


<b>Cover Sheet for Proposals</b> <i>(All sections must be completed)</i> <a href="mailto:a.laws@jisc.ac.uk">a.laws@jisc.ac.uk</a>			
<b>Project Area:</b> <input type="checkbox"/> a) Interoperability Demonstrators <input checked="" type="checkbox"/> b) Data Audit Framework <input type="checkbox"/> c) Pilot Implementations of the data audit framework			
<b>Name of Lead Institution:</b>		University of Glasgow	
<b>Name of Proposed Project:</b>		Data Audit Framework Development Project (DAFDP)	
<b>Name(s) of Project Partner(s):</b>		Digital Curation Centre, University of Edinburgh, King's College London, eCrystals (UKOLN and University of Southampton).	
<b>Full Contact Details for Primary Contact:</b> <b>Name:</b> Professor Seamus Ross <b>Position:</b> Director HATII University of Glasgow & Associate Director Digital Curation Centre (DCC) <b>Email:</b> <a href="mailto:s.ross@hatii.arts.gla.ac.uk">s.ross@hatii.arts.gla.ac.uk</a> <b>Address:</b> George Service House, 11 University Gardens, <i>University of Glasgow</i> , Glasgow G12 8QJ <b>Tel:</b> +44 (0)141 330 3635, <b>Fax:</b> +44 (0)141 330 3788			
<b>Length of Project:</b>		6 Months	
<b>Project Start Date:</b>		1 February 2008	
		<b>Project End Date:</b> 31 July 2008	
<b>Total Funding Requested from JISC:</b>		£100000	
<b>Total Institutional Contributions:</b>			
<b>Outline Project Description</b> <p>Following and building upon the work undertaken with the development of the DRAMBORA methodology for repository assessment the project will develop and provide examples for a Data Audit Framework (DAF) adapted to the current needs of data curation activities in higher education institutions. The project seeks to develop a general solution but also to highlight the differences created by variations across field of research and organisations through case studies. It will provide scenarios to respond to these variations. Adopting a reflective outlook that takes its influence from data audit within a range of domains and disciplines, and that is at all times informed by the expectations and requirements of key stakeholders, the project will iteratively develop and subsequently validate its audit framework throughout a series of pilot assessments in a range of representative HE environments.</p> <p>The project will distribute a software tool intended to support and facilitate data audit and incorporating a registry component intended for recording the results of data audits based on the framework—this will provide a powerful mechanism to enable institutions to know what data they have and where it is, as well as for the UK to create a national perspective of our research data assets.</p>			
<b>I have looked at the example FOI form at Appendix A and included an FOI form in the attached bid (Tick Box)</b>		YES ✓	NO
<b>I have read the Circular and associated Terms and Conditions of Grant at Appendix B (Tick Box)</b>		YES ✓	NO

## Data Audit Framework Development Project (DAFD)

### 1. Introduction

This proposal is in direct response to recommendations made by Liz Lyon, our colleague within the Digital Curation Centre, in the seminal JISC-commissioned report *Dealing with Data*: ‘a framework must be conceived to enable all Universities and colleges to carry out an audit of departmental data collections, awareness, policies and practice for data curation and preservation.’ The recent Report of the OSI e-Infrastructure Working Group presses a similar agenda if the UK is to ensure its research institutions adapt emerging e-infrastructure realities.<sup>1</sup> One of the current challenges for the higher institutions in the UK is their efficient participation in the national knowledge economy. Management and reuse of research data have become critical success factors for excellence in research and they are central to any national e-infrastructure. While research data offer benefits, they also pose risks and managing the risks and reaping the benefits requires knowledge of institutional data holdings. A Data Audit Framework provides the mechanism to collect and manage this knowledge.

**The aim of the DAFD project** is to develop and provide examples for a **Data Audit Framework (DAF)** adapted to the current needs of data curation activities in UK HE institutions. The project will develop a general solution as well as highlight, through case studies, differences caused by research field and organisation type. To achieve this aim, the DAFD foresees active collaboration with stakeholders, whose interest and involvement are summarised in Table 1 below. The DAF will enable institutions to find out what data they hold, where it is located, and who is (or is not) responsible for it, offer a tool for managing this information and sharing it with other institutions in a controlled environment.

If HE institutions are to ensure that they maximise their potential to exploit, manage, and reuse research data, as well as managing risks associated with its long term availability then they must be able to quickly and easily establish an overview of the data collections they hold, what policies and practices are in place within their organisations, and understand how aware staff are of their responsibilities for digital resources.

The DAFD project methodology combines an evidence-based approach with extensive consultation with key stakeholders in the field of research data creation, use, and curation. The major components of the project implementation will be: scoping in consultation with key stakeholders; creating of an initial version of Data Audit Framework; evaluating and testing of the Framework and its further iterative development; creation of a set of case studies; developing website and wiki support facilities and a central resource; and establishing dissemination mechanisms. The Data Audit Framework will be usable by Universities and Colleges without dedicated or specialist staff and its application will require only limited investment of time or effort.

