


Cover Sheet for Proposals <i>(All sections must be completed)</i>			
Project Area: <input type="checkbox"/> a) Interoperability Demonstrators <input type="checkbox"/> b) Data Audit Framework <input checked="" type="checkbox"/> c) Pilot Implementations of the data audit framework			
<hr/>			
Name of Lead Institution:		University of Edinburgh	
Name of Proposed Project:		Edinburgh Data Audit Implementation Project	
Name(s) of Project Partner(s):			
Full Contact Details for Primary Contact:			
Name: Robin Rice Position: Data Librarian Email: R.Rice@ed.ac.uk Address: EDINA, 160 Causewayside, Edinburgh EH9 1PR Tel: 0131 651 1431 Fax: 0131 650 3308			
Length of Project:		6 months	
Project Start Date:		3 March, 2008	
		Project End Date: 29 August, 2008	
<hr/>			
Total Funding Requested from JISC:		£30,000	
<hr/>			
Total Institutional Contributions:		£25,644	
1. Outline Project Description: The issues in preserving Edinburgh's wide range of research data will be similar to the issues faced by all UK HE institutions. Given Edinburgh's extensive research scope and the range and complexity of the data generated, we believe that Edinburgh provides the perfect testbed for a (partial) institutional data audit. Most aspects of data preservation are faced at Edinburgh, so a study based on this institution will provide baseline information across the range of issues involved, and will permit extrapolation of the findings to other research institutions across the UK. Investigating the situation in a particular university will uncover issues that would otherwise not come to light, for example whether the infrastructure needs to be changed.			
I have looked at the example FOI form at Appendix A and included an FOI form in the attached bid (Tick Box)		YES	
I have read the Circular and associated Terms and Conditions of Grant at Appendix B (Tick Box)		YES	

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, or if we are successful in our bid for funding and our project proposal is made available on JISC's website.

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

Background

1. The University of Edinburgh's Information Services' (IS) motivation for bidding as one of the exemplar implementations of the data audit framework relates to its desire to provide better central support for its researchers and data curators throughout the research and data life cycles of their activities. In a related bid, the University of Edinburgh appears as a partner in a bid for the Data Audit Framework Development to be submitted by Seamus Ross, HATII and Digital Curation Centre, University of Glasgow. By participating in the development of the framework, the University will be able to ensure a good fit of the framework methodology with the requirements for institutional support of its researchers and data curators. Through the work in this exemplar implementation strand, the University will be able to extend the audit work to a further number of key schools and research institutes.

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

2. Following a recent restructuring of the University designed to optimise the performance of a research-led institution, Information Services responded with a realigned organisation that is freshly focused on supporting the research agenda. To ensure research community engagement, a root and branch review of IT services was led by the Vice-Principal for High Performance Computing and e-Science. Measures implemented included the creation of three senior liaison posts, each immersed within the research community, to ensure that service planning is fully aligned with the aspirations of the University's three Colleges. The division of Information Services into individual teams with responsibility for resources, infrastructure, applications and service delivery, where previously these were diluted across traditional Library and IT functions, has enabled a new user focus and an opportunity to concentrate technical expertise on promoting innovative solutions. These changes have produced a new working culture for Information Services staff and their direct association with the emerging needs of 21st century researchers. Measurement of benefits has been accomplished through a rigorous process of self-reflective evaluation, and through the scrutiny and endorsement of annual service plans and capital proposals, which must be linked directly to College and institutional strategic plans.
3. One component of the realigned service, the Library, has proved a national leader in the provision of open access repository services. This development demanded intensive advocacy on matters as diverse and contentious as the benefits of open access publishing and the legislative intricacies governing researchers' rights and responsibilities. This exercise demonstrated the critical importance of building a service that reflects user requirements, whether in providing practical tools (e.g. to generate publications lists) or in assuring users that their intellectual output is protected. Success can be judged by a high level of engagement by the research community.
4. Information Services consists of the following divisions: Library User Services, Library & Collections, IT User Services, IT Infrastructure, Applications, EDINA and the Digital Curation Centre, supported by the Vice Principal's Office. Each of these divisions has a unique perspective on the research, computing, and data requirements of the University. This project will capitalise on this diversity to draw strengths throughout Information Services and their contacts with academics in the University to ensure best utilisation of the audit framework.
5. Edinburgh is a research led University; new figures show that the University won a record £210 million to fund its research in 2006/07.¹ Research data are generated by individuals and research groups in all 23 Schools of the three Colleges, spanning a very wide range of disciplines in the Humanities and Social Sciences, Medicine and Veterinary Medicine, and in Science and Engineering. There is an increasing trend for research groups to work as part of

1. _____

¹ <http://www.ed.ac.uk/news/ebulletin/November07/article36164.html>

global collaborations which create substantial amounts of data that need to be accessible to researchers based in different institutions. Research funding agencies now usually stipulate that there should be a plan for sharing the data collected during the lifetime of a research project but they allow no provision for the cost of preserving this data after the end of the contract period.

