

**Repositories and Preservation Proposal Cover Sheet**

<b>Cover Sheet for Proposals</b> (All sections must be completed)		<b>JISC Capital Programme</b>
<b>Name of Capital Programme:</b> Repositories and Preservation Programme		
<b>Bid for Call Area :</b> (Please tick ONE BOX ONLY, as appropriate)		
<b>Tools and Innovation (Strand B)</b>		
<input type="checkbox"/>	<b>Call Area I – Tools and Innovation Projects</b>	Please specify area of proposed project eg ‘ <i>metadata generation and validation</i> ’
<b>Discovery to Delivery (Strand C)</b>		
	<b>Call Area II – Discovery to Delivery Projects</b>	<input type="checkbox"/> a) Version identification framework <input type="checkbox"/> b) Persistent identifier interoperability demonstrator <input type="checkbox"/> c) Federated access management and repositories <input type="checkbox"/> d) Semantic interoperability demonstrator
<b>Repository Start-Up and Enhancement (Strand D)</b>		
	<b>Call Area III – Repository Start-Up and Enhancement Projects</b>	<input type="checkbox"/> a) Repository start-up projects <input checked="" type="checkbox"/> b) Repository enhancement projects
<b>Digital Preservation and Records Management (Strand H)</b>		
	<b>Call Area IV – Digital Preservation and Records Management Projects</b>	<input type="checkbox"/> a) Digital preservation across the lifecycle <input type="checkbox"/> b) Models and implementation of preservation services <input type="checkbox"/> c) Preservation tools development
<b>Shared Infrastructure Services (Strand I)</b>		
	<b>Call Area V – Shared Infrastructure Services Projects</b>	<input type="checkbox"/> a) Pilot implementation of licence registry <input type="checkbox"/> b) Pilot national name and factual authority service <input type="checkbox"/> c) Scoping an architecture to support digital policy management <input type="checkbox"/> d) Scoping a terminology registry
<b>Name of Lead Institution:</b>	EDINA National Data Centre, University of Edinburgh	
<b>Name of Proposed Project:</b>	DISC-UK DataShare	
<b>Name(s) of Project Partner(s):</b>	EDINA, London School of Economics, University of Oxford (Nuffield College), University of Southampton	

**Full Contact Details for Primary Contact:**

**Name:** Robin Rice  
**Position:** Data Librarian  
**Email:** R.Rice@ed.ac.uk  
**Address:** EDINA  
 Main Library Bldg.  
 George Square  
 Edinburgh  
 EH8 9LJ  
**Tel:** 0131 651 1431  
**Fax:** 0131 650 3308

**Length of**            **2 years**

**Project:**

**Project Start**     1 March 2007

**Project End Date:** 27 February 2009

**Date:**

**Total Funding Requested from JISC:**     £ 267,782

**Funding Broken Down over Financial Years (April – March):**

Apr06 – Mar07	Apr07 – Mar08	Apr08 – Mar09
12,227	127,628	127,927

**Total Institutional Contributions:**            269,730

**Percentage Contributions over the  
Life of the Project:**

**JISC**  
50%

**PARTNERS**  
50%

**Outline Project Description**

This bid arises from an existing UK consortium of data support professionals working in departments and academic libraries in universities, and builds on an international network with a tradition of *data sharing and data archiving* dating back to the 1960s in the social sciences that has emerged more recently in other domains (Burnhill et al, 2005). By working together across four universities and internally with colleagues already engaged in managing open access repositories for e-Prints, this partnership will introduce and test anew model of data sharing and archiving to UK research institutions. By supporting academics within our institutions who wish to share datasets on which written research outputs are based, this network of institution-based data repositories develops a niche model for deposit of 'orphaned datasets' currently filled neither by centralised subject-domain data archives/centres/grids nor by e-Print-based institutional repositories (IRs).

**I have looked at the example FOI form at Appendix A and included an FOI form in the attached bid (Tick Box)**

YES

NO

**I have read the Circular and associated Terms and Conditions of Grant at Appendix B (Tick Box)**

YES

NO

## Appendix A

### FOI Withheld Information Form

We would like JISC to consider withholding the following sections or paragraphs from disclosure should the contents of this proposal be requested under the Freedom of Information Act.

We acknowledge that the FOI Withheld Information Form is of indicative value only and that JISC may nevertheless be obliged to disclose this information in accordance with the requirements of the Act. We acknowledge that the final decision on disclosure rests with JISC.

Section / Paragraph No.	Relevant exemption from disclosure under FOI	Justification

Please see <http://www.ico.gov.uk> for further information on the Freedom of Information Act and the exemptions to disclosure it contains.

Example:

Section / Paragraph No.	Relevant exemption from disclosure under FOI	Justification
2.4	s.43 Commercial Interests	Contains detailed description of our proposed system design which would damage our commercial interests if disclosed by making this information available to competitors

## I. INTRODUCTION

*"Data is the next Intel Inside" – Tim O'Reilly*

### A. Rationale

1) This bid arises from an existing UK consortium of data support professionals working in departments and academic libraries in universities, and builds on an international network with a tradition of *data sharing and data archiving* dating back to the 1960s in the social sciences that has emerged more recently in other domains (Burnhill et al, 2005). By working together across four universities and internally with colleagues already engaged in managing open access repositories for e-Prints, this partnership will introduce and test a new model of data sharing and archiving to UK research institutions. By supporting academics within our institutions who wish to share datasets on which written research outputs are based, this network of institution-based data repositories develops a niche model for deposit of 'orphaned datasets' currently filled neither by centralised subject-domain data archives/centres/grids nor by e-Print-based institutional repositories (IRs).

2) For example, the social science survey of the StORe project found that nearly half of respondents had created some form of quantitative dataset (such as surveys), and over a third had produced a qualitative dataset (such as interviews). Yet 71% indicated they had never deposited data in a source repository such as the UK Data Archive (Burton, 2006 pp 13, 17). Similarly, the OpenDOAR registry of repositories indicates which repositories in the UK have a policy to accept datasets, but experience at the partners' institutions shows that in actuality this is a long way from being the case. But the common policy to support only PDF filetypes undermines the ability to deposit datasets in forms that allow re-use in the sense of being able to manipulate the data for analysis in any way. We believe it is safe to say that active advocacy is not geared toward gaining deposit of research datasets in any existing UK university IR.<sup>1</sup>

3) There are of course some notable centralised data archives and centres serving particular disciplines in the UK, such as the UK Data Archive/Economic and Social Data Service (UKDA/ESDS) for the social sciences and the NERC Data Centres for natural and environmental sciences. Other disciplines have created vast online databases on the Internet or over e-Research grid networks, which is the logical place for 'publishing' data outputs in those domains. (Illustrative case studies can be found in the *Joint Data Standards Study*, 2005.) This project does not aim to challenge these existing institutions that have set internationally recognised high standards in data archiving, management and curation, nor the model of centralised data repositories by subject area. It does however, wish to explore the role of filling in the gaps left open by the paucity of coverage of dedicated data archives, and in doing so, gain leverage from being able to work closely and directly with potential depositors at one's own institution. Indeed, the lifecycle approach to data sharing encourages intervention at the earliest stages of a research project to ensure adequate consent, documentation etc., are achieved for the data to be usable by others (Humphrey et al, 2000).

### B. Partnership

4) DISC-UK, Data Information Specialists Committee-United Kingdom, is a support network of data librarians, data managers and social science librarians based at UK research universities formed in early 2004.<sup>2</sup> Founded by members of IASSIST,<sup>3</sup> an international organisation of social science-based data professionals, DISC-UK members meet 2-3 times per year to discuss common issues and concerns. This existing partnership – involving staff at the Universities of Oxford, Edinburgh, Southampton and the London School of Economics—will form the foundation for the project's activity. Each of these universities has a strong commitment to their IRs and are active in enhancing them. DISC-UK members will work closely with staff involved with IR management and development at their own institutions to pilot successful models for incorporating deposit of research data into institutional repositories. In this way, the partner institutions will be early adopters and provide exemplars, particularly but not exclusively for numeric datasets in population-based research areas—this is the domain expertise that DISC-UK partners have in common.

### C. Contribution to JISC Repository Programme

5) Outputs from the project will help to demystify data as complex objects in repositories, and assist other IRs in overcoming barriers to incorporating research data. By building on lessons learnt from current JISC-

<sup>1</sup> Six out of 87 UK open access repositories listed in OpenDOAR record 'datasets' as a content type; aside from eCrystals at Southampton, DSpace@Cambridge and NDAD (National Digital Archive of Datasets) these returned only PDF articles in a search on 'datasets', and had no option to browse by dataset content type.

<sup>2</sup> See <http://www.disc-uk.org/>

<sup>3</sup> International Association of Social Science Information Service & Technology, [www.iassistdata.org/](http://www.iassistdata.org/)

funded data repository projects such as CLADDIER and GRADE<sup>4</sup> the project will help realise the vision of the Digital Repositories Review of a “coherent aggregation of content from a network of institutional repositories” (Heery and Anderson, 2005), and more particularly of the Digital Repositories Roadmap, see for example the milestone under Data, “Institutions need to invest in research data repositories” (Heery and Powell, 2006).

#### *D. From Open Access to Open Data*

6) Open access repositories are still a relatively new feature in scholarly communication. The Open Archives Initiative was only launched in 1999, for example (Suber, 2006), and many of the thorny problems associated with the management of institutional repositories have yet to be solved (including both technical and cultural). Nevertheless libraries and funding agencies in North America have already begun to address the arguably new role of research libraries in the curation of data. (For example, at the recent NSF/ARL “Workshop on New Collaborative Relationships: The Role of Academic Libraries in the Digital Data Universe” in September 2006<sup>5</sup> and the Science Commons Workshop in October 2006.<sup>6</sup>) Similar issues will be explored in an upcoming RIN workshop in London on 5 December.<sup>7</sup> The project seeks to gain knowledge of prior experience from North America and elsewhere.

7) “Open Data” as a concept is quickly gathering momentum as the latest term in the “open” trilogy along with Open Source (for software) and Open Access (for research outputs, primarily e-Prints).<sup>8</sup> It indicates a recognition that there is a rising level of expectation among users for complete access to an intellectual work, not only the final published post-print, but the body of evidence drawn on to create that final output. This is compatible with the scientific method of allowing replication of results by others, and the rich tradition of secondary analysis in the social sciences and other population-based research domains. It is also in line with several recent initiatives to open up publicly-funded research data to public availability.<sup>9</sup>

8) However, there are reasons data are not always provided in a completely open way. Aside from commercial value, these may include confidentiality concerns, copyright questions, and sheer complexity. Many scholars are more comfortable with an informal method of sharing, so they can assess the use to which the data will be put and decide whether to give the requestor access on a case by case basis.<sup>10</sup> This is one of the reasons datasets have not been incorporated into IRs to date; the project will need to address this.

## **II. PROJECT DESCRIPTION**

1) The project, known as the DISC-UK DataShare project, will commence Thursday, 1 March 2007 and end Friday, 27 February 2009, though as the timetable shows there will be related activity before and afterwards.

2) DISC-UK is an existing support network for data professionals working in universities (Data Information Specialists Committee-UK). DataShare is not an acronym, but is descriptive of the essence of the purpose of the project. We wished for the name to be meaningful to academics as well as librarians. It also has a synergy with an earlier successful JISC project in the 5/99 Learning & Teaching Programme, “Using Numeric Data in Learning and Teaching”, known as Datateach, which studied the barriers to the use of secondary data sources from the UKDA, MIMAS and EDINA, and covered issues around institutional support for learners and teachers in quantitative methods within the social sciences and related fields.<sup>11</sup>

### *A. Project Aims*

The project's overall aim is to contribute to new models, workflows and tools for academic data sharing within a complex and dynamic information environment which includes increased emphasis on stewardship

<sup>4</sup> The data cluster of the current Repository Programme projects can be found on the Repository Wiki, [http://www.ukoln.ac.uk/repositories/digirep/index/Category:Data\\_repositories\\_cluster](http://www.ukoln.ac.uk/repositories/digirep/index/Category:Data_repositories_cluster). Also see appendix for brief descriptions of projects related to this bid.

<sup>5</sup> <http://www.arl.org/info/events/nsfworkshop.html>

<sup>6</sup> “Creating a vision for making scientific data accessible across disciplines” <http://www.spatial.maine.edu/icfs/>

<sup>7</sup> “Getting the most out of data, Making the most of research” <http://iceconferences.com/digitaldata/>

<sup>8</sup> For example, a wikipedia entry for the term was added only as recently as October 2006.

<sup>9</sup> For just one example see the 2004 “OECD Ministerial Declaration on access to research data from public funding”. The new Medical Research Council (MRC) data sharing requirements are based on the OECD principles.

<sup>10</sup> Both the StORe and GRADE projects have discovered such evidence and have begun grappling with solutions.

<sup>11</sup> The project director and manager were the same people as in this project, and data librarians from two of the current partner institutions were involved (LSE and Edinburgh).

of institutional knowledge assets of all types; new technologies for doing e-Research; new research council policies and mandates; and the growth of the Open Access / Open Data movement.

3) Objectives:

1. Build capacity of institutional repositories in UKHE to respond to the unique requirements of research datasets as a new 'document type'.
2. Use a range of open source software repository solutions – Eprints, DSpace, Fedora – to provide exemplars and add-on tools for managing datasets as institutional repository items.
3. Produce and disseminate findings – in cooperation with the Repositories Support Project (RSP) and the Repositories Research Team (RRT) - to inform library and repository managers about the organisational and technical issues associated with the deposit of research data.
4. Work with the RSP, Digital Curation Centre (DCC) and others to identify training needs and solutions for increasing skills of information professionals in UKHE for managing research data.

