JISC Digital Repositories Programme - 03/05

Name of lead institution/organisation: Heriot-Watt University

Name of proposed project: Pilot Engineering Repository Xsearch (PERX)

Project partners: Cranfield University, Institution of Civil Engineers (ICE)/Thomas Telford Ltd, Adiuri Systems Ltd, Geotechnical, Rock and Water Resources Library (GROW), Regional Support Centre East Midlands.

Full contact details for primary contact

Name: Roddy MacLeod
Position: Senior Subject Librarian, Heriot-Watt University
Email: R.A.MacLeod@hw.ac.uk
Address: Library, Riccarton Campus, Heriot-Watt University, Edinburgh, EH14 4AS
Tel: 0131 451 3576
Fax: 0131 451 3164

Objective(s) proposal will meet:

The proposal principally addresses objective (iv) by

• Establishing a pilot subject service to support discovery of resources across multiple repositories, and scoping future development of such services.

The project also contributes to objective (ii) by

- Conducting advocacy work with potential repository providers to encourage their participation in digital repository development, and exploring associated cultural, political and commercial issues.
- Exploring the cultural and political barriers to the use of repositories in the learning and research communities.
- Examining how a subject-based resource discovery service can be embedded within different user environments.

Length of project and total cost to the JISC over its life Length: 2 Years, Total Cost to JISC: [Budget Removed from Public Version]

Cost of proposal to the JISC in each year or part year [Budget Removed from Public Version]

Proposed project start and end dates June 2005 –June 2007

Outline project description

To develop a pilot service which provides subject resource discovery across a series of repositories of interest to the engineering learning and research communities. This pilot will be used as a test-bed to explore the practical issues that would be encountered when considering the possibility of full scale subject resource discovery services.

Issues to be investigated include: the range and availability of actual and potential digital repository sources; exploration of cultural barriers to the use of repositories in the subject community, functionality of software tools; advocacy to encourage participation of repository providers; maintenance issues; interactions with infrastructural shared services; enhancing metadata quality; embedding and reuse of resource discovery services; improving search and browse results presentation; service profiling for particular audiences.

1. Introduction

1.1 Outline Project Description

To develop a pilot service which provides subject resource discovery across multiple digital repositories of interest to the engineering learning and research communities. Possible repository content sources include institutional repositories, repositories of technical reports, learning objects, images, interactive multimedia resources, theses, technical data, publishing community sources, professional society sources and funded research collections (see Appendix A).

The pilot will act as a test-bed to explore the practical issues that would be encountered when considering the possibility of full-scale subject resource discovery services. Issues to be investigated include: the range and availability of potential digital repository sources, functionality of software tools, exploration of the barriers to repository use in this subject community, advocacy work with potential repository providers to encourage their participation in digital repository development, maintenance issues, interactions with infrastructural shared services, enhancing metadata quality, embedding and reuse of resource discovery services, improving search and browse results presentation and service profiling for particular audiences.

1.2 Length of the project

Two years - June 05 - June 07

1.3 Proposed start date

Project will commence 1st June 2005

1.4 Summary of Project Outcomes and Benefits.

The pilot will assist in the scoping of a national repository service infrastructure (Activity Area (iv)) by acting as a test-bed to explore practical issues, demonstrating potential and informing future developments. Although the study will focus on resource discovery within engineering, a number of the issues earmarked for investigation will be of relevance to other subject disciplines. The experience gained from the pilot will therefore be a valuable asset for any future subject-based resource discovery service.

