The Journal of Open Archaeology Data and PRIME: Incentivising Open Data Archiving

Brian Hole (UP & UCL)
WAC7, Jordan, 18 January 2013
Overview: why, how, where

1. **Why** share data?
2. **How** to enable and incentivise sharing: The Journal of Open Archaeology Data
3. **Where**: The PRIME project
Dad - how is it that Richard Dawkins can be everywhere?
The Social Contract of Science

- Research requires an effective, efficient distribution model
- Research funders are now demanding this – it will become the main model
Dataset to accompany the excavation report for a 'liberated African' graveyard in Rupert’s Valley, St Helena, South Atlantic

The dataset derives from the excavation and post-exavation of a freed slave ('liberated African') graveyard on the South Atlantic island of St Helena. The report for this site has been produced in two parts: a printed volume published by the Council for British Archaeology (Pearson et al. 2011); and a digital volume deposited with the Archaeological Data Service (ADS). Also held by ADS are supporting materials such as museum reference...

Latest papers

Archaeological Survey and Monitoring Data from the Flower of Ugie, Wrecked 1852 in the Eastern Solent, England

APRIL 5, 2012

This paper discusses the potential future reuse of the archived dataset resulting from the archaeological survey of the shipwreck remains of the Flower of Ugie, a wooden sailing barque wrecked in the Eastern Solent in 1852. Seven years of archaeological survey are represented in the site plans which record the structural details of the vessel along with many of its material characteristics. This record serves to establish the baseline condition of the site, against which future monitoring of the site will be conducted allowing informed management decisions to be made.
Metajournals

Repository

Metajournal

Research articles

Metapaper impact metrics:

- 16 citations
- 3 Wikipedia mentions
- 52 Topsy tweets
- 21 Facebook likes
- 7 Mendeley readers
Data journals

• Data is already citable (e.g. with DataCite DOIs) but this is not something researchers are familiar with doing.

• Even when used, DataCite DOIs cannot currently be used for citation tracking.
What is a data paper?

A data paper...

• ... describes the methodology with which a dataset was created.
• ... describes the dataset itself.
• ... details the reuse potential of the data.
• ... is often authored by a data scientist.
• ... is citable, enabling reuse to be tracked.

A data paper is **not**...

• ... a research paper. A data paper only describes a dataset. But it will reference research papers that are based on the data.
• ... simply replication of the information in a data repository.
Peer review

1. The paper contents
   a. The methods section of the paper must provide sufficient detail that a reader can understand how the resource was created.
   b. The resource must be correctly described.
   c. The reuse section must provide concrete and useful suggestions for reuse of the resource.

2. The deposited resource
   a. The repository must be suitable for resource and have a sustainability model.
   b. Open license permits unrestricted access (e.g. CC0).
   c. A version in an open, non-proprietary format.
   d. Labeled in such a way that a 3rd party can make sense of it.
   e. Must be actionable.
Dataset to accompany the excavation report for a 'liberated African' graveyard in Rupert’s Valley, St Helena, South Atlantic

The dataset derives from the excavation and post-excavation of a freed slave ('liberated African') graveyard on the South Atlantic island of St Helena. The report for this site has been produced in two parts: a printed volume published by the Council for British Archaeology (Pearson et al. 2011); and a digital volume deposited with the Archaeological Data Service (ADS). Also held by ADS are references such as museum reports...

Latest papers

Archaeological Survey and Monitoring Data from the Flower of Ugie, Wrecked 1852 in the Eastern Solent, England

APRIL 5, 2012

This paper discusses the potential future reuse of the archived dataset resulting from the archaeological survey of the shipwreck remains of the Flower of Ugie, a wooden sailing barque wrecked in the Eastern Solent in 1852. Seven years of archaeological survey are represented in the site plans which record the structural details of the vessel along with many of its material characteristics. This record serves to establish the baseline condition of the site, against which future monitoring of the site will be conducted allowing informed management decisions to be made.
PRIME
Publisher, Repository and Institutional Metadata Exchange
PRIME: Project focus

• Developing a system to exchange metadata between:
  • the UCL Discovery EPrints institutional repository
  • the Archaeology Data Service subject repository
  • the Journal of Open Archaeology Data (JOAD)

• Focusing on archaeology data only to pilot the system

• Building on other successful JISC projects:
  • DryadUK
  • REWARD
  • SWORD-ARM
The **REWARD** Project

- Precursor to PRIME
- A 6-month JISC-funded pilot project to introduce RDM to the UCL Institute of Archaeology
- Tested whether using familiar workflows and tools makes archiving data less burdensome
  - Preparing a data management plan
  - Use of an institutional repository
  - Publishing a data paper
**Familiar workflow**

The researcher deposits the data into UCL Discovery, as they are accustomed to doing with their other research outputs. The data and paper are then peer reviewed by the journal.

The researcher submits a paper to JOAD, as they are used to doing with other journals. Following peer review, the researcher gets a citable publication with a DOI.

The researcher creates and manages their data as always, but with assistance from the DMP Online tool to plan and increase efficiency.

**Mechanism**

**UCL Discovery**
Unlocking UCL research

**Eprints**

**Journal of open archaeology data**

**DMP online**
The K.I.D.C. Data Management Planning Tool

**Output**

Openly archived dataset

Published data paper

Data management plan shared at institution
Where is the data?
PRIME: Use Case #1

- A UCL Researcher deposits data in an external subject repository.
- The subject repository sends the metadata and DOI of the data to the UCL institutional repository so that it has a record of the output.
PRIME: Use Case #2

- A UCL Researcher deposits data in their institutional repository.
- The institutional repository sends the metadata and DOI of the data to the appropriate subject repository so that it has a record of the output.
PRIME: Use Case #3

- A UCL Researcher submits an article to a journal, and is asked to archive the data as a precondition of publication.
- The journal sends the metadata to the subject repository so that the author does not have to re-enter it.
- The subject repository sends the metadata and DOI of the data to the institutional repository so that it has a record of the output, and the DOI back to the journal to link the article with the data.
Methodology

1. Metadata profile:
   • Simple at first (DataCite)

2. Metadata exchange:
   • Symplectic Elements (e.g. UCL)
     • Harvesting by Symplectic
     • User is asked to acknowledge ownership
   • Pure Eprints (e.g. University of Southampton)
     • EPrints harvests from others
     • User initiates harvesting

3. Case studies
Questions?

Links

http://www.openarchaeologydata.metajnl.com
http://www.ucl.ac.uk/prime
http://www.ucl.ac.uk/reward
http://www.metajnl.com