



JISC Project Plan

Project Information			
Project Identifier	<i>To be completed by JISC</i>		
Project Title	Data Management Rollout at Oxford (DaMaRO)		
Project Hashtag	#damaroxford		
Start Date	1 October 2011	End Date	31 March 2013
Lead Institution	University of Oxford		
Project Director	Prof. Paul W. Jeffreys		
Project Manager	Dr James A.J. Wilson (Interim PM: Pip Willcox)		
Contact email	James.wilson@oucs.ox.ac.uk , Pip.willcox@bodleian.ox.ac.uk		
Partner Institutions	Oxford University Computing Services, Research Services, and Bodleian Libraries, University of Oxford		
Project Webpage URL	http://damaro.oucs.ox.ac.uk/ , http://blogs.oucs.ox.ac.uk/damaro/		
Programme Name	Managing Research Data Programme (02): 2011-13		
Programme Manager	Simon Hodson		

Document Information			
Author(s)	Pip Willcox		
Project Role(s)	Interim Project Manager		
Date	8 November 2011	Filename	
URL	<i>If this report is on your project web site</i>		
Access	This report is for general dissemination		

Document History		
Version	Date	Comments
0.95	8 November 2011	Interim Project Manager's first draft
1	20 December 2011	The Project Plan: first iteration
1a	20 December 2011	The Project Plan: budget redacted

Table of Contents

N.B. This table of contents 'auto-populates' - to update the table of contents – place cursor in the table of contents, right-click your mouse, click 'update field', select appropriate option

1.	Project Overview	3
1.1	Project Summary.....	3
1.2	Objectives.....	3
1.3	Anticipated Outputs and Outcomes	4
1.4	Overall Approach	5
1.5	Anticipated Impact.....	5
1.6	Stakeholder Analysis.....	5
1.7	Related Projects	6
1.8	Constraints	7
1.9	Assumptions	7
1.10	Risk Analysis	7
1.11	Technical Development.....	8
1.12	Standards	9
1.13	Intellectual Property Rights	10
2	Project Resources.....	10
2.1	Project Partners.....	10
2.2	Project Management	10
2.3	Project Roles.....	11
2.4	Programme Support.....	11
3	Detailed Project Planning	11
3.1	Evaluation Plan	11
3.2	Quality Assurance	12
3.3	Dissemination Plan	12
3.4	Exit and Embedding Plans	14
3.5	Sustainability Plans	14
	Appendices	15
	Appendix A. Project Budget.....	15
	Appendix B. Workpackages	15

1. Project Overview

1.1 Project Summary

The Data Management Rollout at Oxford (DaMaRO) Project is creating a research data management policy for the University and the infrastructure to enable researchers to comply with it.

We will take the outputs of the various research data management projects that the University has been engaged in over the last few years and combine them into a better-integrated suite of tools and discovery mechanisms that will support researchers throughout the data life-cycle, from planning to re-use.

Of particular note is the DataFinder tool that the Bodleian Libraries will be developing as the hub of the DaMaRO systems infrastructure. This will enable the discovery of data hosted in various places around the University and beyond, including the Bodleian Libraries' 'DataBank' (developed through the [Admiral](#) and [DataFlow](#) projects), the Database as a Service (DaaS) system (created during the [Sudamih](#) and [ViDaaS](#) Projects), departmental and other local data stores, the Web 2 research management network Collective Wisdom ([Colwiz](#)). Where possible it will connect this data with research papers and publications held in the [Oxford University Research Archive \(ORA\)](#).

Behind the scenes, DataFinder will be able to automatically gather metadata from each of the tools and repositories it connects to, assign Digital Object Identifiers (DOI) where desirable, and ensure that the metadata complies with the national [DataCite](#) standards. Furthermore, DataFinder will make the metadata it gathers available as linked data, and also map it to the [CERIF](#) standard, so that appropriate alerts can be issued to compatible research management systems.

DataFinder will be designed to be implementable as a hierarchical structure, so that an institutional instance can harvest data from departmental instances and, ultimately, a national instance could harvest the data from institutional instances, forming a UK data discovery tool. It will be a freely-available web service.

Besides software development, the DaMaRO Project will develop data management training and documentation for researchers, expanding and adding to the materials developed during the [Sudamih Project](#) and other projects funded by [JISC Managing Research Data Programme](#). A long-term business plan for providing and maintaining the DaMaRO environment will be developed during the course of the project.

DaMaRO will work alongside the Oxford DMP Online Project [website forthcoming] to ensure that the mechanisms it develops are integrated with data management planning tools used in the University.

1.2 Objectives

Project Identifier: DaMaRO

Version: 1.0a

Contact: Pip Willcox

Date: 20/12/2011

The Project will embed and integrate the outputs of a number of UMF and JISC-funded research data management projects into an enhanced institutional infrastructure, supported by researcher training and guidance, and underpinned by a University Data Management Policy.

The outcome will be an infrastructure capable of supporting aspects of the research data lifecycle from planning to re-use.

The infrastructure at Oxford will be modular and federated, allowing for future growth, development, and adaption as the University and local and national data management environment evolves.