In addition, the proposal will make a contribution to Activity Area (ii) by conducting a review of the digital repository landscape in engineering; by exploring cultural and political issues and acting in an advocacy capacity to encourage participation and promote interoperability with a range of engineering data providers; and by examining how a subject-based resource discovery service can be embedded within the learning process. The proposal will also input into activity areas (v) and (vi). Specific outcomes and benefits of the project include:

- 1. Pilot subject-based resource discovery service (WP2.1, 2.2).
- 2. Elucidation of end-user opinion on the appropriateness and usefulness of a subject-based approach to resource discovery across multiple digital repository collections (WP3.1).
- 3. Advocacy work with potential data providers and production of advocacy materials which will encourage exposure of new digital repositories (WP3.2).
- 4. Analysis of common shared service usage scenarios and demonstration of shared service interactions (WP4.3).
- 5. Analysis of maintenance issues associated with enabling resource discovery across multiple digital repository collections (WP4.1).
- 6. Demonstration of metadata quality augmentation for the improvement of resource discovery (WP4.4).
- 7. Embedding of resource discovery functionality within a VLE (WP4.5)
- 8. Feasibility studies investigating means to improve search and browse filtering techniques and service profiling for particular groups (WP4.6, 4.7).
- 9. Reports which help scope the future development of the subject-based component of a digital repository landscape (WP5).

2. Project description

2.1 Partners

This proposal is led by a team based at Heriot-Watt University comprising staff from Heriot-Watt University Library and the Institute for Computer Based Learning (ICBL). Project partners include Cranfield University, the Institution of Civil Engineers (ICE)/Thomas Telford Ltd, Adiuri Systems Ltd (Cyana Engineering), Geotechnical, Rock and Water Resources Library (GROW), and Regional Support Centre East Midlands.

The proposed project will bring together important players representing various parts of the engineering information community, both within HE and FE, and also within and outwith the UK, including institutional repository providers, publishers, professional bodies, and commercial service providers. It draws upon experience outside the HE/FE education sector. This is particularly important in the subject area of engineering, where professional societies play an important role in the education and career development of engineers, and where numerous relatively small publishing houses produce learning and professional development material. A vision for a national distributed digital repository infrastructure would not be complete without provision for the output of such bodies.

Heriot-Watt University Library is a centre of excellence in the provision of information to engineering disciplines. The Library hosts EEVL, the engineering, mathematics and computing hub of the Resource Discovery Network (RDN), and was a partner in the Subject Portals Project. Recently, the library completed two PALS funded projects investigating the use of OAI-PMH with engineering publishers, and promoting publisher's use of RSS.

ICBL has a long history in studying design, development and evaluation issues in the use of learning technology in higher education, and is committed to supporting the development of open, standards-compliant eLearning solutions for UK Higher Education. ICBL currently coordinates the CETIS Metadata and Digital Repository Special Interest Group, through which it supports the creation and adoption of eLearning standards and the development of the JISC eLearning Framework. Recent projects include: EASEIT-Eng, which evaluated how Engineering academics use digital learning resources; FAILTE, which created a standards-compliant catalogue of engineering learning resources; and an ELF demonstrator project, embedding resource discovery tools within VLEs.

Cranfield University is a leading research university, and host to one of the largest institutional repositories in the UK – QUEprints. Cranfield University also developed the AERADE service, which provides integrated access to a collection of over 13,000 key aerospace and defence resources.

The **Institution of Civil Engineers** (ICE) represents over 70,000 professionally qualified civil engineers worldwide. It publishes the ICE Virtual Library, a digital archive of technical engineering papers from 1836 to the present. **Thomas Telford Limited** provides information for the international construction and engineering markets, and is wholly owned by the Institution of Civil Engineers.

Adiuri Systems Ltd provide information access services that are used to create tailored information environments. Cyana Engineering, one of the services provided by Adiuri, provides engineers with access to validated sources of online information.

The **Geotechnical, Rock and Water Resources Library** (GROW) is a digital library based at the University of Arizona. GROW is part of the National Science Digital Library (NSDL), and has collected 1,800 resources and developed over 200 interactive multimedia educational resources in three sub-disciplines of engineering.

The Regional Support Centre East Midlands is part of the JISC Regional Support Centre Network. It has extensive experience in promoting and evaluating learning and teaching in the Fe and HE community, and regularly conducts workshops for engineering lecturers.

2.2 Background

Heriot-Watt University Library and ICBL have high level expertise in resource discovery, federated searching across multiple content sources, and engineering information sources. Project staff have participated in both the technical development and information architecture of the 05/99 funded Subject Portals Project (SPP) which ended in September 2004.