B. Workpackages and Deliverables

Workpackage	Deliverable
WP1: Project Management	1) Ensure efficient and timely delivery of project; set up distance collaboration tools. Establish website; contact and reporting with JISC; administration. 2) Successful liaison across partner institutions, JISC Repository Programme, RSP, RRT, DCC, in addition to data centres and research councils. 3) Ensure project staff are given training and other professional development opportunities, including liaison with other projects and dissemination activities. 4) Arrange for both formative and summative evaluation. This will include commissioning an external evaluator familiar with UKHE (See Evaluation, p 6).
WP2: Institutional Repository Development	5) Each institution will develop its own workplan as part of the overall project plan. The workplan will cover changes needed to re-align each institution's repository and data services to achieve an embedded framework for deposit, management and preservation of research data, as determined by each institution's strategy and desired scope of repository activity. 6) Each institution will work with the project consultant to improve its processes and workflows, policies, mandates, and technologies towards best practice. 7) Each institution will use either the DCC DRAMBORA tool or the CRL TRAC Criteria and Checklist to mark progress and gaps in its repository development and preservation planning, particularly with regard to research data. 8) Each institution will identify academic champions to assist with take-up of the deposit service locally and determine how to make best use of such allies (e.g. form local project steering committee, hold departmental outreach events, publicise deposits of 'whole works' by such champions to showcase the associated benefits of research exposure).
WP3: Technology	9) Implement existing open source tools for data management within native applications (e.g. EPrints, DSpace, and Fedora implementations at the sites). 10) Customise existing tools to make them useful to data management; provide documentation for changes made; maintain close contact with RRT throughout. 11) Engage with technical consultants as needed. <sup>12</sup> 12) Expose data for different presentational purposes including mash-ups. 13) Scope/specify/test other tools of use to data management if appropriate. 14) Investigate linking between datasets and research papers within repositories or across repositories (incorporate learning from StORe, CLADDIER, eBank) 15) Implement a technical watch by and for project participants on development issues affecting the project but not within scope (including images, version control, other complex objects e.g. learning objects, scientific application profile development, shared services, e-Research and Virtual Research Environments).
WP4: Outreach and Dissemination	16) Develop a dissemination plan with emphasis on deliverables in the second year and based on a complete stakeholder analysis. 17) Cooperate with the RSP and DCC to customise experience-based outputs for dissemination to UKHE community, including tool-kits for data management.

<sup>12</sup> As agreed with JISC programme manager ; up to half of funds identified for external evaluation to be diverted towards project-wide technical consultancy.

	<p>18) Papers will be submitted to relevant conferences and journals by DISC-UK partners. Presentations will be given by request to other meetings.</p> <p>19) Experience-based outputs including use cases, documentation, best practice guidelines, and fact sheets will be written by the project team on a wide range of topics, in consultation with RSP and others as to desired topics.</p> <p>20) A training programme will be designed (but not delivered within the project) to teach skills to information professionals who wish to support staff and students in their use of research data in their institutions for analysis, management and sharing in repositories.</p>
--	--

### C. Methodology and Performance Indicators

4) The DISC-UK bid is based on a distributed model in which each participating partner is responsible for the work on their own repositories, yet experience, support and knowledge are shared in order to increase levels of success. This builds on the existing informal collaboration of DISC-UK members for improving their data libraries and models of data support at each institution. It will also bring academic data libraries in closer contact with e-Prints repository managers and develop new forms of cooperation between these distinct groups of information professionals within academic environments. The advantage for the broader community is to provide exemplars for a range of approaches and policies in which to embed the deposit and stewardship of datasets in IRs. Indeed, among the partners there will be exemplars for the three main repository solutions in the UK: EPrints, DSpace and Fedora.

5) There are a number of difficult issues that must in some way be resolved for any IR to accept responsibility for deposit of datasets in earnest, including ingest, metadata schemas, storage, file format migration, unique identifiers & levels of granularity, applying OAI-PMH to datasets, and making content available for centralised harvesting by data type. Since the DataShare project will in this sense be a first true attempt, gaining understanding from international precedents will be crucial. Contact will be made with particular IR managers known to have experience in managing research datasets, such as at the University of Michigan, Johns Hopkins University, and MIT in the US. A consultant well-versed in these developments in the US—Ann Green from Digital Life Cycle Research & Consulting, New Haven, Connecticut, USA—(Green and Gutmann, 2007) will assist each partner in aligning its development plans with institutional strategies and situations, taking into account: policies, procedures, software choices, workflow, organisational structure, suitable levels of support (e.g. for assisted deposit and data curation), preservation planning and scope. Ms Green will visit each institution early in the project<sup>13</sup> and facilitate meetings of data and repository staff, advising on these areas and training needs of participating staff. After these meetings, full workplans for each institution will be submitted to the project manager, including suitable and measurable performance indicators. Ms Green will revisit the partnership in the middle of the project to assist the sites to test their repositories against the DCC DRAMBORA toolkit and/or the CRL Trustworthy Repositories Audit & Certification: Criteria and Checklist, as well as towards the end to participate in the dissemination strategy.

6) The *project plan*, to be developed in full within the first three months of the project, but with a full draft available within the first month, will be developed to deliver outcomes beneficial to the wider JISC community, as well as to take into account the particular situations at each of the participating institutions.

### D. Institutional Summaries

#### 7) London School of Economics:

*LSE Research Online* - <http://eprints.lse.ac.uk/>

London School of Economics Data Library - <http://www.lse.ac.uk/library/datlib/Default.htm>

*LSE Research Laboratory Data Service* - <http://rlab.lse.ac.uk/data/>

8) The LSE's Data Library is integrated with the British Library of Political and Economic Science. The BLPES has a successful e-Prints repository which has recently begun piloting management of complex objects such as datasets, involving the Data Librarian. The software is based on EPrints; previously the Data Library has experimented with using VDC (The Virtual Data Centre) open source software from Harvard, to manage online datasets using the XML-based DDI standard (Data Documentation Initiative). A second data service serves the LSE Research Laboratory (RLAB), which brings together LSE's leading research centres in economics, finance, industrial relations, social policy and demography. One staff member from each Data Library is involved in DISC-UK and the project.

9) The LSE has led the social sciences survey work for the StORe project, the VERSIONS project, and is a partner of the Higher and Further Education and the Information Management Environment e-Infrastructure

<sup>13</sup> The first meetings have been scheduled for the week of 23 April, 2007 in consultation with JISC.

project and can provide expertise in access management which will be explored for determining levels of open and restricted access to datasets. Participating staff are also actively involved with Neureus, a network of European economics librarians, which will be a target for project dissemination.

10) University of Oxford (Nuffield College):

*Oxford Research Archive* - <http://ora.ouls.ox.ac.uk/>

*Oxford Data Library (Nuffield College)* - <http://www.nuff.ox.ac.uk/projects/datalibrary/>

11) Although based in Nuffield College, the Data Manager of Oxford's Data Library provides data and associated support services for all social scientists in the university. She works closely with researchers in acquisition and management of social science data, in particular survey microdata. She will be the primary contact for the project and is one of the founders of DISC-UK. A separate repository for the Data Library is not envisaged, as Oxford EPrints, run centrally from Oxford University Library Services, intends to include social science numeric datasets in its holdings. This is part of a planned migration of repository software from EPrints to Fedora/VTLS Vital. There are also two departmental e-Print repositories at Oxford registered with OpenDOAR, in the field of Mathematics. The project will work with repository and library staff within OULS, with the primary aim of encouraging and supporting the deposit of datasets from social science researchers, not otherwise deposited in the UKDA.

12) University of Southampton:

*e-Prints Soton* - <http://eprints.soton.ac.uk/>

*University of Southampton Library Data Resources* - [www.soton.ac.uk/library/resources/collections/data/](http://www.soton.ac.uk/library/resources/collections/data/)

13) Southampton is one, if not the, leader in the field of institutional repositories and also is the developer of the well-regarded EPrints software. Their policy is open to deposit of datasets already, however this has not been actively pursued except in speciality repositories such as that belonging to the National Oceanography Centre, Southampton (NOCS). The Social Sciences Librarian, a DISC-UK member, plays the role of data librarian in addition to her other roles, as part of the Hartley Library and maintains an extensive website of pointers to datasets in the social sciences. The Faculty Liaison Librarian for Law, Arts and Social Science will also help to facilitate collaboration between repository and data support staff in this project. Southampton is currently involved with three projects in the data cluster of the JISC Repositories Programme, CLADDIER, R4L and GRADE (the latter through NOCS) which we expect to provide synergy and insight to the project. Through the PRESERV project, Southampton is already using PRONOM as a tool for identification of file formats and preservation through migration. This experience is also of great use to the project.

14) EDINA & Edinburgh University Data Library:

*Edinburgh Research Archive (ERA)* - <http://www.era.lib.ed.ac.uk/>

*Edinburgh University Data Library* - <http://datalib.ed.ac.uk/>

*EDINA* – <http://edina.ac.uk/>

15) As in the former Datateach project, the EDINA National Data Centre will play a coordinating role for the project, which is designed to enhance *institutional capacity* for supporting research datasets.<sup>14</sup> EDINA's experience in repositories through GRADE, JORUM, and PROSPERO/The Depot will help to inform the project. The Data Library is a service for the University of Edinburgh, and while based organisationally with EDINA, will act as a participating partner to enhance repository services for Edinburgh depositors. The two data librarians will collaborate with the ERA manager to set up a seamless workflow for users depositing both EPrints and datasets. The Data Library expects to run a DSpace data repository, which is the same software used by ERA. The new data repository service will be promoted to users as part of Information Services. The Data Library views the development of a data repository as part of its planning to support medical researchers in need of a data sharing plan for MRC research proposals.

*E. Dissemination*

16) A dissemination plan will be elaborated as part of the project plan, based on some of the deliverables above. A two-way process is crucial to the project, to bring greater understanding to both data professionals and repository managers working within universities and research libraries about each other's realms. Demystification will be a goal for both groups, within and outwith the project. Identification of training opportunities and interaction with other projects will be key. A stakeholder analysis will be conducted.

*F. Evaluation*

17) Because this is a new area for UK IRs, an external evaluator familiar with UKHE and the JISC information environment will be chosen early in the project to help create an evaluation plan, covering both

---

<sup>14</sup> Although EDINA is based in Scotland, its Memorandum of Understanding with HEFCE allows it to lead a bid in the Capital Programme.

progress made at each institution, and the impact of the project as a whole on the broader community. Both summative and formative forms of evaluation shall be used.

**G. Sustainability**

18) The funding will greatly assist the partners to get 'up and running' with an ambitious plan. There is a momentum to this project which both precedes and will last beyond the life of the project. Two elements will help to ensure the sustainability of the project's outputs for the partners involved: the DISC-UK network itself and its connections with an international community of data professionals; and the commitment of each partner institution to the goals of the project and to enhancing their repositories for incorporation of research data. The project will work with both the RSP and the DCC to make its external outputs available and useful beyond the life of the project.

**H. Intellectual Property**

19) For external written outputs, copyright will belong to the authors or their institutions, but will be licensed to JISC in perpetuity, and will also be made available open access. Software will be developed within open source environments. There are IPR issues to do with deposit of research data which will be explored as part of the project, with each institution coming up with suitable policies and procedures.

**I. Timetable (pre-funding period is extended; project end date will be set forward depending on start date)**

Activity	2007				2008				2009
	Dec '06- Feb (pre- funding)	Mar- May	Jun- Aug	Sep- Nov	Dec '07 - Feb	Mar- May	Jun- Aug	Sep- Nov	Dec '08- Feb
Relationship-building and consultation with relevant parties (RSP, DCC, RRT, UKDA and academic champions at local institutions)	----->								
Study outcomes of related projects. Technical watch. JISCMail project list & filestore set up for internal communication	----->								
Draft project plan; consortium agreement; set up project website	☺								
Methodology with consultant decided and revised (per institution)		→			→			→	
Collaborative meetings w/ RRT, RSP, others. Training events as identified or requested.			☺ <sup>15</sup>	☺	☺	☺	☺	☺	☺
Specific deliverables written for dissemination by RSP, DCC.					----->				
Performance indicator milestones such as n number of datasets in each repository					☺	☺	☺	☺	
Submit final report and summative evaluation									☺

**J. Risk Analysis**

Risk	Impact / Probability	Mitigation
Delay in start; recruitment problems	High / Low	Identified project officers, liaison roles, and project manager are already in post. Work can be done even if recruiting on technical posts causes delay.
Difficulties in developing a consortium agreement and	High / Low	DISC-UK is an active partnership already. As part of bid preparation, foundation for cooperation between data

<sup>15</sup> The first dissemination event has been scheduled: project team members will present on the role of the data librarian at the SCONUL event, *Support for e-Research: Filling the Library Skills Gap* to be held at the National e-Science Centre in Edinburgh on 14-15 June 2007.

robust partnership		libraries and repository managers has been laid.
Technical solutions not available for identified problems.	Medium / Medium	Include technical staff in project who can develop open source software and common tools; contribute solutions to broader community to reduce technical barriers to acceptance of datasets by other institutions. Maintain technical watch and communication with other projects. If necessary, migrate to one of the other systems (e.g. e-Prints, DSpace, Fedora).
Lack of acceptance of institutional repositories as place for depositing/managing datasets.	Medium / Medium	Be active in repository programme to show how DISC-UK model fits into landscape; communicate with key stakeholders including data centres and research councils; put sufficient effort into outreach with academics at each partner institution; understand their needs; recommend deposit in large data centres such as UKDA whenever appropriate.
Loss of key staff before end of project.	Medium / Low	Ensure embeddedness of project in institutional practices; involve both data library and repository staff in project setup and workflows so that knowledge never rests with one person only.