6. One way to ensure data are preserved is to 'publish' them, that is make them publicly accessible. The University's Data Library is working with the Edinburgh Research Archive² to set up a DSpace digital repository, with standards-compliant discovery metadata. This will preserve and make available data from researchers willing to openly share their data on the Internet who do not already have access to a data archive or repository in their field. Edinburgh DataShare³ pilot service aims to explore options across a number of disciplines within the University as part of the Data Library's goal⁴ to assist Edinburgh-based researchers in managing and sharing their data.
7. An internal Research Computing Consultation which was carried out during 2007 reported that most staff at the University of Edinburgh work with data up to 100 GB, most backing this up by a variety of personal methods that did not involve University servers. Over 25% of the data would be impossible to recreate in case of a loss, with around 50% of respondents noting that data would take months or years to recreate. One third of respondents were required to retain data, with 16% wishing to retain data for over 20 years. A number of research groups generate very large quantities of data including particle physicists, astronomers, systems biologists and medical imaging teams.
8. The Edinburgh Compute and Data Facility (ECDF) was set up in 2006 as a collaboration between Information Services, NeSC (National e-Science Centre) and the e-Research focus within the University, to provide cluster computing together with a Storage Area Network (SAN) for research data. The ECDF team therefore have direct contact with research groups who need to use the ECDF to process and store large quantities of research data.
9. A number of recent JISC-funded projects led within the Information Services of the University of Edinburgh will inform the direction of the work. These include:
 - the Digital Curation Centre's SCARP project⁵ to investigate disciplinary attitudes and approaches to data deposit, sharing and re-use, curation and preservation
 - successful completion of EDINA's ground-breaking GRADE project⁶ for innovations in how to share geospatial data derived from licensed sources
 - the Depot⁷ - a national facility for Open Access deposit of post-prints hosted by EDINA
 - the DISC-UK DataShare project⁸ to set up institutional data repositories at four institutions, including the pilot Edinburgh DataShare service, led by the Data Library
 - Project StORe's⁹ survey and pilot demonstrator (now undergoing a new phase in cooperation with the CLADDIER project) led by the University Library.

Quality of Proposal and Robustness of Workplan

10. Generally, Information Services wishes to contribute this data audit institutional exemplar because of its desire to provide new and better support for data creators, collectors and

1. _____
² <http://www.era.lib.ed.ac.uk/index.jsp>
³ <http://datalib.ed.ac.uk/datashare.html>
⁴ <http://datalib.ed.ac.uk/userserv.html#4>
⁵ <http://www.dcc.ac.uk/scarp/>
⁶ <http://edina.ac.uk/projects/grade/>
⁷ <http://depot.edina.ac.uk/>
⁸ <http://www.disc-uk.org/datashare.html>
⁹ <http://jiscstore.jot.com/WikiHome>

curators within the University so that they can make informed choices about appropriate levels of management and curation their data receives and for how long. It is envisaged that best practice will be advanced so that in general, those holding data in ways identified at the bottom of the 'Data Sharing Continuum' graphic¹⁰ can move up the continuum towards more secure and well-documented forms of storage and preservation, and those with data suitable for sharing will move up the continuum towards open standards-based data publishing. This will be a user-centred approach, focusing not on university-wide blanket policy development, but in improving support, guidelines and tools for principal investigators to make better decisions about data management, sharing, and archiving, with policy decisions located at the research group or school level in the first instance.

