

JISC DEVELOPMENT PROGRAMMES

SWORD Cover Sheet

PROJECT PLAN

Project Acronym	SWORD	Project ID	
Project Title	Simple Web service Offering Repository Deposit		
Start Date	01/03/2007	End Date	31/08/2007
Lead Institution	UKOLN, University of Bath		
Project Director	Rachel Heery		
Project Manager & contact details	Name: Julie Allinson Position: Repositories Research Officer Email: j.allinson@ukoln.ac.uk Address: UKOLN, University of Bath, Bath, BA2 7AY Tel: 0114 2486457 or 01225 386580 Fax: 01225 386838		
Partner Institutions	Intrallect Ltd. University of Southampton, EPrints University of Wales Aberystwyth, Centre for Advanced Software and Intelligent Systems (CASIS)		
Project Web URL	http://www.ukoln.ac.uk/repositories/digirep/index/SWORD		
Programme Name (and number)	Repositories and Preservation Programme, Tools and Innovations, 10/06		
Programme Manager	Phil Vaughan, p.vaughan@jisc.ac.uk		

Document

Document Title	SWORD project plan		
Reporting Period			
Author(s) & project role	Julie Allinson, Project Manager		
Date	2007-04-03	Filename	Sword-project-plan-02.doc
URL			
Access	☐ Project and JISC internal		☒ General dissemination

Document History

Version	Date	Comments
0.1	2007-03-22	Draft project plan for internal project circulation
0.2	2007-04-03	Draft project plan for review by Programme Manager



JISC Project Plan : SWORD

Overview of Project

1. Background

The effective and efficient population of repositories is a key concern for the repositories community. Deposit is a crucial step in the repository workflow; without it a repository has no content and can fulfill no further function. Currently most repositories exist in a fairly linear context, accepting deposits from a single interface and putting them into a single repository. Further deployment of repositories, encouraged by JISC and other funders, means that this situation is changing and we are beginning to see an increasingly complex and dynamic ecology of interactions between repositories and other services and systems. By and large developers are not creating repository systems and software from scratch, rather they are considering how repositories interface with other applications within institutions and the wider information landscape. A single repository, or multiple repositories, might interact with other components, such as VLEs, authoring tools, packaging tools, name authority services, classification services and research systems. In terms of content, resources may be deposited in a repository by both human and software agents, e.g. packaging tools that push content into repositories or a drag-and-drop desktop tool. The type of resource being deposited will also influence the choice of deposit mechanism. If the resources are complex packaged objects then a web service will need to support the ingest of multiple packaging standards.

There is currently no standard mechanism for accepting content into repositories, yet there already exists a stable and widely implemented service for harvesting metadata *from* repositories (OAI-PMH – Open Archives Initiative Protocol for Metadata Harvesting). This project will implement a similarly open protocol or specification for deposit. By taking a similar approach, the project and the resulting protocol and implementations will gain easier acceptance by a community already familiar with the OAI-PMH.

This project aims to develop a **Simple Web-service Offering Repository Deposit (SWORD)** - a lightweight deposit protocol that will be implemented as a simple web service within EPrints, DSpace, Fedora and IntraLibrary and tested against a prototype 'smart deposit' tool. The project plans to take forward the lightweight protocol originally formulated by a small group working within the Digital Repositories Programme (the 'Deposit API' work)¹. The project is aligned with the Object Reuse and Exchange (ORE)² Mellon-funded two-year project by the Open Archives Initiative, which commenced in October 2006. Members of the SWORD project team are represented on its Technical and Liaison Committees. The ORE project has identified a 'Register' service as part of its interoperability layer. Work on the SWORD project fits well with the definitions of the 'Register' service. The SWORD project is not attempting to duplicate work being done by ORE, but seeks to build on existing work to support UK-specific requirements whilst feeding into the ongoing ORE project. This ties in with the activities of the JISC Common Repository Interfaces Working Group³.