### III. BUDGET<sup>16</sup>

Directly Incurred Staff	Mar-07	April 07-March 08	April 08-March 09	Total £
Project Manager, Grade 8, 0.4 FTE	1,754	21,474	20,746	43,974
Tech Officer (Edi), Grade 7, 0.5 FTE	1,538	19,207	19,095	39,840
Tech Officer (Ston), Grade 7, 0.5 FTE	1,071	11,783	14,698	27,552
Tech Officer (LSE), Grade 7, 0.5 FTE	1,494	16,433	19,377	37,304
Tech Officer (Oxf), Grade 7, 0.25 FTE	885	10,839	10,471	22,195
<b>Total Directly Incurred Staff</b>	<b>6,743</b>	<b>79,735</b>	<b>84,387</b>	<b>170,865</b>
<b>Non Staff</b>				
T&S	1,600	9,600	8,000	19,200
Evaluation consultancy	0	5,000	5,000	10,000
Consumables	20	240	220	480
Equipment/PCs	1,400	0	0	1,400
Consultant	0	3,484	3,216	6,700
<b>Total Directly Incurred Non Staff</b>	<b>3,020</b>	<b>18,324</b>	<b>16,436</b>	<b>37,780</b>
<b>Directly Incurred Total</b>	<b>9,763</b>	<b>98,059</b>	<b>100,823</b>	<b>208,645</b>
<b>Directly Allocated</b>				
Directors across partners	576	7,047	6,808	14,431
Edinburgh staff	1,097	13,421	12,966	27,484
Edinburgh staff	410	5,120	5,090	10,619
LSE staff	886	10,684	10,303	21,873
LSE staff	428	5,235	4,009	9,671
Oxford staff	531	6,503	6,282	13,317

<sup>16</sup> Original budget, unaltered. In fact, alterations will need to be made in the project plan depending on start date and diversion of up to half the evaluation funding toward technical consultancy.

Southampton staff	1,071	11,783	14,698	27,552
<b>Directly Allocated Total</b>	<b>4,998</b>	<b>59,792</b>	<b>60,156</b>	<b>124,947</b>
<b>Indirect Costs</b>	<b>8,497</b>	<b>101,960</b>	<b>93,463</b>	<b>203,920</b>
<b>Total Project Costs</b>	<b>23,258</b>	<b>259,812</b>	<b>254,442</b>	<b>537,512</b>
<b>Amount Requested from JISC</b>	<b>12,227</b>	<b>127,628</b>	<b>127,927</b>	<b>267,782</b>
<b>Institutional Contributions</b>	<b>11,031</b>	<b>132,184</b>	<b>126,515</b>	<b>269,730</b>
<b>Percentage contributions over the life of the project</b>	<b>JISC</b>	<b>50</b>	<b>Partners</b>	<b>50</b>

#### A. Qualitative and quantitative benefits

1) Each partner institution will directly allocate staff time to the project (up to 25% per individual), and in return will gain additional effort for technical support (50% at each institution except Oxford, at 25%). Oxford and Southampton will use the additional effort for data curation and advocacy; Edinburgh and LSE will hire or buy out additional technical staff to design and support the repository software itself. Servers (hardware and software) do not need to be purchased as they already exist at each site, though one or two staff machines are included in the budget. The benefits of hiring the external consultant and evaluator will be shared across the partner sites, as will travel expenses and training costs. Opportunities for public speaking will be shared. EDINA will receive additional resources for project management, which will enable the Data Librarian to be seconded at 40%, allowing resource for additional effort in the Data Library to replace her. The nature of other benefits and contributions for each institution can be seen in the institutional summaries.

#### IV. KEY PERSONNEL

##### 1) EDINA & Edinburgh University Data Library

Project Director: Peter Burnhill

Peter was the first full-time manager of Edinburgh University Data Library (1984) and Director of EDINA since 1996. He holds an MSc in Statistics from the London School of Economics and worked for five years as a social survey statistician for the Centre for Educational Sociology at the University of Edinburgh. He is a past President of IASSIST, and participated in relevant international projects since the 1980s such as an initiative for cataloguing machine-readable datasets using AACR2, and member of Working Group 7 and subsequent task force on data and facilities for the UK Inter-agency Committee on Global and Environmental Change in the early 1990s. He was a consultant for Statistics Canada in the late 1990s, which subsequently launched the Data Liberation Initiative, which set up a network of data libraries for supporting statistical data at universities across Canada. As Project Director for the Datateach project he provided administrative support and contributed to the design and quality of the survey work and analysis. Peter was the Director of the Digital Curation Centre during its setup phase from 2004-2005. He has been active in leading several of the EDINA repository-related projects and services, including JORUM and Prospero/The Depot.

##### 2) Associate Project Director: Robin Rice

Robin is the University's Data Librarian and actively participates in EDINA Business Development. She was the project manager of the Datateach project as part of JISC's Learning & Teaching 5/99 Programme, which lay the groundwork for the formation of DISC-UK. She also worked with Peter as Project Coordinator to set up the Digital Curation Centre (2004-05). She participated as a primary author on the Prospero interim repository scoping work for JISC. She coordinates the IASSIST web team and serves on the Administrative Committee as a European member. If DataShare is funded she will be seconded to EDINA for 40% time to be project manager and will also work on the Edinburgh-based repository activity.

##### 3) Project Officer: Stuart Macdonald

Stuart is Assistant Data Librarian for the Data Library and is responsible for the EDINA agcensus service. Along with the Data Librarian he is responsible for providing access to and support for numeric and encoded data for analysis and teaching on behalf of Edinburgh University researchers. He is IASSIST Assistant Treasurer for the European region and a founder member of DISC-UK. He is also editor of Intute: Social

Science Statistics and Data section. Stuart will be responsible for ingested content (including metadata enhancement) and will assist with the design of the Edinburgh-based data repository.

4) IR Liaison role: Dr. Theo Andrew, Electronic Publications Officer, Edinburgh University Library

#### 5) London School of Economics

Project Officer: Luis Martinez

Luis is the Data Librarian at the London School of Economics, responsible for the provision of data services to the LSE research and teaching community. He holds a BSc in Mathematics specializing in Computer Science and Statistics and enrolled in the second year (part-time) of the MSc in Administration, Design and Management of Information Systems at the LSE Department of Management. Luis is a member of the CURL-SCONUL E-research Task Force and for the last two years has been involved in the feasibility study of a repository for datasets, testing DSpace and the Virtual Data Centre. Furthermore, he has been involved in a School wide group that produced recommendations for repository software for the LSE Library senior management.

6) Project Officer: Tanvi Desai

Tanvi is Data Manager of the Research Laboratory at the LSE. She has extensive experience supporting international research groups with a need for comparable micro-data. In particular, data management for the PiEP project<sup>17</sup> gave her an insight into the challenges faced in gaining access to and developing comparable international datasets. The project's data needs led to the development of the PiEP-LISSY System, a secure remote access system to comparable European linked employer-employee microdata. She has advised both the Australian and New Zealand National Statistical Institutes on remote access systems, as well as attending Expert Groups at Eurostat to advise them on their data access strategy. She wrote a 'Review of International Data Needs and Resources' for the Economic and Social Research Council, with the aim of gaining an understanding of the opportunities and the obstacles presented by international data resources.

7) IR Liaison role: Frances Shipsey, VERSIONS Project Manager and Eservices Librarian (from Jan 2007)

#### 8) Nuffield College, Oxford

Project Officer: Jane Roberts

Jane is the Data Manager of Oxford University Data Library, based in Nuffield College. She has a BA in Politics, and worked for many years as a Research Assistant/Officer in university departments. She began the Data Library 18 years ago, and during this time has worked on numerous data collection, management and analysis projects. She is a member of the IASSIST Outreach Committee and the Programme Committee for several recent IASSIST conferences, a member of the Data Support Committee for the Europe-wide EQUALSOC project<sup>18</sup>, and a founder member of DISC-UK.

9) IR Liaison role: Sally Rumsey, Project Manager: Institutional Repository, SERS, Oxford University Library Services

#### 10) University of Southampton

Associate Project Director: Dr Mark Brown

Dr. Brown has been University Librarian at the University of Southampton since 2001. He was Project Director for a range of digitisation projects, including the AHRC EPPI Project and the current CSR2 Digitisation Project, and leader of the digitisation team at Southampton. He has acted as Project Director for the JISC funded TARDIS project in the FAIR Programme to develop an institutional repository at Southampton, and currently leads the Institutional Repository Steering Group, which is coordinating the development of the IR at Southampton. Currently he is Chair of the Consortium of University Research Libraries (CURL), a member of JISC Content and Services Committee, and the JISC IE Working Group, and is the nominated JCCS representative on the EDINA Board.

11) Project Officer: Harry Gibbs

Harry is an Academic Liaison Librarian working with the School of Social Sciences and holding special responsibility for data provision. With regard to data, her role includes maintaining and developing the University's Data Resources Web pages and dealing with user enquires regarding a range of social sciences data. If DISC-UK DataShare is funded, she will work within the University of Southampton to promote and facilitate data sharing.

---

<sup>17</sup> The Pay Inequalities and Economic Performance project was funded by the EC under FP5.

<sup>18</sup> Economic Change, Quality of Life & Social Cohesion, <http://www.equalsoc.org/>

12) IR Liaison role: Wendy White, Repository Manager and Faculty Liaison Librarian: Law, Arts and Social Science

13) Digital Life Cycle Research & Consulting

Ann Green is an independent research consultant focusing upon the digital life cycle of scholarly resources, including their creation, delivery, management, long-term stewardship, and preservation. She has an extensive background in digital archiving and user-driven support services in the social sciences. Ann has participated in the development and promotion of standards for social science statistical metadata and is an active participant in the development of international social science services and initiatives. She is the former director of the Social Science Research Services & Statistical Laboratory (Statlab) at Yale University where she coordinated social science research and instructional technologies, facilities, and services for over 15 years. Ann participated in campus wide information technology planning as well as initiatives in digital library infrastructure and services at Yale. She is the immediate past President of IASSIST and former Chair of the Executive Council of the Inter-university Consortium for Political and Social Research (ICPSR). Current projects include program evaluation of digital archives and related services, as well as consultation on 'cradle to archive' planning for digital collections.

## V. REFERENCES

(2004) "OECD Ministerial Declaration on access to research data from public funding" (Annex 1)  
[http://www.oecd.org/document/0,2340,en\\_2649\\_34487\\_25998799\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/0,2340,en_2649_34487_25998799_1_1_1_1,00.html)

Burnhill, P., Rice, R. and D. Geraci (2005). *The social side of science data sharing: distilling past efforts. Preserving Value Conference Proceedings*. Edinburgh: The Royal Society, 21-23 November 2005. Poster paper: <http://www.ukoln.ac.uk/events/pv-2005/pv-2005-final-poster-papers/040-poster-1.pdf> Poster: <http://www.ukoln.ac.uk/events/pv-2005/posters/burnhill-geraci-rice.pdf>

Burton, G. (2006). *The Source to Output Repository Project: Social Science*. [Unpublished.]  
<http://hdl.handle.net/1842/1418>

The Digital Archiving Consultancy, The Bioinformatics Research Centre (University of Glasgow) and The National e-Science Centre (NeSC) (2005). *Large-scale data sharing in the life sciences: Data standards, incentives, barriers and funding models (the Joint data standards study)*.  
<http://www.mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC002552>

Green, A. and Gutmann M. (2007). *Building partnerships among Social Science researchers, institution-based repositories, and domain specific data archives*. Forthcoming: **OCLC Systems and Services: International Digital Library Perspectives**. 23(1). Preprint: <http://hdl.handle.net/2027.42/41214>

Heery, R. and Anderson, S. (2005). *Digital repositories review*.  
[http://www.jisc.ac.uk/uploaded\\_documents/digital-repositories-review-2005.pdf](http://www.jisc.ac.uk/uploaded_documents/digital-repositories-review-2005.pdf)

Heery, R. and Powell, A. (2006). *Digital repositories roadmap: Looking forward*.  
<http://www.ukoln.ac.uk/repositories/publications/roadmap-200604/>

Humphrey, C.K., Estabrooks, C.A., Norris, J.R., Smith, J.E. and K.L. Hesketh (2000). *Archivist on board: Contributions to the research team. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* [On-line Journal] 1(3). <http://qualitative-research.net/fqs/fqs-eng.htm>

O'Reilly, T. (2005). *What is Web 2.0: Design patterns and business models for the next generation of software*. 30/9/2005 (p 3). <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html?page=1>

Rice, R., Burnhill, P., Wright, M. and S. Townsend. (2001). *An enquiry into the use of numeric data in learning and teaching: Report and recommendations for UK higher education*.  
<http://datalib.ed.ac.uk/projects/datateach/report/>

Suber, P. (2006). *Timeline of the Open Access Movement*. **Open Access News**  
<http://www.earlham.edu/~peters/fos/timeline.htm>



INFORMATION SERVICES GROUP

The University of Edinburgh  
Main Library Building  
George Square  
Edinburgh EH8 9LJ

Telephone +44 (0)131 650 1000  
or direct dial +44 (0)131 650 4977  
Fax +44 (0)131 650 4978  
Email H.Hayes@ed.ac.uk

17 November 2006

JISC  
King's College London  
1<sup>st</sup> Floor, Brettenham House  
5, Lancaster Place  
London  
WC2E 7EN

Dear Sir/Madam

The DISC-UK Datashare project aims to populate partner institutions' repositories with small-scale research datasets, thereby exposing a host of issues for analysis and policy development during the life of the project, leading to deliverables that will help guide other UKHE institutions wishing to engage in this new and exciting area of repository activity.

