

Consortium of University Research Libraries

Revised proposal in response to JISC Circular 1/02: Focus on Access to Institutional Resources (FAIR) Programme

SHERPA: Securing a Hybrid Environment for Research Preservation and Access

1. Summary

The SHERPA Project will aim over the course of three years to create a substantial corpus of research papers from several of the leading research institutions in the UK by establishing 'e-print archives' which comply with the Open Archives Initiative (OAI) Protocol for Metadata Harvesting (OAI PMH) using eprints.org software. The contents of these e-print repositories (and the harvestable metadata) will be freely available to all the HE and FE community and beyond. The creation, population and management of the repositories will be the core of the Project. Learning outcomes from these activities will also be shared with the community in a variety of ways, including the establishment of mechanisms to assist proactively other institutions in setting up and managing repositories.

In addition to the creation of e-print repositories themselves, it is proposed to carry out detailed investigations into digital preservation of e-prints. Particular emphasis will be placed on the practical issues associated with the possible implementation of standards such as the Open Archival Information System Reference Model (OAIS, see Appendix 1).

The SHERPA proposal has been produced under the auspices of CURL (the Consortium of University Research Libraries). CURL is also providing substantial financial support, covering key elements of the digital preservation costs and some hardware costs.

2. Partners

The 'Development Partners' for the Project comprise six university libraries which are members of CURL, plus the British Library (also a CURL institution), the University of York Library (which is not a member of CURL), and the Arts and Humanities Data Service (AHDS). The partners are:

- University of Nottingham (lead site)
- University of Edinburgh
- University of Glasgow
- Universities of Leeds, Sheffield and York ('White Rose' partnership)
- University of Oxford
- British Library
- AHDS

These partners will participate in the Project throughout its three years. They will be involved in the early installation of pilot repositories and also in ongoing work to investigate repository population and management.

In addition, a number of 'Associate Partners' will join SHERPA during the course of the Project. These will comprise seven additional CURL institutions. Associate Partners will receive the full support of the central Project Team but will be partly self-funded (the Project will only contribute to the cost of a server for each). It is expected that their work will however be accelerated by the fact that they can take on board the lessons learned from the Development Partner implementations.

3. Vision

The SHERPA Project will aim to develop and promote an environment in which the research output from several of the UK's largest research-led institutions is freely available on OAI-compliant e-print repositories to the rest of the HE and FE community and beyond. CURL libraries, which are based in these research-led institutions, are in a strong position to achieve this on behalf of the wider community.

The OAI PMH was originally developed to facilitate interoperability between e-print repositories. A number of successful e-print repositories, such as arXiv (<http://www.arxiv.org>) and CogPrints (<http://cogprints.soton.ac.uk>), are now OAI-compliant. However, most of the successful e-print repositories are centralised discipline-based implementations which are limited to a small number of subject communities. There are as yet few major examples of distributed institution-based e-print repositories. Yet the institutional model has the potential to promote the wide multi-disciplinary use of e-prints as a means of scholarly communication, and it therefore needs testing. The SHERPA Project will aim to test the viability of the model in practical ways by deploying and promoting a number of repositories in various research institutions and then analysing the technical, cultural and economic issues that emerge. Attention will be given to applicability of the model to different subject disciplines bearing in mind their varying conventions of communication and publishing.

The institutional repository model will be tested by SHERPA mainly in CURL university institutions. As well as including single-institution repositories, SHERPA will also test out the model where several institutions share a single server. The 'White Rose' partnership of the universities of Leeds, Sheffield and York will provide this dimension to the Project. In addition, the British Library will establish a server to host papers produced by 'non-affiliated' researchers (those using the BL and others working in research institutes or independently).

The work of SHERPA will not, however, be limited to these institutions. It is envisaged that the Project will explicitly investigate issues associated with enlargement. Bringing the Associate Partners on board will be used to test out enlargement models. Further enlargement (beyond the Associate Partners) will also be investigated. In addition, SHERPA will actively encourage other institutional implementations by providing advice and assistance to others in setting up repositories.

Preliminary work carried out by some of the Partners in this bid indicates that e-print repository implementations will only be successful in the context of major institutional advocacy initiatives. Academic authors will need to be persuaded that the e-prints initiative is a useful way to improve scholarly communication. Assistance will then be required to help them format and deposit e-prints. They will also need advice in addressing issues such as IPR and copyright. The SHERPA Project will aim to co-ordinate advocacy initiatives, assist authors in depositing papers and provide practical advice in dealing with IPR and copyright issues. In doing this SHERPA will build on work already being carried out by groups such as the CURL Task Force for Scholarly Communication, the JISC Scholarly Communication Group and the SCONUL Advisory Committee on Scholarly Communications. For example, during the Spring of 2002, CURL mounted local advocacy campaigns directed at academics in at least 14 separate venues. These have been evaluated with a view to informing future activity. The learning experiences and advocacy materials from activities like this in SHERPA will be made freely available to the wider community.

