

<b>Project Proposal : Version for Public Dissemination</b>			
<b>Name of JISC Initiative:</b>		<b>Data Management Infrastructure</b>	
<b>Name of Lead Institution:</b>		University of Oxford	
<b>Name of Proposed Project:</b>		Supporting Data Management Infrastructure for the Humanities (Sudamih)	
<b>Name(s) of Project Partners(s)</b>		<b>Any private sector involvement in the Project</b> NO	
<b>Full Contact Details for Primary Contact:</b> <b>Name:</b> Professor Paul Jeffreys <b>Position:</b> Director of IT <b>Email:</b> paul.jeffreys@odit.ox.ac.uk <b>Tel:</b> 01865 273229 <b>Fax:</b> 01865 273275 <b>Address:</b> Oxford University Computing Services 13 Banbury Road Oxford, OX2 6NN			
<b>Length of Project:</b>	<b>18 months</b>		
<b>Project Start Date:</b>	01 October 2009	<b>Project End Date:</b>	31 March 2011
<b>Outline Project Description</b> The Supporting Data Management Infrastructure for the Humanities (Sudamih) Project aims to address a coherent range of requirements for the more effective management of data (broadly defined) within the Humanities at an institutional level. Whilst the project is fully embedded within the institutional context of Oxford University, the methodologies, outputs and outcomes will be of relevance to other research-led universities, especially but not only, in their support of research within the humanities. The project places emphasis on two particular areas: recognition and support for the “life's work” nature of much of humanities research; recognition and support for the simple and effective creation of online databases for typical data-types within the Humanities (Database as Service for e.g. text, image and geo-data). The Sudamih Project is driven by the requirements of researchers within the Humanities Division at Oxford; will operate as a collaborative project between the research community and institutional service providers; builds on existing internal and JISC-funded strategic activities within Oxford; and will work closely with the Digital Curation Centre(DCC), the Research Information Network (RIN) and the UK Research Data Service (UKRDS) initiative.			

# Supporting Data Management Infrastructure for the Humanities

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

### 1.1 Introduction and Project Motivation

1. The Supporting Data Management Infrastructure for the Humanities (Sudamih) Project aims to address a coherent range of requirements for the more effective management of data (broadly defined) within the Humanities at an institutional level. Whilst the project is fully embedded within the institutional context of Oxford University, the methodologies, outputs and outcomes will be of relevance to other research-led universities, especially but not only, in their support of research within the humanities. The project places emphasis on two particular areas: recognition and support for the “life's work” nature of much of humanities research; recognition and support for the simple and effective creation of online databases for typical data-types within the Humanities (Database as Service for e.g. text, image and geo-data). The Sudamih Project is driven by the requirements of researchers within the Humanities Division at Oxford; will operate as a collaborative project between the research community and institutional service providers; builds on existing internal and JISC-funded strategic activities within Oxford; and will work closely with the Digital Curation Centre(DCC), the Research Information Network (RIN) and the UK Research Data Service (UKRDS) initiative.
2. The Humanities Division in Oxford is amongst the largest centres of humanities research and teaching in the world. The Division supports an extraordinary variety and diversity of research domains, with an equally diverse range of research practice and culture. Research data in the humanities has some distinct differences from data within other disciplinary areas. Primary and secondary sources of data in the humanities often have a long lifespan. Projects such as the Lexicon of Greek Personal Names have been in operation since the 1970s. The Oxyrhynchus Papyri project has enough material for another one hundred years of research. It is clear that research activities, and the associated research data, tend to evolve over decades (and are often intrinsic to the “life's work” of an individual academic). It is often impossible to anticipate at the outset how outputs from humanities research, including so-called 'raw' data, might be re-used. The embracing and extending of methodologies from other disciplinary areas, within and outwith the humanities, and their application to primary sources in order to create new knowledge is one of the hallmarks of humanities research. Of course, this evolutionary process is increasingly dependent on the discoverability and accessibility of the data, but also, from the point of view of the individual scholar or the small research project team, the *manageability* of the data over potentially long periods of time.