I am very happy to support the proposed DISC-UK DataShare project. This project helps to integrate a number of strands for the University knowledge strategy in relation to academic library and data library objectives for enhancing research support.

Yours sincerely

Helen Hayes  
Vice principal for Knowledge Management and  
Librarian to the University

21 November 2006

Ms Amber Thomas  
Programme Manager, Information Environment  
JISC Executive  
University of Bristol  
3rd Floor, Beacon House  
Queens Road  
BS8 1QU

Dear Ms Thomas

**Re. JISC Capital Programme – Circular 04/06: Repositories and Preservation Programme, Strand D (b): Repository Enhancement Projects**

**DISC-UK DataShare**

The University of Southampton is pleased to take part in the DISC-UK DataShare Project as a primary partner.

I confirm that the proposal has the full support of the University of Southampton's senior management and that we agree to make appropriate institutional contributions to the project as documented in the proposal document.

Yours sincerely,



**Professor W A Wakeham  
Vice-Chancellor**



THE LIBRARY  
NUFFIELD COLLEGE  
OXFORD OX1 1NF

Telephone: +44 1865 278547  
Fax: +44 1865 278621  
E-mail: [Elizabeth.Martin@nuffield.ox.ac.uk](mailto:Elizabeth.Martin@nuffield.ox.ac.uk)  
Website: <http://www.nuffield.ox.ac.uk/library/>

21 November 2006

Ms Amber Thomas  
Programme Manager  
Information Environment  
JISC Executive  
University of Bristol  
3rd Floor, Beacon House  
Queens Road  
BS8 1QU

Dear Ms Thomas,

**DISC-UK DataShare Project Proposal**

I am pleased to support the proposal from the Data Information Specialists Committee-UK for funding under the JISC Capital Programme.

The DISC-UK collaboration has provided an effective support framework for the development of the partner data libraries, and the members are ready to work together in this important area of repository enhancement.

We will also be pleased to have the opportunity to work alongside Oxford University Library Services' institutional repository project, with our data librarian's expertise in the creation and management of datasets in the social sciences.

I confirm Nuffield College's institutional endorsement of the proposed project.

Yours sincerely,

Elizabeth Martin  
Nuffield College Librarian

17 November 2006

Ms Amber Thomas  
Programme Manager  
Information Environment  
JISC Executive University of Bristol  
3rd Floor, Beacon House Queens Road  
Bristol BS8 1QU

Dear Ms Thomas

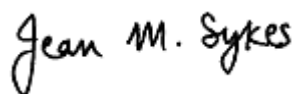
**DISC-UK DataShare Project : Repository Enhancement Project**

I am pleased to support the proposal from EDINA and the other DISC-UK members (Southampton University, Oxford University, LSE) for funding under the JISC Capital Programme. This letter confirms LSE's institutional endorsement of the proposed project.

LSE Data Library and its partner data libraries in the DISC-UK have a two-year track record of collaboration on research data initiatives. We are delighted to have the opportunity to support this proposal which, if successful, will build on work already carried out by the DISC-UK partner data libraries. A successful proposal will bring benefits to LSE, its academics and digital repository communities in the UK as a whole.

I confirm that LSE agrees to be a partner and gives its institutional support to the DataShare project.

Yours sincerely



**Jean Sykes**  
Librarian and Director of IT Services

---

**Robin Rice**  
Data Librarian  
EDINA National Data Centre and University Data Library  
Information Services  
Main Library Building, George Square  
University of Edinburgh  
Edinburgh EH8 9LJ  
R.Rice@ed.ac.uk, 0131 651 1431

---

## Experience

- ◆ Service manager for a University Data Library responsible for staff training and mentoring, range of online services with in-depth user support.
- ◆ Fourteen years working as a data professional in two countries. Strong service orientation within Higher Education.
- ◆ Expertise in UK, US, European and global social datasets/databases, including population census data, large-scale surveys, official statistics, financial data, economic time-series and geo-spatial data, with in-depth awareness of their use in research in a wide range of disciplines.
- ◆ Highly skilled in data analysis applications including SPSS, ArcGIS, Excel and Access. Knowledge of SAS and Stata statistical packages, other database applications, and MapInfo desktop mapping software. Experienced web editor using Dreamweaver and graphics packages; proficiency in Unix.
- ◆ Knowledge of standards relating to digital preservation and online data sharing including XML, the OAIS Reference Model, preservation metadata, Data Documentation Initiative (DDI), Dublin Core metadata, Open Archives Initiative. Awareness of digital library developments, repository software and processes, Grid technologies and e-Science applications, interoperability standards.
- ◆ Demonstrated ability as a project manager, including set-up of a multi-site, academic centre of expertise with complex management structure and budget in excess of one million pounds per annum. Excellent organisational, planning, communication and facilitation skills.

## Employment History

Oct. 1998—present EDINA National Data Centre and University Data Library Information Services University of Edinburgh	<b>Data Librarian (and Service Manager from 2004):</b> Manage overall service provision including user support, promotion activities, collection development and preservation, and website development. Train and direct Assistant Data Librarian in assisting users, teaching training courses, acquiring and documenting datasets, writing web pages and documentation, data management activities, and current awareness service. Direct Documentation Assistant in formatting user guides, documentation and web pages. Provide individual research consultation on use of
--	--

secondary datasets to staff and students in various software environments including statistical packages, database software, and desktop mapping (GIS) applications. Promote and publicise Data Library events and services to staff in all colleges. Develop training courses and learning materials for staff and students. Participate in classroom activities as needed in consultation with course organisers and lecturers. Act as liaison between EDINA division (unit) and Information Services staff. Serve on IS committees as required. Meet with other EDINA service managers at weekly management meetings. Responsible for Data Library strategy with Director, and contribute to EDINA strategy.

Lead EDINA project activity in the areas of data curation and digital preservation. Coordinate EDINA staff contributions to the Digital Curation Centre as a project partner. Develop business strategy and funding opportunities as member of the EDINA Business Development team and through Information Services committee work. Take on project management work from time to time in consultation with EDINA Director.

Feb. 2004—April 2005  
Digital Curation Centre  
University of Edinburgh

**Phase One Project Coordinator (60% secondment):** Work closely with the Director (Phase One) to set up and manage a national distributed centre of expertise with both a service and research orientation across four partner institutions, during the first year of operation. Responsible for writing and coordinating project plan and other reporting documents to deliver to funders on time. Facilitate the Consortium Agreement and other Management deliverables. Set up collaborative, remote work environment through Access Grid meetings, staff email lists, Wiki and Intranet private web pages. Oversee clerical support for meetings. Help to create new job descriptions, hire and train staff. Work with the Outreach team to disseminate project information over the website and at events. Work with Services team to set up DCC Helpdesk system and procedures provided by EDINA. Assist Director and Outreach team in planning official Launch event with 150 invited 'VIP's in the field. Give papers on digital curation or

the DCC at conferences.

- Feb. 2000—Sept. 2001  
Edinburgh University Data Library
- Project Manager (50%):** *The Use of Numeric Data in Learning and Teaching*, JISC-funded project with partners at the British Library of Political and Economic Science (LSE), the Data Archive, MIMAS, and EDINA. Set up a steering committee of UK academics to oversee production of national survey, case studies, and report. Lead author of final deliverable, *An Enquiry into the Use of Numeric Data in Learning and Teaching: Report and Recommendations for UK Higher Education*. Available: <http://datalib.ed.ac.uk/projects/datateach.html>
- Nov. 1992—Sept. 1998  
Data and Program Library Service, University of Wisconsin-Madison, USA
- Special Librarian** (Associate Special Librarian until Jan. 1995): Provided reference and referral to social science students and faculty in finding and using statistical data files and their documentation in all available formats. Developed collection through acquisition of new datasets and reference materials. Produced and edited library newsletter, *DPLS News*. Monitored the Internet for current awareness resources in social science research and instruction. Created and maintained library WWW pages and FTP site. Responsible for preservation activities including migration and refreshment of files and media. Assisted with Windows NT system administration of server and client workstations, including nightly backups.
- July 1991—Nov. 1992  
Data and Program Library Service, University of Wisconsin-Madison, USA
- Library Assistant:** Provided initial reference service to patrons, kept circulation records. Worked with computer files across platforms (mainframe, Unix, DOS) and media (9-track tape, DAT, CD-ROM).
- July 1990—Aug. 1992  
Kurt F. Wendt Engineering Library, University of Wisconsin-Madison, USA
- Indexer:** Organized a special collection (*Value Engineering Reference Center*) on behalf of the Society for American Value Engineers. Designed and entered a computerized index for an archival collection of memorial papers and for the annual *S.A.V.E. Conference Proceedings*, available on the Engineering Library network.

## **Education and Training**

### **Qualifications**

School of Library and Information Studies, University of Wisconsin-Madison, USA. *Master of Arts* (September, 1990-May, 1992). Valmai Fenster Award for "Outstanding promise of exceptional scholarly contribution to the profession."

Clark University, Worcester, Massachusetts, USA. *Bachelor of Arts Cum Laude* in Global History (September 1983-June 1987). Phi Beta Kappa. Elected Commencement Student Speaker.

Park Ridge High School, Park Ridge, New Jersey, USA. Graduated as honors student (1983-1987).

### **Other Relevant Training**

*Various EUCS training courses on desktop software, etc.*

10<sup>th</sup> ECDL, Alicante, Spain.

European Conference on Research and Advanced Technology for Digital Libraries, (18-20 Sept., 2006).

DEFF, Copenhagen, Denmark.

Building the Info Grid: Digital Library Technologies and Services - Trends and Perspectives, (26-27 Sept., 2005).

IASSIST Workshop, Madison, Wisconsin, USA.

Pre-conference half-day training in STATA, SPSS, and SAS statistical software (25 May, 2004).

Erpanet and CODATA Seminar, Biblioteca Nacional Lisboa, Portugal.

Appraisal of Scientific Data (15-17 December, 2003).

Institute for Social and Economic Research, University of Essex.

British Household Panel Study two-day training course using SPSS (18-19 July, 2002).

Fred Pryor Seminars, Holyrood Hotel, Edinburgh.

'How to Supervise People,' one-day seminar on staff management (21 June, 2001).

JISC Project Management Workshop, Kings College London.

One-day course for project managers of Learning and Teaching Programme (27 Sept., 2001).

## Professional Memberships

- Member, British Society of Population Studies (Dec. 2004).
- Founding member, Data Information Specialists Committee—DISC-UK (Feb. 2004).
- Elected to administrative committee and appointed Website Editor (volunteer coordinator) to IASSIST—International Association of Social Science Information Services and Technology (May 2003-present. Individual member since 1998.) Programme and local arrangements committee for 2005 IASSIST/IFDO conference hosted in Edinburgh.
- Three-year term on Radical Statistics Troika (board) (Feb. 2002-2005).
- Member, Social Research Association (2001-present).
- Member, BURISA—British Urban and Regional Information Systems Association (1999-present).
- Member, Association of Survey Computing (1999-present).
- Member, ALISS (formerly ASSIGN) — Association of Librarians and Information professionals in the Social Sciences (1999-present).
- Elected Secretary, Board of Directors, Computer Professionals for Social Responsibility, USA (1997-1998).