Stakeholder	Focus	Involvement in the project
JISC	Contributes to JISC mission ‘to provide world-class leadership in the innovative use of ICT to support education and research’.	Promoter; Funder; Advisory Committee.
Institutional repositories	Improves understanding of place in framework of digital resource provision and management. Introducing the notion of self-audit.	Audit case studies; validation of tool; usage of tool.
UK HE Content-Creator Communities	Improves understanding of data curation lifecycle. Encourages use of standards, a focus interoperability, data quality, and data curation.	Audit case studies; validation of Framework; DAF usage.
Content-user communities	Improves knowledge of and access to digital content.	Dissemination.
Data centres and Information discovery portals	Improves quality of data and the trust to the content offered. Integrate data audit models in acquiring content.	Audit case studies; validation of Framework; DAF usage.
The Research Councils	Promote the development of research data management and reuse in various academic fields.	Advisory Committee; promotion of DAF
UUK	Takes care of excellence in research in universities.	Awareness and promotion of DAF.
SCONUL	Promotes excellence in library services in HE and national libraries; data audit is increasingly essential in this arena.	Awareness and promotion of DAF.
CURL	Aims to build a Distributed and Hybrid Research Library of the future covering various research fields—data audit essential.	Awareness and promotion of DAF.
UCISA	Provides and develops academic, management and	Awareness and promotion of

<sup>1</sup> *Developing the UK’s e-infrastructure for science and innovation: Report of the OSI e-Infrastructure Working Group*, <http://www.nesc.ac.uk/documents/OSI/report.pdf>, (London 2007).

	administrative information systems.	DAF.
RIN	Leads and co-ordinates the provision of research information in the UK	Awareness and promotion of DAF.
DCC	Provides a national focus for research and development into curation issues. Promotes expertise and good practice, both national and international, for the management of all research outputs in digital format.	Project partner; awareness and promotion of DAF; long term management of DAF Tool, and training.
UK Data Centres and Services	Adds to existing expertise and understanding of digital preservation, adds to overall understanding of digital repositories in the UK.	Audit case studies; validation of tool; usage of tool; awareness and promotion of DAF.
Digital Preservation Europe	Data feeds into DPE's existing cataloguing activities, adds to overall understanding of digital repositories in the UK, and knowledge of digital preservation	Awareness and promotion of DAF; support for delivery of <i>DAF Interactive</i> .

**Table 1. Key Data Audit Framework Stakeholders**

## 2. Project Description

### 2.1 Background/rationale

As data is increasingly relied upon to act as an evidential base for analytic work, within a diverse selection of domain areas, the quantity of data stored and relied upon is increasing at a tremendous rate. Assurances of research data management are urgently sought to support continued validation of the conclusions that they yield and to reassure of their continued availability and reusability. The implications of poor data management and quality are widespread and potentially damaging, and may be manifested financially, legally, or even in terms of reputation.

**Locating and Accounting** for data within any organisations is a complex task. Among the most obvious challenges relate to the data collections themselves and the flows that characterise their creation and lifecycle. Factors included the validity and completeness of encoded data; its integrity, correctness and consistency; the adequacy of available representation information or other metadata; and the extent to which conformance with controlled vocabularies or taxonomies are maintained. Other considerations include organisational structures and functions, business processes, users and services and the characteristics of any external sources of data. Still other key factors include authenticity, trustworthiness, fidelity, provenance, and, in some cases even, currency (i.e. freshness). A successful methodology to support data audit must explicitly define each information asset attribute that must be collected. It must be delivered as a set of guidelines and techniques that defines a rational and reproducible process for audit. Institutions need guidance as to what information should be collected in a data audit, where they can find the information and how they might exploit the information and data once they have collected it.

### 2.2 Methodology Development: Lessons from DRAMBORA

The development of the DAF methodology will rely on the experiences gained from developing the DRAMBORA toolkit.<sup>2</sup> This toolkit was developed in collaboration by the DCC and DPE and has undergone a series of pilot audits to develop an improved on-line audit tool DRAMBORA Interactive. The DRAMBORA approach of identifying the organisational context, concrete objectives, tasks and assets to support the risk identification and management exercise has a proved an effective method for gathering knowledge of how repositories work (and do not work) in fulfilling their core objectives. We will apply a similar methodology to create the Data Audit Framework.

### 2.3. The Methodology to Create the Data Audit Framework

The development of the DAF methodology for data audit will evolve in an iterative development process encompassing extensive consultation with key representative stakeholders, in order to establish data and user requirements, as well as a contained series of case studies within four UK data-holding Universities. These institutions will be suitably representative of the wider UK HE data landscape, and offer a valuable degree of diversity in terms of scale, scope, materials and strategy. A DAFD expert advisory panel (see below Sec 2.6.3) will be convened to direct the scoping studies, evaluative research and methodological definition. We will also liaise closely with the five exemplar data audit institutions selected by the JISC through the parallel call.

#### 2.3.1. Evaluative Collaborators

The DAFD project will use specific DAFD partner projects and institutions to provide it with the evidence based needed to define the Data Audit Framework: the University of Glasgow, University of Edinburgh, King's College London, and the eCrystals archive for Crystal Structures at the University of Southampton. By working with data creators, users and owners we aim to ensure that our data audit toolkit is applicable to the wide range of data created by UK research-led institutions from the arts and humanities, to the sciences, and to clinical trials.