### 1.2 Data Management Infrastructure Activities in Oxford

3. The project will build on the data management infrastructure framework developed by an internally funded scoping study<sup>1</sup> and its continuation through the JISC-funded Embedding Institutional Data Services in Research (EIDCSR) project<sup>2</sup>.
4. The former study was a cross-agency collaborative effort to scope the requirements for services to manage and curate research data generated at Oxford and involved interviewing around 40 researchers as well as conducting a consultation exercise with service units across the University. The project also contributed to the UKRDS feasibility study and the piloting of the Data Audit Framework (DAF) through the JISC funded DISC-UK DataShare project.
5. The EIDCSR Project is currently addressing the data management requirements of two collaborating research groups, from medical sciences and computational biology, by scoping their curation and preservation requirements for research data and embedding selected elements of the digital curation lifecycle, including policy, workflow, and sustainability solutions within the research process at an early stage. The project is again an intra-institutional collaboration, bringing together research groups from the Medical Science Division and the Mathematical and Physical Life Sciences Division, the Library (OULS) and Computing services (OUCS), the Oxford e-Research Centre (OeRC), and the Research Service Office

---

<sup>1</sup> The scoping digital repository services for research data management project, <<http://www.ict.ox.ac.uk/odit/projects/digitalrepository/>>.

<sup>2</sup> The embedding institutional data curation services in research project, <<http://eidcsr.oucs.ox.ac.uk/>>.

(RSO). IBM are also contributing to the project.

6. In addition to this, the project will learn from and build on the work undertaken by the Libraries with projects such as Paradigm<sup>3</sup> and FutureArch<sup>4</sup> investigating the management and curation of private archives. Through these activities there is a recognition that early intervention is required to assist authors to manage their own data with a view to future curation. There are a number of lessons learnt and future directions that would be directly applicable to supporting humanities researchers (in particular) to manage the data they create or have responsibility for over (at least) that part of their career they spend in Oxford.
7. The proposed project will also build on other previous work undertaken in the University of Oxford to understand the working practices and requirements from researchers in Humanities and map their activities across the division. Notably these include the JISC-funded Humanities VRE projects<sup>5</sup> and the Digital Humanities at Oxford<sup>6</sup> activities led by the OeRC.

### 1.3 Humanities Research Data Management Requirements in Oxford

8. In preparation for this proposal, an initial requirements analysis was undertaken within the Humanities Division, to establish a baseline set of priorities relating to the data management infrastructure development for scholars within the Humanities at various points in their academic career; undertaking research both individually and as directors or contributors to collaborative research projects.
9. Two separate strands of activity have been identified as a result of an initial high-level requirements analysis; both are underpinned by the "life's work" theme that aims to address longitudinal data management requirements during the course of an academic career.
10. **Data management support and training:** currently there is little formal training and support for data management at any point in the academic career at Oxford. Initial user requirements analysis has established that priority areas for support are likely to include:
  - personal data management (ranging from file/folder naming conventions; version control; synchronisation of working data across devices; to email management and ensuring data remains findable and meaningful many years later);
  - managing the process of sharing data, whether controlling access for a few or ensuring open access for all, and the lifecycle of subsequent enhancement, reuse and distribution (where the agents may range from individual scholars to international collaborative teams);
  - streamlining the process of creating online databases, or migrating from personal desktop databases, and related query interfaces, whether for graduate research or grant-funded projects; and the subsequent sustainability of both the underlying data and the 'interpretative' layers provided by interfaces, and interoperability with other databases;
  - and of continued importance within the humanities, the development and management of hybrid data collections, comprising digital objects of various ages, paper-based files, bibliographic data – all of which may combine both ordered and unordered sub-collections, the definition of which tends to change as the intellectual nature of the research evolves over time, and be dependent on the type of research activity.
11. The development of training and support activities are intended to address the requirements of researchers at various stages, for example: final year undergraduate and first year graduate students; early career researchers; new project teams; and established scholars recently arrived in Oxford. The initial scoping study that informed the EIDCSR Project identified that researchers, whether graduate students or established scholars, were most likely to engage with, and benefit from, data management support activities if those activities formed part of the existing research support services at the divisional level. Therefore, the Sudamih Project will pilot training and support activities for the Humanities Research Support Team within the Humanities Division with the intention of evaluating the feasibility of some element data management support being provided by the team. The training programme will be devised on a modular basis, in collaboration with the DCC, with a view of applying modules as part of graduate research methods training, new staff induction and mentoring, and as part of awareness-raising activities for established staff.

