journals
A team was assembled...

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.
Integrating Open Source Systems

We plan to do this by integrating two well-established open-source systems:

1. Open Journal Systems (OJS) [Willinsky 2005]
2. Dataverse Network [King 2007; Crosas 2011]
Dataverse Network

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.

Note: Major New Version 4.0 Coming Soon (Spring 2014)
Dataverse provides...

✓ Option for backups and replication of data in different locations (LOCKSS) so data is never lost.

✓ Re-format for long term accessibility so data never become obsolete.

✓ Extract Variable, Header (FITS) Metadata from datasets.

✓ Universal Metadata standards (DDI, DC, DataCite).

✓ Interoperability with other systems through standard protocols (such as OAI-PMH, APIs).

✓ Generates a DOI for permanent linking to datasets.

The Dataverse takes care of the archival infrastructure ("plumbing") for you!
Open source journal management and publishing system created by PKP to expand & improve access to research.
Journals using OJS to publish >=10 articles per year. Credit: Juan Alperin, PKP
OJS provides...

✓ Technical infrastructure for an entire editorial management workflow (submission, peer review, indexing).

✓ ‘Plugin' architecture (a la WordPress, Drupal, etc) to easily add new features without the need to change the entire core code base.

✓ Universal Metadata standards (expandable Dublin Core).

✓ Interoperability with other systems through standard protocols (such as OAI-PMH, APIs).

✓ Option for backups and replication of journals in different locations (LOCKSS) so access is never lost.
OK, so what is the integration going to do?

OJS plugin for: Data + supporting files + metadata sent via API (SWORD-based) to Dataverse
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
Progress to-date

1. Compiled a list of potential journals (~100) that we can work with.
2. Publishers reviewed our plugin workflow and mockups pre-development.
3. Completed and released first version of plugin+API.
4. Small sample of journals tested + provided feedback of integration via survey.
5. Some of our partners have begun production testing.
6. Working on improvements to UI and metadata customization for v2 release.
Demo

http://lib-hf1.lib.sfu.ca/ojs-2.4.3
Survey Findings

➔ Overall **positive**: Most testers plan to implement plugin!

➔ OJS doesn’t have the most eye-catching UI. Dataverse link & data citation must be prominent.

➔ But users excited for an **OJS 3 + Dataverse 4 overhaul**!

➔ Default dcTerms metadata defined in the SWORD spec is too limited for all of our users; the ability to customize fields is important.

➔ Plugin should not treat all supplemental files as Dataverse-worthy (e.g., authors’ notes).
Next Steps (Spring 2014)

1. Publish best practices for data sharing policies.
2. Develop & release updated versions of plugin + Data Deposit API.
3. Continue to release all code and documentation :)
4. Expand on Collaboration.
Advantages to Integration

➔ Reference implementation so others can reuse and extend (extension of SWORD).

➔ 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.
Thank you!

Contact
Alex Garnett - garnett@sfu.ca
Eleni Castro - ecastro@fas.harvard.edu

Project Website
http://projects.iq.harvard.edu/ojs-dvn

References


