JISC Grant Funding 07/11

Cover Sheet for Bids *(All sections must be completed)*	**JISC**

Please indicate which strand you are applying for:	**Strand A1**		**Strand B**	
	Strand A2	✓	**Strand C**	_

Name of Lead Institution:	University of Oxford
Name of Proposed Project:	Data Management Rollout at Oxford (DaMaRO)
This Project involves one or more commercial sector partners : NO	**Name(s) of any commercial partner company(ies)**

Full Contact Details for Primary Contact:

Name: Professor Paul Jeffreys
Position: Director of IT, University of Oxford
Email: paul.jeffreys@odit.ox.ac.uk
Tel: 01865 273229
Address: Oxford University Computing Services, 13 Banbury Road, Oxford. OX2 6NN

Length of Project:	18 months		
Project Start Date:	1st October 2011	**Project End Date:**	31st March 2013

Total Funding Requested from JISC:	**£227,374**
Total Institutional Contributions:	**£216,566**

Outline Project Description:

The Data Management Rollout at Oxford (DaMaRO) 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. This 'DataFinder' tool will be able to harvest metadata from compliant systems and data stores, both internal and external to Oxford. It will be implementable as a hierarchical structure, ultimately enabling the provision of a national DataFinder instance that can harvest from institutional instances to form a UK-wide data discovery tool. Each instance can feed higher/lower levels as desired.

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.

I have looked at the example FOI form at Appendix A and included an FOI form in this bid	**YES / ~~NO~~** (delete as appropriate)
I have read the Funding Call and associated Terms and Conditions of Grant at Appendix B	**YES / ~~NO~~** (delete as appropriate)
<u>**For FE institutions only:**</u> **Please tick this box if you are an FE institution in England, please tick this box to confirm that you meet the eligibility requirement of teaching HE to more than 400 FTE**	

Data Management Rollout at Oxford (DaMaRO)

1 Appropriateness and Fit to Programme Objectives and Overall Value to the JISC Community

1. The DaMaRO Project will enhance and extend Oxford's existing pilot research data management infrastructure to enable and encourage the wider re-use and repurposing of research data. It will enable Oxford to comply with policies from research funders that mandate that research data must be deposited in an appropriate repository and made available for public use, interrogation, and scrutiny, for a certain period of time after the end of a project.

2. The research data infrastructure environment developed by DaMaRO will support the full research data lifecycle from planning to re-use. It will build upon and embed experience acquired during the Eidcsr, Admiral, Sudamih, DataFlow, and VIDaaS projects, and draw together the services and training developed during those projects into a coherent and integrated whole. DaMaRO will pay particular attention to implementing and integrating systems to capture metadata (at various stages), link existing repositories, and enable the discovery of datasets. A searchable catalogue of metadata relating to data stored at Oxford (and potentially elsewhere) will comprise one of the key project outputs.

3. The Pproject will start from the institutional data policy drafted under the Eidcsr Project and seek to provide underpinning infrastructure required by researchers to comply with that policy. Each component part of the data management infrastructure developed by DaMaRO will be accompanied by a service-level description indicating the characteristics and requirements of the data that it is designed for, and the security and access levels that are offered. As the supporting infrastructure is implemented, the University will be able to refine its institutional data policy accordingly. The refined policy will be proposed to the University's Research Information Management Sub-Committee to the Research Committee (RIMSC) for consideration and, ultimately, ratification.

4. A key aspect of the project in terms of providing future sustainability will be the development of a long-term business plan defining the future responsibilities of the staff delivering the project strands, their ongoing costs, and the benefits of continued provision. Methods to cover the ongoing costs of provision will be explored.