<sup>3</sup> The PARADIGM project, <<http://www.paradigm.ac.uk/>>.

<sup>4</sup> The FutureArch project, <<http://futurearchives.blogspot.com/>>.

<sup>5</sup> Building a virtual research environment for the Humanities, <<http://bvreh.humanities.ox.ac.uk/>>.

<sup>6</sup> Digital humanities at Oxford, <<http://www.oerc.ox.ac.uk/humanities>>.

12. **Database as a Service (DaaS):** The implementation and piloting of an online database provisioning service (Database as a Service) was identified as one of the highest priorities for data management infrastructure by the Humanities Division. There are currently technical barriers to the creation of databases (broadly defined). Increasingly both individual researchers and project teams want to easily create online databases to assist with the analysis and dissemination of (often) disparate datasets. In this context a database may be viewed as the combination of datatypes, relationships, data itself and the presentation of data through various retrieval interfaces. All of which in their development reflect an intellectual as well as mechanical investment. Raw data is often less useful as a foundation for new research than the combination of data and intellectual interpretation presented through a coherent set of interfaces. Thus, it is no longer unusual for a PhD student in the humanities to develop a database to support their research and then subsequently wish to share both the data and the interfaces that present routes of interpretation into that data. The DaaS will include support for textual, image and geo-data, reflecting the multimedia nature of much of humanities research. The DaaS itself will be built on and extend database provisioning infrastructure developed by OUCS originally for internal purposes (including a review of the Archer eResearch Toolset).
13. Support for *textual data* management will take advantage of the world-class expertise Oxford has in the creation and curation of digital texts (via, for example, the Text Encoding Initiative, and activities associated with the Oxford Text Archive). *Image data* is also a popular data type across humanities disciplines, common to both individual and project data collections. The project seeks to provide guidance on lightweight tools and methods for the management of images on personal computers (including mobile devices) as well as best practice for the integration of images, data and metadata with the DaaS.
14. In common with other disciplinary areas, the *visualisation of data* is increasingly a requirement within the humanities. The construction of maps, for example, is a common task for various purposes (e.g. mapping changes over time and place). Literary studies, archaeology, and area studies all include research activities that involve the identification, analysis and visualisation of location data. As with image data the project will investigate and advise on tools, lightweight server applications and other means of end-to-end interoperability with the DaaS, from personal computing devices to e.g. Google Maps mashups.

#### 1.4 Project Description

15. It is proposed that the project will commence 1 October 2009 and end 31 March 2011.
16. In general the scope of the project includes the following elements:
- will be driven by, and address the requirements, of active and ongoing humanities research.
  - will have a principal focus on the “life's work” (*Lebenswerk*) theme and the embedding of data management infrastructure, tools and processes within humanities research not only at different points in the research lifecycle but also at different points along the academic career continuum.
  - in supporting the various stages of the academic career, the project objectives will include the development and piloting of training and support materials in collaboration with the DCC, for e.g. graduate research methods training; early career staff development; and established scholars.
  - whilst the resource constraints of the project will not enable a deep engagement with more than two or three research activities (reflected in the curricula vitae of the co-investigators), those activities have been selected to include within them a mixture of career stages, research methods, data types, and collaborations; as well as ensuring that the project outcomes are applicable to a broad range of activities both within and beyond the University of Oxford (in this sense complementing the EIDCSR project both in terms of the subject coverage but also the deliverables).
17. The project's objectives include:
- developing institutional services for data management, curation and long-term preservation for selected humanities research activities, but with a view to ensuring the deliverables can then be expanded to other activities within the humanities and beyond;
  - addressing database support needs within the humanities, including support for specific

data types but also sustainability and service costing models;