## Publications

Burnhill, Peter et al (2006). *Perspectives and models relating to a national facility to support deposit of Pre- & Post-Prints under terms of Open Access*. Edinburgh/Nottingham: EDINA and SHERPA, April, 2006.  
Available: <http://edina.ac.uk/projects/prospero/ProsperoAppendixFull.pdf>

Burnhill, Peter, Rice, Robin and Diane Geraci (2005). "The social side of science data sharing: distilling past efforts." Preserving Value Conference Proceedings. Edinburgh: The Royal Society, 21-23 November 2005.  
Available: <http://www.ukoln.ac.uk/events/pv-2005/pv-2005-final-poster-papers/040-poster-1.pdf>

Rice, Robin (2005). "The UK Digital Curation Centre: a working resource for librarians." *Assignment 22* (4) July 2005, pp 11-14. Available:  
[http://www.dcc.ac.uk/docs/publications/The\\_UK\\_Digital\\_Curation\\_Centre-RR.pdf](http://www.dcc.ac.uk/docs/publications/The_UK_Digital_Curation_Centre-RR.pdf)

Williams, Malcolm, Collett, Tracey and Robin Rice (2004). *Baseline study of quantitative methods in British Sociology*. Durham/ Birmingham: BSA/C-SAP, 2004. Available: [http://www.britsoc.co.uk/Library/C-SAP\\_BSA\\_Baseline\\_Study\\_of\\_Quantitative\\_Methods\\_in\\_British\\_Sociology\\_2004.pdf](http://www.britsoc.co.uk/Library/C-SAP_BSA_Baseline_Study_of_Quantitative_Methods_in_British_Sociology_2004.pdf)

Rice, Robin (2003). "The USA PATRIOT Act and American libraries." *Information for Social Change*, No. 16 (Winter 2002/2003). Available:  
<http://www.libr.org/ISC/articles/16-Rice.html>

Rice, R., Burnhill, P., Wright, M. and Townsend, S. (2001). *An enquiry into the use of numeric data in learning & teaching: Report and recommendations for UK higher education*. Edinburgh: Edinburgh University Data Library, Sept. 2001. Available: <http://datalib.ed.ac.uk/projects/datateach/report/>

Rice, R. (Spring, 2001). "Understanding barriers to the use of numeric data in learning and teaching." *IASSIST Quarterly*, (Vol 25:1), pp 5-9. Available: <http://iassistdata.org/publications/iq/iq25/iqvol251rice.pdf>  
[See also several similar project dissemination articles listed at <http://datalib.ed.ac.uk/projects/datateach.html#pubs>]

Rice, Robin (2000). "Providing local support for academic data users." *Data Archive Bulletin*, pp 8-11. Essex: UK Data Archive, May, 2000. Available: <http://datalib.ed.ac.uk/discuk/docs/rice2000.pdf>

Rice, Robin (1996). "Data sources for Library and Information Science research," *LIBRES* 7 (1). Available: <http://libres.curtin.edu.au/libre7n1/rice.htm>

Ed Cortez and Robin Rice (1994). "An investigation into the role of public libraries with online services." *Public Library Quarterly* 14 (2/3) 1994: 7.

## Personal details:

### Stuart Robert Macdonald

88 Blackford Avenue  
Edinburgh  
EH11 1ST

0131-650-3304 (daytime)  
0131-668-3722 (evening)

**Date of Birth:** 17.08.63

**Nationality:** British

**Marital Status:** Married

## Employment record:

**Jan. 1999 – present: Assistant Data Librarian**      **Edinburgh University Data Library,  
George Square, Edinburgh**

- Register and assist Edinburgh University staff & students to access, subset and extract numeric and encoded research data for analysis and teaching from a wide variety of sources. Demonstrate various interfaces to academic and support staff via informal or formal training sessions and presentations in addition to face-to-face consultations to determine research needs and advise on data sources (including those offered by other national services e.g. MIMAS). Supervise 1 documentation assistant.
- Responsible for the migration of a command-driven agricultural census data service to an online web-based visualization and data download service for the JISC-funded EDINA National Data Centre (URL: <http://edina.ac.uk/agcensus/>). Responsibilities include liaising with relevant individuals in order to implement an authorization and authentication system; conduct usability tests (Nielsen-Norman); design and create online documentation; construct a business model; format input data; create metadata records, online help and training materials; negotiate with external organizations license for context mapping data and the receipt of raw input data from government departments. Currently investigating the integration of complimentary gridded data (such as meteorological and climate variables) into the national service.
- Course instructor for Edinburgh University Computing Services courses: *Using Ordnance Survey Data via EDINA Digimap*; *Introduction to Visualising Data in ArcGIS 9*.  
Course Instructor for *UK Population Census Data* at School of GeoSciences
- Chair, present at and attend seminars, workshops and conferences (e.g. International Assoc. for Social Science Information Services and Technology - IASSIST) representing the EDINA National Data Centre/Edinburgh University Data Library and to promote EDINA services (e.g. poster sessions). Presentations include:

*Counting Cows and Cabbages: web-based extraction, delivery and discovery of geo-referenced data* presented at IASSIST 2004, University of Wisconsin-Madison, USA 26 May 2004 (URL: <http://dpls.dacc.wisc.edu/iassist2004/wednesday.html#41>)  
*EDINA agcensus Service* presented at the Association of Librarians in Land-Based Colleges and Universities (ALLCU) conference, Harper Adams University College, July 16<sup>th</sup> 2004  
EDINA Exchange (as part of *Map and Data Place* presentation), National E-Science Centre, Edinburgh, 11 May 2004 (URL: <http://edina.ed.ac.uk/events/edinaexchange.shtml> )

*Introduction to EDINA agcensus* presented at e-Olympics Scotland, National E-Science Centre, Edinburgh, 23 September 2004

(URL: <http://www.rsc-ne-scotland.ac.uk/eolympics/scotland/>).

*The Data Library for Economics Researchers* presented at the London School of Economics Nereus seminar, 19 September 2005

(<http://www.lse.ac.uk/library/datalib/Nereus/>).

*Supporting Users of Electronic Data Resources in Research and Teaching* presented at UKeiG seminar and training event, University of Manchester, 5 July, 2006 (URL: <http://www.ukeig.org.uk/training/july06/eresources.html>).

- Currently participate in the Web Action group for IASSIST which evaluates and develops the IASSIST website (URL: <http://www.iassistdata.org>). I am a founder member and a web-editor for DISC-UK (the Data Information Specialist Committee - URL: <http://datalib.ed.ac.uk/discuk/>) and editor of the Statistics and Data section of Intute- Social Sciences – URL: <http://www.intute.ac.uk/>).
- Contribute articles to University publications including Edinburgh BITs, EDINA Newsline.
- Order, index and catalogue datasets; maintain Data Library web site (URL: <http://datalib.ed.ac.uk/>) and associated links and write data to a range of media (CD-Rom, networked server areas via ftp); locate and annotate relevant external data sources as part of current awareness service.
- Participate in Data Library internal projects as well as undertaking extensive costed consultancy to researchers. An example of the former was converting Scottish Household Survey documentation and data into DDI compliant XML. An example of the latter was collating data sources relevant to developing educational indicators for a Nuffield Review of 14-19 education for a researcher from the Centre for Educational Sociology which was published as an Annex to the final report.

**Nov. 1997 – Jan. 1999: Catalogue Editor**

**Edinburgh University Main Library  
George Square, Edinburgh**

- Created and edited UKMARC format catalogue records using AACR11, LC subject headings and Authority Control. Downloaded and edited records from OCLC, RLIN, CURL etc.

**Nov. 1996 – Oct. 1997: Information Assistant**

**Pre-School Voucher Centre  
Forres Street, Edinburgh**

- Supervised and trained Helpline staff, liaised with Local Authorities and the Scottish Office.

**July 1995 – Nov. 1996:**

- Traveled and worked in Australasia, South East Asia, Nepal and India.

**June 1993 – June 1995: Assistant Librarian**

**Daily Record,  
Anderston Quay, Glasgow**

- Attended to journalistic queries, indexed hardcopy and electronic images (Fastfoto), selected electronic material for use, searched and loaded in-house cuttings database (BRS) and online databases (FT-Profile), picture research.

**Sept. 1989 – June 1993: Various positions**

**National Library of Scotland,  
George IV Bridge, Edinburgh**

- **Cataloguer** - created and edited USMARC format catalogue records (*Bibliographic Services*)
- **Computing Officer** - duties included user support and documentation, main frame operations, installed and tested PC and mainframe software (*Computing Services*).
- **Curator** - duties included readers' services, indexed and catalogued incoming cartographic material (*Map Library*).

**Nov. 1987 – Sept. 1998: Library Assistant**

**Erskine Medical Library,  
George Square, Edinburgh**

- Duties included service desk and circulations, periodical and book processing, binding preparation.

**Sept. 1985 – Sept. 1987: Research Assistant**

**Heriot-Watt University,  
Chambers Street, Edinburgh**

- Duties included experimentation and bio-scientific research on nitrogen content of barley.

## **Education and training:**

**1981 – 1985:**

**Heriot-Watt University**

- BSc Hons degree in Biochemistry.

**1988 – 1989:**

**Strathclyde University**

- Postgraduate Diploma in Information and Library Studies.

## **Training**

- TESOL (Teaching English to Students of Other Languages) Certificate (1991).
- EUCS training courses in Access, HTML, UNIX, SPSS, Excel, MapInfo/ArcGIS
- Netskills XML/XSL

## Skills and interests:

### Professional:

- Assistant Treasurer of IASSIST (Europe - appointed 2006).
- Computing skills: WindowOS, UNIX, HTML authoring tools, SPSS, Access, Excel, ArcGIS, Photoshop, Adobe Acrobat suite, PowerPoint, MS Word, FTP/telnet protocol.

### Personal:

- Rugby (have represented Edinburgh District Union XV).
- Poetry (published in numerous literary magazines and anthologies).
- Reading, traveling and photography.
- Recently worked part-time for SACRO, Edinburgh (aiding rehabilitation of offenders into the community).
- Conversational French.

## Referees:

Mr Peter Burnhill  
Director and Head  
EDINA National Data Centre  
Main Library Building  
George Square  
EDINBURGH  
EH8 9LJ

[Ms Robin Rice  
Data Librarian  
Edinburgh University Data Library  
Main Library Building  
George Square  
EDINBURGH  
EH8 9LJ]

## Publications:

*Data sources which identify a sample of educational indicators.* Annex to a Nuffield Foundation report "The Scope, Feasibility and Organisation of a Nuffield Annual Review of 14-19 Education" by Professor David Raffe, University of Edinburgh, October 2002.

Geoinformatics, Jan/Feb 2005 – *EDINA agcensus: evolution of a web service*

IASSIST Quarterly, Spring 2004, Vol. 28 No. 1 – *Web-based extraction, delivery and discovery of geo-referenced data* - (<http://www.iassistdata.org/publications/iq/iqvol28.html>)

EDINA Newslines, October 2005, Volume 9, Issue 3 - *agcensus service breaks new ground*

Ariadne, Issue 44, July 2005 - *Supporting Local Data Users in the UK Academic Community* - (<http://www.ariadne.ac.uk/issue44/>)

ASSIGN, July 2005, Volume 22, No. 4 – *Looking for Data Directions in the UK?*

IASSIST Quarterly, Summer 2005, Vol. 29 No. 2 – *Local Data Support Landscape in the UK* (<http://www.iassistdata.org/publications/iq/iqvol29.html>)

## **JANE H. ROBERTS**

### **Curriculum Vitae**

78 Hayfield Road  
Oxford OX2 6TU

01865 559484  
07967 977305 (mobile)

jane.roberts@nuffield.ox.ac.uk

Nuffield College  
New Road

Oxford OX1 1NF  
01865 278638

### **Skills and Experience Summary**

**Extensive experience in all aspects of social, economic and health surveys, including**

- research design and planning
- instrument development
- fieldwork and interviewer management
- data preparation and cleaning
- data manipulation, analysis and interpretation, using SPSS
- preparation of reports and publications

**Establishment and management of a Data Library, including**

- searching data sources
- acquiring, mounting and supporting survey datasets
- subsetting and merging data files
- providing consultancy service on all aspects of data management and analysis for graduate students and faculty members

**Qualified practitioner of Traditional Chinese Medicine** – using acupuncture for treating the whole range of physical and emotional conditions, and preventive health care.

### **Employment History**

9/02- **Nuffield College, Oxford**  
Research Officer, Data Services Manager

3/96-4/97 **Dept. of Public Health and Primary Care, University of Oxford**  
Data management consultant (part-time with above)

3/87-9/02 **Computing and Research Support Unit, Social Studies Faculty, University of Oxford**

3/87-3/88 Research Officer, data management and analysis support

4/88-9/02 Research Officer, Data Services Manager

5/79-2/87 **Dept. of Applied Social Studies and Social Research,**

## **University of Oxford**

Research Officer, Fieldwork Manager

- Effects of Changes in Welfare Benefits Policies on Low-Income Households
- Study of Factors Influencing Career Choices of School Leavers

1/76-5/78 **Cancer Research Campaign / Dept. of Psychiatry,**

## **University of Oxford**

Research Assistant, Counsellor

- Evaluation of a Health Education Programme

## **Additional Data Management and Analysis Consultancies**

- Survey of University Teachers 1989 (with Prof. A. H. Halsey, Nuffield College)
- Employment in Britain Survey 1991 (with Prof. D. Gallie, Nuffield College)
- Trends in Chronic Sickness in the UK 1992 (University of Oxford Unit of Health Epidemiology)
- Housing Tenure Patterns in Ethnic Minorities in UK 1992 (with Prof. C. Peach, School of Geography, University of Oxford)
- Survey of Disabled Adults 1993 (Dept. of Applied Social Studies and Social Research, University of Oxford)
- Child Abuse Project 1993 (Dept. of Applied Social Studies and Social Research, University of Oxford)
- Survey of Refugees in Oxford 1996 (Ethnic Minorities Business Service)
- Incidence of Disability in the UK 1995 (UK Further Education Funding Council)
- Children in Care 1996 (with Dr A. Buchanan, Dept. of Applied Social Studies and Social Research, University of Oxford)
- Eastern European Eurobarometer module: Questionnaire design 1996 (Oxford Research International Ltd.)
- Surveys of Staff and Students in Oxford 1997 (University of Oxford Commission of Enquiry)
- International Students Survey 1998 (Dept. of External Relations, University of Oxford)
- British Students Survey 1998 (International Office, University of Oxford)
- Clinical Audit 1998 (College of Integrated Chinese Medicine)
- Origins and Destinations of Nuffield College Alumni 2001 (with Dr D. Butler, Nuffield College)
- Survey of Professors of Sociology in Britain 2001 (with Prof. A. H. Halsey, Nuffield College)
- The Employability and Occupational Attainment of Humanities and Social Science Graduates 2001 (British Academy, with Dr SY. Cheung)
- Ethnic Minorities in the Labour Market: Research Overview 2001 (UK Cabinet Office, with Prof. A. Heath)
- Ethnic Penalties in the Labour Market; Employers and Discrimination 2005 (DWP, with Prof. A. Heath and Dr SY. Cheung)