In particular, the project will focus on development of systems capable of ingesting and exposing metadata, via a semantically-aware catalogue of Oxford datasets, 'DataFinder'.

This tool will be able to harvest metadata from compliant systems and data stores, both internal and external to Oxford.

The Project will start from the institutional data policy drafted under the EIDCSR Project and seek to provide the infrastructure required to enable researchers to comply with aspects of that policy. As the infrastructure is put in place, the policy will be refined and ultimately submitted to the University's Research Information Management Sub-Committee committee for ratification.

The training materials, university-wide research data management policy, and DataFinder tool will be complete by 31 March 2013.

1.3 Anticipated Outputs and Outcomes

Output / Outcome Type <i>(e.g. report, publication, software, knowledge built)</i>	Brief Description
Project website	Publication of up-to-date information and links as the project progresses.
Project blog	Share knowledge and experience gained during the project with at least one blog entry per month.
Training and support plan	Plan for implementing a training and support module for data planning.
Training and support materials	Supporting materials for training colleagues in data discovery through DataFinder, using the DataCite metadata kernel.
Communications plan	Plan for dissemination of project findings within the data management and curation, and research communities,.
Policy documentation	In consultation with University management, formulate an institution-wide plan for implementing data management policy.
Metadata standards knowledge and specifications	A review of metadata standards to maximize data citation and discovery will establish fields to be used by DataFinder.
Software development: metadata capture	Software to assist automated capture of metadata from existing sources, including DaaS, DataStage, Colwiz, external cloud-hosted services.
Software development: data storage	An ingest service for SWORD-compliant datasets to DataBank.
Software development: semantically aware catalogue	DataFinder, the hub of the technical infrastructure, will aggregate metadata from disparate sources, provide dataset registration (through DOIs), and discovery, and location and access details.
Software documentation	Service Level Description for each component part of the DaMaRO project.

Sustainability plan	Cost/benefit analysis and business plan to cost management, preservation and curation of research data.
Workshop	Direct dissemination of the knowledge and technical experience gained in the project.

1.4 Overall Approach

The project is to be conceived and managed in four distinct but interconnected strands under which services are reviewed and developed. Each strand will be driven by an implementation group, with a single individual assigned as the lead with responsibility for delivery of work under that strand. The four strands will continue beyond the life of the project.

1. Ratification of University-wide policy on research data management.
2. Link together the University's technical infrastructure to provide researchers with the support require for managing their research data,
3. Provide training for researchers and data-curators in the use of the services for storage, linking and location of data.
4. Create a business plan for sustaining the infrastructure of research data at Oxford.

1.5 Anticipated Impact

Impact Area	Anticipated Impact Description
Institutional policy	University-wide ratified agreement regarding research data management.
Researcher behaviour	Through the provision of the technical infrastructure and training in its use, researchers' ability to store, locate and reuse data will be facilitated, offering increased possibilities for reproducing research, compliance with funder policies and supporting research excellence.
Data management infrastructure	DataFinder will enable dataset discovery and location, through ingesting and exposing metadata, across a range of pre-existing and potential, future data stores.
Sustainable infrastructure	The business plan will provide guidance of the cost of research data curation for future data-producing projects. It will offer recommendations for the sustainability of the technical and support services being developed as part of the project, suggesting ways DataFinder can be managed and maintained.
Communication of ideas	The research data management community will benefit from the infrastructure model and best practice this project will develop and implement. Developing open-source software will allow community members to contribute to and learn from the process.

Impact Areas: maintain research excellence; maintain teaching & learning excellence; be more effective/save money; have a positive impact on wider society; be ready for technology needs in the future.

1.6 Stakeholder Analysis

Stakeholder	Interest / stake	Importance (H/M/L)
University of Oxford	The systems, policies and practices implemented during DaMaRO will enable the University to manage its	High

	data assets and comply with the requirements of the Funding Councils.	
Academic Researchers at UK HE Institutions	Academic researchers will be able to discover easily, and, where appropriate, retrieve and re-use Oxford-generated data that would not have been known or accessible to them in the past.	High
Research Support Staff at Higher Education Institutions besides Oxford	Research support staff at universities can learn from the experiences and outcomes of the DaMaRO Project when implementing research data management infrastructure at their own universities.	High
UK Funding Councils	The UK Funding Councils will be able to maximize the value of the research they fund by there being systems in place to manage it, cite it, and disseminate it.	Medium
Digital Curation Centre (DCC)	The DCC will be consulted over the training and support work, and the materials developed during this work package will be fed back to the DCC for re-use elsewhere.	Medium
JISC	DaMaRO will in some respects show the work of the Research Data Management programme beginning to reach fruition, as various JISC- and UMF-funded tools and services are integrated into a coherent infrastructure which can be promoted to and used by researchers within an institutional context. It will provide a model to recommend (or otherwise!) to other HE institutions.	Medium
Internal and external graduate training organizations in professional skills	The training and support developed during DaMaRO will help inform broader research skills training	Low
University and Colleges Information Systems Association (UCISA)	UCISA are likely to be interested in the ways in which the data management systems developed by the project can be integrated with academic information system more broadly.	Low
Research Information Network (RIN)	The RIN will be interested in the effects of the policies and infrastructure developed by DaMaRO on researchers and research practices.	Low

1.7 Related Projects

Project Identifier: DaMaRO

Version: 1.0a

Contact: Pip Willcox

Date: 20/12/2011

External

C4D (CERIF for Datasets): <http://www.sunderland.ac.uk/beacon-di/exampleprojects/ceriffordatasetsc4d/>

CERIFy: <http://cerify.ukoln.ac.uk>.