- understanding the training and support requirements for humanities researchers and developing a data management skills course and other support activities, in collaboration with the DCC but tailored to the requirements of the humanities.
- investigate the roles and responsibilities of service providers in Oxford for supporting humanities researchers with the management and curation of research data; develop a deeper understanding of research workflows and how they may interface with institutional services.

### **1.5 Alignment with the Activities and Objectives of Oxford University**

18. The University of Oxford ICT Strategic Plan<sup>7</sup> includes institutional repository services to manage and curate research data as one of the key priorities. Additionally the ICT Subcommittee of the Planning and Resource Allocations Committee, has further refined the University's commitment to the curation of research data:

*"The University of Oxford is committed to supporting researchers in appropriate curation and preservation of their research data, and where applicable in accordance with the research funders' requirements. It recognises that this must be achieved through the deployment of a federated institutional data repository. This repository has to be supported by a suitable business model, and where possible funded through full economic cost recovery, so that the University can guarantee that the data deposited there will be managed over the long term. The data repository will be a cross-agency activity developed and supported by a number of departments within the University and will build, as far as possible, on existing services, including the Oxford University Research Archive (ORA). It will be overseen by a Steering Group which reports to the University Research Committee. The management and curation of research data will be addressed in cooperation with specialist agencies, research funders and other institutions in the UK and internationally. Oxford is committed to playing a significant role within the foreseen UK Research Data Service Pathfinder activities."*

19. Oxford's highly devolved organizational arrangement is mirrored by a devolved, coordinated ICT structure, intended to combine responsiveness to user demands whilst remaining faithful to the principle of subsidiarity. This project aims to work within this structure to combine central and local aspects (whether technical or training). By way of an example, the technical component of this project will take a service-oriented architecture approach to implementation, taking advantage of open APIs where possible.

### **1.6 Value to the JISC Community**

20. This project supports the JISC strategy by developing and integrating ICT infrastructure and services to support research, particularly within the humanities. The outputs from the project will contribute to institutional support, particularly within research-intensive universities, for the management of research data. The project will bring together expertise from different service units within Oxford, and collaborate with the DCC, in particular, to address the needs of research activities in Humanities. The result should highlight areas of successful practice in the provision of data management infrastructure, of relevance to other research institutions wishing to implement infrastructure and training models for the provision of research data management services.

21. The project focuses on a particular research domain but the institutional nature of the project will ensure that the outputs from the project will be transferable to other research domains. The project is part of a wider programme of activities to provide institution-wide data management infrastructure services, significant components of which are the University's continued participation in the development of the UK Research Data Service, and working in collaboration with external agencies such as the DCC and the RIN.

## **2. Quality of Proposal and Robustness of Workplan**

### **2.1 Project Plan**

22. The overall approach follows the three stages process methodology proposed on the Call for Proposal document:

<sup>7</sup> The University of Oxford ICT Strategic Plan, <<http://www.ict.ox.ac.uk/strategy/plan/>>.