There are many scenarios supporting a common Deposit web service, for example:

- 'Easy-Deposit' service. This might be a centralised, or a local, service which would be able to accept deposits and direct them to an appropriate repository or multiple repositories. It might be implemented by institution or by a third-party service such as the JISC-funded Prospero project's "The Depot"⁴.
- Support multiple deposits. In the light of the RCUK statements on Open Access⁵, the need for this facility is clear. Some Research Councils are mandating deposit of research output

¹ http://www.ukoln.ac.uk/repositories/digirep/index/Deposit_API

² <http://www.openarchives.org/ore/>

³ <http://www.ukoln.ac.uk/repositories/digirep/index/CRIG>

⁴ <http://prospero.edina.ac.uk/>

⁵ <http://www.rcuk.ac.uk/access/default.htm>

created out of funded projects into Research Council specified repositories. Authors will also want, or may need, to additionally deposit in their institutional or departmental repository. A multiple deposit facility would be a significant benefit for depositors with limited time to submit material and for repository advocates working hard to sell the benefits of repositories to academics.

- Support transfer of deposits between intermediate hosts, e.g. from central or laboratory repository to another repository, or to a preservation service.
- Support anonymous deposit and more complex, authenticated deposit without the need to integrate repository software with local authentication systems.
- Support acceptance and handling of incomplete records.

Support for a common repository interface would have significant benefits across the repository landscape. Not only would it help facilitate closer collaboration between different repository suppliers and repositories, it would play a key role in promoting the adoption of common standards towards wider interoperability.

2. Aims and Objectives

Aims

- To improve the efficiency and quality of the repository 'Ingest' function
- To diversify and expedite the options for timely population of repositories with content
- To facilitate the creation and use of common deposit interfaces
- To improve repository interoperability as outlined in the Information Environment
- To take a service-oriented approach to development as outlined by the E-Framework

Objectives

- To produce a standard mechanism for depositing content in repositories
- To test and refine the lightweight protocol originally formulated by a small group working within the Digital Repositories Programme (the Deposit API)
- To evaluate existing standards that might be used to offer a deposit web service
- To implement the deposit service in EPrints, DSpace, Fedora and IntraLibrary
- To develop a prototype 'smart deposit' tool
- To disseminate the resulting work and encourage community uptake
- To ensure that the approach developed by this project is cognisant of UK requirements (as defined by the JISC Common Repository Interfaces Group – CRIG) and International work in this area (including the OAI-ORE activity)

3. Overall Approach

The project will draw on the scenarios and use cases collected by the Repositories Research Team⁶ and the documented discussions from the Deposit API work to help identify requirements. The draft serialisations developed by the Deposit API work will be agreed and finalised and will be implemented by the sub-contractors. The project will be structured into five workpackages, running largely concurrently with considerable cross-fertilisation between each. Testing phases will be used to iteratively guide development and will engage additional partners such as the JISC-funded SOURCE project. Throughout the course of the project outputs will be made publicly available in a timely manner to ensure currency of information for the target community. Management and development will be kept lightweight for this short project.

The chief issue for this project is achieving interoperability between repositories, allowing material to be deposited in a standard way into a range of repositories.

The scope of this project is primarily technical development. It is concerned with *initial* deposit and will not consider requirements for updating existing resources, metadata creation or performing

⁶ http://www.ukoln.ac.uk/repositories/digirep/index/All_the_Scenarios_and_Use_Cases_Submitted

duplicate checking. The project will attempt to specify two layers of compliance to the deposit protocol, one enabling repositories to accept deposits based on a set of minimal mandatory requirements and one with more constraints to offer additional security and control. It will also define an explain service to support remote inspection of repository policy. Beyond this, it is anticipated that the project will highlight a range of non-technical issues relating to repository policy, as well as legal, cultural and business issues. These are out of scope, but will be identified and documented for any future work in this area.

Critical success factors include the successful development of a standard deposit mechanism in each software platform and demonstration that an easy-deposit client can push content into each repository to fulfil a range of identified use cases and scenarios. Community acceptance and implementation of the web service by repositories and other deposit agents are also essential as is ensuring that this project remains cognisant of other work around deposit and also that it does not invent new standards or specifications where these are not necessary.