Collective Wisdom (Colwiz): <http://colwiz.com>.

data.bris: <http://data.bris.ac.uk/>.

DataCite: <http://www.datacite.org>.

EuroCRIS CERIF project: <http://www.eurocris.org>.

IRIOS (Integrated Research Input and Output System): <http://www.irios.sunderland.ac.uk>.

IRIOS2: <http://www.irios.sunderland.ac.uk>.

RMAS (Research Management and Administration System): <http://www.exeter.ac.uk/research/rmas>.

SWORD: <http://swordapp.org/>.

Internal

A Data Management Infrastructure for Research Across the Life sciences (ADMIRAL): <http://imageweb.zoo.ox.ac.uk/wiki/index.php/ADMIRAL>.

Building the Research Information Infrastructure (BRII): <http://brii.bodleian.ox.ac.uk/>.

DataBank: <https://databank.ora.ox.ac.uk>.

DataFlow: <http://dataflow.ox.ac.uk>.

Embedding Institutional Data Curation Services in Research (EIDCSR): <http://eidcsr.oucs.ox.ac.uk/>.

Open Source Software Advisory Service (OSS-watch): <http://www.oss-watch.ac.uk>.

Oxford DMPonline Project: <http://cottagelabs.com/projects/oxforddmponline/>.

Oxford Research Archive (ORA): <http://ora.ox.ac.uk>.

Supporting data management infrastructure for the Humanities (SUDAMIH): <http://sudamih.oucs.ox.ac.uk/>.

Virtual Database as a Service (ViDaaS) project: <http://ViDaaS.oucs.ox.ac.uk>.

1.8 Constraints

DaMaRO is working to a budget and timetable agreed by JISC. Beyond the usual limitations associated with these arrangements, the main constraints are:

- Standards compliance.
- Robustness of resulting software.

1.9 Assumptions

We assume that all partners will be able to honour their commitments regarding staff time, budgets, and accounting standards, and that projects whose work DaMaRO will link up will be completed on time and to their stated specifications. Beyond contractual obligations, we are counting on the good will and energy of the staff involved in the project to enable us to find new solutions to any technical limitations encountered. We also assume that we will be able to find, recruit, and work closely with an expanding group of enthusiastic test users who will thoroughly test the system and give useful feedback for its improvement.

We assume that the scope of DataBank (currently in draft) is for smaller “long tail” data, up to about 50Gb.

1.10 Risk Analysis

Risk Description	Probability (P) 1 – 5 (1 = low)	Severity (S) 1 – 5 (1 = low)	Risk Score (PxS)	Detail of action to be taken (mitigation / reduction / transfer / acceptance)
------------------	---------------------------------------	------------------------------------	------------------	--

	5 = high)	5 = high)		
Staffing				
Failure to allocate appropriate staff for the project.	2	4	8	Use existing staff with experience of working on data management infrastructure projects where possible.
Loss of key staff before end of the project.	2	3	6	Ensure regular communication between all project staff, so that processes and progress are clearly understood by the team, and staff may be redeployed to cover different work packages.
Organizational				
Data Management Policy proves controversial and is not accepted by the Research Committee.	3	4	12	Consider lessons learnt from draft policy developed by EIDCSR Project; involve Research Committee members in drafting policy; consult with University of Edinburgh and other universities who have already considered similar policy issues.
Lack of a sustainable business model.	2	4	8	Take advantage of expertise and support materials developed by JISC to ensure that all aspects of costs/benefits are considered.
Expectations mismatch between project and wider research community.	2	3	6	Ensure cross-disciplinary academic representation on Steering Group. Ensure requirements gathering is broad and feeds through to other project strands.
Estimations of time required to complete technical work packages are inaccurate.	2	2	4	Allocate staff members who are familiar with the nature and scale of this type of project. Closely monitor progress on each WP, and allocate resources accordingly.
Lack of coordination between project stakeholders.	1	3	3	Ensure clear reporting and communication lines; take advantage of existing institutional communication structures
Technical				
Project adopts different standards from other UK HEI	1	4	4	Follow JISC guidance regarding standards; consult with other HEIs.

1.11 Technical Development

DataBank (on which DataFinder will be based), is based on standard well-proven, off-the-shelf web components, and provides a simple, generic model that can add value to almost any research

environment. DataBank enables “sheer curation” and “curation by addition”. Consequently DataFinder will provide a sound platform for specialized data discovery and retrieval services. Regular meetings with ViDaaS, DataFlow and OxfordDMP will ensure synchronized co-development of software.