23. **Stage 1. Analysis:** An analysis of the research data management requirements using the latest revision of the Data Audit Framework (DAF) and the DCC Data Management Plan Checklist, at the divisional level in order to define the component services that would be required to support Humanities researchers in Oxford across their research lifecycle, with a particular focus on the creation of accessible, data-rich resources, their ongoing maintenance and the provision of data management training.
24. **Stage 2. Pre-implementation:** The identification of “touch-points” of the requirements gathered with existing services in the University and in particular with those provided at the Faculty-level and those provided centrally. It is expected to develop a 'service catalogue' for activities to readily identify services for research data management at various stages in the research process. This stage will also include the analysis of training requirements for Humanities researchers, identifying the different needs of those early in their careers and those who are well established.
25. **Stage 3. Implementation:** The pilot implementation of an online, low cost database provisioning service to streamline the configuration and building of online database applications that support multimedia (including text and images) datasets often generated within humanities research as well as the capacity to map geo-referenced data. The service will be trialled with selected activities within the Division at different stages in their lifecycle (e.g. a start-up project, an established project; supporting a PhD student, and an established scholar) and with different types of data. The implementation phase also includes the piloting of training modules and other support activities – designed to support the Humanities Research Support Team and devised in conjunction with the DCC (e.g. Using Digital Curation 101 as a starting point).

## 2.2 Project Deliverables

26. The proposed list of project deliverables includes:

- A report detailing the requirements from Humanities researchers in Oxford that also documents the management practices and training needs, based on the DAF methodology. Desktop research will also enable requirements within Oxford to be contextualised within the requirements of the broader UK humanities communities.
- A pilot 'database as a service' (DaaS) system evaluated by early adopters within the humanities that enables the efficient creation of online databases and semi-automation of the development of retrieval interfaces; the database as service will be attuned to good practice in data curation, including the early capture of metadata; and include particular support for image and geo-data as well as text and other popular data types.
- A series of training modules to improve researchers' skills in data management, at various points in the academic career and data lifecycle, covering best practice and also services available to them in the University and elsewhere.
- Project website with information about the project, including a blog to describe the day to day experiences, and RSS-supported bookmarks of relevant activities and publications;
- Publications and presentations in relevant journals and conferences.
- Two project workshops to bring together humanities researchers with experts in data management, whether from within the humanities or as service providers (local and national).
- A final report, describing the process of implementing data management infrastructure for the humanities and making recommendations on how JISC might consider continuing work in this area.

## 2.3 Summary of Workpackages

	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	
	<i>Analysis</i>					<i>Pre-Implementation</i>					<i>Implementation</i>								
<b>WP1</b>																			
<b>WP2</b>																			
<b>WP3</b>																			
<b>WP4</b>																			
<b>WP5</b>																			

	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	
	Analysis					Pre-Implementation					Implementation								
WP6																			
WP7																			
WP8																			
WP9																			

27. **WP1. Project management:** To ensure timely and efficient delivery of project deliverables, produce project plans and progress reports, manage the finances and liaise between project collaborators. Operates throughout project. Tasks will include project planning; JISC reporting; liaising between collaborators.
28. **WP2. Audit and requirements analysis:** To audit data assets and data management practices in the Humanities Division using the DAF methodology and undertake an analysis of researchers' requirements for data management infrastructure services including training; capture attitudes to data sharing and perception of data value. Methods will include interviews and the DCC Data Management Plan Content Checklist. Output will include a report describing data management practices within humanities subjects and requirements, with an emphasis on training and support. Three months, on iterative basis.
29. **WP3. DaaS Infrastructure:** To translate user requirements to a functional specification; undertake evaluation of technologies to enable a Database as a Service infrastructure; recommend best approach; implement within context of data management lifecycle; and pilot with selected early adopter Humanities research activities.. Technologies likely to include PostgreSQL and utilities originally developed for in-house use, coupled together with 'widget' devices and served via a common platform. The project will also review relevant components of the Archer Toolset (e.g. XDMS, Hermes). The pilot includes an evaluation to consider the feasibility, scope and resources required for "database as a service" to be integrated with institutional data management infrastructure. Core development over seven months, augmented by WP4 and WP5.
30. **WP4. Image data management:** Investigate and recommend methods and tools for end-to-end management of image metadata and image object relationships within the context of institutional infrastructure; including image capture (where appropriate), personal filestore, interoperability with the DaaS, and other dissemination routes. Operates in tandem with WP3, over 12 months.
31. **WP5. Visualisation of data:** In particular, investigate and recommend processes and tools for the capture, processing and visualisation of geo-reference data; including interoperability with the DaaS. This WP will build on other relevant activities.<sup>8</sup> Operates in tandem with WP3, over 12 months.
32. **WP6. Data management training:** To analyse data management training and other support needs from WP2; undertake desk research to evaluate existing modules (e.g. DCC 101); develop and pilot training modules and support services, in collaboration with the Humanities Research Support Team and the DCC, as well as institutional service providers. Operates in two three-monthly iterations, coinciding with two academic terms.
33. **WP7. Cost models for data management services:** To develop appropriate cost models for the ongoing support and development of infrastructure to support humanities research (recognising that a significant proportion of humanities research is not grant-funded). This workpackage will be augmented by work undertaken within the EIDCSR Project as well as the current review of the resource allocation model for IT services within the University.
34. **WP8. Dissemination:** To establish dissemination mechanisms and disseminate project outputs and information on relevant websites, at conferences and other events, and in journals and other publications. The project will run two themed workshops in conjunction with the Steering Group and other external parties. Operates throughout the project and expects to make use of existing channels as much as possible (e.g. via the DCC Data Management Forum, RUGIT, arts-humanities.net).

