

Cover Sheet for Proposals <i>(All sections must be completed)</i>			
Name of Capital Programme: Repositories and Preservation Programme			
Bid for Calls : (Please tick ONE BOX ONLY, as appropriate)			
Discovery to Delivery and Interoperability Demonstrators (Strand C)			
Call I – Interoperability Demonstrators		a) Interoperability demonstrators	
Name of Lead Institution: London School of Economics and Political Science			
Name of Proposed Project: Federated Access to Repositories (FAR)			
Name(s) of Project Partner(s): University of Cambridge			
Full Contact Details for Primary Contact:			
Name: John Paschoud			
Position: Infosystems Engineer and Projects Manager			
Email: j.paschoud@lse.ac.uk			
Address: Library London School of Economics & Political Science 10 Portugal Street London WC2A 2HD			
Tel No: 020 7955 6129			
Fax No: 020 7955 7454			
Length of Project: 9 months			
Project Start Date: 01/11/2007		Project End Date: 31/07/2008	
Total Funding Requested from JISC: £82,211			
Funding Broken Down over Financial Years (Mar – Apr):			
Sep07 - Mar08		Apr08 – Mar09	
£45,180		£37,031	
Total Institutional Contributions: £85,977			
Outline Project Description			
<p>A workshop at the OAI5 conference in April 2007 identified a need for improved attribute managed federated access to repositories, and produced a set of requirements. The Federated Access to Repositories (FAR) project would take up this list, and use it to:</p> <ul style="list-style-type: none"> ● Create recommendations for attributes to describe authorisation to repositories; where appropriate, these recommendations will be passed to relevant standards groups ● Develop extensions to the repository software (EPrints and DSpace) so that it can be used with Shibboleth and meet these requirements (building on earlier work where this exists) ● Test existing work integrating Shibboleth and Fedora and recommend to developers what will be needed to meet these requirements ● Install demonstration repositories to show how this can work in practice ● Feed the modifications and full documentation into the appropriate repository software development process to ensure maintenance through future releases of Eprints and Shibboleth ● Develop procedures to consider ways of extending federated access management to new functionality introduced into the repository software products following the conclusion of the project <p>A demonstrator of the software will be produced as well as a final version.</p> <p>This project falls under the Interoperability Demonstrators area of the call.</p>			
I have looked at the example FOI form at Appendix A and included an FOI form in the attached bid		YES (see Appendix A of this document)	
I have read the Circular and associated Terms and Conditions of Grant at Appendix B		YES	

I Introduction - Appropriateness and Fit to Programme Objectives and Overall Value to the JISC Community

1. This proposal is submitted by the London School of Economics and Political Science (LSE) in response to the Interoperability demonstrators strand of the Repositories and Preservation Call. The proposed project will produce software to be incorporated into the EPrints¹ and DSpace² repository software products. This software will enable Shibboleth³ and appropriate attribute repositories to be used to fulfil the requirements⁴ developed by the community at a workshop at the OAI5 Conference⁵ held on 18 April 2007. These requirements address the use of federated access management in the context of the various use cases for repository access, including open access use cases, the special roles for privileged access to an institutional repository, and content-related access restrictions. Important scenarios identified include that of usage by cross institutional collaborations and virtual organisations (such as many research groups) and that of the use of federated access management in conjunction with the embargo function now present in all three products. Cross-institutional collaboration in research groups has been important for many years, and this is therefore one area in which the work of the FAR Project will benefit the JISC community. It continues the work of JISC in extending Federated Access Management take-up in the UK HE community, and so brings together two of JISC's strategic themes, eResearch and Federated Access.
2. Several use cases could be suggested based on the scenarios from the OAI5 conference which would benefit from the type of access management described in the OAI5 requirements list. The exemplary use case for the FAR Project will be that of the distributed research group considered as a virtual organisation. Research groups often contain individuals from several higher education organisations, frequently from different countries. Participation in such a group may be fairly formal or informal (depending on whether they are formed in response to funding awards or consist of researchers with a particular common interest); but in either case it is likely to vary with time. In many cases, such groups will have resources in common, such as published papers, work in progress, or data sets, which could naturally be stored in a repository. Whether this repository is solely for the research group or is a subcollection of an existing institutional repository, these varied items will have complex access requirements (including the possibility of management separate from that of other collections in the repository in the second instance) and a need for users from a variety of institutions, a situation tailor-made for federated access.
3. The FAR Project will also analyse the scenarios presented in the RepoMMAN Project's User Needs Analysis⁶ published in March 2007. Several scenarios described in this document require finely granulated authorisation decisions, though of these some are not applicable with current EPrints/DSpace functionality (such as the annotation of repository objects described in Scenario 6). Another source of requirements to address will be the OAI-ORE Thoughts on Compound Objects⁷ published in May 2007.
4. The project team will not undertake development Fedora⁸ document management software, but will test the software extensions produced by the RAMP project⁹ for Shibboleth integration, and will feed back information about unmet requirements from the OAI5 list to the Fedora community. Working closely with this project and with the RepoMMAN project will be facilitated with money budgeted for Fedora consultancy below.

