# JISC Final Report

<table>
<thead>
<tr>
<th>Project Information</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Project Identifier</strong></td>
</tr>
<tr>
<td><strong>Project Title</strong></td>
</tr>
<tr>
<td><strong>Project Hashtag</strong></td>
</tr>
<tr>
<td><strong>Start Date</strong></td>
</tr>
<tr>
<td><strong>Lead Institution</strong></td>
</tr>
<tr>
<td><strong>Project Director</strong></td>
</tr>
<tr>
<td><strong>Project Manager</strong></td>
</tr>
<tr>
<td><strong>Contact email</strong></td>
</tr>
<tr>
<td><strong>Partner Institutions</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Project Web URL</strong></td>
</tr>
<tr>
<td><strong>Programme Name</strong></td>
</tr>
<tr>
<td><strong>Programme Manager</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Document Information</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
</tr>
<tr>
<td><strong>Project Role(s)</strong></td>
</tr>
<tr>
<td><strong>Date</strong></td>
</tr>
<tr>
<td><strong>URL</strong></td>
</tr>
<tr>
<td><strong>Access</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Document History</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Version</strong></td>
</tr>
<tr>
<td>1.0</td>
</tr>
<tr>
<td>1.1</td>
</tr>
</tbody>
</table>
Table of Contents

1 ACKNOWLEDGEMENTS ........................................................................................................... 3

2 PROJECT SUMMARY ............................................................................................................. 3

3 MAIN BODY OF REPORT ........................................................................................................ 3

   3.1 PROJECT OUTPUTS AND OUTCOMES .............................................................................. 3
   3.2 HOW DID YOU GO ABOUT ACHIEVING YOUR OUTPUTS / OUTCOMES? ....................... 5
   3.3 WHAT DID YOU LEARN? .................................................................................................. 5
   3.4 IMMEDIATE IMPACT ....................................................................................................... 6
   3.5 FUTURE IMPACT ............................................................................................................. 7

4 CONCLUSIONS ....................................................................................................................... 8

   4.1 GENERAL CONCLUSIONS ............................................................................................. 8
   4.2 CONCLUSIONS RELEVANT TO THE WIDER COMMUNITY .............................................. 8
   4.3 CONCLUSIONS RELEVANT TO JISC .............................................................................. 8

5 RECOMMENDATIONS ............................................................................................................. 9

   5.1 GENERAL RECOMMENDATIONS ..................................................................................... 9
   5.2 RECOMMENDATIONS RELEVANT TO THE WIDER COMMUNITY ................................. 9
   5.3 RECOMMENDATIONS RELEVANT TO JISC ................................................................. 9

6 IMPLICATIONS FOR THE FUTURE ...................................................................................... 9

   6.1 GENERAL IMPLICATIONS ............................................................................................... 9
   6.2 RECOMMENDED NEW DEVELOPMENT WORK .............................................................. 9
   6.3 SUSTAINABILITY OF PROJECT OUTPUTS ....................................................................... 10
   6.4 LONG TERM PROJECT CONTACTS ............................................................................. 10

7 REFERENCES .......................................................................................................................... 10

8 APPENDICES .......................................................................................................................... 11

   8.1 APPENDIX 1: CASE STUDIES ......................................................................................... 11
   8.2 APPENDIX 2: INSTITUTE OF ARCHAEOLOGY SURVEY ................................................. 16
   8.3 APPENDIX 3: RECOMMENDATIONS FOR THE INSTITUTE OF ARCHAEOLOGY .......... 24
1 Acknowledgements

REWARD was funded by JISC under the Managing Research Data programme, with thanks to the programme manager, Simon Hodson, whose coordination enabled many valuable exchanges with other projects and institutions. Many thanks to the researchers at the UCL Institute of Archaeology who volunteered their time to complete case studies and attend the workshops, and to staff at UCL Library Services and Ubiquity Press for work on integrating the data journal and institutional repository.

2 Project Summary

A six-month project, REWARD involved modification of researchers’ existing workflows to encourage best practices in data archiving, dissemination and reuse at the UCL Institute of Archaeology. The central premise was that existing processes, resources and incentives require only minor modification in order to achieve this. The project both proved that this could be done successfully, and provided mechanisms for these practices to remain embedded in future. REWARD achieved the following objectives:

1. **RDM best practices**
   - Awareness and embedding of best practices in research data management was encouraged at the UCL Institute of Archaeology. This was accomplished through two well-attended workshops in November 2011 and March 2012, and tutoring on data management planning through use of the DMP Online tool.

2. **Data archiving**
   - Increased data archiving by Institute of Archaeology researchers was promoted. The UCL Discovery repository was modified to accept datasets in December 2011, and through REWARD seven datasets were openly archived.

3. **Data journal use**
   - The Journal of Open Archaeology Data was modified to direct authors to deposit data in UCL Discovery, and overall seven data papers were written by Institute of Archaeology researchers.

4. **Case studies**
   - The experiences of researchers using DMP Online, JOAD and UCL Discovery were tracked in seven case studies (see Appendix 1).

5. **Dissemination**
   - Two workshops were held for REWARD at the Institute of Archaeology in November 2011 and March 2012. Conference papers were presented at the Computer Applications and Quantitative Methods in Archaeology Conference in Southampton in March 2012 and the Association of Learned and Professional Society Publishers (ALPSP) Data Publishing Seminar in London in April 2012. A research paper detailing the outputs of the project has also been submitted to the International Journal of Digital Curation, and the project has been used extensively for two chapters submitted to a forthcoming book on archaeological data.

3 Main Body of Report

3.1 Project Outputs and Outcomes

The majority of REWARD outputs can be downloaded from the project website at: [http://www.ucl.ac.uk/reward/resources](http://www.ucl.ac.uk/reward/resources), while other URLs are given below.