SPP's primary aim was to deliver an aggregated cross search tool which facilitated searching across both JISC supported and non-JISC information resources. Initial focus was on searching via Z39.50 but this was later

expanded to permit the inclusion of resource discovery from OAI-PMH data repositories. SPP resulted in the installation of prototype portal services at various RDN hubs including EEVL, which permitted resource discovery across multiple content sources. The SPP developed software is available as open source and commitment has been made to its ongoing support by JISC until 2008.

Experience gained with involvement in SPP indicates that technically enabling searching across multiple repositories raises a raft of other issues that are crucial to actual service provision. These include; understanding particular digital repository landscapes, advocacy work to encourage participation from potential data providers, maintenance issues, cultural and political barriers to the use of repositories by academics, interactions with infrastructural shared services, enhancing metadata quality, embedding and reuse of discovery services, means of effective results presentation, and service profiling for different user groups.

This proposal aims to utilise the tried and tested SPP outcomes as a platform for deployment of a test-bed service allowing a range of practical issues relating to actual service provision to be explored. In our opinion the developments and achievements of SPP make an ideal starting point for developing a pilot service to provide subject resource discovery across multiple digital repositories. Much of the necessary infrastructure is already in place (e.g. harvesting and indexing of repositories using OAI-PMH, inclusion of Z39.50 targets) and the underlying cross search functionality is proving to be both robust and scaleable. It is pertinent to point out that the focus in this proposal is not on further technical development of the SPP software but is on its utilisation to explore practical service related issues. Other possible software options (e.g. Discovery+, MDC) and their functionality will be utilised within the project in order to gain a better understanding of different approaches and to compile a list of likely functional requirements for a full scale service.

We have deliberately taken a broad-brush approach and are proposing the consideration of multiple issues in this project, as this reflects the real situation that a subject-based resource discovery service would encounter in the current complex environment.

The RDN executive is supportive of this proposal which is in keeping with possible future service development areas of the RDN. The HE Academy Engineering Subject Centre is also supportive of this proposal.

2.3 Rationale for a Subject-based Pilot Service for Engineering

There has so far been ad hoc development of digital repositories, yet the needs of users are not ad hoc, and are very often subject specific. A CTI LTSN evaluation¹ reported that "*Virtually all respondents believe that academic staff identify most readily with their subject*" and this is perhaps particularly true in the case of an applied subject such as engineering. The report suggested that in order to be successful, the implementation and integration of information and communications technologies should be tackled from a subject perspective, one which is in keeping with a subject based approach to resource discovery.

As more institutional repositories develop, the task of subject-based resource discovery will become even more complex. Resource discovery across digital repositories by subject is currently very difficult, browsing is particularly time-consuming, as is search result filtering. The work outlined in this proposal considers these issues and investigates possible means to improve resource discovery from multiple repositories at the subject-based level. This proposal helps address the issue of harvesting using OAI-PMH on a subject basis, which is also recommended by the EPIC Report².

Engineering represents an important and substantial sector within both Higher and Further Education. However, digital repository uptake and usage within the engineering community is disappointingly low. Digital repository work within engineering is not well supported compared with other disciplines.

In contrast to other subjects (e.g. physics - arXiv, and cognitive sciences - Cogprints), no sizeable subject-based repositories exist for engineering. No equivalent to the AHDS exists within engineering, and the appropriate

¹ HEFCE Report 98/47. An evaluation of the Computers in Teaching Initiative and Teaching and Learning Technology Support Network, September 1998

² Delivery, Management and Access Model for E-prints and Open Access Journals within Further and Higher Education. A joint report by The Electronic Publishing Innovation Centre (EPIC) in partnership with Key Perspectives Limited. Loughborough University.

research council (EPSRC) offers no explicit funding for information related developments. There is, therefore, currently little to stimulate the interest of the engineering community in digital repositories. Better subject-based discovery options in engineering would help raise the profile of digital repository-related work within the community, and would encourage growth in the rate that relevant materials are deposited in repositories.

