


|   |  |                          |                  |
|---|--|--------------------------|------------------|
| <b>Cover Sheet for Proposals</b><br><i>(All sections must be completed)</i>   |  <b>JISC Digital Repositories</b><br><b>Programme: Options review of the ways in which institutional and subject repositories can productively interact</b> |                          |                  |
| <b>Name of Bidder:</b>  | <b>Science and Technology Facilities Council</b>   |                          |                  |
| <b>Name(s) of Project</b><br><b>Partner(s):</b>   | <p align="center">—</p>  |                          |                  |
| <b>Full Contact Details for Primary Contact:</b><br><br><b>Name:</b> Catherine Jones<br><b>Position:</b> Library Systems Development Manager<br><b>Email:</b> c.m.jones@rl.ac.uk<br><b>Address:</b> R61 1-02<br>STFC Rutherford Appleton Laboratory<br>Harwell Science and Innovation Campus<br>Didcot<br>OX11 0QX<br><br><b>Tel:</b> 01235 445402<br><b>Fax:</b> 01235 446403  |  |                          |                  |
| <b>Length of Project:</b>   | <b>Five months</b>   |                          |                  |
| <b>Project Start Date:</b>  | <b>March 2008</b>  | <b>Project End Date:</b> | <b>July 2008</b> |
|   |  |                          |                  |
| <b>Total Funding Requested from JISC:</b>   | <b>£41017</b>  |                          |                  |
|   |  |                          |                  |
| <b>Outline Project Description</b><br><p>The project will produce a set of coherent and analysed recommendations for steps that can be taken to improve the interaction between institutional and subject repositories, with a view both to ameliorating problems and taking advantage of opportunities. Close interaction with the stakeholders will ensure focus and clear identification of costs and benefits. Full use will be made of previous work on for example repository interaction and versions identification.</p> <p>The project will identify and survey stakeholders; study the issues arising; propose and analyse steps that make be taken to improve the interaction between repositories; work out preconditions and dependencies between steps; and write a report recommending a concrete set of steps that should be taken.</p> |  |                          |                  |

## **Options review of the ways in which institutional and subject repositories can productively interact**

### **Project scope**

The scope of the project is to survey issues and identify problems arising from the lack of coordination of content in the fast-developing landscape of repositories in the UK. The issues will be identified and assessed with respect to the important stakeholders. The project will identify straightforward practical steps to ameliorate the problems arising, and the results will be written up as report suitable for dissemination and the setting of future work.

The review will be restricted to UK repositories that hold freely available research papers, though it will need to take into account some subject repositories of wider international scope, though cannot make recommendations for these. The main focus will be on journal articles in the various stages of their existence (preprints and postprints); to cover the whole range of material that is found in repositories would be too open-ended a task for this study.

There will be no recommendations on the principles for the relationship between different classes of repositories; rather the emphasis will be on achievable steps that may be taken. As well as the repositories themselves, the study will consider related services such as resource discovery.

### ***Types of steps to be considered***

Within the scope of this study, bearing in mind that it is not concerned with issues of policy or principle as to where material “should” be deposited, there are essentially two broad areas that should be considered:

- steps relating to the deposition process
- steps relating to searching and retrieval

Not to prejudge the results of the initial fact finding, but for each of these areas there should be consideration of:

### **Deposition process**

- Metadata required (including potential enhancements, adoption of standard application profiles and consideration of persistent identifiers eg RIDIR’s work)
- Version identification of items both within and in multiple repositories
- Transfer of content from one type of repository to another
- Multiple deposition in one automated action
- Pointers between repositories and how these might be made (i.e. the item is not held as content in multiple repositories but only in one place, but with explicit links)

### **Searching and Retrieval process**

- Use, and potentially identification of new requirements, of commonly provided services, such as Sherpa RoMEO, ROAR, OpenDOAR, IESR, Intute repository search (not exhaustive list)
- Relationships between research paper repositories and data archives (including exploration of new initiatives such OJMS <http://proj.badc.rl.ac.uk/ojims> and making the LEAP <http://ads.ahds.ac.uk/project/leap/> which blur the journal article/data repository boundaries)
- Relationships between individual repositories, federated/group repositories (such as National ones) and Subject repositories

## Background and rationale

The importance of institutional repositories as a feature of the Open Access landscape is growing. OpenDOAR, the directory of open access repositories, currently lists around 80 UK institutions supporting repositories, including many universities but also other bodies. Institutions have their own policies—in some cases mandates for deposition—while communities of researchers may have established practices, for example depositing preprints in subject repositories as with arXiv for the high energy physics community. It is generally true that the content of institutional repositories is variable across subject areas and between individuals. Indeed the population of repositories of any type obviously depends on the behaviour of individuals, and this is still in a state of flux.