¹<http://www.eprints.org/software/>

²<http://www.dspace.org/>

³<http://shibboleth.internet2.edu/>

⁴<https://gabriel.lse.ac.uk/twiki/bin/view/Projects/AmTutorialOaiCernAmlrUseCases>

⁵<http://oai5.web.cern.ch/oai5/>

⁶<http://www.hull.ac.uk/esig/repomman/downloads/R-D14-full-user-needs1.pdf>

⁷<http://www.openarchives.org/ore/documents/CompoundObjects-200705.html>

⁸<http://www.fedora.info/index.shtml>

⁹<http://www.ramp.org.au/ramp-s-authentication-and-authorization-modules-for-fedora>

5. The objectives of the project will be to:

- Create recommendations for attributes to describe authorisation to repositories; where appropriate, these recommendations will be passed to relevant standards groups
- Develop extensions to the repository software (EPrints and DSpace) so that it can be used with Shibboleth and meet these requirements (building on earlier work where this exists)
- Test existing work integrating Shibboleth and Fedora and recommend to developers what will be needed to meet these requirements
- Install demonstration repositories to show how this can work in practice
- Feed the modifications and full documentation into the appropriate repository software development process to ensure maintenance through future releases of Eprints and Shibboleth
- Develop procedures to consider ways of extending federated access management to new functionality introduced into the repository software products following the conclusion of the project

A demonstrator of the software will be produced as well as a final version.

6. An important aspect of the project will be to aim to ensure that the code enabling Shibboleth integration does not become obsolete with updates to either Shibboleth or the repository software, whether EPrints or DSpace; the project developers will liaise with the teams of repository developers and the Internet2 developers to ensure that this is the case as far as is possible. The project will also maintain contact with a range of national federations to keep the project in line with the latest thinking on cross-federation access to resources.
7. The FAR Project will address a specific area of requirement identified by the DISK-UK DataShare Project¹⁰, in which the LSE is a partner. The FAR Project will focus on the access management requirements needed for the storage and sharing of data sets in today's research environment where the concept of "open data" is still being developed and new solutions are needed. To this end, the project will fund a small proportion of the time of the LSE staff working on the DataShare project in addition to that funded by the DataShare project itself, to enhance collaboration. The DataShare work on the interface to the repository for dataset management would enable the expansion of part of the use case from paragraph 2 by the following. A researcher who produces and deposits data in the repository has the option to specify a set of access conditions for the dataset. Consequently the researcher can make the resource available to another individual registered on the repository or a group of people (e.g. a research group, department or School).
8. It is intended that this project will start on 1 November 2007, and be completed by 31 July 2008.

II Quality of Proposal and Robustness of Workplan

Project Partners

9. The **London School of Economics** project team has a long history of involvement with Shibboleth, dating back to the first UK installation of the software. This work includes several projects in which integration of Shibboleth with other software played a major part. These include the Secure project, which looked at a variety of different tools and how they might be used in conjunction with Shibboleth, including uPortal, Endeavor's library management system, and Internet2's Signet and Grouper.
10. Of particular importance as background to this bid, the team led the ShibboLEAP project. There were two main strands to this project: to help a range of London University institutions install Shibboleth Identity Provider software, and to produce a version of EPrints which

¹⁰http://www.jisc.ac.uk/whatwedo/programmes/programme_rep_pres/repositories_sue/datashare.aspx

Federated Access to Repositories (FAR) Project Proposal

permitted Shibboleth login. The final report on the EPrints strand¹¹ describes how the project was partly successful in that a version of the EPrints 2.1.13.1 software was produced which could use Shibboleth for authentication and authorisation, but it did not address the wider range of use cases identified at the OAI5 workshop and was not integrated into the EPrints development cycle, including the standard EPrints authentication code which has been changed with the release of version 3.0. The current bid will take the development a step further by ensuring that the OAI5 requirements are fully addressed and code produced is fully integrated into the EPrints development lifecycle.