The preservation of the digital content made available by SHERPA is a key concern of the Project. Research will be carried out in this area to examine the feasibility of the application of standards such as the Open Archival Information System (OAIS) Reference Model. In April 2001, Peter Hirtle, Associate Editor of *D-Lib* wrote, "an OAI system that complied with the OAIS reference model, and which offered assurances of long-term accessibility, reliability, and integrity, would be a real benefit to scholarship."¹ The SHERPA Project will investigate how such a system might be created. The work will draw heavily on expertise from the CEDARS Project, and will be influenced by continuing international developments in this area. It will be carried out in partnership with the AHDS which have already developed considerable expertise in this area. It will also be assisted by collaboration with the British Library and also the Research Libraries Group.

More broadly, access to the papers in the SHERPA repositories will be promoted by working with available OAI Service Providers, some of whom (such as ePrints-UK) are also funded as part of the FAIR Programme. Collaboration on metadata standards will for example be essential to facilitate full interoperability.

The thrust of SHERPA is towards a vision of the academic community taking control of its own authored content in order to improve the research, learning and teaching processes. Working towards

¹ Peter Hirtle, 'Editorial: OAI and OAIS: What's in a Name?' *D-Lib Magazine* 7, 4, April 2001
<<http://www.dlib.org/dlib/april01/04editorial.html>>

this vision will involve testing out new business models of publishing which may run in parallel with current commercial models. The aim will be to enhance the effectiveness of research communication by promoting free access to a wider range of material than is presently available. SHERPA will also aim to conform with the ideal of “a true democracy of learning opportunity” identified by JISC in its recent draft *Information Environment: Development Strategy 2001-2005*². By making the content available, ensuring that greater interoperability is achieved, and promoting user-friendly search interfaces, the Project hopes to bring the activity of researchers much more easily into the digital domain of undergraduate students, fusing the worlds of learning and research in ways which could benefit both. It will clearly also assist the *Development Strategy* aim of achieving a sustainable future for scholarly information by taking the digital preservation challenge seriously.

Where possible SHERPA will work in partnership with publishers. We will actively seek publisher partners who wish to work with us to help to achieve a hybrid (that is, commercial *and* non-commercial) environment for research preservation and access. For example, we hope some publishers may agree to copyright transfer arrangements which enable copies of articles published in their journals to be deposited in SHERPA open access repositories, and will also create links to these from their own servers. The American Physical Society provides a model for the type of far-sighted publisher with which we hope to work. The APS has recently shaped its policies to accommodate e-print repositories. Alliances will also be sought with organisations such as SPARC. Our aim in SHERPA is to work towards the *rebalancing* of the ‘for fee / for free’ scholarly information environment, but we expect that academic libraries will continue to work within a hybrid environment of commercial *and* freely published research for the foreseeable future, with the commercial content retaining an advantage in the quality of its presentation and its ‘value-added’ functionality.

The proposed model of the provision of free versions of papers will need to be evaluated. The SHERPA evaluation study will examine the culture change dimensions of the Project, as well as measuring the success of the technical deliverables. It will also look at usage of the services in order to assess the possible the long-term viability of the system. Key to the evaluation will be an assessment of the value of the system to researchers in institutions outside the partner sites, with particular attention paid to the value to new universities and those still building up their research profiles.

4. Objectives

- To construct six exemplar institutional e-print repositories in Development Partner sites using eprints.org software and to share widely the skills gained.
- To augment these repositories with at least seven other institutional repositories during the course of the Project.
- To investigate key issues in populating and maintaining e-print collections and share widely lessons learned.
- To specify and assign high quality OAI-compliant metadata to the items in e-print collections which facilitate wide public access to their contents, and to share the standards used.
- To investigate the feasibility of the long-term preservation of e-prints, with particular emphasis on the application of the OAIS Reference Model, and to share the lessons learned.
- To effect a culture shift among academic staff regarding the deposit of papers, both pre-refereed papers and those already accepted for publication in academic journals.
- To evaluate the process of creating, populating, and managing e-print repositories, and assess its value to the entire UK research community and beyond.

5. Key deliverables

- Thirteen institutional OAI-compliant e-print repositories populated by research materials and freely available on the web.
- Published guidelines for setting up e-print repositories covering issues such as metadata, file format and copyright.
- A set of activities to assist other institutions in setting up repositories for the duration of the Project.
- A detailed report on setting up and populating a standards-based long-term digital archive.
- Publicly available advocacy material for communicating with researchers.