4. Project Outputs

Tangible deliverables:

1. Protocol or specification
2. Evaluation reports on relevant standards
3. Implemented web service in Eprints, Fedora, IntraLibrary and DSpace
4. Deposit client implementation
5. Service expression in the E-Framework
6. Case studies
7. Project plan, progress report, final report, completion report, exit and sustainability briefing
8. Project web site

Intangible knowledge and experience:

1. Understanding of the scenarios and requirements for deposit and multiple deposit
2. Awareness of any issues or further work required in this area
3. Expertise to feed into future JISC projects, the Information Environment and international projects such as OAI-ORE
4. Proof of concept development and testing

5. Project Outcomes

The main outcome of this project is the creation of a standard deposit web service that repositories or other deposit agents can implement in order to facilitate accepting content from remote deposit clients and offer depositors, both human and machine, a more streamlined mechanism for deposit. The impact of this will be to expedite and expand options for getting content into repositories. This will have cross-domain impact, for institutional open access repositories, learning object repositories, any other content or subject-based repositories and also potentially for any other systems that deposit digital content.

The reference implementation will demonstrate how the deposit web service might fulfil some of the identified usage scenarios. The impact of this will be to help encourage other projects and services to make use of the deposit web service and further develop deposit tools.

6. Stakeholder Analysis

Stakeholder	Interest / stake	Importance
Repository software developers	Supporting the deposit web service may be seen as a selling point for their software. Developers involved in the project have a particular interest in its successful implementation in their own product.	High
Repository data providers and implementers (including	They will be able to implement the web service in order to accept	High

institutional, subject and others)	deposits in a standard way and exploit the benefits that a standard deposit web service can offer.	
International projects	The deposit web service may be of interest to projects investigating interoperability issues or wishing to explore new deposit options.	High
Repository managers	They will be able to utilise the web service to accept content from more sources and depositors and to offer simpler deposit options to depositors. There are clear advocacy benefits.	Medium
Funding bodies	The deposit web service can help facilitate multiple deposits into specified repositories.	Medium
Repository depositors and users	They will be offered more standard, simpler deposit options such as an easy deposit desktop client.	Medium
JISC development programmes and projects	The deposit web service may be of interest to projects working on similar tools or wishing to support a range of deposit options.	Medium
Institutional managers	They may be interested in the benefits offered by the deposit web service, in particular if it facilitates additional content deposited into institutional repositories.	Low
Service providers, aggregators etc.	The deposit web service may be used as a mechanism of pushing content back to repositories.	Low

7. Risk Analysis

Risk	Probability (1-5)	Severity(1-5)	Score (P x S)	Action to Prevent / Manage risk
Recruitment difficulties	1	4	4	Project team was established prior to proposal submission; use of subcontractors mitigates against recruitment difficulties
Loss of a team member	2	4	8	Multiple staff at each site have the expertise and skills required.
Failure of partners to deliver work on schedule	2	4	8	Regular communication and use of subcontracts should help reduce any risk here. Partners have access to established networks of developers and may be able to draw on additional resources if any subcontractor is unable to fulfil their commitment.

Changing specification	4	3	12	The specification is likely to change during implementation as issues arise. By making use of the UKOLN Repositories Research Team and the project's technical advisors, the project will aim to create a high quality specification. When changes are required, the experience of the developers and their knowledge of the repository platforms will be relied upon.
Technical knowledge	1	4	4	Risk will be kept low by using experienced developers with knowledge of the domain.
Technical advancement	3	2	6	With repository software changing rapidly, it is important to ensure that the implementations are suitable for inclusion in the base releases of the repository platforms. By using core developers of the platforms, it is anticipated that code will be of suitable quality for inclusion and will be written for the most up to date versions of the repository platforms.
Alignment with UK needs	2	3	6	The Deposit API needs to be aligned with the needs of UK (and worldwide) users. Making use of the UKOLN Repository Research Team and the project's technical advisors will ensure it is targeted appropriately.
Legal issues	1	3	3	The project is committed to making its content available in an open source or open access way. No legal issues are anticipated.