5. The architecture of the planned infrastructure is represented in figure 1, below:
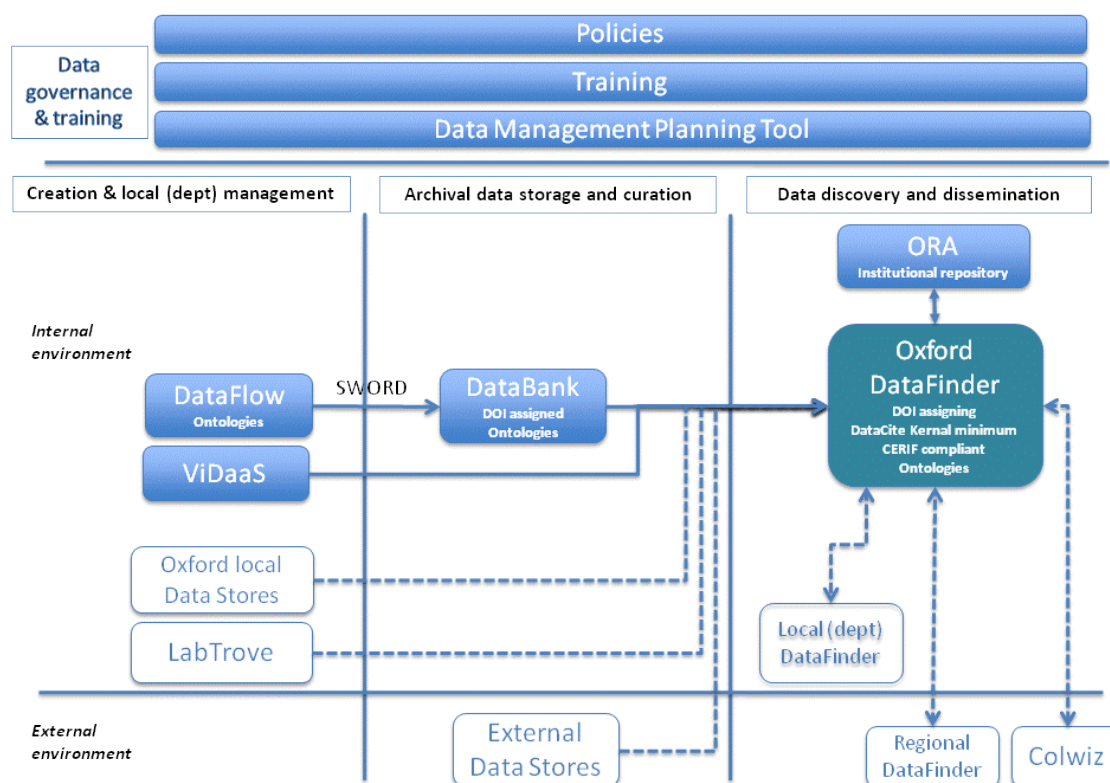

Figure 1: The proposed architecture of the University of Oxford's integrated data management infrastructure

6. The diagram illustrates how the institutional policy defines and rests on top of the facilitating infrastructure. Between the policy and the services diagram sits the training that the institution must provide to academic researchers, both so that they are aware of the overlying policy and of the underlying tools and services that enable compliance with it. This in turn rests upon the data management planning layer; this is the point at which researchers address how they will fulfil policy requirements via the available infrastructure. The DCC's DMP Online tool, which will be customized for local use as part of the related Oxford DMP*online* Project, being submitted to Strand C of this call, is regarded as the medium through which the planning process will be undertaken in future.

7. The services infrastructure section indicates how the DaMaRO Project intends to link together and integrate existing infrastructure – both internal and external – with the new components to enable end-to-end data management. External sources could include cloud-based tiered storage. Data placed into the DataStage system being developed by the UMF-funded DataFlow Project can then be pulled into the University of Oxford's 'DataBank' preservation and curation system. The UMF-funded VIDaaS Database as a Service system has its own mechanisms for preservation. The descriptive metadata enabled by these services will then be pulled into the new Oxford 'DataFinder' service, where it can be searched and discovered.