11. The team has also been involved in a sequence of projects focusing on user attributes and identity management generally, which provide experience relevant to Workpackage 1. This provides background to Workpackage 1 of this proposal in particular. This includes the JISC UK Scoping Study for eduPerson and related schemas, which aimed to establish 'grass roots' requirements in HE, FE and other UK education sectors for schema describing institutional students, staff and other users. More recently and more significantly, the LSE and Cardiff University are currently lead partners in The Identity Project (2006-2007), which is investigating the current state of identity management in UK higher education institutions, with particular attention to determination of whether business practices meet the requirements of federated access to resources.
12. As described in more detail below (in paragraph 24), the LSE also brings considerable experience of working with international colleagues to the project, thus enhancing the wide applicability of the project work and strengthening the reputation of the JISC programme internationally.
13. The LSE maintains an EPrints repository, as mentioned above.
14. The DSpace@Cambridge team, based at the **University of Cambridge**, started as a three-year collaboration between Cambridge University Library, Cambridge University Computing Service, and MIT Libraries, funded by a grant from the Cambridge-MIT Institute. From July 2006 it became a strategic service of the University of Cambridge, and has the objective of developing an institutional repository that will preserve and disseminate digital materials created by or associated with the University of Cambridge. DSpace@Cambridge has been heavily involved in DSpace development during the project and this will continue through the lifetime of the FAR project. Cambridge has already implemented a Shibboleth server for access to e-resources, which will go into production in conjunction with the Shibboleth Athens gateway in September.

Workplan, timetable and deliverables

Work packages are the responsibilities of the indicated partners.

WP	Name and description	Timetable	Deliverables
0	Project Management (LSE) Ensure efficient and timely delivery of project and production of deliverables detailed in Workpackages 1-4. Liaison with JISC; management of website and dissemination	Project months 1-9	Signed Consortium Agreement (0.1), Detailed project plan (0.2), Final Report, which would describe the experience gained in WP1-4 (0.3); this would be the main documented learning outcome of the FAR project, and would contain conclusions about the process
1	Attribute development (LSE, Cambridge) Create recommendations for	Project months 1-4	Report describing attributes and their use in terms of repository access (1.1); LDAP server

¹¹<http://www.angel.ac.uk/ShibboLEAP/deliverables/finalreports/EPrints.pdf>

Federated Access to Repositories (FAR) Project Proposal

	attributes to describe authorisation to repositories. Where appropriate, these recommendations will be passed to relevant standards groups		containing sample users with these attributes (1.2)
2	EPrints demonstrator development (LSE) Develop demonstrator version of the software to extend EPrints so that it can be used with Shibboleth and meet these requirements; work with test users (EPrints installation managers) to ensure that the demonstrator is compatible with the requirements of a variety of EPrints installers	Project months 3-6	Demonstration system running demonstrator software for EPrints (2.1); draft documentation for EPrints (2.2)
3	EPrints final version development (LSE) Enhance functionality of demonstrator to ensure ease of installation and configuration; integration with main EPrints development process; produce and test documentation for installers	Project months 5-9	Software release as part of EPrints or as standard extension (3.1); final version of documentation easily accessible from EPrints web site (3.2); report describing the future place of the software developed in the FAR project in the long term EPrints development programme (3.3)
4	DSpace demonstrator development (Cambridge) Evaluate existing software enabling use of Shibboleth with DSpace; Carry out parallel tasks to those listed for EPrints in WP 2.	Project months 3-6	Demonstration system running demonstrator software for DSpace (4.1); draft documentation for DSpace (3.2)
5	DSpace final version development (Cambridge) Carry out parallel tasks to those listed for EPrints in WP 2.	Project months 5-9	Software release as part of DSpace or as standard extension (5.1); final version of documentation easily accessible from DSpace web site (5.2); report describing the future place of the software developed in the FAR project in the long term DSpace development programme (5.3)
6	Fedora Federated Access Study (LSE) Evaluate existing software enabling use of Shibboleth with Fedora with respect to Federated Access requirements.	Project months 5-9	Report outlining any unmet requirements and suggestions on software modifications that would help meet them (6.1)

Federated Access to Repositories (FAR) Project Proposal

15. The timetable may be subject to change to fit in with the EPrints¹² and DSpace¹³ development cycles which are not yet fixed as far in advance as 2008.
16. A server will be purchased to carry out Workpackage 2, and will be used for the development work in Workpackages 3-7.