11. The **timetable** will be written in consultation with JISC and the team developing the data audit framework, with preparation to begin in February, and this project to begin in March, beginning with development of the project plan, finalizing the local steering group, and selecting the research centres or schools to be audited. Completion is envisaged within six months, e.g. end of August 2008. If a delay to implementation to allow development of the methodology is seen as desirable by the designated teams and funder, then the completed work will be finished no later than March 2009, as indicated in the tender document. (As the designated project staff will be seconded from their permanent posts within Information Services, the timetable can be flexible.)
12. The major **deliverables** will consist of at least five implementations of data audit in research centres or schools in each of the three Colleges and therefore across a range of disciplines. Minor deliverables will include case studies, reporting of 'lessons learned', publications and presentations given to appropriate professional bodies in different fields to help build capacity among academics and support professionals, outputs deposited in the central repository managed by the framework development team, and guidance documentation geared toward discipline-specific and cross-disciplinary data creators and curators. Local deliverables to the University will consist of recommendations for school-based and central policy development, a tool for data managers to conduct a self-audit, guidelines and documentation, and enriched support services. Also, it is envisaged that this project will dovetail with Edinburgh DataShare, by identifying researchers who wish to share their data through the DSpace repository or through web services or 'mash-ups'.
13. **Project management arrangements:** the project director will be Robin Rice, Data Librarian, who is currently the project manager of DISC-UK DataShare project, within the JISC Repositories and Preservation Programme. The project manager, who will contribute her full time to the implementations of the audit for the life of the project, and if funded, will also contribute to the HATII-DCC-led bid for the development of the audit framework for an overlapping period, is Dr Cuna Ekmekcioglu. The line manager of Dr Ekmekcioglu, the Director of Edinburgh Compute and Data Facility, Jean Ritchie, will also be closely involved and help to select the local steering group and the candidate units for audit. The Digital Curation Centre is based within IS at the University of Edinburgh and will have a strong steer on the project, particularly from the DCC Research Director, Prof. Peter Buneman. Prof. Jeff Haywood, the Director of Information Services and Acting Director of the Digital Curation Centre, and Peter Burnhill, Director of EDINA will be involved in oversight of the project and selection of the project committee and candidate units for audit. Susan Graham, Records Management Officer will be on the steering group, as will Chris Adie, Academic Liaison Director for the College of Science and Engineering. If our related bid as a partner in the development of the audit framework is funded, then Seamus Ross will also be a key advisor on the project management arrangements for Edinburgh's implementation of the data audits.

1. _____

¹⁰ Developed as part of the DataShare project by Robin Rice. See appendix, or http://www.disc-uk.org/docs/data_sharing_continuum.pdf.

14. An initial **risk assessment** is below; this will be revisited as part of the development of the project plan.

Risk description	Risk probability	Risk impact	Risk amelioration
Loss of key staff during project implementation	Low	High	Existing staff will be seconded to the project who have a stake in the outcome of the project as part of their normal posts.
Delay in audit implementations due to delay in development of audit framework in related strand.	High	Low	Project staff can be flexible on the timetable; a suitable timetable will be sought with the related project and with JISC.
Lack of cooperation from academics in participating in the data audit	Medium	High	The formation of the local steering group should ensure contacts with key research groups and 'champion' academics are ensured. Incentives – including financial ones, will be considered.
Methodology for audit developed by framework team is found to be unworkable at University of Edinburgh	Low	Medium	This is the reason we would like to be a partner with the framework development team (as in bid submitted by S Ross), so this is less likely to happen. If another team is selected to develop the framework and this risk occurs, we would seek help from the JISC programme manager to help us find a compromise or solution.

15. **Intellectual Property Rights:** For 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 within the author's institutional repositories or the repository/database developed as part of the audit framework. Academic publication of some articles will be pursued. Software will be developed within open source environments. There are IPR issues to do with deposit of research data in open access repositories which will be explored as part of the project. Copyright of any research outputs or data documentation will remain with the researchers' whose data collections have been audited.

16. **Sustainability:** Much of the envisaged 'lessons learned' from conducting the exemplar audits will be about to what extent these activities are sustainable within normal operating procedures and costs of the research groups and schools. Information Services has been re-examining its support structure (see Background section) and will be in a better position to understand what support it can provide centrally after the exemplar audits are complete. It is clear that better support for data curation is a priority for a support service striving to help its community meet its goal of excellence in research within a world-class University. Not only the Data Library, but also the computing and library services will need to be involved to meet this aim.

Engagement with the Community

17. **Preliminary Stakeholder Analysis** – this will be expanded upon as part of the project plan.