## **Education and Qualifications**

- 1970-73 **University of Reading** - BA(Hons) Politics class 2ii
- 1978 **Middlesex Polytechnic** - Diploma in Counselling
- 1990 **University of Oxford** - MA(status)
- 1994-98 **College of Integrated Chinese Medicine** - Licentiate in Acupuncture
- 2002 Elected to **Fellowship of the Royal Statistical Society**

### Publications

Noble M., Smith G., Payne J. and Roberts J. (1988) *The Other Oxford. From Supplementary Benefit to Income Support; Who Gains and Who Loses?* University of Oxford Department of Social and Administrative Studies

Noble M., Smith G., Payne J. and Roberts J. (1989) *The Other Oxford. Low Income Households in Oxford*, University of Oxford Department of Social and Administrative Studies

Heath A., Mills C. and Roberts J. (1992) Towards Meritocracy? Recent Evidence on an Old Problem. In *Social Research and Social Reform* (eds Crouch C. and Heath A.) Oxford: Clarendon Press

Heath A., McMahon D. and Roberts J. (1994) *Ethnic Minorities in the Labour Market* London: JUSST Working Paper Series

Heath A., McMahon D. and Roberts J. (2000) *Ethnic Differences in the Labour Market: A Comparison of the SARs and LFS* Royal Statistical Society, Series A 163

### Personal Details

**Nationality** British    **Date of Birth** 23<sup>rd</sup> January 1952    **Family** 2 adult children

#### **Interests**

- Complementary medicine – Acupuncture practitioner
- Greece - travelled extensively, especially in remote areas
- Editing text for friends, colleagues and OUP
- Sailing, skiing
- Woodcarving

**Tracy A. Gibbs  
(aka Harry)**

*Garden Flat, 22 Lupin Road, Southampton, SO16 3LB  
Tel: 023 80486907*

**Profile**

I am a dedicated and conscientious professionally qualified librarian, with experience of a broad range of library work and an aptitude for working on my own initiative. Since joining the University of Southampton in 2004, I have focussed on developing my knowledge and understanding in the area of data provision.

**Career History**

I am currently employed in the Hartley Library at the University of Southampton.

**2004 – date**

**Academic Liaison Librarian (Social Sciences)**

I am responsible for supporting and developing links with the Divisions of Economics, Politics and Social Statistics, with special responsibility for the provision of social science data. With respect to data, my role includes maintaining and developing the University's Data Resources Web pages and dealing with user enquires regarding a range of social sciences data. In 2005, I became a member of I-ASSIST (International Association for Social Science Information Service and Technology), an organisation of professionals working to provide data services in support of research and teaching. More recently, I have become a member of DISC-UK .

I was employed in the Study Centre at Yeovil College between September 1999 and September 2004. During that time I undertook four roles, including two professional posts.

**2003 – 2004**

**Help Desk Team Leader**

I supervised a successful and cohesive team of three library assistants and oversaw the day-to-day running of the combined issues and IT help desk. I was responsible for many aspects of staff management and for the maintenance of the Study Centre Web pages.

**2001 – 2003**

**Subject Librarian** (Science, Maths, Computing, Sport & Leisure)

As Subject Librarian my tasks included academic liaison, user education, producing promotional materials, holding a resources budget, classification, preparing resource packs, enquiry duty, attending meetings, research and using electronic resources including e-journals, electronic indexes and online reference sources.

**2000 – 2001**

**Library Services Assistant**

My main tasks in this role were in circulations and administration, which incorporated responsibility for periodicals, involvement in the acquisitions process and the co-ordination of distance learning courses. I gained a good working knowledge of the library system, OLIB, and also gained experience of a second system, Heritage.

**1999 – 2000**

**Library Services Assistant – IT**

This role was based at the IT help desk where I was jointly responsible for providing software and basic hardware support. It proved a good opportunity to develop my IT skills, in particular in using Microsoft Word, Excel, PowerPoint and Publisher.

## Previous Experience

- My work experience began with part-time waitressing whilst at school, where I displayed a natural commitment to good customer service.
- As a pupil and student I undertook much child-minding work with children of all ages and spent my gap year teaching in a nursery school in West Africa.
- On graduating, I took a number of full-time jobs as a waitress, shop assistant, production-line worker and carer. Despite, undertaking many menial tasks I always demonstrated a professional attitude towards my work and a commitment to my employers.
- I also undertook further voluntary work in West Africa and for a development charity called Tree Aid, in Bristol. During my time at Tree Aid, I developed an existing interest in libraries by reorganising their small collection of resources.

## Education and Qualifications

2001 – 2003	University of Bristol MSc in Information and Library Management (with Commendation)
1993 – 1996	University of Glamorgan BA (Hons) Environment and Social Values (First Class)
1992	Yeovil College A Levels in Biology, Politics and Chemistry (B,B,C)
1990	Park School, Yeovil College 9 GCSEs (including Mathematics and English)

## Training and Development

I am always keen to accept offers of further training where they arise. Examples of recent training and development events include those covering federated searching, the CLA digitisation licence for HE and the provision of European data.

## Professional Activities

I am a member of the Chartered Institute of Library and Information Professionals (CILIP) and have recently submitted my application for Chartered membership.

## Published Works

'Web design for librarians: an FE student perspective', *Library and Information Update*, April 2004, 28-29.

'Student Preferences in the design of library Web pages: a UK, Further Education Perspective', *New Review of Children's Literature and Librarianship*, 11 (1), 33-45. (April 2005)

## Personal Details

Nationality:	British
Date of birth:	5 <sup>th</sup> November 1973
Marital status:	Single

## Interests

I enjoy walking, cooking, watching sport and film-going. Recently I have developed an interest in open-water swimming – the training keeps me fit and I am proud to have recently completed my first race around Brownsea Island in Poole Harbour.

# JISC Capital Programme – Repositories and Preservation Targeted Curriculum Vitae

**Name:** Wendy White

**Qualifications:** BA, MA, DipLib, MCILIP, ILTM

**Current Employer:** University of Southampton

**Current Post Title:** Faculty Liaison Librarian, Law Arts and Social Sciences

## **Experience and Expertise of Relevance to Project:**

- Manager of the University of Southampton's Institutional Repository e-Prints Soton. The repository is being used for the RAE and this role therefore includes membership of the University's main RAE Planning group chaired by the DVC for Research. There is a responsibility for strategy and policy development, liaison with researchers in all disciplines across the University and management of a team of staff including the metadata editor and software developer.
- Faculty Liaison Librarian, Law Arts and Social Sciences responsible for developing and delivering library support for researchers across the Faculty managing a team of academic liaison librarians including the Data Librarian.
- Member of the Digitisation Services Management Group chaired by the Librarian which oversees projects including the JISC funded 18<sup>th</sup> Century British Parliamentary Papers project to deliver a fully-searchable XML database of up to 900,000 pages of British Official Publications
- Have delivered a series of training courses across the UK for Relay Europe on behalf of the European Commission on sources of European data and statistics.
- Resource evaluator with a period as Section Editor of the statistics and demography section of the Social Science Information Gateway (SOSIG)

## **Publications of Relevance to the Project:**

Hey, Jessie M.N. White, Wendy, Simpson, Pauline, Brown, Mark and Lucas, Natasha (2006) [Fast flows the stream: tackling the workflow challenge with the University of Southampton Research Repository](#), At, *Open Scholarship 2006: New Challenges for Open Access Repositories, Glasgow, UK, 18-20 October 2006*.

Available from: <http://eprints.soton.ac.uk/41913/> [Accessed 20 November 2006]

Harwood, N. and White, W. (2003) *How to trace EC statistics*, 4<sup>th</sup> ed. Manchester: European Information Association.

Petit, S. Cochrane, J. and White, W. [Exploring Planet SOSIG: Law, Statistics and Demography](#). *Ariadne* 17

Available from: <http://www.ariadne.ac.uk/issue17/planet-sosig/intro.html> [Accessed 20 November 2006]

# Luis Martínez Uribe

Email: l.martinez@lse.ac.uk  
Nationality: Spanish  
Date of Birth: 15<sup>th</sup> August 1977

## Education

---

Oct'05 – Sep'07 **MSc Analysis, Design and Management of Information Systems**  
**Department of Management, London School of Economics, United Kingdom.**

Part-time MSc whilst working full time as the Data Librarian at the LSE.

Courses covered:

- IS Management
- Development of IS
- e-Government
- Security in IS for Organisations

Oct'95 – Jun'01 **BSc in Mathematics specialized in Statistics and Computer Science.**

**Department of Mathematics, Universidad Complutense de Madrid, Spain.**

One year of Erasmus Program at the University of Strathclyde (Computer Science and Statistics Departments)2000/01. Glasgow

Courses included:

- Java Programming
- Databases
- Time Series Analysis
- Regression Analysis

## Employment History

---

Sep '01 – Date **Data Librarian.**

**Information Services Department, LSE Library, London.**

Responsible of the provision of data library services to the LSE research and teaching community and the Library's general interest in electronic statistics resources. Main responsibilities include :

- Maintenance of the expanding Data Library collection
- Development and maintenance of Library web pages and collaboration with IS colleagues in the integration of data resources on the electronic library
- Provision of support to postgraduate students and staff in finding and using electronic quantitative resources
- Liason and communication both internally ,with individuals and departments, and externally with data providers, national data centres and data specialists.
- Management of the data resources budget with over £ 150K

The role also has given me the opportunity to participate in various activities such as:

- The Library/ITS DataLab project to develop a room for teaching with datasets
- Member of national groups such as CURL/SCONUL Task Force on eResearch or Data Information Specialist Committee DISC-UK

Sep'02 – Jun03 **Research Assistant.**

**Interdisciplinary Institute of Mangement, London School of Economics, London.**

Data management responsibilities as part of a research project involving the manipulation and merging of datasets.

## Other Skills

---

### ICT Experience

- Operating Systems: DOS, Windows (95, 98, 2000, ME, NT, XP), UNIX, LINUX
- Programming languages: Basic, Pascal, Fortran, Visual Basic, Perl, C++, Java
- Mathematical Software and Statistical Packages: Mathematica, Maple, Matlab, SPSS, SAS, Stata, S-Plus
- Web technologies: HTML, XML, XSLT, JavaScript, Cgi-Perl, PHP, ASP

### Languages

- **Spanish** Mother tongue
- **English** Fluent (written/ spoken)
- **Italian** Intermediate(spoken).

## Publications

---

- with Stuart Macdonald "The Local Data Support Landscape in the UK" *IASSIST Quarterly Volume 29 No 2 (2005)*
- with Stuart Macdonald "Data Resources in the UK" *ASSIGNATION Volume 22 No 4 (July 2005)*
- with Stuart Macdonald "Supporting Local Data Users in UK Academic Community" *Ariadne Issue 45 (2005)*

## Invited Presentations

---

- "Supporting users of electronic data resources in research and teaching" UKeIG Workshop, University of Manchester July 2006
- "Data sans Frontiers" 2<sup>nd</sup> Nereus Data Workshop, Oxford University May 2006
- "Looking for data directions?- Ask the Data Librarian" IASSIST/IFDO Conference 2005, Edinburgh 26 May
- "Using ESDS International Databanks" MIMAS Open Forum 2005, Manchester 4th July

## Short Courses

---

- Sep'06 **Negotiation Skills** – One day course at LSE.
- Sep'05 **Introduction to XML Development** – Three day course by gbdirect.
- Jan'03 **Databases on the World Wide Web** – One day course by Netskills.
- Jan' 02 **Spatial Query and Analysis using GIS** – Three day course at LSE.
- Nov'01 **Market knowledge and Datastream Advance** – Two days course by Thomson Financial.

## Membership of Groups and Associations

---

### Groups

- Member of the CURL-SCONUL Joint e-Research Task Force
- LSE Data Advisory Group
- LSE Library Meridian Implementation Group

### Associations

- Member of IASSIST (International Association of Social Science and Information Service Technology) <http://www.iassistdata.org/>
- Member and cofounder of DISC-UK (Data Information Specialist Committee). <http://datalib.ed.ac.uk/discuk/>

## CURRICULUM VITAE - TANVI DESAI

### RELEVANT EMPLOYMENT HISTORY:

**June 2006-:** **Study and conference on European Labour Market Analysis using Firm-level Panel Data and linked Employer-Employee Data**

*Duties:* Coordinator for Task 4: Development of Linked and Panel Datasets for European Labour Market and Social Policy Analysis; Negotiation with European National Statistical Institutes for Data Access; Management of a secure remote access system; A survey of legal rules governing access to confidential business data in European Member States; Analysis of resource availability and suitability for international comparative research.

**April 2005-Jan 2006: ERSC Review of International Data Resources and Needs**

To assess and report on the nature of the resources required to undertake international and comparative research, in terms of the resources currently available, the quality of such resources and their accessibility; To make recommendations for strategies to improve resources for international research.

**July 2004-July 2005: Data Consultant, LoWER Network**

*Duties:* Advising network on possible data related deliverables; Advising on strategies to improve knowledge sharing among Network members; Support for locating and accessing data.