Approaches adopted for DataStage and earlier IBRG projects (e.g. ADMIRAL, FlyWeb) will be continued in DaMaRO, including:

- **Sheer curation**, employing familiar tools, working practices and dataset formats.
- **Agile development** techniques and **user-led design**, with ongoing requirements analysis.
- **Test-led development**, capturing requirements as automatically testable functions.
- Third-party **open source software**, and adoption of **standards** wherever possible.
- Loose coupling of services with programmatic access, using a **‘RESTful’ approach**.

For the DaMaRO Project, as we move to creating industrial-strength code, we will also adopt a **continuous integration** approach. Nightly code check-in and builds will lead to robust, reliable and trustworthy code, and will develop coherence within the team.

We will coordinate the software development via Github, and the test suite via Jenkins.

1.12 Standards

Name of standard or specification	Version	Notes
SWORD2-compliant clients and servers	V2	As advised by JISC, this protocol will be employed to ingest data. All the major repository systems (DSpace, EPrints, Fedora, etc) are SWORD-compliant servers, while a few systems, such as the Open Journals System interface and the BioMed Central repository depositor provide SWORD clients. We will develop DataBank as a SWORD 2 compliant server and DataStage as a SWORD 2 client.
DataCite	5	Metadata will adhere to DataCite standards as a minimum.
DataCite Metadata Kernel		These terms have already been mapped to RDF as part of the JISC Open Citations Project, as described in DataCite2RDF – Mapping DataCite Metadata Scheme Terms to Ontologies.
Direct Object Indicator (DOI) minting and assignment for DataBank deposited items		Via the DataCite API, currently in development at the British Library.
Object name creation and validation for persistent URLs		
CERIF	Most recent version	The DataFinder tool will be CERIF-compliant for future data exchange. It is a complex standard, beyond the scope of this project to implement in its entirety, but a key standard for the description and management of research information, which is gaining ground with both funding agencies and institutions. DaMaRO will work to identify the subset of CERIF that would enable DataFinder to generate a CERIF-compatible

		notification when a project output is made available.
Metadata standards		To include RDF, DC, as appropriate.
Configurable SOLR indexes for searching content		
Linux		Ubuntu Linux has been chosen as the most widely-used, well-supported Linux, based on the respected Debian Linux distribution.

1.13 Intellectual Property Rights

Any IPR resulting from this project will remain the property of the organization generating it. Results arising from projects funded by the JISC at Oxford would therefore usually be owned in the first instance by the University as the employing institution. It is proposed that software outputs developed within this project will be released under an OSI-approved open source software licence.

2 Project Resources

2.1 Project Partners

The DaMaRO project is based wholly at the University of Oxford, in the departments of Oxford University Computing Services (<http://www.oucs.ox.ac.uk/>), the Bodleian Libraries (<http://www.bodleian.ox.ac.uk/>), and Research Services (<http://ox.ac.uk/research/>).

2.2 Project Management

A Steering Group and an Advisory Board comprising internal and external representatives from the key stakeholder communities will support and guide the work of the project. The Steering Group will meet twice a year for the duration of the project, and will raise awareness of the project's work at the policy-making levels of the University, and beyond. The Advisory Board will meet four times a year, and will adopt a directing role, communicating expertise into the project, and raising awareness of the work of the project within the research community.

Monthly project working group meetings will be held. There will be reporting lines within the Bodleian Libraries and Oxford University Computing Services, and to the PICT (PRAC (Planning and Resource Allocation Committee) ICT (Information and Communication Technology) sub-committee).

We have a relatively flat management structure controlled by the Software Development Manager and Project Manager who operate under the guidance of the Principle Investigators, and based on open and frequent communication between team members, related to achievement of particular goals. Communication bottlenecks and strategic decisions are dealt with by the group meetings, composed of the Principal Investigators, Project Managers and Software Development Manager. The project working group meets face-to-face once a month (agendas circulated in advance – all project members entitled to attend). The software developers are in frequent contact with the Software Development Manager (at least once a week) and also meet in person every other month.

For each task in our work plan, responsibilities are decided at project group meetings, or in software development meetings, enabling clear accountability and instigation of appropriate communication between team members. Communication between DaMaRO and other related projects (ViDaaS, DataFlow and OxfordDMP) will be ensured by regular meetings.