The information landscape for engineering is particularly complex, and encompasses, in addition to learning materials and scholarly articles, resource types such as standards and specifications, design data, technical reports, trade news and patents. All of these form important parts of the engineer's information landscape. At the same time, there is a large number of publishers and national bodies in engineering. Professional societies play a particularly important role in the STM publication landscape - "Society and other non-profit titles account for 85% of the top 20, three quarters of the top 200 and two thirds of the top 500 ISI ranked titles"³.

Engineering is therefore an ideal test-bed sector for the analysis and study of cultural and political issues surrounding a subject-based digital repository landscape. Issues studied would include how professional society and other repositories containing a variety of different resource types can fit alongside institutional repositories in the information landscape of the future. Though particularly pertinent to engineering, many of the findings would map to other subject areas.

2.4 Work Plan

The work plan below divides the effort into five broad work packages covering: a review of the digital repository landscape; set up of a test-bed pilot service; working with end-users and data providers; pilot enhancement; and project management.

Work Package 1: Review of Digital Repository Landscape and Identification of content sources in Engineering.

High Level Aim: Consideration of the digital repository landscape with particular reference to engineering and engineers information needs. Selection of appropriate digital repository sources and identification of gaps in digital repository provision.

WP 1.1. Rev	iew of Digital Repository Landscape within the Engineering Discipline
Description	of Work: The project team will conduct a review of the existing published work relating to the digital
repository lar	ndscape within engineering. Consideration will be given to engineering curriculum and research areas
within HE & I	E, the information needs of engineers, cultural issues specific to engineering and differing resource
discovery rec	uirements within the engineering discipline. Attention will be paid to any relevant UK studies (e.g.
Didet) and to	relevant overseas developments (e.g. NSDL).
Tasks	 Review of existing projects, services and initiatives.
	 Literature search of paper and electronic published materials.
	Gather and collate information.
Outcomes	Engineering Digital Repository Landscape analysis that will help inform the pilot development. The
	analysis will be made available on the project web site.
WP 1.2 Iden	tification of Digital Repository Sources within Engineering
WP 1.2 Iden Description	tification of Digital Repository Sources within Engineering of Work: Identification of potential engineering digital repository sources and their means of
WP 1.2 Iden Description interoperabili	tification of Digital Repository Sources within Engineering of Work: Identification of potential engineering digital repository sources and their means of ty (e.g. OAI-PMH, z39.50, and SRW). Directories of Repositories (e.g. DOAR, OAIster, Grainger
WP 1.2 Iden Description interoperabili Engineering	tification of Digital Repository Sources within Engineering of Work: Identification of potential engineering digital repository sources and their means of ty (e.g. OAI-PMH, z39.50, and SRW). Directories of Repositories (e.g. DOAR, OAIster, Grainger Repositories Listing) and the IESR will be utilised.
WP 1.2 Iden Description interoperabili Engineering Tasks	 tification of Digital Repository Sources within Engineering of Work: Identification of potential engineering digital repository sources and their means of ty (e.g. OAI-PMH, z39.50, and SRW). Directories of Repositories (e.g. DOAR, OAIster, Grainger Repositories Listing) and the IESR will be utilised. Identification of currently available and relevant Digital Repositories (See Appendix A for initial
WP 1.2 Iden Description interoperabili Engineering Tasks	 tification of Digital Repository Sources within Engineering of Work: Identification of potential engineering digital repository sources and their means of ty (e.g. OAI-PMH, z39.50, and SRW). Directories of Repositories (e.g. DOAR, OAIster, Grainger Repositories Listing) and the IESR will be utilised. Identification of currently available and relevant Digital Repositories (See Appendix A for initial listing).