<table>
<thead>
<tr>
<th>Output / Outcome Type (e.g. report, publication, software, knowledge built)</th>
<th>Brief Description and URLs (where applicable)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCL Discovery modification</td>
<td>The UCL Discovery repository (<a href="http://discovery.ucl.ac.uk">http://discovery.ucl.ac.uk</a>) was modified to accept datasets with appropriate metadata including DOIs (e.g. <a href="http://dx.doi.org/10.5334/data.1394754978">http://dx.doi.org/10.5334/data.1394754978</a>).</td>
</tr>
<tr>
<td>Internal workshops</td>
<td>Two internal workshops were held, in November 2011 (reports: <a href="http://www.ucl.ac.uk/reward/resources">http://www.ucl.ac.uk/reward/resources</a>).</td>
</tr>
<tr>
<td>Open workshop participation</td>
<td>REWARD also took part in the JISC MRD Developer workshop/hack</td>
</tr>
<tr>
<td><strong>Data management plans</strong></td>
<td>Six data management plans were created using the DMP Online tool (<a href="http://www.ucl.ac.uk/reward/documents/REWARD_DMPs.zip">http://www.ucl.ac.uk/reward/documents/REWARD_DMPs.zip</a>).</td>
</tr>
<tr>
<td>--------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Integration of JOAD with UCL Discovery</strong></td>
<td>Instructions for depositing content in UCL Discovery were added to JOAD, and refined with feedback from the case studies (<a href="http://openarchaeologydata.metajnl.com/repositories/#ucldiscovery">http://openarchaeologydata.metajnl.com/repositories/#ucldiscovery</a>).</td>
</tr>
<tr>
<td><strong>Data papers</strong></td>
<td>Seven data papers were published in JOAD:</td>
</tr>
<tr>
<td></td>
<td>Moshenska, G. 2012. Selected Correspondence from the Papers of Thomas Pettigrew (1791-1865), Surgeon and Antiquary. <em>Journal of Open Archaeology Data</em> 1(2), DOI: <a href="http://dx.doi.org/10.5334/4f913ca0cbb89">http://dx.doi.org/10.5334/4f913ca0cbb89</a></td>
</tr>
<tr>
<td></td>
<td>Palmisano, A. 2012. Diachronic and Spatial Distribution of Khabur Ware in the Early Second Millennium BC. <em>Journal of Open Archaeology Data</em> 1(2), DOI: <a href="http://dx.doi.org/10.5334/4f8d6ed49bd54">http://dx.doi.org/10.5334/4f8d6ed49bd54</a></td>
</tr>
<tr>
<td></td>
<td>Crema, E.R. and Nishino, M. 2012. Spatio-Temporal Distributions of Middle to Late Jomon Pithouses in Oyumino, Chiba (Japan). <em>Journal of Open Archaeology Data</em> 1(2), DOI: <a href="http://dx.doi.org/10.5334/4f8eb4078284b">http://dx.doi.org/10.5334/4f8eb4078284b</a></td>
</tr>
<tr>
<td></td>
<td>Uribe Villegas, M.A. and Martín-Torres, M. 2012 Typology, Technology, Composition and Context of Muisca Metalwork (Colombia, AD 600-1800): a Database. <em>Journal of Open Archaeology Data</em> 1(1), DOI: <a href="http://dx.doi.org/10.5334/4f60dd6baa298">http://dx.doi.org/10.5334/4f60dd6baa298</a></td>
</tr>
<tr>
<td></td>
<td>Bevan, A. and Conolly, J. 2012. Intensive Survey Data from Antikythera, Greece. <em>Journal of Open Archaeology Data</em> 1(1), DOI: <a href="http://dx.doi.org/10.5334/4f3bcb3f7f21d">http://dx.doi.org/10.5334/4f3bcb3f7f21d</a></td>
</tr>
<tr>
<td></td>
<td>Harrington, S. and Brookes, S 2012. ASKED – the Anglo-Saxon Kent Electronic Database. <em>Journal of Open Archaeology Data</em> 1(1), DOI: <a href="http://dx.doi.org/10.5334/4f33a7b040dd1">http://dx.doi.org/10.5334/4f33a7b040dd1</a></td>
</tr>
<tr>
<td><strong>Case studies</strong></td>
<td>Seven case studies were carried out, recording the experiences of researchers using DMP Online, JOAD and UCL Discovery (see appendix 1).</td>
</tr>
<tr>
<td><strong>Recommendations for the UCL Institute of Archaeology on Research Data Management</strong></td>
<td>Three main recommendations were produced (see appendix 3).</td>
</tr>
<tr>
<td><strong>Final project report</strong></td>
<td>This document.</td>
</tr>
<tr>
<td><strong>Research paper</strong></td>
<td>A research paper detailing the outputs of the project has also been submitted to the journal International Journal of Digital Curation, and when published will also be available here: <a href="http://www.ucl.ac.uk/reward/resources">http://www.ucl.ac.uk/reward/resources</a></td>
</tr>
<tr>
<td><strong>Conference presentations</strong></td>
<td>REWARD was presented at the following conferences:</td>
</tr>
<tr>
<td></td>
<td>1. Computer Applications and Quantitative Methods in Archaeology, Southampton, March 2012 (<a href="http://www.ucl.ac.uk/reward/resources">http://www.ucl.ac.uk/reward/resources</a>)</td>
</tr>
<tr>
<td></td>
<td>2. Association of Learned and Professional Society Publishers</td>
</tr>
</tbody>
</table>
3.2 How did you go about achieving your outputs / outcomes?