8. Standards

Name of standard or specification	Version	Notes
WebDAV (http://www.webdav.org/)		Review
JSR 170 (http://www.jcp.org/en/jsr/detail?id=170)		Review
JSR 283 (http://www.jcp.org/en/jsr/detail?id=283)		Review
SRW Update (http://www.loc.gov/standards/srw/)		Review
Flickr Deposit API (http://www.flickr.com/services/api/)		Review
Fedora Deposit API (http://www.fedora.info/definitions/1/0/api/)		Review
OKI OSID (http://www.okiproject.org/)		Review
ECL (http://ecl.iat.sfu.ca/)		Review
Atom Publishing Format and Protocol (http://www.ietf.org/html.charters/atompub-charter.html)		Review

Other standards may be identified and reviewed in the course of the project. In general, the project will work with components specified by the W3C web architecture in line with the agreements reached by the OAI-ORE Technical Committee⁷.

9. Technical Development

The project will take a lightweight approach to technical development. Development partners will be given the autonomy to use their own development methodologies, with the emphasis being on iterative and agile practices. Throughout the life of the project, lessons learnt and results of user testing will feed into iterative revision and development.

10. Intellectual Property Rights

All background information and know-how used in connection with the project shall remain the property of the party introducing the same. All material generated during the course of the project will remain the property of the creator, with the proviso that all outputs, including documentation and code, created as part of this project will be made available, free at the point of use, to the UK HE and FE community in perpetuity and may be disseminated widely in partnership with JISC.

Project Resources

11. Project Partners

UKOLN, University of Bath (Lead)

Role: Lead partner and project management

Contact: Julie Allinson, j.allinson@ukoln.ac.uk

The following partners will contribute the project:

EPrints, University of Southampton (subcontractor)

Role: EPrints technical development work and contribution to the review of existing standards and specification of the deposit web service

Contact: Les Car, lac@ecs.soton.ac.uk

Intrallect Ltd. (subcontractor)

Role: IntraLibrary technical development work and contribution to the review of existing standards and specification of the deposit web service

Contact: Martin Morrey m.morrey@intrallect.com and Sarah Currier s.currier@intrallect.com

Centre for Advanced Software Intelligent Systems (CASIS), University of Wales Aberystwyth (subcontractor)

Role: DSpace, Fedora and reference implementation technical development work and contribution to the review of existing standards and specification of the deposit web service

Contact: Neil Taylor nst@aber.ac.uk

Note: due to the short-term nature of this project there will be no consortium agreement; rather, the project partners will be engaged as subcontractors.

12. Project Management

Project management will come from UKOLN, University of Bath and be carried out by Julie Allinson (UKOLN) with leadership and direction from Rachel Heery. This project will complement and be informed by UKOLN, and partner, involvement in other repositories projects.

This is a relatively small, short-term project and to reflect this, project management will be kept lightweight. Project communication and decision-making will be undertaken through an email discussion list and project meetings, likely to be 2 face-to-face meetings and a number of telcon/online meetings. Sub-contractors will be responsible for completion of the technical development work according to their own working methodology. There will be no management

⁷ <http://www.openarchives.org/ore/documents/OAI-ORE-TC-Meeting-200701.pdf>

committee, although a number of experts in the area have agreed to act as advisors to the project and will be invited to contribute to the project mailing list and meetings.

Project team:

- Julie Allinson, UKOLN, University of Bath, Project Manager *
- Rachel Heery, UKOLN, University of Bath, Project Director**
- Les Carr, University of Southampton, leading the EPrints development
- Chris Gutteridge, University of Southampton, EPrints developer
- Martin Morrey, Intrallect Ltd, leading the IntraLibrary development
- Sarah Currier, Intrallect Ltd, leading the IntraLibrary development
- Neil Taylor, CASIS, University of Wales Aberystwyth, managing the CASIS sub-contracts
- Stuart Lewis, DSpace and reference client developer
- Richard Jones, DSpace developer
- Glen Robson, Fedora developer

* the project manager will spend approximately 20% of allocated time on project management

** Liz Lyon will act as project director for the initial period of the project

Advisors:

- Richard Green, University of Hull (Fedora)
- Jim Downing, University of Cambridge (DSpace)
- David Flanders, Birkbeck, Bloomsbury Colleges (SOURCE project)

No training needs have been identified. If knowledge or training is required in any of the technologies and/or standards used across the project, an appropriate training course or advice from an expert or community will be sought, e.g. through a relevant JISC development project or via an email discussion list.

13. Programme Support

The project will benefit from help in connecting with other projects working in similar areas.