This situation may cause confusion to authors and readers of research material. Authors might be uncertain as to the most suitable place to deposit their material, or might resent any obligation to place the same paper in two or more different repositories. Readers performing searches might retrieve versions of the same paper and be unsure if they are really identical or what the relationship is between them. From an organizational point of view, there might be under-reporting, where a piece of work should (under the individual repositories' remit) be in two repositories but only appears in one.

While the problems arising from the current uncoordinated situation are clear, there may well be opportunities waiting. Readers accustomed to searching in subject repositories might find more varied material of interest in institutional repositories holding some of the same material. Different classes of repositories might be suitable for different kinds of quantitative assessment.

## Methodology

The project will adopt the following step-by-step approach.

1. Survey the existing state of institutional and subject repositories and the relationships between them. This will be a brief survey, building on the existing knowledge of the project team and work that has previously been conducted. The work of the team on the JISC-funded projects CLADDIER and VIF (Versions Identification Framework) will be especially relevant. We will map this to the funding bodies to get a wide view of important institutions in different subject areas. This survey will produce themes that the project team will explore further with the stakeholders identified.
2. Identify stakeholders and establish links with them. The key types of stakeholders are listed at paragraph 12 of the ITT. This task will, working closely with JISC, identify the most effective routes of approach given the limited timescale and effort. Stakeholder involvement will be encouraged by anticipating issues that are likely to engage the stakeholders, such as under-reporting mentioned above, of interest to institutions and research funders.
3. Survey the stakeholders to identify the perceived issues arising from the coordination of repositories. The intention is not to carry out an exhaustive survey, but rather to focus on key players and draw out their perception of the problems and possible opportunities arising from repository coordination/interaction. The presence of a Particle Physics Department at STFC will give a convenient way into that community, for example. As the UK IR managers are an easily identifiable group, we will get a general overview of their concerns by use of an online questionnaire. Specific key IR managers will be identified and approached for a more in depth view.
4. Prioritise the issues

5. Identify preliminary set of steps that may be taken  
Here the aim is to generate a first set of steps that can be analysed and prioritized later.
6. Assess the benefits and costs and work out the implications of and preconditions for the steps: technical/legal/... and identify cross-dependencies between steps
7. Discuss the outcomes of the stakeholder interviews and analysis with key stakeholders, to get feedback on whether the proposed steps resolve their issues. Steps 4-7 are very likely to be undertaken more than once.
8. Draw together and propose a small coherent set of steps for practical implementation  
These will be presented in terms of strengths and weaknesses (if any) for the different stakeholders.
9. Adjust following the comments of the Review Group.

Regarding project management, this is a small study so there will not be a heavy management structure. There will be one STFC project manager, responsible for the day-to-day operation of the project, plus an STFC project director for a small fraction of his time, who will contribute to the project work so as to have visibility and make input, as well as being the point of contact in the unlikely event of any difficulties between the project manager and JISC. As this is not a project with multiple partners, there is no need for any representative management committee structure, and it is expected that close interaction with the JISC Programme Manager will be sufficient to ensure that expectations are met on all sides.

As requested, an interim oral report will be presented to JISC, and a final report preceded by a draft.

## Workplan

CMJ: Catherine Jones, SCL: Simon Lambert; LDG: Linda Gilbert and RMD: Robert Darby

### Work package 1: Project management

|                       |  |
|-----------------------|--|
| <b>Work package 1</b> | <b>Project management and reporting (10% of budget)</b>  |
| Objective:            | To manage the entire project and to ensure that the JISC Project Management guidelines are adhered to. To liaise with stakeholders. To evaluate the success of the project against preset criteria   |
| Deliverables:         | Project plan; core project documents; evaluation report.   |
| Tasks                 | <ol style="list-style-type: none"> <li>1. Write detailed project plan</li> <li>2. Manage finances</li> <li>3. Manage project to the plan</li> <li>4. Report back to JISC and other stakeholders throughout the project</li> <li>5. Evaluate success</li> </ol> |
| Timescales:           | Months 1-5   |

### Task outlines

| Task                                     | Start date | End date  | Allocation | Comments    |
|--|------------|-----------|------------|-------------|
| Write detailed project plan              | 28/4/2008  | 2/5/2008  | CMJ        | <i>Done</i> |
| Manage finances & project to plan        | 28/4/2008  | 15/9/2008 | CMJ        |             |
| Report back to JISC – half way reporting | 1/7/2008   | 31/7/2008 | CMJ        |             |
| Evaluate success & completion reports    | 1/9/2008   | 15/9/2008 | CMJ        |             |