WP 1.2 Iden Description interoperabili Engineering Tasks	 tification of Digital Repository Sources within Engineering of Work: Identification of potential engineering digital repository sources and their means of ty (e.g. OAI-PMH, z39.50, and SRW). Directories of Repositories (e.g. DOAR, OAIster, Grainger Repositories Listing) and the IESR will be utilised. Identification of currently available and relevant Digital Repositories (See Appendix A for initial listing). Identification of potential Digital Repository sources/providers which are currently not available
WP 1.2 Iden Description interoperabili Engineering Tasks	 tification of Digital Repository Sources within Engineering of Work: Identification of potential engineering digital repository sources and their means of ty (e.g. OAI-PMH, z39.50, and SRW). Directories of Repositories (e.g. DOAR, OAIster, Grainger Repositories Listing) and the IESR will be utilised. Identification of currently available and relevant Digital Repositories (See Appendix A for initial listing). Identification of potential Digital Repository sources/providers which are currently not available for federated searching. These sources/providers may be approached in WP3.2
WP 1.2 Iden Description interoperabili Engineering Tasks	 tification of Digital Repository Sources within Engineering of Work: Identification of potential engineering digital repository sources and their means of ty (e.g. OAI-PMH, z39.50, and SRW). Directories of Repositories (e.g. DOAR, OAIster, Grainger Repositories Listing) and the IESR will be utilised. Identification of currently available and relevant Digital Repositories (See Appendix A for initial listing). Identification of potential Digital Repository sources/providers which are currently not available for federated searching. These sources/providers may be approached in WP3.2 Identification of obvious gaps in Digital Repository provision within Engineering.
WP 1.2 Iden Description interoperabili Engineering Tasks Outcomes	 tification of Digital Repository Sources within Engineering of Work: Identification of potential engineering digital repository sources and their means of ty (e.g. OAI-PMH, z39.50, and SRW). Directories of Repositories (e.g. DOAR, OAIster, Grainger Repositories Listing) and the IESR will be utilised. Identification of currently available and relevant Digital Repositories (See Appendix A for initial listing). Identification of potential Digital Repository sources/providers which are currently not available for federated searching. These sources/providers may be approached in WP3.2 Identification of obvious gaps in Digital Repository provision within Engineering. Listing of currently available digital repository sources, potential digital repository sources and gap
WP 1.2 Iden Description interoperabili Engineering Tasks Outcomes	 tification of Digital Repository Sources within Engineering of Work: Identification of potential engineering digital repository sources and their means of ty (e.g. OAI-PMH, z39.50, and SRW). Directories of Repositories (e.g. DOAR, OAIster, Grainger Repositories Listing) and the IESR will be utilised. Identification of currently available and relevant Digital Repositories (See Appendix A for initial listing). Identification of potential Digital Repository sources/providers which are currently not available for federated searching. These sources/providers may be approached in WP3.2 Identification of obvious gaps in Digital Repository provision within Engineering. Listing of currently available digital repository sources, potential digital repository sources and gap areas. Details will be made available on the project web site

Work Package 2: Set up Test-bed Pilot Service

High Level Aim: Set up a pilot service offering basic resource discovery functionality across multiple digital repositories, in accordance with the JISC Integrated Information Environment and including available and relevant software toolkits for resource discovery.

³ Morris, S. and Olivieri, R. The secret life of STM publishing. <u>Serials</u> 17(2), July 2004 p.116

WP 2.1 Pilot	/P 2.1 Pilot Resource Discovery Service Set up	
Description	of Work: Set up a software framework and inclusion of selected repositories.	
Tasks	Setting up of a resource discovery service using SPP software.	
	Addition of identified repository targets.	
	Ongoing support, development and debugging.	
Outcomes	Basic pilot service operational.	

WP 2.2 Plug-in of Software Tools for Resource Discovery in Digital Repositories

Description of Work: Addition of software tools & toolkits which enable resource discovery within digital repositories. Relevant tools should support resource discovery via Metadata harvesting (OAI-PMH), Cross Searching (Z39.50, SRW) or via RSS. UK toolkits such as Discovery+ (D+) and Middleware for Distributed Cognition (MDC) will be considered along with other potential software tools (e.g. OAIster Middleware (Digital Library eXtension Service)).