REWARD was a six-month project, so we were careful to avoid scope-creep and to ensure that the project aims and tasks were realistic. Central to achieving best practices in data archiving was to enable researchers to plan for this from the outset, ensuring that they would be well prepared and have set aside time for this at the appropriate stages in their respective projects. This was achieved by holding two internal workshops, and by guided production of data management plans with researchers who volunteered to take part in case studies.

The two workshops, held at the UCL Institute of Archaeology, were very well attended by a range of researchers, ranging from masters students to senior staff members. They began by introducing the reasons for sharing research data. They then looked at the benefits data management planning, and introduced the DMP Online tool. The range of repositories available for archaeological data was then surveyed, and the benefits of data publication, focusing on the Journal of Open Archaeological Data, were covered. The workshops also involved a large amount of discussion, and a lot of feedback was collected on researcher motivations, experience, and perceived barriers to data sharing.

Seven researchers from the Institute of Archaeology then agreed to undertake case studies, which began with the guided creation of data management plans with DMP Online. The focus of plan creation was to establish a clear plan for collecting and storing their data so that it could be archived with minimal effort in the UCL Discovery institutional repository, along with publication of a data paper in the Journal of Open Archaeological Data. The institutional repository was chosen because it is a low barrier, easy to use archiving solution that researchers are already familiar with, while the data journal provided a familiar means of motivation by providing a citable paper. The case studies recorded the experiences and challenges of the researchers in each of these areas (see appendix 1).

An online survey of researchers at the Institute was also carried out in order to assess their attitudes to data sharing, as background for the case studies. The results are presented in appendix 2. The results of the survey and the case studies are analysed in more detail in the paper submitted to the International Journal of Digital Curation.

The project then undertook to embed these methods within the Institute of archaeology, whereby the data management plans were made available for future use, and recommendations were made to the Institute for improving management of its data, including a proposal for an intranet page with links to relevant RDM resources (see appendix 3).

3.3 What did you learn?

The key learnings from the REPORT project can be broken down by the main activity areas as follows:

Repository integration

- Went well: We found that while it was relatively simple to modify the UCL Discovery repository to accept datasets, and researchers involved with the project found this a familiar and comfortable place to archive their data. Five of the seven data papers published during the project resulted in UCL Discovery data deposits, while the remaining two chose to use the Archaeology Data Service.
- Difficulty encountered: We found that repository staff had little experience with datasets, and therefore lacked confidence when approached by researchers. It was necessary for example
for the REWARD team to assign DOIs to UCL Discovery records as the staff themselves were not able to do so.

- Difficulty encountered: It was not possible to obtain DataCite DOIs from the British Library for the project, which was unexpected but due to their internal concerns about the suitability of institutional repositories for holding data. Crossref DOIs were therefore used instead.

- Interesting outcome: We found that researchers decided where to deposit their data based on the size of the dataset and its curation requirements. Where the datasets involved were relatively small, researchers were very happy to use UCL Discovery. For the two larger datasets involved, there was however a strong preference for the Archaeology Data Service subject repository. This suggests that offering researchers both kinds of archiving could be the strategy that is most successful.

Data journal usage

- Went well: We found that data citation was seen as a significant incentive for researchers at the Institute of Archaeology to share data (see appendix 1). Seven data papers were published during the project with 2,629 views, 485 downloads, and positive altmetric scores (see section 3.4)

Case studies

- Went well: We found researchers at the Institute of Archaeology were very open to learning and applying RDM best practices, although there was clearly more the case with those at earlier career stages (see appendix 1).

- Difficulty encountered: Researchers at the Institute of Archaeology were generally unaware of UCL policy in regard to open access, research codes of conduct and data management, so information on this needed to be added to workshop presentations and case study briefings.

3.4 Immediate Impact

REWARD has had a direct impact on the approach taken to data management by researchers at the Institute of Archaeology, both those who took part in the case studies and others who attended the workshops and published data papers separately.

The following table details the impact of the data papers published in JOAD within one year of publication. While only one paper has been cited in a research article, one would not yet expect significant citations to have accrued yet. 2,629 paper views and 485 downloads indicate strong interest in accessing this data by researchers who otherwise would not have known how to access it. The majority of papers have also scored well in terms of ImpactStory altmetrics, indicating that they have been discussed in social media and referenced by the public, e.g. on Wikipedia.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Title</th>
<th>DOI</th>
<th>Citations</th>
<th>Views</th>
<th>Downloads</th>
<th>ImpactStory altmetrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moshenska</td>
<td>Selected Correspondence from the Papers of Thomas Pettigrew (1791-1865), Surgeon and Antiquary</td>
<td><a href="http://dx.doi.org/10.5334/f913ca0cbb89">http://dx.doi.org/10.5334/f913ca0cbb89</a></td>
<td>0</td>
<td>428</td>
<td>58</td>
<td>highly discussed; cited</td>
</tr>
<tr>
<td>Palmisano</td>
<td>Diachronic and Spatial Distribution of Khabur Ware in the Early Second Millennium BC</td>
<td><a href="http://dx.doi.org/10.5334/f8d6ed49bd54">http://dx.doi.org/10.5334/f8d6ed49bd54</a></td>
<td>0</td>
<td>545</td>
<td>134</td>
<td>highly discussed; saved;  cited</td>
</tr>
<tr>
<td>Crema &amp; Nishino</td>
<td>Spatio-Temporal Distributions of Middle to Late Jomon Pithouses in Oyumino, Chiba (Japan)</td>
<td><a href="http://dx.doi.org/10.5334/f8eb4078284b">http://dx.doi.org/10.5334/f8eb4078284b</a></td>
<td>0</td>
<td>451</td>
<td>94</td>
<td>highly discussed; saved;  cited</td>
</tr>
<tr>
<td>Uribe</td>
<td>Typology, Technology.</td>
<td><a href="http://dx.doi.org/10.5334/f8eb4078284b">http://dx.doi.org/10.5334/f8eb4078284b</a></td>
<td>0</td>
<td>210</td>
<td>54</td>
<td>saved;</td>
</tr>
</tbody>
</table>
Researchers who published with JOAD responded very favourably to the response to their data papers, demonstrating the motivational role they can play in achieving data archiving. For example, in an email to the project team at the end of his case study, Enrico Crema wrote: "It has been a very useful experience and I'll definitely carry on publishing more data papers in the future."