14. Budget

See Appendix A

Detailed Project Planning

15. Workpackages

See Appendix B

16. Evaluation Plan

Timing	Factor to Evaluate	Questions to Address	Method(s)	Measure of Success
Ongoing	Efficacy of deposit web service	Does the deposit web service support the identified scenarios?	User testing, iterative development; QUALSERV	Positive feedback and successful deposit
Ongoing	Evaluation reports on standards or specification	Does the selected standard or specification meet the requirements for deposit?	Peer review	Selection or development of appropriate standard or specification
Ongoing	Community uptake of the web service	Has the web service been implemented by repositories and deposit agents?	Usage logs	Implementation and use

Evaluation will be by means of iterative testing and development mechanisms, and through community engagement and feedback. The project will also take part in the wider JISC evaluation of the Programme, in particular through the involvement of the Repositories Research Team. Part of the project management tasks will be to produce an Exit and Sustainability Briefing to identify further activities and suggestions of how these might happen.

17. Quality Plan

Output	Protocol or specification				
Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
Months 1-5	Fitness for purpose	Comparison against scenarios and requirements	Documentation	Project Manager and Developers	
Output	Implemented web service				
Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
Months 1-6	Adherence to protocol or specification	Validation against documentation	Successful validation	Developers	
Months 1-6	Fitness for purpose	User testing and development	Successful deposit and acceptance of deposits against specified criteria	Developers	
Output	Deposit client implementation				
Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
Months 1-6	Adherence to protocol or specification	Validation against documentation	Successful validation	Developers	
Months 1-6	Fitness for purpose	User testing and development	Successful deposit	Developers	

18. Dissemination Plan

Timing	Dissemination Activity	Audience	Purpose	Key Message
March 2007	Announcement of project on various web sites, blogs, mailing lists and newsletters	All stakeholders	Awareness-raising and promotion	The project is happening
Ongoing	Updates to ORE-TC	OAI-ORE project	Join-up with international activity	SWORD may be of interest, use and feedback is encouraged
Ongoing	Announce availability of outputs to various web sites, blogs, mailing lists and newsletters	All stakeholders	Awareness-raising and promotion, community acceptance and uptake, feedback	Use and feedback is encouraged
Ongoing	Ad-hoc promotion of project through email discussion lists, f2f discussions etc.	All stakeholders	Awareness-raising and promotion, community acceptance and uptake, feedback	SWORD may be of interest, use and feedback is encouraged

Ongoing	Journal articles and presentations, as opportunities arise	All stakeholders	Awareness-raising and promotion, community acceptance and uptake, feedback	SWORD may be of interest, use and feedback is encouraged
---------	--	------------------	--	--

19. Exit and Sustainability Plans

Project Outputs	Action for Take-up & Embedding	Action for Exit
Reports	Publish on project web site; deposit into an appropriate repository	Repository policy should guarantee a period of retention
Protocol or Specification	Publish documentation on project web site; deposit into an appropriate repository	Repository policy should guarantee a period of retention
Implementations (code)	Developers have direct links to the development community to ensure appropriate maintenance	Maintenance passes to the development community or key software developers
Project web site	Housed on RRT wiki	Maintenance passes to RRT
Intangible knowledge and expertise	Most staff working on other JISC-funded projects or within software development communities.	Staff will have re-usable knowledge.

Project outputs that may have potential to live on after the project ends:

Project Outputs	Why Sustainable	Scenarios for Taking Forward	Issues to Address
Deposit protocol	Uptake and use of the protocol will demonstrate its future usefulness	Continued hosting on UKOLN web site or RRT wiki	Updates and maintenance
Reference implementation	Although it is being developed as a proof of concept, this implementation may prove useful as a deposit tool	Additional funding to develop this client into tool; possible release as an open source tool	Updates, hosting and maintenance
Web service in EPrints, Fedora, DSpace and IntraLibrary	Community uptake of will demand continued availability of the web service	Embed into the software, either into the core release or as a plug-in	Updates, hosting and maintenance
Case studies	Useful for stakeholders	Continued hosting on UKOLN web site or RRT wiki	Currency of information
Evaluation reports	Useful for stakeholders	Continued hosting on UKOLN web site or RRT wiki	Currency of information

Appendixes

Appendix A. Project Budget

Appendix B. Workpackages