2.3 Project Roles

Team Member Name	Role	Contact Details	Days per week to be spent on the project
Professor Paul Jeffreys	Principle Investigator	Paul.jeffreys@oucs.ox.ac.uk ; 01865 273200	Support as needed.
Dr James Wilson	Co-PI, Project Manager	James.wilson@oucs.ox.ac.uk ; 01865 613489	4.5 (April 2012-March 2013, 0.5 FTE)
Pip Willcox	Interim Project Manager	Pip.willcox@bodleian.ox.ac.uk ; 01865 280026	2.5 (1 November 2011-31 March 2012, 0.5 FTE)
Asif Akram	Senior Developer	Asif.akram@oucs.ox.ac.uk ; 01865 273200	2.5 (April 2012-March 2013, 0.5 FTE)
Alexander Huber	Metadata Specialist	Alexander.huber@bodleian.ox.ac.uk ; 01865 280032	0.1 (October 2011-March 2012, 0.07 FTE)
Neil Jefferies	Development Manager	Neil.jefferies@bodleian.ox.ac.uk ; 01865 280588	Support as needed.
Dr Meriel Patrick	Analyst	Meriel.patrick@oucs.ox.ac.uk ; 01865 273200	1.6 (April 2012-March 2013, 0.33 FTE)
Sally Rumsey	Digital Collections Development Manager, member of Advisory Group	Sally.rumsey@bodleian.ox.ac.uk ; 01865 283860	Support as needed.
To be appointed	Systems Developer		5 (February 2012-February 2013, 1 FTE)

2.4 Programme Support

We would appreciate support in coordinating the standards we are using with other projects in related fields, and in providing opportunities for sharing the knowledge and expertise gained during this project.

3 Detailed Project Planning

3.1 Evaluation Plan

Timing	Factor to Evaluate	Questions to Address	Method(s)	Measure of Success
First quarter, 2013	Integration of DataFinder and DataBank	Does a deposited item's metadata appear in DataFinder?	Test deposit.	Presence of metadata.
First quarter, 2013	Integration of DataFinder and ViDaaS	Does metadata for an item in ViDaaS appear in DataFinder?	Test deposit.	Presence of metadata.
First quarter,	Integration of DataFinder and	Can external research networks, such as	Check sums to validate data.	Successful import and export of data, and

2013	external research networks, e.g. Colwiz	Colwiz, share data with DataFinder?		reliable DataFinder search results returned.
March 2013	Institutional approval	Do members of RIMSC agree to a University-wide data management policy?	Submission of draft policy to RIMSC.	Ratification of institutional policy.
March 2013	Overall success of DaMaRO	Is data management rolled out at Oxford?	Post-implementation review by Steering Group.	Implementation of DataFinder as a service.

3.2 Quality Assurance

When will QA be carried out?	Who will carry out the QA work?	What QA methods / measures will be used?
Output / Outcome Name	Project website and blog	
October 2011, and ongoing	Project managers	Follows W3C standards, and passes validation checks. Reviewed by Project Working Group.
Output / Outcome Name	Software	
January 2012-March 2013	Software developers	Software will be assessed automatically through the use of continuous integration.
Output / Outcome Name	Training and support materials	
April 2012-March 2013	Project managers	Training sessions' efficacy will be assessed by questionnaire.
Output / Outcome Name	Project workshop	
TBD	Project Managers	Workshop's usefulness will be assessed by questionnaire.
Output / Outcome Name	Final business case	
March 2013	Project Managers	The business case is clear about anticipated costs and benefits and clearly explains next steps and future responsibilities. We will draw on expertise from OSS Watch in assessing the costs and benefits, and use existing tools where available. It will be approved by the Steering Group and Advisory Board.
Output / Outcome Name	Publications and presentations	
Throughout	Project Managers and Project Working Group	Acceptance by peer-review panels. Reviewed by Project Working Group.

3.3 Dissemination Plan

Timing	Dissemination Activity	Audience	Purpose	Key Message
--------	------------------------	----------	---------	-------------

Throughout	Project website	The data management community, researchers within and beyond the University.	To communicate the project's objectives and achievements, providing links to other institutions and relevant information.	To cement DaMaRO's place as a key project in the field of research data management.
Throughout	Project blog	Project partners, people working on related projects in the field.	To communicate the ongoing technical development and other project-related work, e.g. conferences attended, presentations given.	Communication of ongoing project activities.
Throughout	Project twitter feed	People working on related projects in the field; researchers	To provide relevant updates of project activities, and links to more information.	Informal communication with interested parties.
March 2012 onwards	Publicizing DaMaRO within Oxford: keep ITLP and BDLSS informed of progress; produce related informative materials; investigate divisional and departmental channels for reaching policy-makers.	Policy-makers	Make colleagues aware of the project's work, and the role DataBank will play in managing research data. Use data collected from test-users to inform ongoing software development.	Raise awareness amongst colleagues who advise on data management.
March 2012 onwards	Publicizing DaMaRO within Oxford: recruit test-users; produce related training materials; advertise hands-on training sessions; investigate divisional and departmental channels for reaching researchers.	Department and division heads, researchers.	Make colleagues aware of the project's work, and the role DataBank will play in managing research data. Use data collected from test-users to inform ongoing software development.	Raise awareness and train: researchers; colleagues who guide research activities.
March 2012 onwards	Publicizing DaMaRO within Oxford: keep ITLP and BDLSS informed of progress; produce related training materials; advertise hands-on training sessions.	OUCS, Research Services and Bodleian Libraries staff.	Make colleagues aware of the project's work, and the role DataBank will play in managing research data.	Raise awareness and train: colleagues who advise on aspects of research activities; colleagues who