<sup>2</sup> The methods and processes that we used to create DRAMBORA are described in the introduction to DRAMBORA so we will not re-describe them here. See <http://www.repositoryaudit.eu>

### 2.3.2 An Evidence Based Methodology for Constructing the Data Audit Framework

The proposed DAFD Team is extremely conscious of the value of usable tools, particularly following the experiences accumulated developing DRAMBORA. The Data Audit Framework we are proposing will be firmly practice-oriented and intuitively applicable. Following the initial scoping exercise and an analysis of wider disciplinary approaches to data audit, a series of pilot evaluative data audits within UK University data holders will enable the evolution of a draft data audit framework. The very first assessments will utilise the scoped stakeholder and community requirements and input as their intellectual basis. The iterative development process (see Figure 1 above) will lead to the release of a draft framework following the second pilot audit, which will be subsequently utilised and validated during the remaining two audits. Critical in this process will be to understand what information is required for different research communities to document their data and what questions and approaches institutions need to use to find the data. This will provide the version of tool that participants in Exemplar Implementations of the Data Audit Framework will use.

Taking a few sample units we will define a Data Audit Framework which will enable users to be pointed where to look for data, to know how to describe it, and to have pointers for the kinds of documentation to look for. We will construct the audit process as data are not often where researchers claim and believe they are and they are often not documented in the same way researchers ‘remember’ them to be. In contrast, although experimental or observational evidence is costly and time consuming to obtain, it presents a definitive and indisputable insight into data realities. A further pressing challenge for audit is to determine an optimal way of specifying desirable data specific characteristics in a communicable fashion that can be related to the practical activities of auditors and data managers.

### 2.3.4 The Data Audit Framework Release

The initial release version of DAFD will be utilised in the *Exemplar Implementations of the Data Audit Framework* Projects will be available as a toolkit accompanied by data collection and analysis tools. It will describe the steps necessary to undertake a data audit within a range of UK HE research data environments. A second release will be online, accessible via the web. The development of this version will benefit from the experience we have gained constructing *DRAMBORA Interactive*, (release date 31 January 2008). *DAF Interactive* will facilitate aspects of data audit to ensure the creation of a central audit registry at Institutional level which can be federated at national level.

## 2.4 Timetable

Work Package	M1	M2	M3	M4	M5	M6
<b>WP1. Management</b>						
Task 1.1 Organisational management						
Task 1.2 Financial management						
Task 1.3 Technical management						
<b>WP2. Scoping and research activities</b>						
Task 2.1 Case study						
Task 2.2 Research, requirements and route map						
Task 2.3 Scoping study						
<b>WP3. Design and Implementation</b>						
Task 3.1 System design and development						
Task 3.2 Documentation and license						
Task 3.3 Consultation with stakeholders						
<b>WP4. Testing and evaluation</b>						
Task 4.1 Functionality test						
Task 4.2 User testing						
Task 4.3 Sustainability, scalability and business model						
<b>WP5. Communication</b>						
Task 5.1 Advisory Committee Set Up						
Task 5.2 Advisory Committee Activities						
Task 5.3 Advisory Committee Report						
Task 5.4 Project wiki						
<b>WP6. Dissemination</b>						
Task 6.1 General dissemination						
Task 6.2 Project publicity including website						
Task 6.3 Contribution to DCC summer school (of at least preliminary findings)						

Table 2: Data Audit Framework Development Project Timetable

### 2.5 Appropriateness and value of the project

The DAFD Project will lay the ground for data audit activities within UK’s research data creation, use and curation communities. It will provide a methodological framework utilising the evidence-based approach and facilitate the

building consensus within the stakeholder communities. The case studies conducted during this project will offer model examples and good practices in data audit—where to look for data, what kinds of documentation to expect, and what kinds of descriptions are likely to be available.

The centralised data registry which will be developed within the project is a core practical outcome. This registry, in line with the JISC mission, is foreseen as a national reference point enhancing its usefulness with time as it ingests increasing number of audit data long after the project ends. While we hope that all institutions using DAF will feel wish to deposit their audit data in this central repository, our experience with DRAMBORA shows that some institutions for a range of reasons (e.g. contractual, security, privacy) wish to maintain their audit data locally. We will aim to define a solution that supports the creation of a central registry and yet enables institutions to maintain local instances of the registry containing their information collected during their local audit.

The project will contribute to the strategic and practical development of the e-infrastructure within the UK. Its strategic added value is in the contribution towards the knowledge economy where management and re-use of research data collections are of critical significance. DAFD will strengthen the leading role of UK in the European research and development in the field of data audit. This trend of work will increase its influence within the Digital Library Initiative which is one of the flagship initiatives of i2010<sup>3</sup>, the EU policy framework for the information society and media. The project will help to better understand data curation lifecycle and improve performance in the institutions which are responsible for managing data collections. An added value of this project will be a better understanding of needs in professional training for data scientists if they are to supply information about their data.

## 2.6 Project partners and key personnel

### 2.6.1 Project lead and management

Overall responsibility for the project will rest with the Digital Curation Centre (DCC). Overall successful delivery of the project will be the responsibility of Prof Seamus Ross.