² <http://www.jisc.ac.uk/dner/development/IEstrategy.html>

- An evaluation report on the technical, cultural and economic implications of the Project.
- Contributions to the ongoing debate on the future of scholarly communication within the information and academic communities.

6. Assumptions

1. SHERPA does not intend to act as an OAI Service Provider. It is assumed that Service Providers will be available to work with SHERPA (some also funded as part of the FAIR programme) in order to ensure that the created repositories are widely searchable. Preliminary discussions have already been held with the EPrints-UK Project also funded by FAIR.
2. Acquiring content to populate e-print collections will be a significant challenge. It will require a culture shift among existing academic staff who will need to be persuaded that e-print repositories are the right way forward. A considerable part of the Project resource will be concentrated in this area.
3. Current estimates for the rate of deposit for each participating institution are based on a rough understanding of campus outputs of published articles at Edinburgh University of between 2,000 and 3,000 per year.

Year One	Year Two	Year Three	End-of-project total
50	300	400	750

Thus, it is assumed that the six e-print repositories produced by the Development Partners will collect at least a total of 4,500 papers over the three-year project period. In addition, further papers will be added by Associate Partners.

4. Expertise acquired in the course of the CEDARS and Camileon digital preservation projects will be utilised in at least three ways:
 - a. in developing a metadata approach that will simultaneously meet the requirements of OAI-compliant e-print repositories and standards-based digital preservation archives;
 - b. in developing a selection methodology for long-term retention across the e-print collections;
 - c. in developing preferred deposit formats in order to facilitate long-term retention.

7. Project outline

The Project will be carried out over three years and consist of the following:

Work Package 1: Project management

- Objectives To provide management for the complete Project.
To liaise with Project stakeholders.
- Deliverables Production of key project management documents, formation of alliances with OAI Service Providers, final Exit Strategy.
- Timescale Years 1-3
- Issues Issues such as communication between partner sites, the relationship with publishers, the development of mechanisms to advise others and a clear exit strategy will be priorities.

Work Package 2: Setting up the repositories

- Objective To install and customise six institutional e-print repositories in Development Partner institutions using e-prints.org software.
- Deliverables Six institutional OAI-compliant e-print repositories, customised to local requirements. Publicly available documentary materials relating to copyright and IPR issues, metadata standards, and collection development policies and procedures.
- Timescale Year 1
- Issues Key issues to do with metadata, IPR, and collection development will be investigated.

Work Package 3: Populating the repositories

- Objective To acquire content for the repositories set up in WP2.
To carry out an advocacy campaign aimed at local researchers.
- Deliverables Content from a number of disciplines in all of the institutional repositories.
Publicly available advocacy materials.

Timescale Years 1-3
Issues Developing effective communication channels with academics and working with different publication conventions will be given priority.

Work Package 4: Digital preservation: proof of concept

Objective To carry out a major study investigating the key issues in the long-term digital preservation of e-prints, including economic models, selection, ingestion, storage and retrieval.
To carry out 'proof of concept' work in this area.
Deliverables Report on the feasibility of various models.
Some exemplar preservation activity.
Timescale Years 1-3
Issues Developing models which can be scaled up and applied in 'real life' situations will be particularly important considerations.

Work Package 5: Enlargement

Objective To set up at least seven new repositories in Associate Partner institutions.
To populate these repositories.
Deliverables Populated repositories in the Associate Partner institutions.
Timescale Year 2-3
Issues Effectively transferring the lessons learned from Development Partner installations.

Work Package 6: Dissemination

Objective To share learning outcomes widely within the UK HE/FE community and beyond.
To establish mechanisms to advise and assist other institutions in setting up repositories,
Deliverables Dissemination activities including publications, presentations, etc.; including documentation used by the Project Team, such as advocacy material.
Activities providing practical help and support to other institutions.
Timescale Year 1-3
Issues Reaching all stakeholders.

Work Package 7: Assessing the Impact

Objective To carry out formative and summative evaluation of the Project, assessing its impact culturally, organisationally and technically, using extensive feedback from participant sites, and statistical data.
To evaluate the general SHERPA model and its future.
Deliverables Ongoing internal reports plus a final report to JISC.
Timescale Years 2-3
Issues Cultural, economic and management issues will be assessed. Consideration will be given to the viability of the e-prints model in the context of different subject disciplines (and their different conventions of communication) and different types of institution.