<sup>8</sup>E.g. Hestia project, <<http://hestia.classics.ox.ac.uk/project/index.html>>; Erewon Project, <<http://erewon.oucs.ox.ac.uk/>>.

35. **WP9. Evaluation:** To commission a formative evaluation of the project approximately half-way through the course of the project and implement recommendations. The WP will run between the final stages of the pre-implementation phase and the initial implementation phase.

## 2.4 Standards

36. The metadata required to describe, administer and preserve the research data will comply with research community standards as well as those available at the JISC Standards Catalogue. Technical components will be developed in line with a service-oriented architecture approach and use, where possible, open standards, such as Web Services. The project website will comply with relevant accessibility, encoding, and dissemination standards.

## 2.5 Project Governance and Management

37. The project will be led by OUCS and report to the Oxford Digital Repositories Steering Group. This group was established to ensure coordination, communication and collaboration between repository activities across the University and which has a reporting line to the University's Research Committee. In addition to this, the project will be supported by a Steering Group comprising key stakeholder representatives, both from within Oxford and from the national community. The Steering Group will be chaired by a senior humanities academic from a research-led university (not Oxford).

38. A smaller working group comprising members of the project team and other Humanities academics, as well as OUCS, RSO, and OULS, will be established in order to help facilitate an agile and iterative approach to the development of the project outputs.

## 2.6 Risk Assessment

39. 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
<b>Staffing</b>				
Failure to recruit appropriate staff for the project.	3	4	12	Some sharing of staff between EIDCSR and Sudamih should be possible. Else, redeployment of existing staff will be investigated as well as timely recruitment of new staff.
Loss of key staff before end of the project.	2	3	6	Embed project in institutional practices, ensuring a number of individuals have expertise and willingness to assume responsibilities.
<b>Organisational</b>				
Expectations mismatch between project and research activities.	2	3	6	Ensure continued engagement via project's working group and Steering Group to maintain initial buy-in and momentum.
Lack of coordination between project stakeholders.	1	3	3	Ensure clear reporting and communication lines; take advantage of existing institutional communication structures.
<b>Technical</b>				
Development of online database provisioning system proves to be too complex.	2	3	6	WP builds on existing and planned internal infrastructure; staff have requisite technical skills.

## **2.7 Intellectual Property Rights**

40. Any IPR resulting from this project will remain the property of the organisation generating it. Under the University of Oxford's policy on intellectual property (which covers all University employees and students), the University claims ownership of a range of intellectual property rights with commercial potential. The University does not assert any claim to the ownership of copyright in artistic works, books, articles or lectures, apart from those specifically commissioned by the University. 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. The University seeks to maximise the commercial potential of its intellectual property through its wholly-owned technology transfer company, ISIS Innovation Ltd. In accordance with the desires of the JISC Research Data Programme, however, it is proposed to release project deliverables under either a Creative Commons license or, in the case of software, under an open source software license to maximize the benefit for the wider community.