			Use data collected from test-users to inform ongoing software development.	advise on data management.
Throughout	Publicizing DaMaRO beyond Oxford: participate in JISC events; establish contact with other research data management projects and liaise as appropriate; identify key stakeholder groups and compile contact list; circulate significant news.	The research data management community.	Sharing knowledge, software, pitfalls and successes with interested parties.	Raise awareness beyond the University of the project's work.
TBD	Project workshop	Researchers, colleagues within the data management community, policy makers.	Discuss face-to-face the project's usefulness to the institution and individual researchers and research groups.	Information-sharing with target users and policy-makers.
April 2012-March 2013	Publications and presentations	The research data management community.	Formal discussion of the project's work, outputs and outcomes to participate in the wider conversation on managing research data.	Recording project expertise and experience.

3.4 Exit and Embedding Plans

Project Outputs/Outcomes	Action for Take-up & Embedding	Action for Exit
Delivery of DaMaRO as a service	Transition from project to service.	Release DataFinder to open source community.
Draft data management policy	Submit draft policy to RIMSC for ratification.	Ratification of data management policy.

3.5 Sustainability Plans

OSS Watch hopes to maintain a limited ongoing user support service, to the limit of its unfunded ability to do that for this project.

Project Outputs	Why Sustainable	Scenarios for Taking Forward	Issues to Address
Software documentation	Openly available.	Ensure persistence.	If any further development of the software occurs

			beyond the funding period, the documentation should be updated.
DaMaRO blog and website	Openly available.	Ensure persistence.	The contents of the project website and blog will continue to be curated by OUCS.
Training materials	Openly available.	Test thoroughly, update after trialling. Publish online.	Ongoing preservation and maintenance of the project website.
Reports: requirements analysis, business plan, progress, evaluation and final report	Openly available.	Deliver to appropriate individuals, publish online where appropriate, deposit copies in Oxford Research Archive (ORA).	Ongoing preservation and maintenance for ORA and project website.
DaMaRO as a service	The business plan, constructed during the project, will address means of sustaining the work beyond the funding period.	Develop the business plan, in consultation with interested parties, to define ongoing support levels and extract maximum value from DaMaRO's work.	Ensure all interested parties can comply with undertakings in the business plan, that actions can realistically be undertaken and maintained at the stated costs.

Appendices

Appendix A. Project Budget

The budget has been redacted from this version of the project plan.

Appendix B. Workpackages

WP1 – Project Initiation
 WP2 – Project Management
 WP10 – Communication & Dissemination

Strand 1 Research data management policies (Glenn Swafford)	Strand 2 Training, support and guidance (James A J Wilson)	Strand 3 Technical development and maintenance (Neil Jefferies)	Strand 4 Business plan for sustainability (Paul Jeffreys)
WP 3 Institutional policy	WP 4 User training and advice	WP 5 Semantic infrastructure WP 6 Provision of services (ViDaaS, DataFlow, Colwiz) WP 7 Data storage WP 8 Dataset discovery and access	WP 9 Business case and cost models

Work Packages	2011			2012												2013		
	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M
WP1: Project Initiation	█	█	█															
WP2: Project Management	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█
WP3: Institutional Policy									█	█	█	█	█	█	█	█	█	█
WP4: User Training & Support						█	█	█	█	█	█	█	█	█	█	█	█	█
WP5: Metadata Standards	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█
WP6: Metadata Capture						█	█	█	█	█	█	█	█	█	█	█	█	█
WP7: Data Storage				█	█	█	█	█	█	█	█	█	█	█	█	█	█	█
WP8: Data Discovery & Access				█	█	█	█	█	█	█	█	█	█	█	█	█	█	█
WP9: Business Case & Cost Models																		
WP 10: Dissemination	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█

<p>WP1: Project Initiation Set up the project and tools to assist management and monitoring. Involves: recruiting any outstanding members of staff; building and publishing the project website and blog; producing a detailed project plan with requirements definitions, technical architecture, detailed stakeholder analysis, communications plan, and risk analysis.</p>	<ul style="list-style-type: none"> - Establish project working environment & management tools (e.g. SharePoint, etc.). [October 2011] - Publish project website & blog. [October 2011] - Undertake detailed risk analysis. [November 2011] - Undertake detailed stakeholder analysis. [November 2011] - Complete project plan. [November 2011] - Recruit staff to any positions not already filled at start of project. [December 2011]
Project Manager	
<p>WP2: Project Management Ensure timely and efficient delivery of project deliverables; update plans as required; write regular progress reports both internally and for JISC; manage the finances and resource allocation; liaise between project collaborators.</p>	<ul style="list-style-type: none"> - Create detailed workpackage chart, including dependencies. [January 2012] - Write monthly OUCS & Bodleian Libraries progress reports. - Hold monthly working group meetings. - Write termly PICT reports.