### 2.6.2 Partners Consortium and Project staff

DAFD has four core partner institutions: (1) University of Glasgow (Dr Milena Dobрева, Perla Innocenti, Andrew McHugh, Prof Seamus Ross); (2) University of Edinburgh (Dr Cuna Ekmekcioglu, Robin Rice), (3) Kings College London (Mark Hedges, Stephen Grace), (4) eCrystals (represented by UKOLN (Dr Liz Lyon) and the University of Southampton (Dr Simon Coles)

For details of the expertise and previous experience of these organisations and their staff involved in DAFD see Section 5. Appendix B includes the Letters of Support.

### 2.6.3 Advisory Group

At the outset of the project a DAFD Expert Advisory group will be set up, in consultation with the JISC, consisting of eight specialists from within the community who have expertise in data creation and reuse, audit, intellectual property rights, and digital curation. In addition to two meetings during the project, regular emails, the project wiki, and conference calls will be used to ensure regular input from the Advisory Group.

## 2.7 Project plan, Work Packages and deliverables

The project will be developed in 6 work packages, detailed in the next sections, each with associated deliverables.

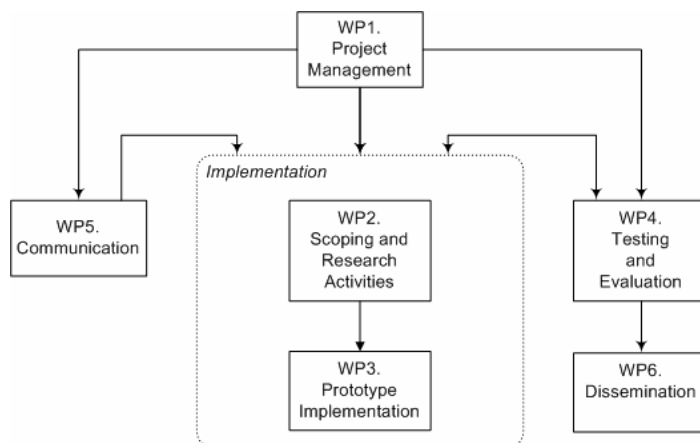


Figure 1: DAFD Workpackage Workflow and Interdependencies

<sup>3</sup> [http://ec.europa.eu/information\\_society/eeurope/i2010/index\\_en.htm](http://ec.europa.eu/information_society/eeurope/i2010/index_en.htm)

## 2.7.1 Work Packages Description

<b>WP1</b>	<b>Project management</b>		
<b>Objectives</b>	This workpackage aims to (a) ensure that all the work packages are managed coherently and that all project outputs are delivered within the agreed timeframe; (b) manage the administrative and functional coordination between the partners (c) manage the financial responsibilities of the partners; and, (d) manage the technical tasks of the project.		
<b>Brief Description</b>	Task 1.1 will involve organisational management and will (a) execute and update the project plan and support the local and work package managers; (b) keep the liaison with the responsible project officer at the JISC, and co-ordinate the project's administration, including the related meetings; (c) trouble-shoot possible conflicts among partners and among partners and third parties; and ,(d) edit and assure the quality control of the reports and the deliverables in general, and keep the project's archive and the project's handbook up to date. Task 1.3 focuses on Technical management including managing the higher-level co-ordination of the technical work packages and organising the technical meetings.		
<b>Duration</b>	m1-m6	<b>Dependencies</b>	All workpackages
<b>Milestones</b>	Successful launch of the project (m1). Successful completion of the project (m6)		
<b>Outputs</b>	Including (1) a detailed work plan, (2) production progress reports to JISC, and (3) Quality assurance.		
<b>Coordinator</b>	Seamus Ross, HATII at the University of Glasgow		
<b>WP2</b>	<b>Scoping and Research activities</b>		
<b>Objectives</b>	Workpackage 2 covering Scoping and Research activities covers (a) the gathering and selection of case studies/examples of successful user requirements studies and open source software solutions for data audit; (b) an investigation of IPR issues associated with the development and deployment of a Data Audit Framework; (c) Analysis of the case studies to enable the development of specifications for the functional, technical and content requirements of the Data Audit Framework; and, (e) the provision of a roadmap and feasibility plan for long-term Data Audit Framework developments, user studies and services.		
<b>Brief Description</b>	This workpackage consists of three tasks. In Task 2.1, Case Studies, we will undertake case studies to accumulate information about data in different research communities and organisational settings; (b) consider how to motivate researchers to contribute to the audit process; (c) consider the minimum amount of information that we need to collect in a data audit for it to be useful; and (d) examine what successful open source solutions for data audit are available and which might be incorporated in the framework. In Task 2.2 on Research, requirements and route map we will (a) define the functional, technical and content requirements of the Data Audit Framework development; (b) provide data audit collection scenarios, and, (c) deliver a roadmap and feasibility plan for long-term Data Audit Framework developments, user studies and services. In Task 2.3, The Scoping Study, we will deliver a scoping study including the methodology adopted, research activities and findings.		
<b>Duration</b>	m1-m2	<b>Dependencies</b>	WP3, WP4, WP6
<b>Milestones</b>	Delivery of the Scoping Report (m4).		
<b>Outputs</b>	(1) a user requirements overview, (2) a statement of the kinds of information to be collected during a data audit and where it comes from; and (3) the detailed scoping study including data audit collection scenarios.		
<b>Coordinator</b>	Milena Dobрева, HATII at the University of Glasgow		
<b>WP3</b>	<b>Design and Implementation (Prototype Implementation)</b>		
<b>Objectives</b>	In Workpackage 3 we will (a) develop the Data Audit Framework in the context of the Scoping Report Findings; (b) define a robust and scalable software architecture; (c) establish a usable and user-oriented interface for the DAF; (d) produce software and document it; (e) draw up recommendations for the enhancement of DAF; and (e) release the DAF under a open source license.		
<b>Brief Description</b>	Task 3.1, System Design and Development, involves the (a) definition of system architecture and requirements; (b) the definition of the user interface and services; (c) the implementation of the prototype version of the Data Audit Framework <sup>4</sup> ; and, (d) iterative testing with regular user test sessions. In Task 3.2, Documentation and License Development, we will (a) prepare documentation of the prototype; and, (d) adopt an open source license. In Task 3.3, which covers Consultation with Stakeholders, we will promote		