8. Track records

8.1. CURL

CURL, the Consortium of University Research Libraries comprises 26 full members, associates and partners, including the British Library. Its vision embraces the aspiration to provide strong leadership and opportunities for innovation for the scholarly library and research communities. CURL has a successful track record in offering services and running development projects on behalf of the HE community. These include COPAC (the CURL OPAC), the HE Archives Hub, and CEDARS.

8.2. The University of Nottingham (lead site)

Information Services (which includes library, computing and media services) at the University of Nottingham has considerable experience in handling major projects. Nottingham is currently the lead site for JISC-funded BIOME Subject Hub, and the RSLP EGIL project. The University was involved as one of two test sites with the European Union funded Liberation project which developed online

learning packages. The institution has also received substantial Non-Formula Funding to develop a number of projects within the Department of Manuscripts and Special Collections. Nottingham is currently a partner in several projects, including the JISC-funded British Education Portal, and RSLP-funded Ensemble, Revelation, Mine of Information, and COCOREES. Nottingham has already developed an experimental e-print repository.

8.3. Other institutional partners

The other institutional partners all have extensive experience of managing and contributing to project work. Several have already gained some expertise in OAI and setting up pilot e-print repositories. Most are members of CURL and therefore have experience of working together in delivering services and managing projects.

8.4. The British Library

As the national library, the British Library has a very extensive digital collection policy, through purchase, voluntary deposit, pilot harvesting of UK domain sites, and digitisation of its own collections (more than forty projects, several with significant external funding). It is developing in parallel its integrated digital architecture and infrastructure, including a multi-million pound OAI-compliant digital library store through a partnership with IBM. It is very active nationally and internationally on the digital preservation and digital standards fronts and is prototyping a mechanism for metadata access and harvesting using Cheshire. It is a lead partner in the EU-funded The European Library (TEL) project, which is developing an international model for the hybrid national library. The British Library has recently signed a strategic alliance with HEFCE, a concordat with AHRB and is working in ever closer co-operation with JISC.

8.5. AHDS

The Arts and Humanities Data Service (AHDS) is a national service funded by the Joint Information Systems Committee (JISC) and the Arts and Humanities Research Board (AHRB), to collect, manage, catalogue, preserve and promote the use of digital resources in research, teaching and learning in the arts and humanities. The AHDS provides advice and guidance in the creation of digital resources to quality standards that ensure their suitability for use in research and teaching and their long-term viability. The AHDS identifies and accessions a wide range of digital resources including texts and e-books, and evaluates, validates, adds metadata, and incorporates the collections into its resource discovery, delivery and preservation systems. In addition to developing significant in-house expertise in preservation issues and processes, the AHDS has been at the forefront of research into digital preservation undertaking two key research projects: an LIC funded project to establish guidelines for digital archiving, and a follow-up BL funded project to create a workbook for the preservation management of digital materials. The latter resulted in the highly acclaimed *Preservation Management of Digital Materials Workbook* published by the British Library. Following significant involvement in the CEDARS project the AHDS is intending to establish a central preservation facility based upon the OAI model. To facilitate this process the AHDS is undertaking a review of its current distributed practice and research into a number of key aspects, including life-cycle management, metadata schema and naming architecture. The outputs from the review will provide guidance and recommendations on key elements to be included in the AHDS preservation facility. Outputs will be published and will also form the basis of a published 'real-life' case study in establishing an OAI preservation archive.

9. Programme of work

It is envisaged that the Project will last for three years, commencing in November 2002. A detailed programme of work including details of deliverables and timescales is given in Appendix 2.

10. Project management

10.1. General

SHERPA will be managed in accordance with JISC project management protocols. It will be informed by formal project methodologies and will conform with recognised good practice in the field.

10.2. Project staff

10.2.1 Project Director

The Project Director will oversee the work of the Project and act as an advocate for it within CURL and other strategic fora. The person identified for this role is Stephen Pinfield, Assistant Director of Information Services at the University of Nottingham, which will be the lead site for the Project.

10.2.2 Project Manager: AR4, 1 FTE (36 months)

Based in the University of Nottingham Information Services, the Project Manager will co-ordinate all Project activity. He or she will manage the Project on a day-to-day basis and have responsibility for budgeting, administration, communication, dissemination, deliverables, and time schedule. The Project Manager will have a good knowledge of e-library and scholarly communication issues and will also have management experience. This post will be a secondment if possible.

10.2.3 Project Officers (Development Partners): AR 2, 1 FTE (36 months)

Each university institutional partner (including the White Rose partnership as a single partner) will host a Project Officer who will work on SHERPA one day a week. The primary responsibility of these Project Officers will be to co-ordinate advocacy work in their own institutions. They will also be involved in general Project dissemination. They will act as the key contacts for the Project in the partner sites.