Operates throughout project.	<ul style="list-style-type: none"> - Write termly JISC reports. - Hold Steering Committee meetings as appropriate. - Produce draft final JISC report. <i>[February 2013]</i> - Produce final JISC report. <i>[March 2013]</i>
Project Manager	
<p>WP3: Institutional Policy</p> <p>Construct an institutional policy, informed by the infrastructure developed earlier in the project, ensuring researchers have the tools required to follow and adhere to it. Based on the institution-wide data management policy developed during the EIDCSR Project, and using the ten points of the University of Edinburgh policy, refine, seek management support, and plan implementation of an institutional data management policy across the University.</p> <p>The refined policy will be detailed, explained, and disseminated via the University's RDM web pages (http://www.admin.ox.ac.uk/rdm/), amongst other channels.</p> <p>Departmental approaches to implementing policy will be explored to help direct future support.</p>	<ul style="list-style-type: none"> - Meetings to discuss policy with stakeholders. <i>[from March 2012]</i> - Produce new draft institutional policy. <i>[September 2012]</i> - Seek ratification of policy from RIMSC (Research Information Management Sub-Committee to the Research Committee). <i>[December 2012]</i> - If necessary, revise and re-submit policy. <i>[January 2013]</i> - Publicize policy to divisional and departmental representatives. <i>[January 2013]</i> - Update Oxford's Research Data Management website to fit the new policy. <i>[January 2013]</i> - Seek feedback on divisional and departmental approaches to implementing policy, and offer advice as necessary. <i>[March 2013]</i>
Glenn Swafford; Advisory Group; Project Manager	
<p>WP4: User Training and Support</p> <p>In collaboration with the DCC, provide user training, guidance and documentation in order that the policies (WP3) can be implemented effectively, and to facilitate use of systems described in subsequent WPs. Training will explicitly build on, expand, and sustain the training and support materials developed by the Sudamih Project and the other JISC MRD-funded training projects, and will be undertaken in close collaboration with Research Services.</p> <p>Although the emphasis of this work package will be on written documentation in order to minimize support costs, the project will also improve and expand existing data management training courses and assess the value of introducing new one-hour training courses, both general and adapted to particular academic disciplines.</p>	<ul style="list-style-type: none"> - Draft detailed training & support implementation plan. <i>[April 2012]</i> - Consult with DCC over training and support plan, including taking across content from ViDaaS <i>[Apr.2012]</i> - Survey data management training materials produced by JISC MRD-funded projects and identify those that should be adapted for Oxford use. <i>[May 2012]</i> - Adapt training material developed for Sudamih to reflect new policy. <i>[July 2012]</i> - Build training & support module for data planning and costing, integrating DMP Online. <i>[September 2012]</i> - Provide updated induction training for new researchers. <i>[October 2012]</i> - Update DaaS and DataStage training to support outputs of WP6. <i>[November 2012]</i> - Produce training for data discovery, using DataFinder. <i>[January 2013]</i> - Revise training & support materials according to user feedback. <i>[March 2013]</i>
James A J Wilson; Analyst; DCC	
<p>WP5: Descriptive and Discovery Metadata Standards</p> <p>Establish metadata foundations for data management at Oxford. Standards will be determined by requirements for citability/discovery of data, and verification of funding body data mandates. Key metadata standards in these areas are DataCite (http://www.datacite.org) and EuroCRIS's CERIF</p>	<ul style="list-style-type: none"> - Confirm appropriateness of standards and review conformity/compatibility with those used in other UKHEIs. <i>[November 2011]</i> - Establish fields to be used by DataFinder, ensuring they match DataCite elements and may be published as open linked data using the DataCite2RDF mapping undertaken as part of the JISC Open Citations Project.2 <i>[December 2011]</i> - Enable DOI assignment in DataFinder. <i>[January</i>