<sup>4</sup> We have not discussed here the implementation, but we have the capabilities among the team to adopt any of a wide range of programming languages and application solutions. All open source tools.

	(a) engagement of key stakeholders in the system design, development and licensing		
<b>Duration</b>	m2-m6	<b>Dependencies</b>	WP2, WP4
<b>Milestones</b>	Beta Data Audit Framework (m4), Data Audit Framework (m6), Prototype documentation and license (m4)		
<b>Outputs</b>	(1) Beta version of the DAF prototype, (2) Prototype documentation, (3) DAF License.		
<b>Coordinator</b>	Andrew McHugh, HATII at the University of Glasgow		
<b>WP4</b>	<b>Testing and Evaluation</b>		
<b>Objectives</b>	Workpackage 4 on Testing and Evaluation aims to control the quality and appropriateness of outputs of the project by developing a testing and validation plan and assuring its application. It will ensure that formative evaluation takes place throughout the project and is reported to JISC, and a summative evaluation is performed to consider recommendations of the project.		
<b>Brief Description</b>	A test plan will be developed and applied to the initial Data Audit Framework prototype and then iteratively to subsequent versions.. Testing will include all of the integrated software components that have successfully passed integration testing and also the software system itself integrated with any applicable hardware system(s). In Task 4.1, Functionality Test, we will (a) test the system's compliance with its specified requirements; and, (b) test the Data Audit Framework Prototype for throughput, reliability, and consistency. In Task 4.2 we will conduct User Testing involving iterative online and field user testing with partners, Expert Advisory Committee and stakeholders involved.		
<b>Duration</b>	m4-m6	<b>Dependencies</b>	WP3
<b>Milestones</b>	Functionality and user testing of the Data Audit Framework prototype (m5).		
<b>Outputs</b>	(1)Prototype evaluation, and (2) Report to DAF Expert Advisory Group and JISC on the value of the project, with recommendations for future development work.		
<b>Coordinator</b>	Mark Hedges, Deputy Director of the Centre for e-Research, King's College London		

<b>WP5</b>	<b>Communication</b>		
<b>Objectives</b>	This Workpackage involves the establishing of an Advisory Committee composed of individuals or representatives of relevant bodies or organisations in consultation with the JISC and in laying down channels of communication with the Exemplar Institutions. The Expert Advisory Committee will meet physically twice and engage in regular communication via email, our wiki and conference calls to support the delivery of the tool. It will also involve the creation of the project wiki for use by the project team to support the development of DAF and the documentation of it and the activities surrounding its development and use.		
<b>Brief Description</b>	Task 5.1, Expert Advisory Committee, will involve setting up of the advisory committee, made of key individuals and representatives of critical organisations. In Task 5.2, Advisory Committee Activities, we will handle the promotion of the programme for the Advisory Committee, comprising of two regular meetings, communication by email, conference calls if necessary, and engagement with the DAF Wiki space. In Task 5.3, Advisory Committee Report, will report on the conclusions and assessment of the expert advisory committee, its results and possible follow-up actions. In Task 5.4 the project wiki will be established and mechanism for its maintenance put in place. This wiki will also be available to five institutions funded to carryout <i>Exemplar Implementations of the Data Audit Framework</i> . In Task 5.5 we will put mechanisms to support the five institutions funded to carryout the Exemplar Implementations and to enable engagement with their application of the Framework.		
<b>Duration</b>	m1-m6	<b>Dependencies</b>	All workpackages.
<b>Milestones</b>	Project Wiki (m1), Project Expert Advisory Committee (m1), Brief Report on Expert Advisory Committee and Input Exemplar Implementations (m6)		
<b>Outputs</b>	(1) Creation of and support by Expert Advisory Committee, (2) deploying the project wiki, and (3) establishing formal links with the five Exemplar Implementation Projects.		
<b>Coordinator</b>	Robin Rice, University of Edinburgh		
<b>WP6</b>	<b>Dissemination</b>		
<b>Objectives</b>	Workpackage Six aims to promote the Data Audit Framework to ensure that it is taken up and used by higher education communities in the UK.		
<b>Brief Description</b>	In Task 6.1 General Dissemination the project will prepare a dissemination plan and motivate and support the preparation of technical reports, posters and demonstrations for scientific and general conferences. In Task 6.2 Project Publicity we will (a) define the project's logo and graphical image, (b) establish and arrange the project website which will provide access to the methodological framework and to the central registry of audit data, (c) promote DAF at scientific conferences and in journals. (Most of this will be		