The project also demonstrated that making data openly available does increase its discoverability by the wider public. As a result of publishing the data paper in JOAD as part of his case study, Marcos Martinón-Torres was approached by a television production company that had found it, and asked to contribute to a documentary on the Muisca tribe for the Smithsonian Channel in the US and BBC Worldwide. He emailed the project team to say that "JOAD is doing its job – long live open access!"

Other researchers at the Institute of Archaeology who did not take part directly in case studies did however choose to publish data papers in JOAD. Andrew Bevan and James Connolly’s paper was immediately discovered by the community and reused for teaching, as described in the following quote from Dr Stuart Dunn at Kings College London, which demonstrates the value of the data journal approach:

“I used the data published by Andrew Bevan and James Connolly in the latest issue of JOAD in an Archaeology and Material Culture class in KCL’s MA in Digital Humanities programme… All in all, the ease with which the data could be accessed, its availability in CSV format and the clarity of the metadata schema meant that I had to put little effort into preparing it as a teaching material: it truly was plug and play. In fact I found myself having to warn the class that archaeological data is most often *not* available in such a user friendly format. The world would be a better place if more AR archaeological data were available in this way.”

The impact of the project has been further analysed in the following paper:


### 3.5 Future Impact

REWARD can be expected to have continuing forms of impact. Firstly, it is expected that the researchers who carried out the case studies and attended the workshops will have benefited from...
the best practices learned, and will continue to manage their data accordingly. Additional data papers have already been promised post-project.

Similarly, now that it has been established and publicised that UCL Discovery is a suitable repository for smaller datasets, it is hoped that its use for this purpose will continue to grow.

All resources produced by the project have been made available to researchers at the Institute of Archaeology on an on-going basis. The main resources in particular have been recommended for a page on the institutional Intranet (see Appendix 3).

The Journal of Open Archaeology Data has proved very successful and will continue in operation. At the same time the access and impact metrics for the data papers published in REWARD will continue to be monitored and reported on.

Finally, the article to be published in the International Journal of Digital Curation will broaden awareness of the project and its methods.

4 Conclusions

4.1 General conclusions

4.1.1. Researchers involved in the project were open to data sharing, and while wary of time and cost constraints, did consider a citable data paper as adequate incentive to do so (see appendix 1).

4.1.2. Most researchers within archaeology have not considered issues such as data preservation, licensing and citation before (see appendix 1).

4.1.3. Researchers at the Institute of Archaeology are generally unaware of UCL policy in regard to open access, research codes of conduct and data management.

4.2 Conclusions relevant to the wider community

4.2.1. Institutional repositories are useful for archiving smaller data sets, particularly where no subject repository exists, or is to expensive for some researchers.

4.2.2. The DMP Online tool was found by all researchers involved in the project to be useful, though some felt that it was slow to use, and were not always aware of the support features available (see appendix 1).

4.3 Conclusions relevant to JISC

4.3.1. Institutional repositories, including those utilising the EPrints system such as UCL Discovery, require only minimal modification in order to be suitable for holding smaller datasets.

4.3.2. To be able to effectively advise on deposition and curation of even smaller datasets, institutional repository staff require at least a minimum amount of training in order to have the ability and confidence to do so.

4.3.3. Institutions such as UCL and the current DataCite organisation in the United Kingdom are not yet prepared for the provision of DOIs within institutional repositories. This is unfortunate, as institutional repositories will otherwise remain ‘second class’ in terms of ability to reference and cite data.

4.3.4. Institutional repositories are well suited for archiving smaller data sets, particularly where no subject repository exists, or is to expensive for some researchers. Institutions also benefit greatly from being able to showcase more of their outputs in their repositories. At the same time, where they exist, specialised subject repositories will often be the better place to ensure the optimal curation, dissemination and preservation of larger datasets. Ways need to be found to enable both kinds of repositories to coordinate archiving so that the majority of research data can be appropriately archived.
5 Recommendations

5.1 General recommendations

5.1.1. Repositories should support data journals as an effective means of incentivising archiving, more widely disseminating their holdings, and encouraging reuse (see conclusion 4.1.1).

5.1.2. The UCL Institute of Archaeology should produce an internal intranet page providing researchers with guidance on institutional policies and issues and resources for research data management (see conclusions 4.1.2 and 4.1.3 and appendix 3).

5.2 Recommendations relevant to the wider community

5.2.1. Institutions and subject repositories should form a coordinated strategy to ensure that research datasets of all sizes and from all disciplines have an appropriate archive location (see conclusion 4.2.1).

5.2.2. The DCC should continue to maintain the DMP Online tool, and look to improve its speed and support features in line with feedback (see conclusion 4.2.2).

5.3 Recommendations relevant to JISC

5.3.1. JISC should encourage the use of institutional repositories for holding smaller datasets, in order to provide a quick and inexpensive data archiving option for researchers who otherwise don’t have access to archives in their subject area, or cannot afford the time and cost to submit to them (see conclusion 4.3.1).

5.3.2. JISC should look at funding basic data curation training courses for institutional repository staff (see conclusion 4.3.2).

5.3.3. It is recommended that JISC find a way of enabling institutions to obtain DOIs. One option would be for JISC to become a UK DataCite managing agent. An alternative would be to mediate between the current managing agent and institutions, providing a mechanism for negotiating access and costs (see conclusion 4.3.3).