Tasks		•	Plug-in of relevant toolkits into the pilot service infrastructure.
		•	Set-up of repository targets to enable them to be utilised by plugged in toolkits.
		•	Evaluation of the functionality of the software toolkits to feed into their future development.
Outco	mes	Ov	rerview of functionality offered by range of software tools. Listing of functional requirements for full
		SC	ale subject-based resource discovery service.

Work Package 3: Working with End-Users and Data Providers

High Level Aim: Solicit feedback from a range of end-user groups at different stages of the project. Work with a range of different types of data providers relevant to engineering (e.g. institutional repository providers, professional societies, publishers, and national bodies) to encourage participation.

WP 3.1 End-	User Evaluation of Pilot
Description	of Work: Solicit end-users' views and opinion on the pilot service, the range of materials available
and the appro	opriateness of a subject-based approach. Focus groups will be conducted after the initial pilot setup
and after con	npletion of the pilot enhancements (WP4). Continual and ongoing feedback from project partners.
Tasks	Identify stakeholder groups (e.g. Lecturers, Researchers, Students, FE/HE, information
	professionals).
	Focus groups and interviews with users.
	Document user experience & views.
Outcomes	Establish end-user opinion on the pilot to inform pilot enhancement work (WP4).
	Establish end-user opinion on: the appropriateness of a subject-based approach to resource
	discovery; the range of different types of materials available in the pilot service; and identification of
	gaps in coverage.
L	

WP 3.2 Adv	ocacy and Encouraging Participation from Data Providers
Description	of Work: Work with potential data providers to understand their needs and motivation, encourage
participation	and promote interoperability. Experience from previous JISC/PALS ⁴ projects indicates that this is an
effective app	roach.
Tasks	• Produce advocacy materials which clearly explain the approaches to, and benefits of, enabling
	resource discovery in multiple repositories.
	• Work with a range of potential data providers (e.g. institutional repository providers, professional
	societies, engineering publishers, EPSRC, etc) to better understand issues.
	• Investigate potential solutions to minimise barriers for data providers (e.g. the OLAC approach ⁵)
Outcomes	Advocacy materials relating to the exposure of Digital Resource collections available on project
	website. Clearer understanding of the cultural, political and commercial issues involved in digital
	repository exposure from various data providers perspectives. New relevant digital repositories may
	consequently be made available.

⁴ The JISC/PALS funded Engineering Trade Information Metadata in RSS Project involved collaboration and advocacy with a number of engineering publishers and resulted in effective promotion and uptake of RSS (see http://www.eevl.ac.uk/projects_503.htm).

⁵ Simons G. and Bird S. Building an Open Language Archives Community on the OAI Foundation. <u>Library Hi Tech</u>, June 2003, vol. 21, no. 2, pp. 210-218(9)

Work Package 4: Pilot Enhancement

High Level Aim: Build on the basic functionality established in work package 2.1. This includes; addressing maintenance issues, addition of new repository sources, interactions with infrastructural shared services, enhancing metadata quality, embedding and reuse, improving search and browse results presentation and service profiling for particular audience groups.

WP 4.1 Addressing Maintenance Issues

Description of Work: There are maintenance issues associated with enabling resource discovery across multiple digital repository collections. Data sources may be updated frequently necessitating reharvesting, distributed search targets may change or become unavailable, and standards may change. This package addresses the practical concerns of maintaining an operational and up to date search service.

Tasks	 Investigate available tools for automatically reharvesting of OAI Content (e.g. NSDL automated ingestion system).
	Implement automatic harvesting mechanism for OAI repositories.
	• Consider maintenance issues relating to distributed search targets (e.g. Z39.50).
	• Work with IESR to utilise up to date target information from central registry (see WP 4.3).
	Investigate appropriate error reporting mechanisms, both to administrators and end-users.
	 Quantify and document maintenance effort required during the pilot.
	 Ongoing monitoring of current trends in interoperability standards.
Outcomes	Automatic harvesting implemented for OAI repositories, and error reporting implemented for
	distributed search targets. Better understanding of maintenance issues and quantification of
	maintenance effort.