Stakeholder	Interest	Importance
JISC, Research Councils,	Funders of research and	High

other research funders	development (and data collection/creation)	
UCISA, SCOUNL, CURL, RIN	Professional organisations developing policies for institutional support for data curation	Medium
University of Edinburgh senior academics and management	Potentially interested in articulating university policies about data management/curation	High
University of Edinburgh researchers and research organisations across 3 colleges	Data creators who are responsible for data management/curation; currently do not receive much central support; practice informed by disciplinary norms	High
National Services: e.g. DCC, RSP, JISC Legal, EDINA, MIMAS, UK Data Archive, AHDS, CETIS, RRT, JORUM	Can both input expertise into project and help disseminate outcomes to broader community	Medium
UoE Information Services	May need to change levels and types of support to researchers; responsible for communicating lessons learned to peers at other universities.	High

18. One of the first activities that will need to be done if funded is to undertake a user requirements analysis. This will be done in tandem with Edinburgh DataShare, which as part of the DISC-UK DataShare project, needs to find 'early adopters' to contribute to an open access pilot institutional data repository. Data owners in various disciplines will be consulted as part of this exercise.

19. Members of the project team will be available to meet with JISC and service providers to discuss findings, evaluate outcomes, and to write and present findings to relevant publications and conferences during and after the length of the project.

Budget

20. Only the Project Director and Project Manager are costed in the budget, although as noted a range of participants will assist in the steering of the project, not to mention the researchers in the candidate research organisations and schools. Additional qualitative benefits to the University of Edinburgh will be in the shape of new evidence for policy-making and a chance to test the framework to see if it works within its organisational setting. The DataShare project expects to benefit from exposure to more research groups as potential data depositors.

Directly Incurred Staff	1 March 2008 – 31 August 2008
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]

[REDACTED]	[REDACTED]		
[REDACTED]	[REDACTED]		
[REDACTED]	[REDACTED]		
[REDACTED]	[REDACTED]		
[REDACTED]	[REDACTED]		
[REDACTED]	[REDACTED]		
[REDACTED]	[REDACTED]		
[REDACTED]	[REDACTED]		
[REDACTED]	[REDACTED]		
Amount Requested from JISC	£30,000		
[REDACTED]	[REDACTED]		
[REDACTED]	[REDACTED]		
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

Previous Experience of the Project Team

- 21. A local steering group will be created in the first month of the project that will ensure leadership from representatives of Library, Computing, Data Library, and Records Management central services as well as ‘champion’ academics who are experienced and concerned with data curation across the three Colleges: Humanities & Social Science; Medicine & Veterinary Medicine; Science & Engineering. (See also paragraph 17.)
- 22. Dr Cuna Ekmekcioglu has worked in the fields of digital libraries, development of teaching & learning materials and learning technologies, and staff development. She has led a range of research projects and presented and published widely in these areas. Her current research interests centre on research computing, and in September 2007 she presented a draft ‘Research Computing: consultation report’ to the University of Edinburgh’s research computing strategy group.
- 23. Robin Rice is the project manager of DISC-UK DataShare, a JISC-funded project within the Repositories and Preservation Programme setting up institutional data repositories at four UK HEIs. As Data Librarian, she is the service manager of the University’s Data Library, and part of the EDINA management group.
- 24. Jean Ritchie is the Service Director of the Edinburgh Compute and Data Facility which is a research compute cluster and SAN. She developed the University’s Knowledge Management Strategic Plan (2004). Working for the Scottish Funding Council (1996 - 2002) she managed two successive programmes of projects across Scotland carrying out work on infrastructure, middleware, e-Learning and staff development.
- 25. Dr Chris Adie is currently Academic Liaison Director for the College of Science and Engineering and the Information Services group in the University of Edinburgh. He provides the key strategic link between the College and Information Services, and is tasked with improving communications and adding value to the relationship between IS and the College. He is a member of the IS Strategic Management Group. Between 2002 and 2006, Chris headed up the Specialist Services Division of Edinburgh University Computing Services.



INFORMATION SERVICES

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 jeff.haywood@ed.ac.uk

11 January 2008

JISC
Northavon House
Coldharbour Lane
BRISTOL
BS16 1QD

Dear Sir/Madam

I am very happy to support the proposed bids from Edinburgh and partners for the Data Audit Framework Development and the Exemplar Implementatino of the Data Audit Framework.

Information Services at the University of Edinburgh have been involved with a number of related activities, and we are interested in pursuing further support of our researchers' data management, sharing and preservation needs centrally. For example, the DISC-UK DataShare funded by JISC as part of the Repositories and Preservation Programme aims to establish new services for helping researchers share data online as appropriate, and in line with new requirements by funders. We see this activity as following on from involvement and interest in the Digital Curation Centre, and in keeping with the goals of our Knowledge Management Strategy.

Both strands, if funded, will be supported by a range of interested advisors around the university, to ensure the outputs will be useful for University activities and goals.

Yours sincerely,

Prof Jeff Haywood
Vice Principal Knowledge Management, CIO & Librarian to the University