Helping Journals to Upgrade Data Publications for Reusable Research

Sonia Barbosa (Project Manager)
Eleni Castro (Project Coordinator)

Institute for Quantitative Social Science (IQSS)
Harvard University

@thedataorg
Purpose
We propose to enable journals to:

Seamlessly manage the submission, review, and publication of data associated with published articles

The project addresses the needs of journal publishers and editors rather than focusing exclusively on researchers and archivists.

The project focuses on the integration of data into the journal submission and publishing process, rather than on earlier stages of data creation and use.

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

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

The result will be to increase the replicability and reusability of published work in social science by improving the infrastructure for, practice of, and incentives related to data publication and citation.
A Dataverse is a virtual data archive with its own branding.

A Study describes and holds the Data.

Dataverse

Collections

Study

Study

Metadata

Data Files
The Dataverse Network is a repository for research data that takes care of long term preservation and good archival practices, while the researchers and data producers keep control of and get recognition for their data.
Why Connect Published Work to Data

Published Results + Data + Metadata + Supporting Files (documentation, code)

= A third party can replicate or reuse, thus validate and advance science
The Centralized Data Repository Provides:

✓ 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 Metadata from data sets
✓ Universal Metadata standards (DDI, Dublin Core)
✓ Inter-operability with other systems through standard protocols (such as OAI-PMH, APIs)

The Dataverse takes care of the archival infrastructure ("plumbing") for you!
Open Journal Systems (OJS)

• Open Journal Systems is an open-source software for the management of peer-reviewed academic journals, created by the Public Knowledge Project (PKP), released under the GNU General Public License.

• The Public Knowledge Project is a non-profit research initiative of the Faculty of Education at the University of British Columbia, the Canadian Centre for Studies in Publishing at Simon Fraser University, the Simon Fraser University Library, and Stanford University.

• focused on the importance of making the results of publicly funded research freely available through open access policies, and on developing strategies for making this possible. It seeks to improve the scholarly and public quality of academic research through the development of innovative online environments.
Open Journal Systems

• Provides the technical infrastructure not only for the online presentation of journal articles, but also an entire editorial management workflow, including article submission, multiple rounds of peer-review, and indexing.

• Relies upon individuals fulfilling different roles, such as the Journal manager, editor, reviewer, author, reader, etc. It has a module that supports subscription journals.

• Has a 'plugin' architecture, similar to other community-based projects such as WordPress, allowing new features to be easily integrated without the need to change the entire core code base.

• Open Journal Systems is also LOCKSS-compliant, helping to ensure permanent archiving for ongoing access to the content of the journal.
Progress to-date

Compiled a list of potential journals (~100) that we can work with. Contacted a small handful to be our 1st round of pilot testers. Including:

• Medical, Science & Social Science Publishers
• Academic Publishers

Three publishers reviewed our plugin workflow and mockups to provide feedback prior to us beginning development on the technology.
High-level Plugin Workflow

**Journal Setup**
- Setup Dataverse repository (DV)
- Journal sets up OJS-DVN plugin

**Submission**
- Author submits Article + Data in OJS
- Data + metadata automatically deposited into DV

**Review**
- Article + Data approved

**Publication**
- Article + Data published

**Post-Publication**
- Article linked to Data from OJS to DVN
- Data Linked to Article from DV to OJS
- Data Usage Stats

Flexible on this workflow
Mockup of Data Deposit (in OJS)

Step 4. Uploading Data Sets

You may provide access to the data sets and other supplementary files related to your research by either uploading them to the journal's dataverse instance, or by linking to the published data if already online.

You may upload any relevant data sets (or other supplementary files) below. These files are stored in a Dataverse Network specifically for this journal. More information on Dataverse can be found here. Clicking on the link below will open the {DVN Name} website in a new window; you will be returned to this submission once complete.

Upload to {DVN Name}

<table>
<thead>
<tr>
<th>ID</th>
<th>TITLE</th>
<th>ORIGINAL FILE NAME</th>
<th>DATE/uploaded</th>
<th>ACTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Untitled</td>
<td>19-282-1-PF.pdf</td>
<td>12-11</td>
<td>EDIT/DELETE</td>
</tr>
</tbody>
</table>

Link to your Data

 Optionally, if your data is already in a Dataverse or other repository, please provide the link below. Please note that for review purposes your data must be published and/or otherwise downloadable.

Link to data: [example1.com](http://example1.com) [example2.com](http://example2.com) [example3.com](http://example3.com)

Save and continue | Cancel
Published Article + Link to Data

Dataverse Demo Journal

Moving toward a recovery-oriented approach in the Swedish mental health system
Carolina Klockmo

Data

Moving toward a recovery-oriented approach in the Swedish mental health system

1. Replication data for: Moving toward a recovery-oriented approach in the Swedish mental health system

Data Set
Carolina Klockmo, 2012, "Moving toward a recovery-oriented approach in the Swedish mental health system", hdl:99999/00000

CoAction Dataverse Network
INDEXING METADATA
ANALYZE
DOWNLOAD

Close
What’s next?

1. Beginning development of plugin and API (SWORD2-compliant) this April.

2. Testers (+4) will provide feedback during Summer 2013.

3. OJS+Dataverse teams: Provide guidance on best practices for Data Sharing Policies and Data Citation.
## Project Roadmap

<table>
<thead>
<tr>
<th>Phase</th>
<th>Description of Phase</th>
<th>Timeframe</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Conduct outreach to OJS journals/publishers in order to have them test out the OJS-DVN plugin and provide feedback on the workflow for how the OJS-DVN plugin should be set up.</td>
<td>Fall 2012-Winter 2013</td>
</tr>
<tr>
<td>2</td>
<td>OJS and DVN teams will begin development work on the plugin</td>
<td>Spring 2013</td>
</tr>
<tr>
<td>3</td>
<td>Once the plugin is ready for testing contact group of publishers/journals from Phase 1 in order to test the plugin in a demo journal instance. All the materials needed for testing will be provided.</td>
<td>Summer/Fall 2013</td>
</tr>
<tr>
<td>4</td>
<td>Gather feedback from pilot testers to ensure that the plugin works as advertised and is ready to be used in a production setting.</td>
<td>Fall 2013</td>
</tr>
<tr>
<td>5</td>
<td>Once testing of the plugin has been completed, the OJS-DVN plugin will be made available for other journals to install and use. Instructions, best practices and other documentation will also be made available.</td>
<td>Fall 2013-Winter 2014</td>
</tr>
<tr>
<td>6</td>
<td>API will be available for other Journal Management /Publishing Systems</td>
<td>Spring 2014</td>
</tr>
</tbody>
</table>
Contacts and Sites

Sonia Barbosa, sbarbosa@hmdc.harvard.edu
Eleni Castro, ecastro@fas.harvard.edu

Contact for issues/questions on Dataverse:
dvn_support@help.hmdc.harvard.edu

Web sites:
http://thedata.org
http://projects.iq.harvard.edu/ojs-dvn (Project specific)
Thank you!

• **Alfred P. Sloan Foundation:** for generous grant support on this project

• Jean Liu: from the [Altmetric blog](https://altmetric.com) for the “*Journal and Data*” illustration