8. The infrastructure developed at Oxford will be modular and federated, allowing for future growth and adaptation as the environment evolves. Systems and services may be slotted in, removed, or replaced as required. For example, the University of Southampton's planned LabTrove service could in future be plugged in to the infrastructure without modifying the underlying environment.

1.1 DataBank

9. Databank is the Bodleian Libraries' archival-standard semantically-aware data management storage, management, and curation system. The EIDCSR, Admiral, and DataFlow projects have all contributed to the development of DataBank, and the DaMaRO Project will bring it to operational maturity by adding a search and access Web front end and ensuring that it offers a SWORD-compliant ingest service for datasets from other services offering SWORD compliance. DataBank will also be able to import or export data to external repositories, such as the UK Data Archive.

1.2 DataFinder

10. Whilst DataBank will act primarily as a data preservation and curation system, DataFinder will offer a semantically-aware catalogue of Oxford datasets. DataFinder will act as a metadata aggregator, providing a means for research dataset registration, discovery, and access. As such it forms the hub of the technical infrastructure developed by DaMaRO.

11. Metadata may be uploaded to DataFinder from multiple sources as required. Initial sources of metadata will include DataBank, DataStage instances, ViDaaS, and the Web 2 research management network Colwiz[1]. Local stores and external data stores (including national and other data services) can be added later as required. Each entry will comprise metadata based on the DataCite 5 mandatory properties. DOIs will be assigned as required to datasets registered in DataFinder (WP7). This will alleviate the problems arising when datasets change location: DataFinder provides the means to keep track of datasets, and provides a resolver service for discovery and access. The identifier field will, wherever possible, resolve to a URL and provide a location for the data. The location might comprise a splash page with a direct link to open access data, or ownership and access information if there are restrictions.

12. DataFinder will be designed so as to be implementable as a hierarchical structure. For example, there can be an institutional instance (University of Oxford DataFinder), which harvests from secondary local instances (e.g. departmental DataFinder). Ultimately, a national DataFinder could be implemented to harvest from institutional instances to form a UK data discovery tool. Each instance can feed higher/lower levels as appropriate.

13. Where appropriate, data records will cite and link to related publications in Oxford's institutional repository, ORA (Oxford University Research Archive). DataFinder will be indexed by SOLO (the Bodleian Libraries' catalogue), using an OAI-PMH pipe. Colwiz will be able to use the same mechanism to make DataFinder resources discoverable within its own search facility, and provides an easy-to-use mechanism for making data publically available.

---

[1] http://www.colwiz.com

14. DataFinder will be provided as a freely available Web-service. Google Analytics will be installed to enable usage statistics to be gathered for datasets. It will also be CERIF-compliant, so that it can handle future data exchange requirements.

1.3   Value for Money

15. The DaMaRO Project will act as a demonstrator of how the various aspects of infrastructure to support data management research can be brought together to create a flexible institutional system for the management of research data, from creation to re-use.

16. The software developed during the course of the project will be documented and available under an open-source licence and placed in an appropriate open repository. This will enable other institutions to take the code and adapt it to their own purposes.

2   Quality of Proposal and Robustness of Work Plan

2.1   Approach

17. DaMaRO is a collaborative project involving the Computing Services, Bodleian Libraries, Research Services, academic departments at the University of Oxford, and Colwiz.

18. 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.

19. Early prototyping with representative users of the DaMaRO infrastructure will be undertaken to ensure that the finished system meets requirements. Training and support materials will be trialled with researchers to ensure that the infrastructure is comprehensible and usable.