Project Management and Governance

17. The FAR project will be managed and supported administratively by a project team based at the London School of Economics. Close co-operation will be maintained throughout the project with the development teams for the selected repository software products. A Linux server will be purchased, installed and set up at the LSE library, which will be used by the developers of the software for the project, which may include members of the main repository software development team in addition to staff from project partners. Software installed on this server will include Shibboleth (both identity provider and resource provider), the repository servers and an LDAP server. The documentation will be developed on a subsection of the LSE Projects Team WIKI¹⁴, and will be transferred to the appropriate section of the various repository software WIKIs upon completion of the project. Project management software already in use by the LSE Projects Team will be used to organise and monitor the tasks within the project.
18. The Project Manager will carry out the administrative management of the project and the Senior Programmer the technical management. The DSpace programmer will continue to be line managed by Dspace@Cambridge staff, but the time s/he contributes to the project will be directed by the Senior Programmer.
19. The governance of the project will be through a board appointed from senior staff from each partner and those involved in working on the project. This group will meet at least twice during the project and will be asked to sign off the deliverables.
20. A consortium agreement will be drafted to be signed by each partner.

Risks

Risk	Action to prevent/manage risk
Staffing: Unavailability of team members	The development process will be specified in detail, which will enable programmers to be replaced quickly if necessary
Organisational: One of the repository software projects engaged with by FAR release timetable changes in a way that is incompatible with project timetable	Ensure close links with relevant development teams so that FAR project has ample warning and can amend timetable accordingly
Organisational: completed software is unmaintainable after project conclusion	Maintaining close links with repository development teams and paying attention to details of extension mechanisms within the software together with high quality documentation should make it maintainable by respective communities.
Organisational: completed software proves incompatible with Shibboleth developments	Developing the software to interoperate with Shibboleth 2.0 (expected release date during 2007) should ensure long term compatibility; the LSE Projects Team has and will maintain close links with the Internet2 Shibboleth developers and with the UK Access Management

¹²<http://wiki.eprints.org/w/Plan>

¹³<http://wiki.dspace.org/index.php/RoadMap>

¹⁴<https://gabriel.lse.ac.uk/twiki/bin/view/Main/WebHome>

Federation, which means that they will be aware of the trends in Shibboleth development and UK usage of the software
--

Intellectual Property and Accessibility

21. All published material produced by the project will follow current JISC (or JISC-endorsed) guidelines for accessibility or will adopt those used by the relevant software projects. Copyright of the appropriate partner institution will be asserted but all material will be covered by the appropriate Creative Commons licence allowing free non-commercial re-use. Software will be released with the appropriate open source licence, which is the GNU General Public Licence version 2¹⁵ in the case of EPrints and the BSD Licence¹⁶ in the case of DSpace.

Sustainability

22. The final release of the software will be fed into the EPrints/DSpace development programme as appropriate, and will be sufficiently well documented that maintenance by the EPrints/DSpace community will be possible. The contacts that the Project team have and will make within those communities will be used to ensure the maintenance of the software. (See also the Risk section.)

III Engagement with the Community

Consultation and Co-operation with Stakeholders

23. The EPrints development team at Southampton University have been involved in the drafting of this bid, and close contact will be maintained throughout the FAR Project lifespan. Additionally, the part of the proposed budget for the project which is set aside for evaluation will be used to fund time from members of the repository communities or other qualified independent persons to determine that the software produced by the FAR Project is fit for purpose.