10.2.4 Technical Officer: AR 2, 1 FTE (24 months)

The Technical Officer will have responsibility for the key technical aspects of the Project. He or she will be involved in the initial installation and customisation of the institutional repositories in partnership with local technical staff. The post will have responsibility for co-ordinating developments on metadata schemas and input procedures. The Technical Officer will also be the key contact with Service Providers (particularly ePrints-UK) and will lead on ensuring the effective exchange of metadata. In addition, the Technical Officer will conduct the research work in digital preservation and will also be involved in carrying out some 'proof of concept' activity in liaison with AHDS.

10.2.5 Project Assistant: Clerical C, 0.5 FTE (36 months)

Providing administrative support for the Project, the Project Assistant's duties will include producing documentation, organising events and carrying out office management.

10.3. Project governance

10.3.1 Management Group

Project developments will be the responsibility of a small Management Group, meeting approximately once every four months. This will be chaired by the Project Director and will comprise the Project Manager and senior representatives of selected partner institutions who have expertise in this area, including John MacColl, Sub-Librarian (Online Services) at the University of Edinburgh.

10.3.2 JISC requirements

SHERPA will comply with JISC requirements for project governance. It is understood that the Project will be monitored by the JISC Executive and steered by the FAIR Programme Advisory Board. The FAIR Advisory Board will report to the JISC Committee for the Information Environment.

The Project will play an active part in its FAIR project 'Cluster' and will make it a priority to establish links with related projects both within and outside FAIR.

10.4. Communication

Communication within the Project team will be facilitated by a combination of face-to-face meetings, telephone conversations and email discussions. To enhance the last of these a closed email discussion list will be set up on JISCmail to include all Project stakeholders. A Project intranet will also be set up to facilitate the sharing of internal documentation.

10.5. Documentation

Internal Project documents and public reports will be produced according to a 'documentation template'. This will be based on the successful CEDARS model and will include standardised layout and design for documents. This will be particularly important for public reports which may be written by different participants.

11. Consultation

Consultation between Project staff and key experts within and outside the UK will occur regularly throughout the funded period to ensure the best advantage is taken of solutions, tools, and insights developed elsewhere. This will include consultation between AHDS and other providers of archiving services and between the Project Manager and other developers of institutionally-based e-print repositories.

12. Promotion

12.1. Definitions

Promotion within the Project is divided here into two main categories: 'advocacy' and 'dissemination'. 'Advocacy' is used here of the promotion of new forms of scholarly communication to academic staff, encouraging them to contribute to e-print repositories. 'Dissemination' is the sharing of the learning outcomes of the Project to the wider HE and FE community and beyond. The two are of course closely related but will need to be planned and managed bearing in mind their different aims.

12.2. Advocacy

One of the key challenges for SHERPA will be getting hold of the content. This will require a significant amount of resource to be used in liaising with authors in the various partner institutions. Strategies of advocacy will be designed both to discuss general scholarly communication issues with staff and to encourage them specifically to contribute content. It will be important to identify immediate benefits for academics in contributing to repositories: for example the setting up of value-added services such as reports on hit counts, and citation analyses (including the use of tools developed by the JISC-NSF-funded OpCit project) will be considered. SHERPA will be informed in these developments by other FAIR projects.

12.3. Dissemination

A key element of the Project will be the sharing of learning outcomes with the wider HE and FE community. Dissemination will be an ongoing activity throughout the Project and will take place within the framework of a Dissemination Strategy devised in the early stages of the Project. Dissemination will be co-ordinated by the Project Manager but will be carried out by all main participants in the Project. Key dissemination activities will include:

- **Project web site:** the web site will be regularly updated by the Project Manager and Project Officers and will include news of new developments. All public project documents will be presented on the web site. In addition, the web site will include links to related sites.
- **Project leaflet:** a leaflet or flyer will be produced explaining the aims and objectives of the Project. This will be appropriate for distribution within institutions and at conferences and meetings.
- **Email announcements:** important milestones will be announced on email discussion lists such as lis-link and the various OAI / digital preservation lists.
- **Publications:** publications will be written for the professional and academic literature.
- **Conference presentations:** various key conferences will be targeted for dissemination opportunities.
- **Project conference:** to be held in the final stages of the Project in Year 3. This will be based on the model of the CEDARS conference held in York in 2000. It will include an **Evaluation Event**.

Efforts will be made to disseminate findings outside of the library and information profession. Opportunities for dissemination amongst archivists, institutional administrators, and academics will be proactively sought. In particular, SHERPA will seek opportunities to discuss issues raised by the Project with publishers.