**Nov 2000-Feb 2005: Data Manager, Pay Inequalities and Economic Performance project**

*Duties:* Data support for international research team; Negotiations for data access with European National Statistical Institutes and Eurostat; Development of secure remote data access system, leading to advising Eurostat and the Australian and New Zealand Statistical Institutes on their remote access policy; Management of confidential data access.

**Jan 1999- :** **Data Manager, London School of Economics Research Laboratory**

*Duties:* Facilitating data use for all research staff ; Managing electronic library of primary and secondary data holdings; data and software support, including development of documentation; Planning RLab data strategy and budget management; Representing academic data interests; for more information see <http://rlab.lse.ac.uk/data/>

**Aug 1998-Dec 1999: Research Assistant for Centre for Economic Performance, London School of Economics.**

*Duties:* Creating usable, coherent data sets from primary data collected; producing tables from national and international data sources; Assistant editor on State of Working Britain;

**Aug 1997-Feb 1998: Research Associate for Institute for Research into International Competitiveness (IRIC), Curtin Business School, Curtin University, Perth, Western Australia.**

**Duties:** Project Management; conducting interviews and focus groups ( including interviews of college chancellors and students in Singapore and Malaysia); questionnaire design; data collection, formatting and analysis of data including development of coding systems.

***Project managed was a Finalist in the Economic Development category of the Premier's Awards for Public Sector Management***  
(from among 133 applications for this category)

#### **AFFILIATIONS:**

IASSIST [International Association of Social Science Information Service and Technology]  
DISC-UK [Data Information Specialists Committee – UK]

#### **PUBLICATIONS & CONFERENCE PAPERS:**

Desai, T. (2003) "Remote Access to Confidential Data" presented at the International Conference on the Comparative Analysis of Enterprise Microdata, September 15-16, 2003, Cass Business School, London, UK.

Desai, T. (2003) "Providing Remote Access to Data: The Academic Perspective" presented at the Joint UNECE/Eurostat work session on Statistical Data Confidentiality, April 7-9, 2003, Eurostat, Luxembourg.

Desai, T, et al (1999) "Gender and the Labour Market" in Gregg, P. and Wadsworth, J. (ed.) The State of Working Britain, Manchester University Press

#### **LANGUAGES:**

French (3S 2W)  
Italian (2S 1W)

#### **PERSONAL DETAILS:**

**Date and place of birth:** 12<sup>th</sup> April 1972, London.

**Contact details:** Centre for Economic Performance  
London School of Economics  
Houghton Street  
London WC2A 2AE

+44(0)207.955-7740

[t.desai@lse.ac.uk](mailto:t.desai@lse.ac.uk)

## **ANN GERKEN GREEN**

61 Loomis Place  
New Haven, Connecticut USA  
06511-2222  
green.ann@gmail.com www.dlifecycle.net

Ann Green is an independent research consultant focusing upon the digital life cycle of scholarly resources, including their creation, delivery, management, long-term stewardship, and preservation. She has an extensive background in digital archiving and user-driven support services in the social sciences. Ann has participated in the development and promotion of standards for social science statistical metadata and is an active participant in the development of international social science data services and initiatives. She is the former director of the Social Science Research Services & Statistical Laboratory (Statlab) at Yale University where she coordinated social science research and instructional technologies, facilities, and services for over 15 years. Ann participated in campus wide information technology planning as well as initiatives in digital library infrastructure and services at Yale. She is the immediate past President of IASSIST and former Chair of the Executive Council of the Inter-university Consortium for Political and Social Research (ICPSR).

### **Employment**

September 2005-date: Self employed Consultant, Digital Life Cycle Research & Consulting  
Engaged in a variety of contracts providing consultation, research, and writing services for digital initiatives at academic and research institutions in the US and UK. Current work includes: research and writing projects on the partnerships among social science researchers, institutional repositories, and data archives; program development and evaluation of digital archives; consultation on the creation and curation of digital collections. Active in conference planning and the administration of international data organizations.

2004-2005: Senior Research Analyst, Integrated Access Council, Yale University Library  
Initiated and supported activities that enhanced the understanding of end-user needs for digital library infrastructure and services. Specific projects included investigating institutional repository activities at several academic research institutions from the perspective of actual and potential faculty requirements. Produced high level recommendations and assessment documents with pros and cons of institutional repository development for delivery to numerous policy and resource allocation units on campus. Developed and supported the communication plans about digital library initiatives on campus by raising awareness about digital library issues, and developing internal and external collaborative relationships supporting digital library development. Participated in numerous Library and campus wide committees related to digital preservation policies, metadata, portal opportunities, and user needs assessment.

1996-2005: Director, Social Science Research Services & Statistical Laboratory (Statlab), Yale University  
Managed, developed and aggressively pursued innovative solutions in research and instructional support for social science computing. Administered all Statlab services including consulting and instruction to faculty, students, and researchers on statistical computing and information resources. Collaborated with libraries, professional schools, and Information Technology Service units in developing state of the art technical solutions in the areas of data and content management as well as Internet based delivery and analysis of social science resources. Represented the interests of social science researchers and the development of digital archives in international professional and standards organizations.

1989-1996: Data Consultant and Acting Director, Social Science Statistical Laboratory, Yale University

1985-1988: Data Archivist, UC Data Program, Survey Research Center, University of California at Berkeley

1982-1985: Data Archivist/Associate Librarian, Cornell Institute for Social and Economic Research

1976-1981: Reference Librarian, Olin Graduate Library, Cornell University Libraries

### **Professional Activities**

Executive Council of the Inter-university Consortium for Political and Social Research, Member 2000-2004; Chair, 2002-2004.

Data Documentation Initiative Committee and Alliance, Member 1995 to 2005. International committee developing metadata standards sponsored by the Inter-University Consortium for Political and Social Research (ICPSR). Co-developer of study description standards and applications manual (Tag Library). DDI Alliance Steering Committee, 2003-2005.

International Association for Social Science Information Service and Technology (IASSIST), President 2001-2005, Vice President 1997-2001, Member since 1982.

Roper Center for Public Opinion, Board of Directors, 2003 to date.

### **Advisory Committees**

UK Economic and Social Research Council, Economic and Social Data Services

Roper Center for Public Opinion

Murray Research Center, Radcliffe Institute for Advanced Study

National Historical Geographic Information Service, University of Minnesota

Digital Governments Research Center, Columbia University and ISI

Center for Arts and Cultural Policy Studies, Princeton University

European Union, Fifth framework Research and Development program: MetaNet (Working Groups 1 and 2)

Canadian Foundation for Innovation

United Nations, Statistics Division, data products review committees

### **Education**

MA University of Michigan, Ann Arbor, MI  
Master of Arts in Library and Information Science, 1975

BA Colgate University, Hamilton, NY  
Bachelor of Arts, cum laude, 1974

### **Grants**

The Andrew W. Mellon Foundation. Supporting Economic Development Research: Economic Growth Center Digital Library. Co-PI with Professor Christopher Udry and Sandra Peterson, Director Social Science Libraries and Information Services. February 2003 through February 2005.

Commission on Preservation and Access. Preserving the Whole: A Two-Track Approach to Rescuing Data and Metadata. Co-PI with JoAnne Dionne. July 1, 1996 through September 30, 1997.

### **Publications**

Green, Ann and Myron Gutmann. "Building Partnerships Among Social Science Researchers, Institution-based Repositories and Domain Specific Data Archives." Forthcoming: OCLC systems & services: international digital library perspectives. v. 23, no. 1. Ed: Bradford Lee Eden. Bradford, England: Emerald Group Publishing, c2007. Preprint: <http://hdl.handle.net/2027.42/41214>

Linden, Julie and Ann Green. (2006). "Don't leave the data in the dark: Issues in Digitizing Print Statistical Publications." D-Lib Magazine. Vol. 12, No. 1. <http://www.dlib.org/dlib/january06/linden/01linden.html>

Green, Ann. Review of Digital Repositories. (2005) Report to the Integrated Access Council, Yale University Library. [http://www.library.yale.edu/iac/documents/DR\\_Review\\_final\\_27Sept05.pdf](http://www.library.yale.edu/iac/documents/DR_Review_final_27Sept05.pdf)

Green, Ann, Sandra K. Peterson, and Julie Linden. (2005) Supporting Economic Development Research: A Collaborative Project to Create Access to Statistical Sources Not Born Digital. A Report to the Andrew W. Mellon Foundation. New Haven: Yale University.  
[http://ssrs.yale.edu/egcdl/Yale\\_EGCDL\\_report\\_0505.pdf](http://ssrs.yale.edu/egcdl/Yale_EGCDL_report_0505.pdf)

Green, Ann, and Jean-Pierre Kent. (2002) "The Metadata Life Cycle," MetaNet Work Package 1: Methodology and Tools. In: MetaNet: A Network of Excellence for Harmonising and synthesizing the development of statistical metadata. The MetaNet Project.  
[http://www.epros.ed.ac.uk/metanet/deliverables/D4/IST\\_1999\\_29093\\_D4.pdf](http://www.epros.ed.ac.uk/metanet/deliverables/D4/IST_1999_29093_D4.pdf)

"The challenges of statistical digital libraries" (2002) In: Proceedings of the 2nd ACM/IEEE-CS Joint Conference on Digital Libraries. NY: ACM Press. With Cathryn Dippo, Patricia Cruse, and Carol Hert.  
<http://doi.acm.org/10.1145/544220.544311>

"One size does not fit all (but a good model might): the academic media & technology pyramid." (2002) In: Proceedings of the 30th annual ACM SIGUCCS conference on User services. NY: ACM Press. With Philip E. Long, Jonathan Lizee, Edward Kairiss, and Charles Powell.  
<http://doi.acm.org/10.1145/588646.588662>

Capps, Cavan, Ann Green and Mark Wallace. (1999) "The Vision of Integrated Access to Statistics: the Data Web," Of Significance... A Topical Journal of the Association of Public Data Users. 1:2, p. 42-47.

Green, Ann, JoAnn Dionne and Martin Dennis. (1999) Preserving the Whole: A Two-Track Approach to Rescuing Social Science Data and Metadata. Washington: Council on Library and Information Resources. (CLIR Publications 83) <http://www.clir.org/pubs/reports/pub83/contents.html>

Green, Donald Philip, and Ann E. Gerken. (1988). "Self Interest and Public Opinion toward Smoking Restrictions and Cigarette Taxes", Public Opinion Quarterly. 53:1-16.

Gerken, Ann E. (1988) "What is a Data Archive and What Should the Information Specialist Know about Managing Locally Maintained Numeric Data Files?" Database Magazine.

### **Professional Presentations**

iPRES: International Conference on Preservation of Digital Objects. October 2006. "Building Partnerships among Social Science Researchers, Institution-based Repositories, and Domain Specific Data Archives." With Myron Gutmann, ICPSR.

<[http://dspace.library.cornell.edu/bitstream/1813/3692/4/Ann\\_Green-ICPSR.pdf](http://dspace.library.cornell.edu/bitstream/1813/3692/4/Ann_Green-ICPSR.pdf)>

Digital Library Federation, Fall Forum, October 2005. "Economic Growth Center Digital Library: Creating Access to Statistical Sources Not Born Digital." With Julie Linden, Yale University.

<<http://www.diglib.org/forums/fall2004/greenlindenpeterson1004.htm>>

Annual Conference of IASSIST ([www.iassistdata.org](http://www.iassistdata.org)):

- May 2006. "Institutional Repositories and Social Science Data: Supporting the Data Life Cycle."
- May 2004, "Economic Growth Center Digital Library: Creating Access to Statistical Sources Not Born Digital."
- June 2002, "StatCat: Building a Statistical Data Finder."
- May 1997, "Preserving the Whole: A Two Tiered Approach to Rescuing Data and Documentation."

- May 1996, "Study-level elements in the Data Documentation DTD in SGML."
- May 1996, Workshop: "Advanced Use of the WWW as a Means of Delivering Data."
- May 1995, "Improving Access to International Statistics: A Local Scenario."
- May 1994, "Introduction to the Management of Data Archives and Data Libraries."

Yale University Libraries Standing Committee on Professional Activities:

- May 2004. "Economic Growth Center Digital Library." With Sandra Peterson and Julie Linden.
- May 2001. "Preserving Social Science Data and Metadata."

Joint Conference on Digital Libraries, Portland OR, June 2002. The Challenges of Statistical Digital Libraries. Panelist.

Simmons College, Graduate School of Library and Information Science, Continuing Education course, April 2002. "Finding, evaluating, and using numeric data." With Julie Linden, Yale University.

ACRL Information Technology Interest Group, Providence, RI, June 2001. "StatCat: Building a Statistical Data Finder." With Julie Linden, Yale University.

Annual Conference of the Association of Public Data Users (APDU):

- October 2001, "Helping to Build the Data Web."
- October 1999, "Why the Data Documentation Initiative is so Important."
- October 1996, "The Data Documentation Initiative: Building an International Standard."
- October 1987, "CD ROM and the Data Archive."
- October 1985, "Data Distribution Activities of Data Archives and Data Librarians."

FASTER Workshop, Voorburg, The Netherlands, April 2000. "The Data Documentation Initiative: Current Status, Future Plans, and Structure." With Peter Jofitis, ICPSR, University of Michigan.

Second Annual Connecticut State Data Conference, Hartford CT, June 2000. "Building the Data Web: the Importance of Intelligent Data."

University of California, Digital Library Initiative, Workshop on the DDI, Oakland CA, September 2000. "Data Documentation Initiative and Aggregate Data." With Wendy Thomas, University of Minnesota.