24. Test users will be recruited as widely as possible; interest from the OAI5 workshop mentioned above (which is an important piece of formative user needs analysis for the FAR project) suggests that repository managers worldwide will be interested in the FAR project's progress. The LSE is also a founder member and lead site of the Nereus consortium¹⁷ of leading European economics information providers, and led the recently completed VERSIONS project¹⁸ which had researcher input from the other consortium members. The other partners in the Nereus consortium will be stakeholders in the outcome of the FAR project and will provide test users for the demonstrator. Additionally, the LSE library has close links with Monash University library in Australia, which leads the Arrow repository project¹⁹ using Fedora software, and they will be asked for input into each workpackage but particularly into Workpackage 6.

25. These internationally recruited test users are intended to comment on how well the software demonstrator matches their expectations of a system which is intended to meet the federated access requirements from the workshop, and so it will be appropriate to recruit repository managers to this role. Such a wide body of test users should enable the FAR project to look at the access requirements of workflow in a variety of disciplines. Since some of the requirements need inter-federation co-operation, international contacts will be important to provide test usage which enable the procedural and technical barriers to such

¹⁵<http://www.gnu.org/copyleft/gpl.html>

¹⁶<http://www.opensource.org/licenses/bsd-license.php>

¹⁷<http://www.nereus4economics.info/>

¹⁸<http://www.lse.ac.uk/library/versions/about.html>

¹⁹<http://www.arrow.edu.au/>

Federated Access to Repositories (FAR) Project Proposal

use to be tested realistically. The additional funds requested for evaluation will be used to ensure that the software meets technical standards: that it is fit for purpose.

26. Contacts will be maintained with the product software development teams, to ensure that ideas developed in the FAR project will be applicable in each repository community. The Australian Partnership for Sustainable Repositories²⁰ has an interest in the integration of Shibboleth and repository software, and communication will be maintained with their project team throughout the project.
27. The LSE Projects Team has a long history of working with Shibboleth, and has close links with the Shibboleth developers at Internet2, the UK Access Management Federation and with Shibboleth federations and projects across the world. This provides a pool of expertise that the FAR Project would be able to call upon for advice if Shibboleth issues developed beyond the (considerable) knowledge of the project team. In particular, the development of federation policy covering inter-federation co-operation and related technical advances will be closely followed.
28. Contact will be maintained with other projects funded under the call to ensure that the lessons learnt in other projects are applied by the FAR team.
29. The LSE Library runs a repository which uses the EPrints software²¹, and the library staff who manage the repository will be closely involved with the work of the FAR project.

Dissemination

30. Dissemination will be principally to appropriate repository and digital library mailing lists. Newsletters will be published to appropriate mailing lists at at least four times during the project lifetime.
31. Information about the project will be published, and regularly updated, on a website. The main deliverables of the project will also be linked to from the website (some, as noted above, are likely to appear on the EPrints website or WIKI). At least one paper describing the project will be submitted for publication to an appropriate journal, or given as a presentation at an appropriate conference.
32. The LSE Projects Team has a lengthy history of working on dissemination activities in partnership with JISC, and will continue to do so for the FAR Project. The project will also aim to present at other relevant conferences and seminars throughout the project timespan.

IV Budget

Directly Incurred Staff	April 07– March 08	April 08– March 09	TOTAL £

²⁰<http://www.apsr.edu.au/index.html>

²¹<http://eprints.lse.ac.uk/>

Federated Access to Repositories (FAR) Project Proposal

Percentage Contributions over the life of the project	JISC 49%	Partners 51%	Total 100%
--	---------------------	-------------------------	-----------------------

Institutional contributions

33. The partners expect to contribute to the project in the form of full indirect overhead costs of their staff involved, with salary costs requested from JISC.

V Key Personnel

34. **John Paschoud** (LSE Library projects team) would act as Project Director and Manager, contributing also to project management tasks, liaison with other projects, JISC and partner institutions, and dissemination on behalf of the project. He is the InfoSystems Engineer and Projects Manager at the Library of the London School of Economics & Political Science (LSE). The Library Projects Team which he heads undertook the original evaluation for the UK JISC of potential new Access Management technologies, recommending that attention should be focussed on the SAML standard and the Shibboleth implementation by Internet2. The Team subsequently undertook a similar evaluation for SURF in The Netherlands. LSE had the first operational Shibboleth installation outside of North America, and the Library Projects Team has been fully involved in subsequent JISC programmes to develop a new Access Management infrastructure for UK education. John is a member of the UK Federation Technical Advisory Group, and is also involved in liaison on behalf of the UK Higher Education community with national bodies developing Access Management infrastructures in other countries. He is currently part-seconded as a member of the JISC Outreach Team supporting institutions, publishers and technology vendors in their adoption of Federated Access Management, and is also responsible for part of the JISC-funded *Identity Project*, developing models for auditing Identity Management practices in HE institutions.