13. Evaluation

Project evaluation is the subject of an identified Work Package within the Project. Evaluation will be both formative and summative. Formative evaluation will consist of ongoing analysis of the various Project services, and formal consultation with users (library and academic staff in the partner sites). An Evaluation Event will be held in Year 3 as part of the conference which the Project will host. Progress against Project targets will also be reported in each six-monthly report to JISC.

The summative evaluation will employ a range of measures, including interviews with academic authors and a statistical evaluation of the impact of the Project. The study will assess the success of the Project in meeting its own objectives and targets, as well as its wider impact on scholarly publishing. It will also assess the SHERPA model: institutional repositories complemented by standards-based digital preservation activity.

It is likely that the Project evaluation will be carried out by an external consultant. The work will be carried in the context of wider evaluation of the FAIR and other JISC Programmes.

14. After the Project

CURL regards this Project as an opportunity to undertake crucial development and advocacy work, with the opportunity of a longer term, sustainable service continuing after the period of funding. At the end of the 3-year period, the model of distributed repositories complemented by standards-based digital preservation activity will have been tested and, it is hoped, validated. Lines of communication with both academics and publishers will also have been established. CURL will have undertaken this activity on behalf of the research and learning community. It is hoped that the resulting SHERPA deliverables can then form a plank of the emerging Information Environment and can be scaled up as appropriate. An Exit Strategy, with a costed plan for a sustainable service, will form one of the deliverables of the Project.

15. Budget

[This section of the proposal has been removed from the public version of the document]

16. Evaluators

[This section of the proposal has been removed from the public version of the document]

17. Key contacts for Project proposal

Dr Paul Ayris
Director of Library Services
Education and Information Support Division
University College London
Gower Street
London WC1E 6BT
Tel: 020 7679 7834
Fax: 020 7679 7373
Email: p.ayris@ucl.ac.uk

Stephen Pinfield
Assistant Director of Information Services
Research and Learning Resources Division
Hallward Library
University Park
University of Nottingham
Nottingham NG7 2RD
Tel: 0115 951 5109
Fax: 0115 951 4558
Email: Stephen.Pinfield@nottingham.ac.uk

Version 2: 15 October 2002

Revised for public dissemination: 6 November 2002

Appendix 1: Open Archival Information System

The Open Archival Information System (OAIS) Reference Model began as a response to NASA's need for a digital archive for its space data. Though it originated in the space community, the fact that it has always been an 'open' system meant that other communities, such as traditional archives and libraries, have always actively participated in its development. A series of international workshops and active involvement of a broad cross section of the international archiving/user communities means that it has been widely adopted by the digital preservation community. It has always been stressed that OAIS is a reference model. It provides a framework for understanding and applying concepts needed for long-term digital preservation. It does not specify an implementation but it does supply the terminology and framework to enable a variety of implementations. The framework specifies generic components required to manage both the digital objects and their associated metadata. A number of implementations now exist, for example, the Royal Dutch National Library used OAIS as the basis of its Deposit System for Electronic Publications (DSEP). The British Library is also using it for its own Digital Library Store, and the CEDARS project has used it as the basis for its metadata specification and the architecture for its proof-of-concept archive. This means that there is a growing pool of practical expertise which can be tapped into and a growing recognition of the importance of the OAIS model. In addition, an RLG/OCLC Working Group is preparing a report specifying the minimum requirements of an OAIS-compliant digital repository.

Appendix 2: Programme of Work

Work Package 1: Project management

Key activities	Duration	Deliverables	Milestones	Completion
Production of Plan (in line with the Project Management Protocol)	Months 1-3	Project Plan	Project Plan complete	End of Month 3
Creation of the Project documentation template	Month 1-3	Documentation template	Template complete	End of Month 3
Production of key project documentation including Quality Assurance guidelines, Consortium Agreement, Risk Strategy etc.	Months 1-4	Quality Assurance guidelines, Consortium Agreement, Dissemination Strategy	Project documentation complete	End of Month 4
Development of the Project web site and intranet	Months 1-3	Project web site and intranet	Project web site and intranet launched	End of Month 3
Management of Project meetings structure, including the Management Group, and participation on Programme meetings	Months 1-36	Formal records of meetings	Meetings	End of Month 36
Negotiations with OAI Service Providers, including ePrints-UK	Months 1-36	Agreements with Service Providers with metadata being harvested	Agreements in place and search services set up	End of Month 36
Project documentation as specified by JISC	Months 1-36	Twice yearly progress reports and final report	Delivery of reports to JISC	End of Month 36
Manage the Enlargement process	Months 18-36	See WP 4	See WP 4	End of Month 36
Develop Exit Strategy	Months 25-36	Exit Strategy document	Exit Strategy complete	End of Month 36