5.3.4. JISC should consider funding ways to improve coordination and communication between institutional and subject repositories. A mechanism for metadata exchange for example could ensure that any relevant repository could have a metadata record including a DOI for a dataset, while that dataset would be located, curated and preserved in the most appropriate place for those activities.

6 Implications for the future

6.1 General implications

Managers of institutional repositories should be considering encouraging research data archiving, and aiming to present themselves as an important part of the solution alongside subject-based repositories. Wherever possible they should also seek to work with data journals to incentivise researchers to participate.

6.2 Recommended new development work

The number of data repositories in the UK has grown steadily due to consistent demand over the past decade, including both institutional [2] and subject-specific kinds [3]. Institutional repositories are vital for showcasing a university’s research outputs of all types [4], and being able to present a comprehensive overview of such outputs is of growing importance for exercises such as the Research Excellence Framework (REF) [5]. Benefits for researchers are also increasingly recognised, with the UCL Publications Policy for example stating that the repository should also “provide each researcher with a central hub for a comprehensive personal record of his/her outputs” [6]. At the same time, higher education institutions do not always have the resources and skills to curate and preserve datasets [7], and several recent studies have concluded both that there is “a clear value in aggregating a large number of data sets by subject or discipline” [3], and a need for more such repositories to cater for all disciplines [8].
It is recommended that development be undertaken to pilot the exchange of metadata between institutional and subject-based repositories. This would ensure that while the data itself was held in the place most appropriate for its curation and preservation, it would also be locatable from other repositories of relevance that have an interest in listing metadata records about it (e.g. a researcher’s home institution).

6.3 Sustainability of project outputs

Project outputs will be maintained as follows:

- **Journal of Open Archaeology Data**: Will continue to be published by Ubiquity Press. All papers are archived with the CLOCKSS system, which guarantees their long-term preservation and open access status.
- **Institute of Archaeology Data Management Plans**: Will be kept in the DMP Online system maintained by the Digital Curation Center, and available through the Institute of Archaeology intranet.
- **Case study and other project reports**: Will continue to be available on the REWARD project website at UCL (http://www.ucl.ac.uk/reward/resources).
- **Project results article**: Will be published as open access in the International Journal of Digital Curation, with a copy archived in UCL Discovery and linked to from the project homepage.

6.4 Long term project contacts

- General project enquiries and the Journal of Open Archaeology Data:
  - Brian Hole
  - Ubiquity Press
  - brian.hole@ubiquitypress.com

- **UCL Discovery**:
  - Paul Ayris
  - UCL Library Services
  - p.ayris@ucl.ac.uk

- **JISC**:
  - Simon Hodson
  - MRD Programme Manager
  - s.hodson@jisc.ac.uk

7 References


8 Appendices

8.1 Appendix 1: Case Studies

Researchers to participate in REWARD were identified through a group email to the Institute of Archaeology staff and research students mailing lists and individual emails to potentially interested researchers recommended by Research Data Management champions in the Institute. Initial meetings were set up on individual basis with interested respondents to inform them of the project and of what participation in REWARD entailed and arrange, if they agreed, the first interview.

Five case studies were initially selected and two more were taken up in the process to represent the full range of different research within archaeology conducted in the Institute of Archaeology.

During the initial interview the ways the researcher dealt with data creation, use, documentation, protection and storage, assessment, ownership, preservation, re-use, licensing, sharing, data management, as a distinct stage in his/her work, citation and other requirements, such as ethical approval and personal data, if relevant, were discussed. After the interview the researcher went away with the task to complete the DMP Online tool using a common account set for all of REWARD’s case studies and use the logbook to monitor this task. In the meantime, the Data Researcher emailed further and more specific information on Research Data Management in relation to the kind of research, type of data and data formats used. Furthermore, institutional requirements on UCL’s Open Access Mandate and the Code of Conduct for Researchers were also communicated to them. Before the final interview, the Data Researcher reviewed the participant’s DMP Online tool. Comments and additional information on specific issues encountered were discussed during the final interview. During the final interview, the researcher’s experience, benefits, challenges, issues with research data management were discussed with the aim to determine whether the data management process can be effectively integrated into the researchers’ existing workflows. Finally, a data paper and data deposition walkthrough was conducted to ensure that most, if not all issues were resolved. At the final stage, participants were asked to submit their data paper to the Journal of Open Archaeology Data and their datasets to UCL Discovery. Further support was provided during this process.

Case Study 1
Spatial Approaches to the Political and Commercial Landscapes of the Old Assyrian Colony Period
By Alessio Palmisano, PhD student

Alessio is using spatial data in widely used proprietary formats and open data formats that the UK Data Archive accepts for deposition. Use of Open Source Software would take him closer to best practices. Although he does not have the knowledge and training required to use them, he uses open data formats and therefore his data is easily deposited and shared.

Alessio believes that ‘if you are confident, you can [share]’ and that ‘it would be best for everyone [if everyone shared data]’. He has had bitter experience of trying to access other researchers’ datasets. He told us that he contacted approximately 40 to 50 scholars but only three of them shared their datasets with him. He has also had good experience of acquiring data sets from research institutes, such as the Chicago Oriental Institute.

He was particularly concerned with the difficulty of standardisation of data in archaeology and he brought to the attention of REWARD the Arcane project (see List of Resources), which exemplifies the effort to standardise chronologies in Eastern Mediterranean and Near Eastern archaeology. He also discussed researchers’ inclination to leave documentation for a later stage of research work, such as publication, when documentation becomes unavoidable.