### Workpackage 2: Mapping the information landscape

|                       |   |
|-----------------------|---|
| <b>Work package 2</b> | <b>Mapping the information landscape (step 1 above; 10%)</b>  |
| Objective:            | To ensure all relevant aspects of the current information landscape regarding repositories are covered.   |
| Deliverables:         | Internal documentation; identification of key issues and particular stakeholders; formulation of stakeholder questionnaires for use in the consultation phase.  |
| Tasks                 | <ol style="list-style-type: none"> <li>1. Fact finding phase</li> <li>2. Co-ordination of key issues</li> <li>3. Planning for stakeholder liaison</li> <li>4. Formulation of stakeholder questionnaires</li> <li>5. Trial run of the stakeholder survey to fine tune the questions</li> </ol> |
| Timescales:           | Months 1-2  |

### Task outlines

| Task         | Start date | End date  | Allocation | Comments   |
|--------------|------------|-----------|------------|--|
| Fact finding | 1/5/2008   | 31/5/2008 | All        | To be divided up amongst the team; each member to take a particular subject area on board. Split to be decided at first team meeting (8/5) |

|  |           |          |         |  |
|--|-----------|----------|---------|--|
| Key issues identification for the stakeholder groups | 21/5/2008 | 4/6/2008 | All     | Each member to bring back their perspectives & general discussion to produce top issues & further areas for research.  |
| Design of IR questionnaire                           | 29/5/2008 | 2/6/2008 | RMD     | This questionnaire is designed to be completed by Repository Managers (primarily IR managers; this will be followed up by the workshop). Other stakeholders will be approached in other ways – see workpackage 3 |
| Test run of IR questionnaire                         | 3/6/2008  | 9/6/2008 | RMD/LDG | Use Linda as first test run as she is part of the IR manager stakeholder group.  |

### **Workpackage 3: Stakeholder identification and consultation**

|                       |  |
|-----------------------|--|
| <b>Work package 3</b> | <b>Stakeholder identification and consultation (steps 2–3 above; 40%)</b>  |
| Objective:            | To ensure a representative sample of stakeholders are identified and consulted on all the identified issues.   |
| Deliverables:         | Internal records of the consultation phase.  |
| Tasks                 | <ol style="list-style-type: none"> <li>1. Arranging stakeholder liaison</li> <li>2. Workshop to gather feedback from a wider community – possibly using UKCoRR or RSP to gather repository managers together.</li> <li>3. Undertaking liaison</li> <li>4. Evaluating the response rate against acceptable level</li> </ol> |
| Timescales:           | Months 2-4   |

| Task   | Start date | End date  | Allocation | Comments  |
|--|------------|-----------|------------|---|
| Identifying key stakeholder groups and deciding the approach | 1/5/2008   | 16/5/2008 | All        | <p>Initial thoughts:</p> <ul style="list-style-type: none"> <li>• IR managers<br/><i>Through general mailing lists, specific people &amp; workshop.[LDG]</i></li> <li>• Subject repository managers<br/><i>Contacting each identified repository. [RMD]</i></li> <li>• Funders – policy makers<br/><i>Contacting specifics (list confirmed with Neil); possible workshop attendees?[CMJ]</i></li> <li>• Additional bodies, such as BL &amp; JISC<br/><i>Contacting specifics (list confirmed with Neil); possible workshop attendees?[CMJ]</i></li> </ul> |

|  |           |           |         |  |
|--|-----------|-----------|---------|--|
|  |           |           |         | RMD, LDG & CMG assigned a stakeholder group to investigate.  |
| Run survey for IR managers             | 9/6/2008  | 30/6/2008 | RMD     |  |
| Organise and run workshop(s)           | 23/6/2008 | 04/7/2008 | CMJ/LDG | Would only be 0.5 day but in this time frame; might do two: North & South  |
| Arrange 1-1 meetings with stakeholders | 19/5/2008 | 27/6/2008 | All     | In a six week period; workshops to be held at the end.<br><br>Team member to be assigned for each stakeholder once timings agreed.                                 |
| Evaluate stakeholder liaison           | 1/7/2008  | 15/7/2008 | CMJ     | Ensure each of the stakeholder groups have an acceptable response rate.<br><br>Feed the results of the surveys; soundings & research into the start of the report. |