Work Package 2: Setting up the repositories

Key activities	Duration	Deliverables	Milestones	Completion
Installation of six prototype repositories	Months 1-6	Six prototype repositories	Prototype repositories installed	End of Month 6
Customise repository interfaces	Months 1-6	Customised interfaces	Interfaces complete	End of Month 6
Populate each repository with test papers	Months 3-8		Test papers in place	End of Month 8
Examine and report on copyright/IPR issues; establish copyright/IPR guidelines	Months 3-12	Report on copyright/IPR issues; public copyright guidelines	Report complete; public guidelines available on the Project web site; policy in place	End of Month 12

Examine and report on metadata format/quality issues, including consultation with key Service Providers (such as ePrints-UK)	Months 3-12	Report on metadata issues; public guidelines	Report complete; public guidelines available on the Project web site; policy in place	End of Month 12
Establish collection development processes	Months 3-12	Collection development policy and procedures	Policy and procedures in place; available on the web site	End of Month 12
Installation of six production repositories	Months 12-18	Six production repositories	Production repositories installed	End of Month 18

Work Package 3: Populating the repositories

Key activities	Duration	Deliverables	Milestones	Completion
Development of advocacy materials	Months 3-12	Advocacy materials (made publicly available)	Suite of advocacy materials available on Project web site	End of Month 12
Advocacy activities in partner institutions	Months 4-36	Ongoing advocacy activities		End of Month 36
Discussions with publishers	Months 4-36	Agreements with publishers	Agreements in place	End of Month 36
Population of repositories with c. 50 papers	Months 4-12		c.50 papers in place	End of Month 12
Continued population of the repositories	Months 13-36	Repositories containing 4,500 papers	Papers in place	End of Month 24 End of Month 36

Work Package 4: Digital preservation: proof of concept

Key activities	Duration	Deliverables	Milestones	Completion
Analysis of preservation standards	Months 6-12	Internal report	Policy on appropriate standards complete	End of Month 12
Specification of metadata quality standards for archives	Months 6-18	Metadata quality standards document	Document made available on the Project web site	End of Month 18
Development of preservation selection criteria (including estimate of expected total paper count)	Months 3-12	Criteria specification document	Document complete and available on the Project web site	End of Month 12

Development of full standards-based compliance specifications (including view of document life cycle)	Months 12-24	Specification document	Document complete and available on the Project web site	End of Month 24
Development of cost models	Months 12-24	Costing document	Document complete and available on the Project web site	End of Month 24
Development of proof of concept demonstrator	Months 12-30	Demonstrator	Demonstrator complete	End of Month 30
Specification on scaling up preservation activity	Months 24-36	Specification document	Document complete and available on the Project web site	End of Month 36

NB The precise level of work carried out in this area will depend on whether any additional funding is available from the forthcoming JISC Digital Preservation Programme.

Work Package 5: Enlargement

Key activities	Duration	Deliverables	Milestones	Completion
Identification of Associate Partners (in partnership with the CURL Board)	Months 1-10		Partners identified	End of Month 10
Induction of Associate Partners into Project and agreement on support/communication arrangements and content targets	Months 12-18	Expanded Consortium Agreement	Expanded Consortium Agreement in place	End of Month 18
Installation and customisation of Associate Partner repositories	Months 18-30	Seven new repositories	Repositories set up	End of Month 30
Population of repositories according to agreed targets (see WP 3)	Months 18-36	Populated repositories	Repositories populated in line with agreed targets	End of Month 36
Planning for possible further enlargement	Months 30-36	Plan on further enlargement	Enlargement plan publicly available on the Project web site	End of Month 36

Work Package 6: Dissemination

Key activities	Duration	Deliverables	Milestones	Completion
Development of a Dissemination Strategy	Months 1-3	Dissemination Strategy document	Dissemination Strategy completed	End of Month 3
Development of the web site as a dissemination tool	Months 3-36	Up-to-date web site	Web site available	End of Month 36
Development of Project dissemination material e.g. leaflets	Months 3-24	Materials	Materials available	End of Month 24
Carrying out various dissemination activities e.g. publications, conference papers etc.	Months 1-36	Publications etc	Various activities completed.	End of Month 36

Develop mechanisms for advising other institutions on setting up repositories	Months 13-36	Mechanisms and materials for advising other institutions	Service in place and carrying out work	End of Month 24 (continuing to Month 36)
Project conference – planning and hosting	Months 20-36	Conference	Conference completed	End of Month 36