Alessio was unaware of the Database Right and he also required further information on DOIs and clarification on the reuse of existing datasets in relation to some of the questions in the DMP Online tool.
Case Study 2
Spatial and Temporal Models of Jomon Settlements
Enrico Crema, PhD student

Enrico’s research project uses spatial data and it is highly demanding in terms of detail and accuracy, which explains the lack of trust to other people’s data, he exhibited, and his insistence on working with raw data that do not include any interpretation or subjectivity. His research project also takes a significantly different theoretical approach to settlement development from most of the projects in his field thus making his data even more particular.

His principles as a researcher are already well informed by best practices. In order to explain his insistence in raw data he said, ‘I want to understand the structure’, ‘I need to be able to create it from scratch because otherwise, if there is any problem, I get lost’. This is also the reason why he uses Open Source Software and open data formats.

Issues he has had with his data related to the incapability of some software to deal with Japanese characters, to CRM authorities outsourcing database construction to companies that do not employ archaeologists and therefore mismanage archaeological data by adding all the important information in the notes field and thus not standardising them and rendering them unusable.

Enrico is preparing a research grant proposal for the Leverhulme Trust that has absolutely no research data management requirements.

Case Study 3
The Relationship Between Popular Science and Public Archaeology in Early Nineteenth Century Britain
Dr Gabriel Moshenska, Leverhulme Trust Early Career Research Fellow

Gabriel is conducting research on the history of public archaeology through the investigation of the correspondence of people who initiated public demonstrations of mummies. His work for this project has involved expensive trips to libraries oversees and transcription of material from hard copy to digital format, already demonstrating the value of open data. He had not really realised that he was reusing existing datasets, as he was literary transcribing them in digital form. The cost and effort he has invested in this research has made him reluctant to assess and discard data once this has been collected, especially since storage is getting cheaper and greater in capacity. Easiness, speed and convenience of access are key factors in his decisions on how to store and back up his data.

An issue regarding the copyright of his material is that this has been collected from different collections that might not have the same regulations in place. His own understanding is that the material (dated to the first half of the 19th century) is too old to be protected by copyright. However this shows a partial understanding of intellectual property rights and copyright.

Gabriel skimmed through the DMP Online tool with great easiness and took a very straightforward and uncomplicated approach to the management of his research data during REWARD.

Case Study 4
Peter Coe, PhD student

Peter is a mature student with great experience in archaeology and in work in the broader public sector. This has granted him with an understanding of public administration and its requirements and the necessity to abide with them. His responsible approach to research data management has been enhanced by the fact that his laptop has been stolen twice in combination with the fact that the second time his external hard drive had failed to backup his data as he thought that it had, so he has had personal experience of painful efforts to retrieve data. Also, because of the nature of his research topic, he is particularly sensitive to issues regarding Freedom of Information, Personal Data and data anonymisation.
Possibly because of his lack of training in IT Peter is not very familiar with software, data types, data formats, storage and computer file management, all challenging issues when it comes to managing an extensive archive of images. He has however quite a thorough documentation routine, although this might also fail him from time to time.

For the above reasons, Peter exhibited a particular dependence to the use of familiar ways to work, e.g. to backup. Finally, and commonly to other researchers, he exhibited a resilience to discarding data. As he said, ‘researchers are collectors by nature’.

**Case Study 5**

*Tourism and the Economic Capital of Archaeology: Measurement and Management for Preservation*

Paul Burtenshaw

Paul’s data consist of notes taken during informal and unstructured interviews. He developed his research methods as he went, through the reasonable fluidity and lack of experience that characterise any PhD project. This has resulted in some issues in his data management, such as using recording equipment with limited capabilities and uncertainty about the use of software for qualitative analysis.

Paul was unaware of UCL and funding (AHRC) requirements. Even though he has acquired ethical approval from the UCL Research Ethics Committee, he expressed the need for more guidance on what he should be doing with his data because he does not use consent forms as most researchers do.

Paul realised at some point half through the project that he would prefer to embargo his data, an option provided for by UCL Discovery. Paul did not complete his case study because of time pressure to move on with his research and lack of adequate time to familiarise with his data before REWARD was over.

**Case Study 6**

*The Muisca Metallurgist in Context*

Dr Marcos Martinon – Torres, Senior Lecturer

Marcos’ research project constitutes a very interesting case of data reuse. He and his colleague use data collected for insurance purposes of artefacts for archaeological purposes, that is, in order to identify the provenance of the material. They use Excel spread sheets because they are familiar with them and simple to use. Marcos acknowledged that ‘we all have our own ways but that doesn’t mean that they are the best ones’, thus acknowledging that he lags behind in terms of how systematic he is in applying a systematic strategy for backing up all of his work regularly. He also said that he would be ‘happy to change formats depending on how much I need to do to get there, how much work it will require for me to competently use this software, how much more will I get out of it’, which are common and reasonable concerns of researchers who are asked to.

In archaeometallurgy, lab work includes a lot of documentation in the duration of the project (lab-books) and is a process well embedded in this subfield’s research routine. Established routines in the field also promote a more rigorous assessment of data.

Finally, Marcos suggested considering making raw data openly available as they are generated from SEM software.

Marcos’ DMP demonstrated careful and thought through engagement with the tool. His case study was the first one to be completed.

**Case Study 7**

*Communicating Archaeology: Public Perceptions and Experience in a Changing Media Environment*

Chiara Bonacchi, PhD student

Chiara’s research combines quantitative and qualitative data in the field of archaeology communication. Chiara emphasised on the balance between outcomes achieved and economy of
resources in terms of adjusting to a different way of work. The way she works and the software she uses has been determined by what was available to UCL research students and what her tutor advised her and taught her to do.

Chiara has taken a very systematic approach to documentation, storage and backup, even to assessing and discarding earlier versions of her work. As other researchers in heritage management, she does not really know what the total size of her data is.