#### **Workpackage 4: Conclusions and recommendations**

|                       |   |
|-----------------------|---|
| <b>Work package 4</b> | <b>Conclusions and recommendations (steps 4–7 above; 40%)</b>   |
| Objective:            | To identify the 10 key recommendations which are the most feasible steps and to produce a formal report outlining these.  |
| Deliverables:         | Formal report.  |
| Tasks                 | <ol style="list-style-type: none"> <li>1. Draft final report</li> <li>2. Have report reviewed by JISC (&amp; potentially other stakeholders)</li> <li>3. Complete final report</li> </ol> |
| Timescales:           | Months 4-5  |

The only intermediate milestone will be at Month 4 on conclusion of the stakeholder identification and consultation.

| Task                   | Start date | End date  | Allocation | Comments   |
|------------------------|------------|-----------|------------|--|
| Draft the final report | 15/7/2008  | 1/8/2008  | All        | CMJ as main coordinator  |
| Review period          | 4/8/2008   | 2/9/2008  | CMJ        | Review group given a month elapse time to look at it as this is prime holiday time.<br><br>Need to review the risks and appropriate/acceptable response levels<br><br>Depending on the responses, may revisit & reissue amended report during this period. |
| Complete final report  | 3/9/2008   | 15/9/2008 | CMJ        |  |

## Risks

| Risk  | Probability (1-5) | Severity (1-5) | Score (P x S) | Action to Prevent/Manage Risk   |
|---|-------------------|----------------|---------------|---|
| Key project staff temporarily unavailable due to unplanned absence (illness, parental leave, etc) | 2                 | 4              | 8             | Redefine or reduce non-core objectives, within existing budgets; consider secondment of other Library staff to project        |
| Problems meeting the Stakeholders in the timescales for the project.                              | 3                 | 5              | 15            | Try different forms of communication (email, telephone rather than face to face meetings). Identify alternative stakeholders. |
| Difficulty in identifying concrete steps and no resolution to issues.                             | 2                 | 4              | 8             | Identify issues in identifying steps. Open a dialogue with JISC on the best way forward.                                      |

## Organisation

The Science and Technology Facilities Council was an early adopter of institutional repositories, with the ePubs service which has been live since May 2004. It is implemented with in-house software that uses the IFLA Functional Requirements for Bibliographic Records model as the conceptual basis, placing the notion of versions of a particular work at the centre. STFC, in its previous organisation of CCLRC, has also supported the work of the NISO/ALPSP working group on Journal Article Versions by providing a member of the working group. STFC/CCLRC was involved in the CLADDIER project, which investigated issues surrounding data publication and linking data to publications and vice versa. ePubs is an important part of the STFC Library and Information Service Strategic Plan 2006-2010 as STFC is heavily reliant on journals both for information retrieval and dissemination. Version identification is a strong strand of this strategic view.

## Staff

The staff involved will be:

### Catherine Jones

Catherine is the Library Systems Development Manager for STFC Libraries and in this role is responsible for the institutional repository ePubs. She was the project manager for ePubs from its inception in 2003. She was a project team member in the JISC funded CLADDIER project and is a team member of the JISC funded Versions Identification Framework project. Her degree is in Computing and Communication Systems. She has written and presented about institutional repositories and is the author of a book entitled "Institutional Repositories: Content and Culture in an Open Access Environment" which was published by Chandos in September 2007.

### Simon Lambert

Simon is the head of Information Services Group within e-Science Department, which runs the STFC libraries at Rutherford Appleton Laboratory and Daresbury Laboratories. He is an experienced project manager with particular expertise in areas such as knowledge management and digital preservation. Most of these project have been national or international collaborations with multiple partners. He has also had a role in the ePubs institutional repository, in formulating policies and development strategies.

### Linda Gilbert

Linda is the Partnership Liaison Assistant Librarian for the STFC Libraries and in this role she is responsible for the development & day to day management of the institutional repository ePubs. Her degree is in Information Management.

**Robert Darby**

Robert is the Systems Support and Development Assistant Librarian for the STFC Libraries and is responsible for journals management and IT systems. He has a degree in English Literature and DPhil in Modern Languages (French Literature) and widens the team's experience from just Physical Sciences into Humanities.

**Costs**

**Budget**

|                                   |  |               |
|-----------------------------------|--|---------------|
| <b>Amount Requested from JISC</b> |  | <b>£41017</b> |
|-----------------------------------|--|---------------|

**Timetable and start date**

As requested by JISC, the STFC team would be able to start work in March 2008, and the duration will be five months. A later start date would also be acceptable.