Work Package 7: Assessing the impact

Key activities	Duration	Deliverables	Milestones	Completion
Negotiations with consultants	Months 6-12	Agreement	Agreement in place	End of Month 12
Design of evaluation specification	Months 13-16	Evaluation specification document	Specification document in place	End of Month 16
Formative evaluation activity	Months 16-30	Evaluation reports	Reports completed	End of Month 30
Summative evaluation activity	Months 24-33	Summative report	Report completed	End of Month 33
Writing of final report to JISC	Months 33-36	Report to JISC	Report approved by CURL Board and JISC	End of Month 36

Appendix 3: Profiles

Paul Ayris

Dr Paul Ayris is Director of Library Services at University College London. He is also chair of the CURL Task Force for Scholarly Communication, a group which has co-ordinated the SHERPA proposal. Dr Ayris previously worked at Cambridge University Library in a number of roles, including Head of IT Services in the Automation Division. There, he was responsible for developing the Union Catalogue project into a full university-wide library automation system, open to all departmental and college libraries. Dr Ayris is an Honorary Lecturer at the School of Library, Archive and Information Studies in UCL. He also has a research interest in late medieval and early modern English history. He is Secretary of the Society for Reformation Studies and Executive Editor of *Reformation & Renaissance Review*, which is published by Sheffield Academic Press.

John MacColl

John MacColl is Director of the SELLIC Project (Science & Engineering Library, Learning & Information Centre) and Sub-Librarian (Online Services) at the University of Edinburgh. He has worked in various roles in academic libraries, computing services and converged information services in the Universities of Glasgow, Aberdeen, Abertay Dundee and Edinburgh. He was Project Officer for Project Jupiter from 1989-1990, promoting the JANET network to university libraries. From 1995-1997 he was Project Director and Managing Editor of the eLib *ARIADNE* Project and publication. Since 2000 he has been a Co-Director of the JISC-funded ANGEL Project which is developing middleware to integrate learning environments and digital library systems. He is currently a member of the JUGL Committee and serves on the CURL Task Force on Scholarly Communication. He has published a number of articles and reviews in the professional press, and co-edited *Delivering the electronic library: an ARIADNE reader* (ARIADNE Project, 1999).

Stephen Pinfield

Stephen Pinfield is Assistant Director of Information Services at the University of Nottingham where he has already managed the setting up of an experimental e-print server. Before coming to Nottingham, he was Project Leader on the eLib *BUILDER* Project at the University of Birmingham. Whilst working at Nottingham, he has taken on the role of part-time JISC Consultant with responsibility for the eLib programme. This has included work to capture, communicate and further develop key issues from eLib. He has produced a number of reports and publications on this and other topics. He is a member of the CURL Task Force for Scholarly Communication.

Appendix 4: Key Acronyms and Initiatives

AHRB – Arts and Humanities Research Board: <http://www.ahrb.ac.uk/>
APS – American Physical Society: <http://www.aps.org/>
Archives Hub: <http://www.archiveshub.ac.uk/>
arXiv: <http://arxiv.org/>
BIOME: <http://biome.ac.uk/>
Camileon – Creative Archiving at Michigan and Leeds: Emulating the Old on the New:
<http://www.si.umich.edu/CAMILEON/>
CEDARS – CURL Exemplars in Digital Archives: <http://www.leeds.ac.uk/cedars/>
CogPrints – Cognitive Sciences E-print Archive: <http://cogprints.soton.ac.uk/>
COPAC – CURL Online Public Access Catalogue: <http://www.copac.ac.uk/copac/>
CURL – Consortium of University Research Libraries: <http://www.curl.ac.uk/>
Dublin Core: <http://dublincore.org/>
eprints.org: <http://www.eprints.org/>
HEFCE – Higher Education Funding Council for England: <http://www.hefce.ac.uk/>
IPR – Intellectual Property Rights
OAI – Open Archives Initiative: <http://www.openarchives.org/>
OAIS – Open Archival Information Systems: <http://ssdoo.gsfc.nasa.gov/nost/isoas/>
OpCit – the Open Citation Project: <http://opcit.eprints.org/>
RDN – Resource Discovery Network: <http://www.openarchives.org/>
RLG – Research Libraries Group: <http://www.rlg.org/>
RSLP – Research Support Libraries Programme: <http://www.rslp.ac.uk/>
SCONUL – Society of College, National and University Libraries: <http://www.sconul.ac.uk/>
SPARC – Scholarly Publishing and Academic Resources Coalition: <http://www.arl.org/sparc>
UKOLN – UK Office for Library and Information Networking: <http://www.ukoln.ac.uk/>