United Nations/ECE Statistical Division, Workshop on Statistical Metadata (METIS). Washington DC, November 2000. "What Metadata When? Ideas from the User Perspective."

Social Science Research Council, New York, NY, November 1999. Workshop: Whither Social Science Data and Data Infrastructure. "Perspectives on Metadata and Preservation."

United Nations, Department of Economic and Social Affairs, Expert Group Meeting on International Statistical Database, May 1997. "Archival Concerns of Digital Preservation, Data Integrity, and Maintaining Access over the Long Term."

ICPSR Biannual Meeting:

- October 1997, "Where in the World is the Codebook? Comments on Social Science data documentation production and access."
- October 1991, "Machine-readable Documentation: Formats and Standards."

## Appendix for the DISC-UK Datashare proposal.

### JISC-funded Repository Projects (from the Data Cluster):

#### **CLADDIER\***

Project Description - (URL: <http://claddier.badc.ac.uk/trac>)

This project was funded by the JISC for support under the Call for Projects in Digital Repositories (March 05)

The University of Southampton and CCLRC institutional repositories are exemplars of institutional repositories in the UK. The British Atmospheric Data Centre (BADc), likewise, is an exemplar of a discipline based data archive. This two-year project will build and deploy a demonstration system linking publications held in the institutional repositories with data holdings in the BADc. As well as the demonstration system, a workshop disseminating information about the project will be held for the environmental science community, and four significant reports will be produced:

1. User Experience of the CLADDIER System (written by active environmental scientists based on their experiences)
2. Identifier Migration Issues for Repositories
3. **Recommendations for data/publication linkage** (based on lessons learned, and a review of the literature)
4. **Methodologies and Practices for Data Publication.**

The CLADDIER system will be a step on the road to a situation where (in this case, environmental) scientists will be able to move seamlessly from information discovery (location), through acquisition to deposition of new material, with all the digital objects correctly identified and cited. The lessons learned will be of applicability for the relationships between other discipline based repositories and institutional repositories.

*\*Citation, Location, And Deposition in Discipline and Institutional Repositories*

## **GRADE\***

Project Description - (URL: <http://edina.ac.uk/projects/grade/>)

This project was funded by the JISC for support under the Call for Projects in Digital Repositories (March 05)

GRADE will investigate and report on the **technical and cultural issues around the reuse of geospatial data within the JISC IE** in the context of media-centric, informal and institutional repositories.

The aim of GRADE is to lay the foundations for a sustainable infrastructure (both cultural and technical) that underwrites the communities' substantial and ongoing investment in the utilisation of geospatial resources within the learning, teaching and research environments of UK academia.

The project will consist of a consortia led by EDINA, with primary partners:

- the Arts and Humanities Research Council (AHRC) Research Centre for Studies in Intellectual Property and Technology Law
- the National Oceanography Centre, Southampton

In addition there are a number of Associate Partners.

Due to limitations of (JISC-supported) repository software packages (DSpace, Fedora, e-Prints) – GRADE team are investigating GeoNetwork (which deals specifically with the complexities of geospatial data metadata – although not repository software per se)

*\*Geospatial Repository for Academic Deposit and Extraction*

## StORe\*

Project Description - (URL: <http://jiscstore.jot.com/ProjectDescription>)

This project was funded by the JISC for support under the Call for Projects in Digital Repositories (March 05)

This project seeks to **address the area of interactions between output repositories of research publications and source repositories of primary research data**. Through user surveys the functionality required by researchers in both types of repository will be determined, with the aim of identifying options for increasing the value of using primary data in source repositories as well as at the point where researchers submit to or download papers from output repositories.

This two year project is multidiscipline in scope, embracing seven scientific domains: archaeology, astronomy, biochemistry, biosciences, chemistry, physics and social sciences. With the aid of the survey results, general principles for middleware development to link source and output repositories together will be formulated, and a business analysis will be performed. Consequent to this analysis a pilot demonstrator will be developed in one of the domain areas.

This project is international in context, with UK and US institutional partners. The lead partner is the University of Edinburgh. Investigation of the astronomy domain is being led from John Hopkins University, Baltimore. The business analysis and pilot middleware build are being undertaken by the UK Data Archive at the University of Essex.

- University of Edinburgh (lead)
- University of York, representing *White Rose Partnership*
- University of Birmingham
- London School of Economics
- University of Manchester
- Imperial College London
- University College London
- UK Data Archive (UKDA)
- Johns Hopkins University

\* *Source to Output Repositories*

## **R4L\***

Project Description – (URL: <http://r4l.eprints.org/about.html>)

This project was funded by the JISC for support under the Call for Projects in Digital Repositories (March 05)

This project, based at the University of Southampton in partnership with the Association of Learned and Professional Society Publishers), will address the area of interactions between repositories of primary research data, the laboratory environment in which they operate and repositories of research publications into which they ultimately feed (through documented interpretation and analysis of the results and in explicit linking and citation of the data sets). It will develop prototype services and tools to address the issues of working with, disseminating and reporting on experimental data.

*\*Research for the Laboratory*

## SPECTRa\*

Project Description - (URL: <http://www.lib.cam.ac.uk/spectra/>)

This project was funded by the JISC for support under the Call for Projects in Digital Repositories (March 05)

The SPECTRa project will address the provision of Open Access to primary research data in experimental chemistry through the use of institutional repositories. It will build on experience already acquired in the Chemistry departments at Cambridge University and Imperial College London, **linking these with the development of the same universities' institutional strategies for library-managed OAI-compliant repositories**. The project will initially study the needs of researchers and scope their data-handling needs. It will use the findings of this study to develop automated tools so that high-volume data can be identified, extracted, and ingested to repositories, where it will be preserved and accessible for use to support research and teaching. The study's outcomes will be analysed to provide generic relevance to the chemistry research community, and its methodology will be formulated to provide generic guidance for similar studies in other sciences. The tools subsequently developed will be available as Open Source code designed for use with the DSpace repository platform.

\* *Submission, Preservation and Exposure of Chemistry Teaching and Research Data*

## Related projects / services

### **eBank**

eBank UK is a JISC-funded project which is a part of the Semantic Grid Programme. The project is being led by UKOLN in partnership with the Combechem project at the University of Southampton and the PSIGate Physical Sciences Information Gateway at the University of Manchester. This initiative is set in the context of the JISC Information Environment, JISC funded development supporting end-users to discover, access, use and publish resources as part of their teaching, learning and research activities. eBank UK brings together chemists, digital librarians and computer scientists in an interdisciplinary collaboration which explores the potential for integrating research datasets into digital libraries by using common technologies such as the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH).

In phase 1 (2003-2004) an OAI-PMH repository was created at the University of Southampton using a modified version of the eprints.org software. A number of datasets from the chemistry sub-discipline of crystallography were deposited in the archive, which offers a local browse and search interface, tailored to crystallography. A metadata schema, based on Dublin Core, was then designed to describe the datasets. Metadata harvesting was carried out and a demonstrator was built to show the potential for linking the datasets to publications.

Phase 2 of the project has now been funded to research the use of complex digital object descriptions and their application in describing scientific data. Another strand of work will explore the pedagogical potential of linking research datasets with e-learning..

### **DART\*** - (URL: <http://dart.edu.au/about/>)

The (Australian) DART project is a proof-of-concept project to develop tools to support the new collaborative research infrastructure of the future. **The project aims to enable researchers and reviewers to access original and analysed data, collaborate around the creation of research outputs, stored publications, plus add content, annotations and notes.** It will also look at the collection of large datasets, including the remote control and automated data collection.

Partners:        Monash University, Melbourne (lead partner)  
                      University of Queensland, Brisbane  
                      James Cook University, Townsville

\*Dataset Acquisition Accessibility & Annotation e-Research Technologies

## **Related Repository Projects**

**IRIScotland\*** – (URL: <http://www.iriscotland.lib.ed.ac.uk/>)

The purpose of the project is – with a view to supporting the Scottish research agenda at both institutional and national (Scottish) level – to:

1. Explore - in collaboration with university senior managers and researchers - ways of bringing about cultural and organisational change by developing institutional research publication policies, procedures and mechanisms, as well as workflows to assist individual researchers, conducive to the promotion of self-archiving in institutional open access repositories;
2. Develop a broad framework for a distributed institutional repository infrastructure for Scottish research and experiment with both a collective hosting repository, in particular for smaller institutions that may not wish to set up their own institutional repositories, and a cross-repository search facility capable of dealing with a wide range of research and research-related digital objects;
3. Identify what can be more effectively done centrally – and whether this should be done at a national Scottish level or a national UK-wide level – or locally at institutional level, taking account of relevant international developments to ensure that the Scottish infrastructure is globally interoperable.

*\* Institutional Repository Infrastructure for Scotland*

**Institutional Repositories Search Service** – (URL: <http://irs.ukoln.ac.uk/>)

**Partners:** University of Manchester - Intute/MIMAS (lead)  
University of Bath - UKOLN  
University of Nottingham - SHERPA

Content deposited in institutional repositories is growing, however there is no comprehensive and easy way to search and retrieve this content. In partnership with UKOLN and SHERPA, Intute has been commissioned by JISC to develop a repository search infrastructure. This development will build on the **ePrints UK** project, and aims to facilitate the discovery, access and retrieval of material.

This project will run for three years, and the initial phase will deliver a search facility for UK institutional repositories.

**Non-UK Data Repositories** (source: *JISC Digital repository Wiki*)

**CURATOR\*** - (URL: [http://mitizane.ll.chiba-u.jp/curator/index\\_e.html](http://mitizane.ll.chiba-u.jp/curator/index_e.html))

(c. 2004) CURATOR (Chiba University's Repository for Access to Outcomes from Research) captures, preserves and makes publicly available intellectual digital materials from research activities on Chiba University campuses, including peer-reviewed articles, theses, preprints, **statistical and experimental data**, course materials and softwares.

CURATOR is intended to function as the portal for the outcomes from Chiba University's research activities.

*\*Chiba University's Repository for Access to Outcomes from Research*

**Demetrius** – (URL: <http://dspace.anu.edu.au>)

(Aug. 2005) **Demetrius** is the name of the Australian National University's Institutional Repository. It provides a stable, secure and accessible repository for the long-term preservation of digital resources created by members of the ANU community.

Demetrius preserves and distributes digital collections including; research reports, papers, projects, data sets, image collections, and more.

**IAI-DIS** – (URL: <http://www.iai.int/>)

A Brazilian data repository based around the scientific data produced within the Inter-American Institute for Global Change Research programs

The DIS is an Internet-based data and information system with the following objectives in focus: 1) Dissemination of data created within the IAI scientific projects by using a metadata creation and management process; 2) Data discovery; and 3) Contribute for the standardization and exchange of scientific data between investigators and institutions.

**Leiden University Repository** – (URL: <https://openaccess.leidenuniv.nl/dspace/>)

Established in Apr. 2004 the Leiden University Repository collects, preserves, indexes and distributes digital scholarly output from researchers including research datasets. Repository software used is DSpace.

**Deep Blue** – (URL: <http://deepblue.lib.umich.edu/index.jsp>)

Deep Blue is the University of Michigan's permanent, safe, and accessible service for representing its intellectual environment. Deep Blue supports a variety of scholarly formats, including data, images, audio and video files, etc. Deep Blue uses DSpace.

**John Hopkins University Digital Knowledge Center** – (URL: <http://ldp.library.jhu.edu/units/dkc>)

The Digital Knowledge Center (DKC) is the digital library research and development department within the Sheridan Libraries of The Johns Hopkins University. The DKC research agenda focuses on the ingestion of materials into digital libraries to ensure access and preservation. The projects emphasize the development of automated tools, systems, and software to reduce the costs and resources associated with converting the vast knowledge within print, audio and video materials into digital form, and with curating content that is "born digital", such as large-scale scientific datasets.

**DSpace at MIT** – (URL: <https://dspace.mit.edu/>)

DSpace at MIT is MIT's online institutional repository - built to save, share, and search MIT's digital research materials. Materials include preprints, technical reports, working papers, theses, conference papers, images, datasets etc.

## **Other notable UK data repositories**

### **UK Data Archive**

<http://www.data-archive.ac.uk/>

### **ULCC: National Data Repository**

<http://www.ulcc.ac.uk/digital-preservation/current-activities/ndr.html>

### **NDAD – The National Digital Archive of Datasets**

<http://www.ndad.nationalarchives.gov.uk/>

### **Arts and Humanities Data Service**

<http://ahds.ac.uk/>

## **NERC Data Centres**

- **British Atmospheric Data Centre (BADC)**

<http://badc.nerc.ac.uk/home/>

- **National Geoscience Data Centre (NGDC)**

<http://www.bgs.ac.uk/programmes/infoserv/ngdc/home.html>

- **NERC Earth Observation Data Centre (NEODC)**

<http://www.neodc.rl.ac.uk/>

- **British Oceanographic Data Centre (BODC)**

<http://www.bodc.ac.uk/>

- **Antarctic Environmental Data Centre (AEDC)**

<http://www.antarctica.ac.uk/Resources/AEDC/>

- **Environmental Information Centre (EIC)**

<http://www.ceh.ac.uk/data/>

- **National Water Archive (NWA)**

<http://www.ceh.ac.uk/data/NWA.htm>

- **NERC Environmental Bioinformatics Centre (NEBC)**

<http://nebc.nox.ac.uk/>