Chiara perceives her work and her research through the lens of the commonly identified ‘digital divide’. She compared what her data looks like to what case studies in GIS look like and thought that her data did not look as professional and scientific and that she did not want to publish them so that she does not dumb down her work.

Findings:
Several discrepancies were found among various sub-fields of archaeology:

• The ‘digital divide’ was confirmed even within archaeology. Researchers in heritage studies were not as familiar with handling terminology around data types and formats as the ones in material science and computational modelling. At least one researcher also expressed concerns that quantitative data derived from social research did not appear as ‘scientific’ as GIS data. Documentation and documentation standards were not as embedded in research routine in heritage studies as in material science.

• Other discrepancies related to the stage of their career researchers were in. Researchers in early stages of their career placed more value on collection of new data than on the reuse of existing datasets. It was extremely difficult to attract the attention of more senior researchers unless they were already research data management champions.

• Overall, researchers in computational modelling and heritage studies were more likely to share data due to familiarity with data sharing within their own community/with research participants.

• Social data were considered as less likely to be of value for other than historical reasons in the long run because they constitute snapshots of social phenomena only and thus from a point in time onwards their scientific value mainly derives from their historical value.

Contradictory behaviours were also noted:

• Although all researchers demonstrated great reluctance to permanently discard data, they all admitted that they had not considered long term preservation, its cost and how to cover it, until they participated in REWARD.

• Although all researchers were convinced of the value of their work to the extent that their data needs to be preserved infinitely, they had difficulties in discussing the broader relevance of their research.

Areas were identified where researchers needed more guidance:

• In spite of the long institutionalisation of Data Protection and IPR, researchers were still unsure of necessary requirements and of ways to incorporate them in their work.

• There was general lack of awareness regarding data licenses.

• Researchers were similarly unaware of the institution’s requirements on research data management (e.g. the Open Access Mandate and the Code of Conduct for Researchers) and funders’ requirement, when there a funder was involved.

Overall, standards upon which researchers were found to rely on in their decision-making consisted of a balance between validity, ease of access, familiarity and convenience:

• Researchers stated that they do not use the institutional servers to store and backup their work any more because they find them complicated and difficult to access but are increasingly using web-based storage solutions, e.g. Dropbox

Finally, in relation to the main question this project seeks to answer:

• All researchers were willing to share and considered citation as an adequate incentive to do so.
Regarding the DMP Online tool:

- All researchers agreed on the usefulness of data management planning from an early stage of research and of the tool as a systematic approach to it.
- Some researchers found that the DMP Online tool would not be useful at PhD level because of the great changes that a PhD project is likely to undergo and because it cannot meet the particularities of specific research field. However, they thought it was useful for later stages in researchers’ careers, when more confidence in the development of a research project can be expected and definitely in the case of collaborative projects.
- Others found the DMP Online tool very useful as a reflective process that even a PhD student should be made to go through early on, even before the first year review, when he/she is more likely to have more time to inform him/herself in order to best prepare for issue that he/she should anticipate.
- Some researchers found that it was slow.
- Very few researchers used the support provided and as a result there was some misunderstanding of questions. This could potentially be a bigger problem with foreign researchers.
8.2 Appendix 2: Institute of Archaeology Survey

The following online survey was carried out with researchers at the Institute of Archaeology in October 2011 in order to gather background information on their attitudes to data sharing. The survey involved 15 closed questions (yes/no, multiple choice and tick-boxes). It was open for nine days and received 32 responses.

What is your attitude towards data sharing?

This survey aims to understand researchers’ attitudes to data sharing. By data sharing we mean making your data freely available to other researchers, in order for them to validate, reuse and remix it, as long as they cite you as the source.

1. What is your gender?
   a. Male
   b. Female

Male 53%
Female 47%

2. What is your age group?
   a. 20-24
   b. 25-29
   c. 30-34
   d. 35-39
   e. 40-44
   f. 45-49
   g. 50-54
   h. 55-59
   i. 60-64
   j. 65-69
   k. 70-74
   l. 75-79
   m. 80-84

Bigger groups of responses:
   1. 38% 25-34 years old
   2. 28% 35-44 years old
   3. 19% 45-54 years old
3. **What is your current academic role?**

(Choose only one)

**Academic staff:**
- a. Professor
- b. Reader
- c. Senior lecturer
- d. Lecturer
- e. Teaching Fellow

**Research staff:**
- f. Senior Research Fellow
- g. Research Fellow
- h. Research Associate
- i. Research Assistant
- j. Visiting/Honorary:
- k. Research student:
- l. Other (specify):

Bigger groups of responses:
1. 44% Research students
2. 13% Lecturers
3. 9% Senior lecturers
4. 9% Professors

4. **What is your main research area?**

(Choose only one)

**a. Environmental Archaeology**
**b. Archaeological Material Science**
**c. Computational Modelling**
**d. Museum Studies**
**e. Conservation**
**f. Cultural Heritage Studies**
**g. Public Archaeology Social, Historical and Cultural Archaeological Studies**
**h. Other (specify)**

Bigger groups of responses:
1. 38% World Archaeology
2. 31% Archaeological Sciences
3. 24% Heritage
5. **How important do you think the following reasons for preserving digital data are?**
   [Very important/Important/Slightly important /Not important]

   If research is publicly funded, the results should become public property.

   ![Bar chart showing 50% Very important, 41% Important, 6% Slightly important, 3% Not important]

   It will stimulate the advancement of science (new research can build on existing knowledge).

   ![Bar chart showing 75% Very important, 25% Important, 0% Slightly important, 0% Not important]

   It may serve validation purposes in the future.

   ![Bar chart showing 58% Very important, 36% Important, 6% Slightly important, 0% Not important]
It allows for re-analysis of existing data.