	possible only after the project ends and will be done as part of the sustainability actions.), and (d) design and release a leaflet promoting the Data Audit Framework. In Task 6.3 we will make a contribution to the DCC Summer School with a module on the Data Audit Framework, and we will launch a half-day tutorial on the Framework for Universities— How to use it and how to evaluate the results.		
<b>Duration</b>	m1-m6+	<b>Dependencies</b>	ALL Workpackages.
<b>Milestones</b>	Project Website (m1), DAFD Leaflet (m6), Announcement (m6), Summer School (m6)		
<b>Outputs</b>	(1) Project website; (2) articles and papers at scientific conferences and journals; (3) DAFD Project leaflet; (4) Announcements of DAF target at different communities (e.g. Principals, Library Directors, University VPs for Research, Internal Audit in Universities); and (5) a training module for use in the DCC Summer School of DAF.		
<b>Coordinator</b>	Perla Innocenti, HATII at the University of Glasgow		

## 2.8 Risk Table

<b>Risk</b>	<b>Probability</b>	<b>Impact</b>	<b>Risk amelioration/Contingency</b>
Lack of availability of staff to handle design, development, deployment, and dissemination of the Framework	Low	High	HATII, DCC, and its partners have the staff resource available to be redeployed to deliver this project on time.
Loss of a team member	Low	Medium	Multiple staff at each site have the expertise and skills required
Lack of agreement between key stakeholders on data audit coverage and procedures.	Medium	Medium	Advisory Committee is foreseen as a body which should reach a consensus.
Lack of Stakeholders Joining the Committee	Low	Low	Community need for the Framework is widely
Data audit in different environments may require different approaches.	Medium	High	JISC seeks for a common methodology and framework, but also to address the specific needs. The project is designed according to this understanding. A general framework will be developed, and case studies will highlight differences when such differences occur.
Lack of coherence between practices in different institutions.	Medium	Low	The project aims to address as much as possible the common needs. It will also contribute to introduce a common approach in the UK practice.
Lack of understanding of data audit importance.	Low	Low	The project has an important role in increasing the awareness and level of knowledge on data curation.
IPR problems not resolvable within project	High	Medium	Right issues for data audit activities need to be clarified (e.g. IPR issues arising when creating a central registry of data audit data is deployed).

*Note: At the outset of the project we will assign each risk an owner to ensure risks are effectively managed.*

## 2.9 IPR and sustainability

### 2.9.1 Project documentation outputs

The project will comply with the terms of the JISC Funding Agreement. Its outputs (methodology, scoping report, Data Audit Framework, final report) will be openly available (with Creative Commons licenses where appropriate), on the project website, any intellectual property from resulting research outputs, will be subject to the Copyright, Designs and Patents Act 1988. All deliverables will be made available to the JISC community in accordance with the JISC Project Management Guidelines. The one exception to this *may* be the unanonymised audited materials collected as part of developing the Data Audit Framework which may be governed by confidentiality agreements between the project partners and the organisations at which data collections had been audited. Anonymous versions of these reports will be available, however.

### 2.9.2 Project Software Licensing

Our aim is to find the most suitable license solution for the Data Audit Framework Methods and Software, not only in terms of performance of available technologies but also in relation to their cost-effectiveness and their longer term maintenance. Our intention will be to release the Data Audit Framework as open source in keeping with the JISC's Policy on Open Source Software for JISC Projects and Services.





archives to interoperate with international subject archives (IUCr and CCDC) and other third party harvesters; develop approaches to the preservation and curation of scientific data in open repositories.

The **Centre for e-Research (CeRch)** was set up by **King's College London** on 1st October 2007. The Centre's remit within King's is to work across all departments in the College to support existing e-research, to extend e-research in new areas, and to encourage innovation in the use of ICT in research. A major element of this is to define strategies and develop an infrastructure for the management and preservation of the large amounts of digital data produced by research across the College. The CeRch team is drawing on the knowledge and experience gained from the activities of the Arts and Humanities Data Service (AHDS), as the Centre incorporates the staff and the research activities of the AHDS Executive, hosted at King's.

## 5.2 Key personnel short bios

**Dr. Milena P. Dobрева** (Researcher Digital Curation, HATII) contributed to the development of the DELOS Digital Library Reference Model (2007) playing a key role in research in the areas of quality, policy, and architecture. She participated in research behind and creation of the DigitCULT series of Technology Watch Reports (2003-4).<sup>5</sup>

**Perla Innocenti** (Researcher in Digital Preservation, HATII) is Co-Investigator of FP7 EU-funded project SHAMAN (Sustaining Heritage Access through Multivalent) and contributing developer to DRAMBORA.<sup>6</sup> Perla has researched repository design, audit research and digital collections usage as part of the HATII team working on EU-funded FP6 projects DigitalPreservationEurope (DPE) and PLANETS.