## **2.8 Exit and Sustainability Plans**

41. The University of Oxford recognizes the need to have a technical infrastructure supported by policies and procedures in place to deal with the management and curation of research data. There is top-level commitment (see 1.5 above) to continue to support and fund activities in this area to ensure sustainable services to support research.
42. The UKRDS is working towards the submission of a proposal to the HEFCE Strategic Development Fund. The University of Oxford is one of the original case study sites and has been allocated funding to provide a liaison role between Oxford and the UKRDS, especially in connection with the EIDCSR Project. The Sudamih Project has the support of the UKRDS and its activities will be fully compatible with proposed UKRDS Pathfinder activities.

## **3. Engagement with the Community**

### **3.1 Stakeholder and Practitioner Engagement**

43. Project stakeholders and practitioners will be engaged throughout the life of the project via the Steering Group and project working group. The project website and the planned workshops will also help ensure ongoing communication between the project and stakeholders.

### **3.2 Dissemination Plans**

44. Internal and external dissemination will run throughout the project. Dissemination will comprise multiple channels of communication:
- A project website, blog and bookmarking site will be set up and maintained with regular updates; the project will also make use of other JISC, RUGIT and digital humanities dissemination channels;
  - An events diary will be developed, with participation in conferences, JISC programme meetings, and other events;
  - At least two articles about the project's findings will be submitted to relevant publications;
  - Two themed workshops will be organized to stimulate discussion amongst stakeholders. Internal dissemination will be facilitated through additional informal events as well as through the proposed stakeholder engagement structure;
  - Project staff will continue to engage with the DCC Research Data Management Forum as part of the communication channels with the JISC community. The project will contribute to other JISC-related activities as appropriate.
45. A more complete dissemination plan will be created with the overall project plan in accordance with JISC requirements to ensure that information and outputs arising from the project are effectively communicated.

## **4. Impact**

### **4.1 Review of Current Starting Point**

46. Research data management is a strategic priority for the University. As noted above the Sudamih Project is a continuation of internally-resourced activities and intended to

complement the existing EIDCSR Project, both in proposed infrastructure and subject domains. Oxford is a key stakeholder in the UKRDS initiative. The Humanities Division is currently under-served with respect to centrally-provided services for the curation and sustainability of research data.

## 4.2 Stakeholder Analysis

Stakeholder	Interest / stake	Importance
University of Oxford – Humanities Division and service units with responsibility for supporting research	The University of Oxford is committed to the project; the participating Humanities research activities are representative with well defined data management infrastructure needs; the service units have defined their interlocking roles and responsibilities to support researchers to manage their data more effectively.	High
JISC	The curation of research data is a strategic priority for both the (now) JISC Infrastructure and Resources and the Support of Research committees.	High
Wider Humanities research community	Interest particularly in the management of research data over the course of an academic career; as well as support for specific data types.	High
UK Research Data Service	The University of Oxford has a formal liaison role with UKRDS to plan Pathfinder services and submit a business plan and proposal to HEFCE.	High
The Digital Curation Centre	The DCC is engaging with staff in HEIs with an interest in data management through the Research Data Management Forum and is involved in providing data curation training.	High
The Research Information Network	Research data is one of the main themes for RIN activities and they have published several influential reports on data management and data sharing.	High
Other relevant projects and services	Other projects investigating issues related to the management and curation of research data may provide useful tools and information, and deploy some of the outputs of this project.	Medium

## 4.3 Impact on Wider Community

47. The outcomes envisaged at the end of the project include:

- Evidence of data management requirements and cultural considerations specific to the Humanities.
- A better understanding of possible institutional models for dealing with the management and sustainability of long-life span data generated during the course of Humanities research activities.
- An improved awareness about the benefits of actively supporting researchers to manage their data in HE institutions.
- The development of services that have the potential to be expanded to other research domains (within and beyond the University of Oxford), providing sustainable research data management capabilities.