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

2.2   Project Plan

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	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

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.	• 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	
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. Operates throughout project.	• Write monthly OUCS & Bodleian Libraries progress reports. • Hold monthly working group meetings. • Write termly PICT reports. • Write termly JISC reports. • Hold Steering Committee meetings as appropriate. • Produce draft final JISC report. *[February 2013]* • Produce final JISC report. *[March 2013]*
Project Manager	
WP3: Institutional Policy	
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. The refined policy will be detailed, explained, and disseminated via the University's RDM webpages (http://www.admin.ox.ac.uk/rdm/), amongst other channels. Departmental approaches to implementing policy will be explored to help direct future support.	• Produce new draft institutional policy. *[September 2012]* • Seek ratification of policy from RIMSC (Research Information Management Sub-Committee to the Research Committee). *[December 2012]* • If necessary, revise and re-submit policy. *[January 2013]* • Publicize policy to divisional and departmental representatives. *[January 2013]* • Update Oxford's Research Data Management website to fit the new policy. *[January 2013]* • Seek feedback on divisional and departmental approaches to implementing policy & advise as necessary. *[March 2013]*
Glenn Swafford; Advisory Group; Project Manager	
WP4: User Training and Support	
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. 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.	• Draft detailed training & support implementation plan. *[April 2012]* • Consult with DCC over training and support plan, including taking across content from VIDaaS *[Apr.2012]* • Survey data management training materials produced by JISC MRD-funded projects and identify those that should be adapted for Oxford use. *[May 2012]* • Adapt training material developed for Sudamih to reflect new policy. *[July 2012]* • Build training & support module for data planning and costing, integrating DMP Online. *[September 2012]* • Provide updated induction training for new researchers. *[October 2012]* • Update DaaS and DataStage training to support outputs of WP6. *[November 2012]* • Produce training for data discovery, using DataFinder. *[January 2013]* • Revise training & support materials according to user feedback. [March 2013]
James A J Wilson; Analyst; DCC	
WP5: Descriptive and Discovery Metadata	• Confirm appropriateness of standards and review conformity/compatibility with those used in other UK

5

Standards 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 (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.	HEIs. *[November 2011]* • 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] *[December 2011]* • Enable DOI assignment in DataFinder. *[January 2012]* • 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	
WP6: Capture of Metadata in the Day-to-Day Creation and Management of Research Data 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: **Database as a Service**, a deployable research database system, being supported by the UMF-funded VIDaaS Project. **DataStage**, a simple data file management system with access control, backup and Web access, being supported by the UMF-funded DataFlow Project. **Colwiz**[3]**,** 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.	• 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 cloud-hosted 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	
WP7: Data Storage 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.	• 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	

---

[2] http://opencitations.wordpress.com/2011/06/30/datacite2rdf-mapping-datacite-metadata-scheme-terms-to-ontologies-2/
[3] http://www.colwiz.com/

WP8: Data Discovery and Access	
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 details. It will be based on registry technologies developed in the JISC-funded BRII Project[4]. 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.	• Develop DataFinder to index metadata held in the Oxford federated data repositories. *[February 2012]* • Incorporate BRII technologies for data categorization and matching. *[March 2012]* • Implement SWORD for metadata deposit *[April 2012]* • Implement OAI-PMH for metadata dissemination/harvest. *[May 2012]* • Set up metadata acquisition links with DataStage (via DataBank), VIDaaS and Colwiz. *[July 2012]* • Map RDF to CERIF and enable CERIF export. *[August 2012]* • Publish metadata as linked data. *[September 2012]* • Enable DOI assignment. *[September 2012]* • Design and develop a search and access Web front end for DataFinder. *[December 2012]* • Design and develop a metadata deposit Web interface for DataFinder for *ad hoc* addition of items. *[February 2013]* • Open pipe to SOLO. *[March 2013]* • Install Google Analytics. *[March 2013]*
Neil Jefferies; Senior Developer; Systems Developer	
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.	• Baseline costs report of current (typical) data management lifecycle costs. *[June 2012]* • Benefits identification & assessment report. *[September 2012]* • Produce outline business plan template to define required estimations. *[October 2012]* • Collect ongoing costs estimations from other project strands. *[November 2012]* • Produce costs/benefits analysis. *[December 2012]* • Produce business plan & ROI for ongoing service. *[February 2013]*
Paul Jeffreys; Project Manager	
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.	• Produce detailed Communications Plan. *[November 2011]* • Update project website as outputs delivered. *[Ongoing]* • Post at least one blog entry per month, plus other material of interest to an external audience. *[Ongoing]* • Stage project workshop. *[July 2012]* • Participate in other events in collaboration with JISC. *[Ongoing]*
Project Manager; Analyst	