35. **Simon McLeish** (LSE Library projects team) would act as the senior programmer on the project. He has acted as Chief Technical Officer for ShibboLEAP, SECURe, PERSEUS and other major projects dealing with the management of resources, users and access. He is maintaining author of the draft 'UKeduPerson' schema, and is a contributing participant in the Open Source Software development projects for Shibboleth, Signet and uPortal. Simon was responsible for implementing the first working Shibboleth Origin outside of North America.

36. **Frances Shipsey** (LSE Library, eServices Librarian) would provide expertise in digital repositories as an institutional repository manager. She has acted as Project Manager for the JISC-funded VERSIONS Project and is a member of the Version Identification Framework Project Steering Committee. She has been LSE's project officer for Economists Online, a digital repositories project conducted by Nereus, the European consortium of economics research libraries.

37. *The contribution of Frances Shipsey to the project would be part of the LSE institutional contribution. The librarian post would be filled partly by secondment from a member of the repository management team reporting to Ms Shipsey and partly by LSE staff working on the Datashare project (see paragraph 5)..*

38. Personnel to fill the project administrator role will be seconded from existing staff or advertised if necessary.

VI Outcomes and benefits

39. The principal outcome of the project will be the software extensions to the repository products and the accompanying documentation. The other deliverables, particularly deliverable 1.1, are also items which are likely to be of benefit to the repository community. There are clear benefits to those institutions which are intending to install or which have

Federated Access to Repositories (FAR) Project Proposal

installed Shibboleth for access to electronic resources as these installations will provide single sign on with their repositories, and allow them to use Federated Access Management for cross-institutional access requirements for repositories (e.g. for storage of documents related to collaborative projects).

VII Appendix A: Freedom of Information

- A1. We would like JISC to consider withholding the following sections or paragraphs from disclosure, should the contents of this proposal be requested under the Freedom of Information Act, or if we are successful in our bid for funding and our project proposal is made available on JISC's website.
- A2. We acknowledge that the FOI Withheld Information Form is of indicative value only and that JISC may nevertheless be obliged to disclose this information in accordance with the requirements of the Act. We acknowledge that the final decision on disclosure rests with JISC.

Section / Paragraph No.	Relevant exemption from disclosure under FOI	Justification
None.		

Federated Access to Repositories (FAR) Project Proposal



Library

United Library of Political
and Economic Science

21 June 2007

Neil Jacobs
Programme Manager
JISC Executive

10 Portugal Street
London WC2A 2HD
tel: +44 (0)20 7955 7218
fax: +44 (0)20 7955 7454
email: j.sykes@lse.ac.uk
www.library.lse.ac.uk

Jean Sykes
Librarian and Director
of IT Services

Dear Mr Jacobs

Re: Project FAR, JISC Repositories and Preservation Programme

I am delighted to support this proposal for a project to be carried out at the LSE, for JISC funding to fulfil needs identified by the repositories community worldwide for federated access to repositories.

By improving the means by which researchers can collaborate across institutions and control access to the material they have deposited in institutional repositories generally, this project will have clear benefits for our own institution.

Equally clearly, the fact that these requirements were generated by the research community indicates that meeting them will be of benefit throughout this community, including the UK HE community served by JISC.

The Projects Team at the LSE Library is exceptionally experienced in work of this nature, having already played a large part in the development of Federated Access Management in UK HE, laying the foundations for the next generation access to research material which is the aim of the FAR Project.

The LSE can also bring considerable experience of working with international colleagues to this project, thus enhancing the wide applicability of the project work and strengthening the reputation of the JISC programme internationally.

I urge JISC to support this project.

Yours sincerely

Jean Sykes
Librarian and Director of IT Services



INVESTORS IN PEOPLE

The London School of Economics and Political Science is a School of the University of London. It is a charity and is incorporated in England as a company limited by guarantee under the Companies Act (Reg. No. 73527)