![Bar chart showing importance percentages](chart1.png)

- 69% Very important
- 31% Important
- 0% Slightly important
- 0% Not important

It may stimulate interdisciplinary collaborations.

![Bar chart showing importance percentages](chart2.png)

- 47% Very important
- 41% Important
- 13% Slightly important
- 0% Not important

It potentially has economic value.

![Bar chart showing importance percentages](chart3.png)

- 9% Very important
- 31% Important
- 31% Slightly important
- 25% Not important
It is unique.

44% Very important
22% Important
25% Slightly important
9% Not important

6. **Have you ever shared data through:**
   - Personal contact (e.g. an email request) from another researcher?
     a. Yes
     b. No
   - Making it available online?
     a. Yes
     b. No
   - The Institutional Repository?
     a. Yes
     b. No
   - An external Research Data Centre?
     a. Yes
     b. No

Bigger groups of responses:
96% - has shared data through email contact.
43% - have made it available online.
21% - Research Data Centre.
14% - Institutional repository.
7. **Have you experienced or would you expect to experience any of the following problems in sharing your data? (multiple answers possible)**

- Loss of competitive advantage
- Incompatible data types
- Restricted access to a data archive either for depositing or accessing the data
- Legal issues
- Lack of technical infrastructure
- Misuse of data
- Lack of financial resources
- Lack of time to prepare it properly
- No problems foreseen
- Other (please specify)

Bigger groups of responses:
- 63% - Lack of time to prepare data properly
- 50% - Incompatible data types
- 41% - Lack of technical infrastructure
- 34% - Misuse of data
- 31% - Restricted access to data archive
- 28% - Legal issues
- 25% - Lack of financial resources
- 22% - Loss of competitive advantage
- 6% - No problems foreseen

8. **To which of the following facilities would you be willing to submit digital research data in the near future? (multiple answers possible)**

- Digital archive of organisation (i.e. UCL Discovery/IRIS)
- Digital archive (data centre) of your discipline (e.g. The Archaeology Data Service)
- Publisher (as supplementary data with research paper).
- External web service
- I do not want to submit digital research data to external facilities
- Other (please specify)

Majorities of responses:
- 81% - Research Data Centre
- 78% - Institutional repository
- 75% - Publisher
- 44% - External web service
9. Have you ever used data from other sources?
   a. Yes
   b. No

10. Have you ever requested data from another researcher?
    a. Yes
    b. No

11. Have you ever published in an open access journal?
    a. Yes
    b. No

12. Do journals to which you typically submit your work require you to include relevant digital research data (i.e. data used to create tables, figures, etc.)?
    a. Yes
    b. No
13. Do you think it is useful to link a research paper with the data that underlies it?
   a. Yes
   b. No

94% Yes
3% No

14. Would you want to be credited when your digital research data was used by others?
   a. Yes
   b. No

97% Yes
0% No

15. Would other incentives such as course reductions, additional sabbatical leave dedicated to research, financial support for research related travels, and support for seminars and financial incentives motivate you to share research data?
   a. Yes
   b. No

88% Yes
9% No
8.3 Appendix 3: Recommendations for the Institute of Archaeology

As a result of feedback from the REWARD workshops and case studies, there are three main recommendations that the project makes to the Institute of Archaeology:

1. Reintroduce data management course at graduate level.
2. Fund a research data teaching position. A part-time role where someone is available to answer researcher questions about RDM, and to guide them on the use of appropriate resources.
3. Create an internal intranet page with the following RDM resources:

   **Introduction**
   Covers what RDM is, what UCL and funder requirements are, why data management plans are necessary, benefits of publishing and archiving data.

   **UCL policies**
   - UCL research ethics information: [http://ethics.grad.ucl.ac.uk](http://ethics.grad.ucl.ac.uk)

   **Funder OA policies**
   Full list available at SHERPA/JULIET: [http://www.sherpa.ac.uk/juliet](http://www.sherpa.ac.uk/juliet)

   **Resources for Data Management Plans**
   - DMP Online tool: [https://dmponline.dcc.ac.uk](https://dmponline.dcc.ac.uk)
   - Instructions for getting UCL plans

   **Related Institute of Archaeology courses**
   Interpreting Archaeological Data (undergraduate): [http://www.ucl.ac.uk/archaeology/studying/undergraduate/courses/ARCL2037](http://www.ucl.ac.uk/archaeology/studying/undergraduate/courses/ARCL2037)

   **Places to publish data**
   - Journal of Open Archaeology Data: [http://openarchaeologydata.metajnl.com](http://openarchaeologydata.metajnl.com)

   **Repositories for archiving your data**
   (see for more info: [http://openarchaeologydata.metajnl.com/repositories](http://openarchaeologydata.metajnl.com/repositories):
   - UCL Researchers: UCL Discovery: [http://discovery.ucl.ac.uk](http://discovery.ucl.ac.uk)
   - UK Data: Archaeology Data Service: [http://archaeologydataservice.ac.uk](http://archaeologydataservice.ac.uk)
   - US/International data: tDAR: [http://www.tdar.org](http://www.tdar.org)
   - Netherlands data: DANS: [http://www.dans.knaw.nl](http://www.dans.knaw.nl)
   - General: Figshare: [http://figshare.com](http://figshare.com)

   **Information about data licensing**
   List at JOAD: [http://openarchaeologydata.metajnl.com/faq/#q5](http://openarchaeologydata.metajnl.com/faq/#q5)

   **Information on data citation**
   - DCC guide: [http://www.dcc.ac.uk/resources/how-guides/cite-datasets](http://www.dcc.ac.uk/resources/how-guides/cite-datasets)

   **Link to REWARD project page** [http://www.ucl.ac.uk/reward](http://www.ucl.ac.uk/reward)