Open Journal Systems and Dataverse Collaboration 1 Year Later
+ Dataverse 4.0 Updates

NETSL 2014 Lightning Talk - 04/11/2014
Sonia Barbosa & Eleni Castro
Institute for Quantitative Social Science (IQSS)
Harvard University
@thedataorg
Motivation

Photo: Jean Liu
Overview

2 year Sloan Foundation grant (2012-2014):

**Public Knowledge Project (PKP)**

**Open Journal Systems (OJS)**

- Simon Fraser University
- Stanford University (John Willinsky)

**Dataverse Network Project**

- Harvard’s Institute for Quantitative Social Science (IQSS) (Gary King & Merce Crosas)
- Micah Altman – Director of Research at MIT
OK, so what is the integration going to do?

OJS Journal

Harvard Dataverse

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

1. Compiled a list of interested journals (~100) that we can work with.
2. Publishers reviewed our plugin workflow and mockups pre-development (Spring ‘13).
3. Completed v1 of plugin+API (Fall ‘13).
4. Small sample of journals tested + provided feedback of integration via survey (Winter ‘14).
5. Released v2 of plugin based on survey (Spring ‘14).
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 more 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

1. Publish best practices for data sharing policies.
2. Based on ongoing feedback: develop & release updated versions of plugin + API.
3. Continue to release all code and documentation on github :)
4. **Expand** on Collaboration.

Photo: Flickr Commons
Major Dataverse Release
Coming Soon

(Goodbye 3.6.x; Hello 4.0!)
Version 4.0 Highlights

1. Major User Interface Redesign with Simpler Workflows
2. Dataverses All the Way Down (Object Model Re-Design)
3. New & Improved Browse/Search Features
4. New Metadata Support
5. New Interactive Data Exploration/Analysis Tool (d3.js)
6. Tabular Files Will Now Include Support and Processing of Excel and CSV Files
Major User Interface Redesign with Simpler Workflows
Dataverses All the Way Down (Object Model Re-Design)
Browse/Search Features
New & Improved

- Ability to search through Dataverses, datasets, and files with the basic search bar
- Browse Dataverses, datasets, and files through facets
- Metadata fields found in results based on the search query will appear in the search results card to show the relevancy of the result
- A user will be able to access all public data from the homepage (the search bar and content area of the homepage)
New Metadata Support

- Consulting with specialists in other domains (starting with biomedical and astrophysics) to extend metadata support.

- Domain specific metadata will be mapped/exported to VO format (astrophysics), and ISA-Tab format (biomedical) with OBI and other ontologies.

- To encourage interoperability between different systems we are expanding our compliance and general support of metadata schemas, which will include mapping/exports to Dublin Core Terms, DataCite 3.0, and an upgrade to DDI-Codebook 2.5).
New Interactive Data Exploration/Analysis Tool (d3.js)

- Our current data analysis tool will be replaced with a more interactive data exploration/analysis tool that will provide histograms, cross tabulations, enhanced descriptive statistics, and model selection.
Additional Support and Processing of Excel and CSV Files

• Supporting additional processing of excel, and csv (without the need for a control card)
Contact
Sonia Barbosa: sbarbosa@hmdc.harvard.edu
Eleni Castro - ecastro@fas.harvard.edu

Websites
http://projects.iq.harvard.edu/ojs-dvn
http://theddata.org

References