48. The outputs and outcomes of the project will be of wider benefit to other HE institutions. The collaboration in Oxford between service units to provide the data management infrastructure and training provision will provide a useful, reusable framework.

## 4.4 Evaluation

49. The project intends to develop a detailed evaluation plan as part of the overall project plan. The evaluation plan will take into account the baseline review at the start of the project. The project has set aside funds for a formative evaluation around half-way through the project. The evaluation will help ensure that the deliverables serve the purpose for which they are intended; that the project has developed an appropriate framework that integrates the requirements of the research groups; the implementation of an online database management system; and the development of a data management training for humanities researchers.

## 5. Project Budget

50. It is proposed, where possible, to join-up posts such as the project manager, analyst and systems programme with similar posts within the EIDCSR Project, in order to form a coherent and balanced team. Similarly, the resource allocated to OULS (data analyst) and the Humanities Division Research Support Team (data support) will be fully embedded with those units.
51. Dissemination via two workshops (at an average of £1,100 each) and community engagement.
52. Hardware budget includes a contribution to the server and storage infrastructure envisaged for the DaaS during the course of the project. The development of any pilot into a production service will be subsumed into the Department's standard planning and resource allocation cycle.

[Budget removed for public dissemination version]

## 6. Roles and Previous Experience of the Project Team

53. **Professor Paul W. Jeffreys** – Principal Investigator; Director of IT for the University of Oxford. He has served as Director of the Oxford e-Science Centre and established the e-Research activities within the University. His research interests were in the field of particle physics, but they are now focused in e-Research and the strategic direction of information technology.
54. **Dr Ian Archer** – Co-Investigator; Fellow and Tutor in Modern History. His primary research interests lie in the history of early modern London and he teaches a broad range of early modern British and European courses. Since 1999 he has been the General Editor of the Royal Historical Society Bibliography, a major project funded by the AHRC to make bibliographic data on British History electronically available. He is also the Editor of the Fell-funded Oxford Holinshed Project which aims to produce an electronic edition making possible comparison of the two versions (1577 and 1587) of the massive Chronicles, the crowning glory of Elizabethan historiography.
55. **Dr Michael Fraser** – Co-Investigator; Head of the Infrastructure Systems and Services Group at Oxford University Computing Services and Director of the Embedding Institutional Digital Curation Services in Research (EIDCSR) Project. Previous project areas have included e-infrastructure, digital repositories, humanities computing, and virtual research environments. He has an academic background in the humanities.
56. **Professor Andrew Wilson** – Co-Investigator; Professor of the Archaeology of the Roman Empire. Current research interests include the economy of the Roman empire; ancient technology; ancient water supply and usage; Roman architecture; Roman North Africa, and field survey. He is currently co-directing a major AHRC-funded 5-year research programme, The Oxford Roman Economy Project, which aims to collate and analyse a series of quantifiable indicators, from papyrological, epigraphic and archaeological sources, on the growth and contraction of the Roman economy between 100 B.C. and A.D. 300.
57. **Dr Andrew Fairweather-Tall** – Research support consultant; Humanities Division Research Co-ordinator. He supervises the setting up and coordination of information systems for disseminating research funding, manages information flow to/from faculties, and liaises with external funders. He co-ordinates training and development for grant holders and is secretary to the Division's research and resources committee.
58. **Ben O'Steen** – Metadata consultant; Software Engineer for the Oxford University Research Archive. He has expertise in semantic web technologies, large-scale storage and preservation, and web-service creation. He has designed and is implementing a service-orientated, distributed, loosely-coupled and highly robust architecture for the archive, based around semantic web, library and internet standards, virtualized hardware and open source software.