WP 4.2 Addition of New Digital Repository Sources

Description	of Work: The digital repositories landscape is dynamic. This package encompasses the
identification	and addition of new repository sources as well as the means to include sources via emerging access
methods (e.g	g. Web Services).
Tasks	Addition of new repository sources resulting from advocacy work (see WP 3.2)
	Ongoing monitoring of digital repositories landscape within engineering.
	• Pilot enhancement to permit inclusion of repository sources via emergent access methods (e.g.
	SRW).
Outcomes	New repository sources of added to pilot, and means of inclusion methods expanded

WP 4.3 Interactions with Infrastructural shared services

Description of Work: This work package will ensure that the pilot service makes the best possible use of the emerging JISC infrastructural shared services. A range of useful shared services are becoming available for interaction with the pilot resource discovery service. Possible infrastructural shared services of interest include: Service Registries (e.g. IESR), Resolver Services (e.g. Edina OpenURL Router), Terminology Services (e.g. HILT M2M Demonstrator), Harvesting Services, and Metadata Schema Registries. The work package will analyse potential usage scenarios of relevant infrastructural services and implement usage of example shared services on a test basis.

Tasks	 Identify all relevant infrastructural shared services & develop potential usage scenarios.
	Work with 'mature' services including Service Registries and Resolver Services.
Outcomes	Infrastructural shared services usage scenarios.
	Experimental usage of minimum of 2 shared services (e.g. IESR & Resolver services).

WP 4.4 Enhancing Metadata Quality

Description of Work: Metadata from harvested repositories is of highly variable quality and may contain a wide variety of errors and inconsistencies. This work package investigates approaches to enhancing metadata quality with a view to improving resource discovery.

Tasks	•	Review metadata augmentation techniques, particularly NSDL work ⁶ .
	•	Identify and implement safe metadata transformations which are applicable to all collections (e.g
		removal of empty metadata elements, double XML encodings, etc).
	•	Consider options for metadata augmentation on an individual collection basis.

⁶ Hillmann, D. I., Dushay, N. and Phipps, J. Improving Metadata Quality: Augmentation and Recombination. Paper presented at the DC2004 Conference (Shanghai, China, Oct. 2004).

Outcomes	Demonstration of metadata augmentation functionality. Improved understanding of metadata
	augmentation issues and recommendations on appropriate approaches.

WP 4.5 Embedding and Reuse of Resource Discovery Services

Description of Work: Exploration of the issues involved in attempting to embed and reuse resource discovery services within different user environments. This will include consideration of relevant interoperability standards including RSS, JSR 168 and IMS specifications.

Tasks	 Embed basic pilot resource discovery functionality within a VLE at host institution on a proof of concept basis⁷.
	Add functionality to pilot to enable users to save search results in formats appropriate for reuse.
Outcomes	Examination of the issues surrounding embedding and reuse and report documenting the viability of
	various approaches.

WP 4.6 Improving Results Presentation

Description of Work: Current federated search tools tend to present results from multiple repositories separately in an unfiltered manner. This package examines the possibilities for enhancing the pilot service by employing search and browse results filtering techniques.

Tasks
Outcomes
Outcomes

WP 4.7 Service Profiling

in control i ching						
Description of Work: Profiling the basic resource discovery service based on the needs of particular audiences or						
groups (e.g. FE, HE, students, lecturers, researchers, etc).						
Tasks • Consider the needs and typical usage scenarios of various user groups (relates to V						
	•	Analysis of the collections offered, collection groupings, terminology and branding styles that would be appropriate for particular user groups.				
	•	Investigate the potential of using filtering techniques (e.g. Adaptive Concept Mapping) to provide contexts based on user groupings.				
	•	Develop instance of profiled service for one particular user group.				
Outcomes	Separate, trial instance of the pilot service based on a particular user group's needs.					

Work Package 5: Project Management & Participation in DR Programme

High Level Aim: Well-managed project contributing to, and learning from, the overall Digital Repositories programme. The JISC programme management guidelines will be adopted.