<p>(http://www.eurocris.org) respectively. To maximize the potential for linking and re-use, as much metadata as possible will be expressed in linked data form. DataFinder should also be able to ingest and publish domain-specific linked data. The DataFinder service will support the fuller DataCite field set internally, but require only the minimal kernel in order to issue DOIs and thereby render a resource citable.</p>	<p>2012]</p> <ul style="list-style-type: none"> - Identify subset of CERIF and enable DataFinder to generate a CERIF-compatible notification when a project output is made available. [March 2012] - Contribute information to training and education outputs to promote DataCite metadata kernel mandatory set across the University as the standard for data storage and citation. [December 2012]
Neil Jefferies; Metadata Specialist	
<p>WP6: Capture of Metadata in the Day-to-Day Creation and Management of Research Data</p> <p>Capture metadata relating to data stored in systems/services/appliances both on Oxford's internal infrastructure and on the HE cloud (initially Eduserv). Initial services consist of:</p> <p>Database as a Service, a deployable research database system, being supported by the UMF-funded ViDaaS Project.</p> <p>DataStage, a simple data file management system with access control, backup and Web access, being supported by the UMF-funded DataFlow Project.</p> <p>Colwiz³, an R&D collaboration and productivity platform for HE researchers that will enable data sharing and discovery. This will be linked to DataFinder as an exemplar external cloud-based system used by researchers to store and manage research data.</p>	<ul style="list-style-type: none"> - Synchronize sources (DaaS, DataStage and Colwiz) with metadata standards determined in WP5. [May 2012] - Develop systems to assist automated capture of metadata from sources. [July 2012] - Extend automated metadata-capture systems to cloudhosted services external to Oxford. [September 2012] - Compile feedback for ViDaaS and DataFlow development teams relating to metadata capture support. [October 2012] - Explore possibility of extracting research metadata from research(ers) using SharePoint. [December 2012] - Investigate possibility of automated capture of metadata from JISC UMF-funded Software as a Service outputs, including LabTrove. [February 2012]
Neil Jefferies; Senior Developer; Analyst	
<p>WP7: Data Storage</p> <p>Design and develop a search and access Web front end for DataBank. DataBank is the Bodleian Libraries' emerging archival-standard semantically aware system for the storage, management, curation and publication of research data. It is designed to store 'small' datasets (up to c. 50MB). It forms part of Oxford's planned federated research data repositories: other data stores will be added or incorporated into the federated data repositories as required. DataBank will provide a SWORD-compliant ingest service for datasets together with their metadata from DataStage and other similar SWORD-compliant clients.</p>	<ul style="list-style-type: none"> - Design and develop a search and access Web front end for DataBank. [February 2012] - Conduct User Acceptance Testing of Web front end and modify accordingly. [March 2012] - Implement ingest via DataStage using SWORD (assuming this is not already implemented via DataFlow Project). [May 2012] - Produce DataBank Service Level Description. [June 2012] - Create a search and retrieval API for Databank, enabling the interrogation of Databank by DataFinder and other Web services. [June 2012] - Produce user documentation for service. [July 2012] - Advise on estimated ongoing and support costs of data storage, to feed into WP9. [August 2012]
Neil Jefferies; Systems Developer	
<p>WP8: Data Discovery and Access</p> <p>Create a semantically-aware catalogue of Oxford datasets: 'DataFinder', to aggregate metadata from multiple disparate sources. DataFinder will form the hub of the technical infrastructure, providing a means for research dataset registration, discovery and location and access</p>	<ul style="list-style-type: none"> - Develop DataFinder to index metadata held in the Oxford federated data repositories. [February 2012] - Incorporate BRIL technologies for data categorization and matching. [March 2012] - Implement SWORD for metadata deposit [April 2012]

<p>details. It will be based on registry technologies developed in the JISC-funded BRIL Project4. Datasets and/or their metadata may be stored in a variety of locations: local departmental or research institute, central University (e.g. DataBank), external (e.g. UKDA or Colwiz). DataFinder will provide an interface to upload metadata, and for metadata to be automatically harvested. Metadata will be stored and indexed, and will be retrieved via a searchable interface. Export tools such as RSS feeds will be provided as appropriate.</p>	<ul style="list-style-type: none"> - Implement OAI-PMH for metadata dissemination/harvest. <i>[May 2012]</i> - Set up metadata acquisition links with DataStage (via DataBank), ViDaaS and Colwiz. <i>[July 2012]</i> - Map RDF to CERIF and enable CERIF export. <i>[August 2012]</i> - Publish metadata as linked data. <i>[September 2012]</i> - Enable DOI assignment. <i>[September 2012]</i> - Design and develop a search and access Web front end for DataFinder. <i>[December 2012]</i> - Design and develop a metadata deposit Web interface for DataFinder for <i>ad hoc</i> addition of items. <i>[February 2013]</i> - Open pipe to SOLO. <i>[March 2013]</i> - Install Google Analytics. <i>[March 2013]</i>
Neil Jefferies; Senior Developer; Systems Developer	
<p>WP9: Business Case and Cost Models Develop a business plan for providing and maintaining the DaMaRO enterprise environment for managing, preserving, and curating research data, informed by the data management policy. The work package will consist of three aspects: an initial benchmarking exercise to establish current practices and costs; a costs/benefits analysis to assess the likely impact of the project; and a business case considering the ongoing costs of the service and the likely return on investment.</p>	<ul style="list-style-type: none"> - Baseline costs report of current (typical) data management lifecycle costs. <i>[June 2012]</i> - Benefits identification & assessment report. <i>[September 2012]</i> - Produce outline business plan template to define required estimations. <i>[October 2012]</i> - Collect ongoing costs estimations from other project strands. <i>[November 2012]</i> - Produce costs/benefits analysis. <i>[December 2012]</i> - Produce business plan & ROI for ongoing service. <i>[February 2013]</i>
Paul Jeffreys; Project Manager	
<p>WP10: Communication & Dissemination The project will participate in programme-wide dissemination activities as well as devising effective communication channels both locally and nationally to ensure awareness and take-up of the services to be provided.</p>	<ul style="list-style-type: none"> - Produce detailed Communications Plan. <i>[November 2011]</i> - Update project website as outputs delivered. <i>[Ongoing]</i> - Post at least one blog entry per month, plus other material of interest to an external audience. <i>[Ongoing]</i> - Stage project workshop. <i>[July 2012]</i> - Participate in other events in collaboration with JISC. <i>[Ongoing]</i>
Project Manager; Analyst	