

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

PKP

PUBLIC
KNOWLEDGE
PROJECT

The
**Dataverse
Network**[™]
Project 

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]



Dataverse Network



Share, Cite, Reuse, Archive Research Data
Scientific data for reproducible research

Harvard Dataverse Network

POWERED BY THE **Dataverse Network** PROJECT v. 3.6.2



[Advanced Search](#) [Tips](#)

The Harvard Dataverse Network is open to all scientific data from all disciplines worldwide. It includes the world's largest collection of social science research data. [Learn more about the Dataverse Network.](#)

Dataverses

[Create Dataverse](#)

625 Dataverses

i A **Dataverse** is a container for research data studies, customized and managed by its owner.

RECENTLY RELEASED DATAVERSES

Jesse C. Johnson	Feb 19, 2014
Elections Project	Feb 4, 2014
Blanca de Lizaur	Feb 3, 2014

Studies

[Create Study + Upload Data](#)

52,980 Studies, **733,029** Files, **945,810** Downloads

i A **study** is a container for a research data set. It includes cataloging information, data files and complementary files.

RECENTLY RELEASED STUDIES

Replication data for: Dynamic Hierarchical Factor Models by Emanuel Moench; Ng, Serena; Potter, Simon	Feb 20, 2014
Impact Lite Dataset by Silvestri, Silvia; Douxchamps, Sabine; Teufel, Nils; Quiros, Carlos F.; Singh, Dhiraj; Ndungu, Anthony; Mutie, lanetta; Ndiwa, Nicholas; Van Wijk, Mark; Rufino, Mariana; Herrero, Mario; Kiplimo, Jasper	Feb 20, 2014

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 Journal Systems (OJS)

The screenshot shows the homepage of the Open Health Data journal. The header features the 'UPmetajournals' logo, the journal title 'open health data', and the 'ubiquity press open access' logo. A navigation menu includes links for Home, About, Editorial Board, Contact, Repositories, FAQ, Articles, and Archive. On the left, there are options to log in, register, follow on Twitter, and follow via RSS, along with social media icons. The main content area contains a description of the journal, a cartoon illustration of a scientist and animals on a book, and sections for 'Most recent articles' and 'Most popular articles'. A sidebar on the right features an advertisement for 'eifl Open Access in developing countries'.

UPmetajournals

open health data

ubiquity press
open access

Home About Editorial Board Contact Repositories FAQ Articles Archive

Log in | Register

Follow on Twitter

Follow via RSS

Tweets

UCL Open Access @UCLoA_funding

UCL researchers: check out our wonderful Open Access FAQs, they explain how to use central funds to pay for your oa: <http://t.co/PIIOFIKRL1>

Open Health Data is a peer-reviewed journal publishing papers that describe openly-accessible datasets with high reuse potential. We work with trusted data repositories to ensure that associated datasets are professionally archived, preserved, and openly available.

Open Health Data was previously published as the *Journal of Open Public Health Data*. The title and scope of the journal were amended on 1 January 2014 to support a broader range of health research.

Most recent articles

Nagulapalli
Data from "Social determinants of unmet hospitalisation need..."
04 Nov 2013

Most popular articles

Alexander, Wint
Projected Population Proximity Indices (30km) for 2005, 2030 & 2050
11 Jul 2013

Ads

eifl
Open Access in developing countries

Open source journal management and publishing system created by PKP to expand & improve access to research.

The Result?

New England Journal of Coffee

[HOME](#) [ABOUT](#) [USER HOME](#) [SEARCH](#) [CURRENT](#) [ARCHIVES](#)

Home > No 1 > **Author**

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 cataloging information is also updated.

Editor edits metadata. Study cataloging information updated again.

Full Text:

[PDF](#)

DATA CITATION

Author, Jen, "Author submission with supplementary files", <http://dx.doi.org/10.5072/FK2/MKCH6> V2 [Version]

KEYWORDS

[keyword 1](#) [keyword 2](#)
[keyword 3](#) [keyword 4](#)

USER

You are logged in as...

admin

- [My Journals](#)
- [My Profile](#)
- [Log Out](#)

NOTIFICATIONS

- [View](#) (27 new)
- [Manage](#)

JOURNAL CONTENT

Search

 ▾

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,

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.

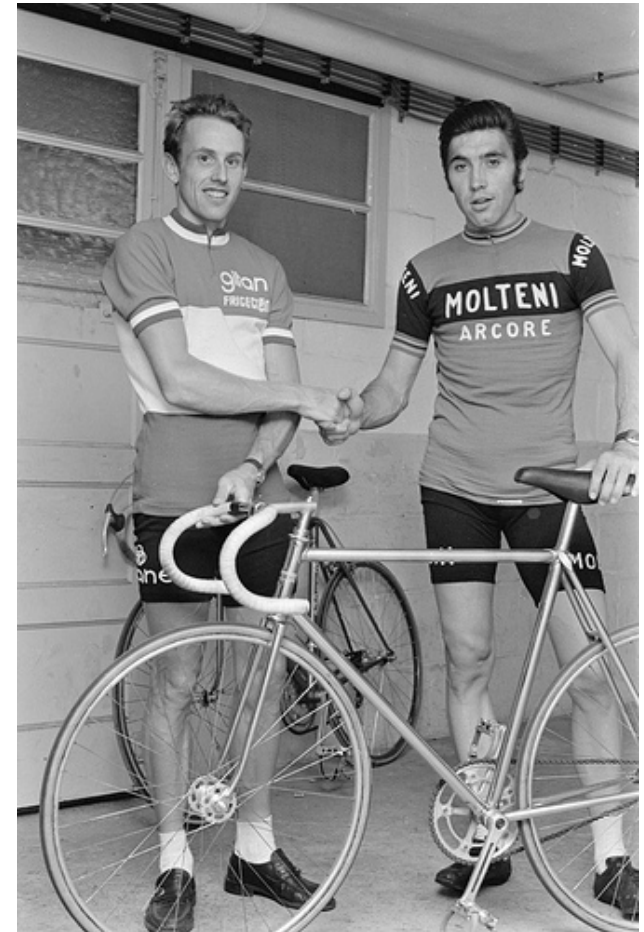


Photo: [Flickr Commons](#)