**Andrew McHugh** (DCC Advisory Services Manager) was lead the developer of DRAMBORA (Digital Repository Audit Method Based on Risk Assessment), and designed and implemented the software behind DRAMBORA Online. In the Digital Curation Centre he is Advisory Services Manager. He also lectures on multimedia systems and design on the MSc in Information Technology run by the Computing Science Department at Glasgow.

**Prof Seamus Ross** is Professor Director of HATII (Humanities Advanced Technology and Information Institute) at the University of Glasgow and Associate Director of the Digital Curation Centre. Other projects on which he was either the co-PI or PI include DELOS, DigitalPreservationEurope (DPE), ERPANET, and DigiCULT.

**Raivo Ruusalepp** of the Estonian Business Archives Consultancy has collaborated with the Digital Curation Centre on audit and certification services in the EU-funded project DigitalPreservationEurope and is technical coordinator in the FP-7 EU-funded digital preservation project Protage. In the past he has been involved in producing the JISC funded reports on the "Feasibility and Requirements Study on Preservation of E-Prints" (2003) and the "Assessment of UKDA and TNA Compliance with OAIS and METS Standards" (2005).

**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.

**Robin Rice** is the project manager of 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.

**Jean Ritchie**, Service Director of the Edinburgh Compute and Data Facility, 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.

**Dr Mark Hedges** is Deputy Director of the Centre for e-Research (King's College London) and Technical Manager of the Arts and Humanities Data Service. He has extensive experience of software development, technical R&D, technical management and project management, gained from 17 years work in the IT industry. He is now also responsible for the development of the e-research infrastructure across King's.

**Stephen Grace** is Preservation Manager at CeRch and the AHDS. Previously he has worked extensively on digitization and digital preservation in the cultural heritage sector, most recently at the Imperial War Museum.

**Dr Simon Coles** is the Manager of the internationally renowned UK National Crystallography Service and has led the its involvement in the UK e-Science program from the outset of the CombeChem project. SJC is a founder member of the CrystalGrid Collaboratory (e-Science Sisters project) and a co-investigator on the digital repository based projects eBank-UK, Repository for the Laboratory and eCrystals Federation (JISC) and the JISC VRE project myExperiment.

**Dr Liz Lyon** is the Director of UKOLN at the University of Bath UK, where she leads work to promote synergies between digital libraries and open science environments. She has led the eBank UK project, which built links between research data, scholarly communications and learning in the crystallography/chemistry domain. She is Associate Director (Community Development) of the UK Digital Curation Centre, in which UKOLN is a partner.

<sup>5</sup> <http://www.digicult.info/pages/techwatch.php?PHPSESSID=1928b267ea416e44a798d7080dd5146>

<sup>6</sup> <http://www.repositoryaudit.eu>

## **Appendix A**

### **FOI Withheld Information Form**

We do not request that the JISC consider withholding any material 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.

Appendix B: Letters of Support



SPB/SR/GCB

11<sup>th</sup> January 2008

Neil Jacobs  
JISC  
Brettenham House (South Entrance)  
5 Lancaster Place  
London  
WC2E 7EN

Dear Neil Jacobs and Colleagues

**RE: Institutional Support for Data Audit Framework Development Project**

As Vice-Principal for Research at Glasgow University and current Chair of the Digital Curation Centre (DCC) Steering Committee I commend the Data Audit Framework Development Project proposal to you. This bid led by the University of Glasgow is from a partnership that also includes the Digital Curation Centre (DCC), the University of Edinburgh, and Kings College London, in response to the JISC Circular 5/07, Calls for Projects in Digital Repositories (November 2007). Glasgow's participation is led by our Humanities Advanced Technology Information Institute (HATII). HATII is exceptionally well-placed to offer the expertise necessary to lead this project and as such, I whole-heartedly endorse HATII's participation in this proposal.

Institutions within the UK educational community have come broadly to recognise that research data assets have long-term value if they are effectively managed in ways that keep them accessible and ensure they do not suffer a loss of authenticity. HATII has built up significant expertise in the realm of data curation and preservation through participation in numerous international projects over the last nine years. It was the lead partner for the highly successful ERANET initiative and is currently a primary partner in the JISC-funded Digital Curation Centre (DCC). HATII has recently been awarded European Union funding under the Sixth Framework Programme to participate in three cutting-edge digital curation research projects; PLANETS, CASPAR, and DPE.