2.3   Project Governance and Management

20. The DaMaRO Project will be overseen by a Steering Committee chaired by the Pro Vice Chancellor (Research, Academic Services and University Collections) or representative. Strategy and policy recommendations which arise will be made through the Research Information Management Sub-Committee to the Research Committee (a principal committee of Council). The Business Plan for the ongoing provision and development of services will be considered by the Research Committee and the Planning and Resource Allocation Committee.

21. A Project Advisory Group comprising senior members of the Research Services, Bodleian Libraries, and Computing Services, plus academic representatives, will advise the Project Team and take particular responsibility for the enhancement of institutional data management policy. Members of the Project Advisory Group will also take responsibility for leading and guiding the four strands of DaMaRO Project work.

---

[4] http://brii.bodleian.ox.ac.uk/

22. The Project Team and members of the Advisory Group will meet at least once per month to ensure that outputs are delivered to an appropriate standard and that issues arising from the work are understood and resolved in a timely manner.

23. The Project Manager will have responsibility for the coordination of day-to-day activities.

2.4 Risk Assessment

24. The risk analysis presented below is only an initial assessment of project risks; a more complete risk analysis will be completed for the overall project plan.

Risk	Probability (1-5)	Severity (1-5)	Score (PxS)	Mitigation
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 HEIs	1	4	4	Follow JISC guidance regarding standards; consult with other HEIs.

2.5 Standards

25. It is an important goal of the project to ensure that the standards adopted are the same as, or compatible with, those adopted by UK HE more broadly. As advised by the JISC, the SWORD protocol will be employed to ingest data. Metadata will adhere to DataCite standards as a minimum, and the DataFinder tool will be CERIF-compliant for future data exchange.

26. The DataCite Metadata Kernel terms have already been mapped to RDF as part of the JISC Open Citations Project, as described in DataCite2RDF[5] – Mapping DataCite Metadata Scheme Terms to Ontologies. CERIF is a much more complex standard, which is beyond the scope of this project to implement in its entirety. However, it is a key standard for the description and management of research information which is gaining ground with both funding agencies (for publishing and activity reporting) and institutions (in the form of Research Information Management (RIM) systems) in the UK and internationally. 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. This could then be picked up by either an institutional RIM system or a funding agency.

---

[5] http://opencitations.wordpress.com/2011/06/30/datacite2rdf-mapping-datacite-metadata-scheme-terms-to-ontologies-2/

2.6 Intellectual Property Rights

27. 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 open source software licence.

2.7 Exit and Sustainability Plans

28. The Business Plan will clearly identify ongoing costs and benefits, making a case for the continued support for each aspect of infrastructure developed, within the context of the infrastructure as a whole (the whole, mostly likely, being greater than the sum of the parts). The ongoing costs of provision will be addressed and a solution found.

29. The division of the project into four 'strands', each of which will be lead by a different individual, is intended to aid sustainability by avoiding the situation where people need to be identified to take elements of the project forward only once the end of the project funding is imminent. Each strand is expected to continue under its current 'ownership' beyond the project itself, ensuring continuity.

30. The software developed during DaMaRO will be open source and made publicly available via an open-source repository such as Google Code. The Software Sustainability Institute (SSI) will be consulted regarding long-term sustainability strategies for the software outputs.

3 Engagement with the Community

3.1 Stakeholder and Practitioner Engagement

31. Project Stakeholders will be engaged throughout the life of the project via the Steering Committee and targeted dissemination activities. There will be emphasis on sharing software and applications. The following stakeholders have been identified from the outset, and a more detailed analysis of stakeholders will be undertaken during the initiation stage of the project.