In the last two years, through the Digital Curation Centre, DigitalPreservationEurope and DELOS Digital Preservation Cluster projects, the University of Glasgow has established itself at the very forefront of digital curation repository audit. The development of the DRAMBORA toolkit has involved extensive studies of data holdings, and HATII and its associates within the bid now boast four individuals who have accumulated internationally

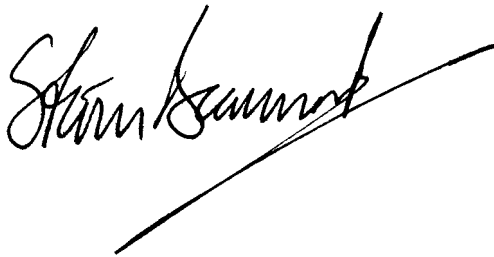
From Professor Steven P Beaumont OBE CEng FREng FRSE  
Vice Principal (Research & Enterprise)

Office of VP (R&E), Research & Enterprise, No. 10 the Square,  
University of Glasgow, Glasgow G12 8QQ  
Tel: +44 (0)141 330 3582 Fax: +44 (0)141 330 5856

unique experience and expertise in data collection and analysis within preservation environments. Able to operate concurrently, the team is peerlessly equipped to develop, apply and validate a suitable framework to meet the varied and pressing needs of the higher education research data community.

Given the significance of the problem, the established strengths of HATII in the area of digital curation and preservation, HATII's leadership in the development of an already highly successful audit framework, and the high quality and experience of the proposed project team, I can, without hesitation, fully endorse both the Data Audit Framework Development Project bid and HATII's involvement in it.

Yours sincerely

A handwritten signature in black ink, appearing to read 'Steven Beaumont', with a long, sweeping horizontal line extending from the end of the signature.

**Professor Steven P Beaumont OBE CEng FEng FRSE**  
Vice Principal (Research & Enterprise)

Enc.

From Professor Steven P Beaumont OBE CEng FEng FRSE  
Vice Principal (Research & Enterprise)

Office of VP (R&E), Research & Enterprise, No. 10 the Square,  
University of Glasgow, Glasgow G12 8QQ  
Tel: +44 (0)141 330 3582 Fax: +44 (0)141 330 5856

14 January 2008

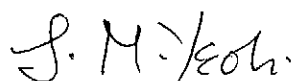
To whom it may concern,

Re: JISC Digital Repositories call - DAFDP Proposal

I write to confirm, on behalf of King's College London, our full support for the proposal submitted by the University of Glasgow entitled 'DAFDP (Data Audit Framework Development Project)', in which King's College London is a project partner. The proposal is being submitted under the JISC call 'Digital Repositories, Theme (b) Data Audit Framework (JISC Circular 5/07).

This project is a significant initiative in enabling the curation and preservation of the valuable digital material produced by research in higher education institutions, and we recognise its great importance both to King's and the wider higher education community.

Yours sincerely,



Jean Yeoh  
Acting Director ISS & Director of Corporate Services  
Information Services and Systems  
King's College London



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](mailto:jeff.haywood@ed.ac.uk)

11 January 2008

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

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

**Vice Principal Knowledge Management, Chief Information officer and Librarian to the University**  
**Professor Jeff Haywood**



School of Chemistry

**Dr S J Coles**  
EPSRC X-Ray Crystallography Service

Highfield,  
Southampton, SO17 1BJ,  
United Kingdom.  
Tel: (+44) 023 80596722  
Fax: (+44) 023 80596723  
e-mail: [S.J.Coles@soton.ac.uk](mailto:S.J.Coles@soton.ac.uk)  
<http://www.ncs.chem.soton.ac.uk/simon.htm>

12<sup>th</sup> January 2008

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

Dear Sir/Madam

I would like to express my support and willingness to collaborate with the proposed work from DCC and colleagues in response to the JISC calls for Data Audit Framework Development and an Exemplar Implementation.

I particularly wish to lend support as a principal investigator on the eCrystals project, which is implementing a federation of data repositories in the subject domain of crystallography. A Data Audit Framework is a particularly important consideration in this area and our work would be particularly complimentary to that proposed. I therefore confirm our willingness to collaborate and act as a testbed for the development and evaluation of the proposed framework.

I believe this work will greatly benefit the eCrystals Federation and its partners, which constitutes members from the crystallographic, chemistry, digital libraries, chemical information, publishing, learned societies, computer science and data management communities and will therefore be very wide-reaching.

Yours sincerely,

A handwritten signature in blue ink, appearing to read "Simon J Coles", written over a light blue horizontal line.

Dr. Simon J Coles.

Manager, EPSRC National Crystallography Service and Principal Investigator, eCrystals Federation project.

**UKOLN**

University of Bath  
Bath BA2 7AY • United Kingdom



JISC  
Brettenham House  
5 Lancaster place  
Kings College  
London  
WC2E 7EN

14<sup>th</sup> January 2008

Dear Sir/Madam,

Data Audit Framework Project Proposal

I am pleased to write in support of the Data Audit Framework Project proposal and Exemplar development led by the Digital Curation Centre, with UKOLN as a partner, being submitted to the JISC.

The topic being explored in this project, has recently been highlighted as an important area for further investigation in the *Dealing with Data Report* from UKOLN. We are a core partner in the eCrystals Federation and have led eBank data repository development for the past five years. This proposal builds on the very strong DRAMBORA Toolkit development approach and will provide a robust framework for institutions to adopt, in order to assess their readiness for data curation.

The Digital Curation Centre is well-placed to lead this initiative and UKOLN will be able to contribute to both strategic and technical aspects of the work plan.

Yours sincerely,

Dr Liz Lyon  
Director