Stakeholder	Interest / Stake	Importance
University of Oxford	The systems implemented during DaMaRo will enable the University to manage its data assets and comply with the requirements of the Funding Councils.	High
Academic Researchers at UK HE Institutions	Academic researchers will be able to discover, retrieve, and re-use 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.	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-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
Vitae	The training and support developed during DaMaRO will help inform broader research skills training.	Medium
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

3.2 Dissemination Plan

32. DaMaRO will work with researchers to ensure they understand what is being created and why, whilst collaborating with the JISC and the DCC to disseminate plans and findings within the data

management and curation communities. Besides the usual dissemination tools such as a website and blog, DaMaRO will produce a detailed communications plan during the first couple of months of the project, proposing activities to specifically reach representatives of the various stakeholder groups. Training materials developed in the project will be disseminated through the DCC.

4 Budget

[budget not included in this version]

5 Previous Experience of the Project Team

The DaMaRO core Project Team consists of staff from the Computing Services and Bodleian Libraries:

1. **Professor Paul W Jeffreys**, Principal Investigator (Computing Services) – Paul is the Director of IT at the University of Oxford and has been leading Oxford's drive to establish a research data management infrastructure. He was PI of the Sudamih and VIDaaS Projects.
2. **Dr. James A J Wilson**, co-PI and Project Manager (Computing Services) – James has significant experience of managing the development of training and tools to support research data management. He has been Project Manager of the Eidcsr, Sudamih, and VIDaaS Projects.
3. **Dr. Meriel Patrick**, Analyst (Computing Services) – Meriel has acted as the analyst on the Sudamih and VIDaaS Projects, and is an experienced teacher and trainer.
4. **Asif Akram**, Senior Developer (Computing Services) – Asif has been the primary developer of Oxford's *Database as a Service* software during the Sudamih and VIDaaS Projects, and has worked with the Bodleian Libraries DataBank system during Eidcsr.
5. **[Systems Developer t.b.c.]** (Bodleian Libraries) – The Bodleian Libraries will assign a systems developer with knowledge of the Libraries' systems and services to the project.
6. **Alexander Huber**, Metadata Specialist (Bodleian Libraries) – Alexander has considerable experience in metadata management for digital collections.
7. **Sally Rumsey,** Digital Collections Development Manager (Bodleian Libraries) – Sally leads the group responsible for development of the Bodleian Libraries' internal born-digital and digitized collections and manages the Oxford University Research Archive (ORA).She was PI for the BRII (Building the Research Information Infrastructure) Project.
8. **Neil Jefferies**, Development Manager (Bodleian Libraries) – Neil has been instrumental in the development of DataBank and worked with the BRII and Eidcsr Projects.

Drs Wilson and Patrick, and Mr Akram are all assigned to the VIDaaS project until 31st March 2012; it can be stated definitely that no VIDaaS deliverables will be compromised in any way. They will join DaMaRO on the 1st April 2012. A 0.5 FTE Interim Project Manager will guide DaMaRO through the first six months, with James providing guidance and oversight as a co-PI. DaMaRO can also be seen as delivering part of the vision of the UK Research Data Service (national registry for research data).

The Advisory Group will consist of the PIs and directly allocated staff mentioned above, plus: Dr. Sarah Thomas (Bodley's Librarian); Richard Ovenden (Deputy Librarian, Bodleian Libraries); Dr. Glenn Swafford (Director of Research Services); Dr. Andrew Richards (Associate Director Operations and Services, Oxford e-Research Centre); Dr. Michael Fraser (Head of Infrastructure, OUCS); the JISC Programme Manager; researchers representing the academic divisions at Oxford; Kevin Ashley and/or Joy Davidson representing the DCC; a representative from RCUK; and Pamela Reid (Wellcome Trust, Manager – Committees and Coordination).