WP 5.1 Project Management						
Tasks		•	Production of detailed project plan.			
		•	 Internal communication, progress reports, meetings information dissemination, etc. 			
		•	Project website with up to date details of project progress.			
		•	Final Report documenting particular issues raised in pilot development.			
Outcom	ies	Well run project reporting in a timely fashion.				

WP 5.2 Participation in Digital Repositories Programme						
Description of Work: Active involvement in the DR Programme						
Tasks	Participation in Digital Repositories Programme meetings/events.					
	Input into relevant standards, specifications, protocols and frameworks activities.					
Outcomes	Contribution to successful programme outcomes.					

⁷ This task builds on work currently being undertaken as part of a JISC funded ELF demonstrator project at ICBL

Figure 1 Project Timescales



Key: Bars show timescale of work package. **Milestones: MS1** Basic Pilot Service Operational (start Sept 05), **MS2** Pilot enhancements investigated/delivered (start April 07). **MS3** Project Completed, all project outcomes delivered, Final Report Submitted (end June 07).

2.6 Risk Analysis

Risk	Probability (1-5)	Severity (1-5)	Score (PxS) (1-5)	Action to prevent/manage risk
Loss of Staff	1	3	3	Staff associated with this proposal are currently in post and are available from June 2005. At worst, loss of staff would impact on the timely running of the project, as staff would have to be replaced. Alternative staff have been identified. Likelihood of affecting the successful completion: negligible.
Technical failure	1	1	1	The hardware to be used is well managed and maintained, and data is regularly backed-up. The software is kept up-to- date with security patches. Contingency machines are available in the event of hardware failure. Likelihood of affecting successful completion: negligible.
Software failure	2	2	4	JISC have funded maintenance of SPP software until 2008.
Organisational problems	2	1	2	Lead site has project management experience. Letters of commitment have been provided by partners in support of project. Partners will be required to sign a Consortium agreement.

2.7 IPR and Sustainability Statements.

Sustainability Statement - The project will give consideration to the sustainability issues relating to fully-fledged service provision. Analysis of various sustainability models for full service provision will be provided in the final report. Possible models include JISC funded service, commercially sponsored service, subscription based service or via combined approaches. Early indications suggest that a future fully-fledged service might be at least partially sustainable through sponsorship from commercial publishers or via subscriptions.

Intellectual Property Rights Statement

All software and technologies developed by the project will be open source. The use of third party data or software will comply with IPR legislation, and if necessary appropriate agreements will be made for their usage in the pilot service.

3. Budget

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

4. Key personnel

Staff associated with this proposal are currently in post and are available as of July 05. No advertising for additional staff is currently envisaged.

Michael Breaks. Michael Breaks is University Librarian at Heriot-Watt. He has been Director of EEVL, a Hub of the Resource Discovery Network, since 1996. He is currently chair of the Archives Hub Steering Committee, and is a member of the Eduserv Board of Trustees. He is past President of IATUL (International Association of Technological University Libraries).

Roddy MacLeod. Roddy MacLeod has extensive experience of subject-based information retrieval in engineering, and recently edited the 4th edition of "Information sources in engineering" Munchen: KG Saur, 2005. He has been EEVL Manager since its inauguration in 1996 and was responsible for EEVL's contribution to the SPP. He has experience of managing distributed project partnerships and partnerships with publishers.

Roger Rist. Dr Roger Rist has directed the Institute for Computer-Based Learning (ICBL) at Heriot-Watt University since 1993. Under his direction ICBL has taken a leading role in over thirty research and development projects. These projects have focused on design and evaluation issues relating to the use of information and learning technology in education. Dr Rist is a member of the JISC Learning and Teaching Committee and chair of the Steering Group for the UK Centre for Educational Technology Interoperability Standards (CETIS).

Phil Barker. Dr Phil Barker has worked supporting the use, description and evaluation of computer-based learning resources for ten years. In the past he has worked for the Computers in Teaching Initiative (CTI) Centre for Physics, the Learning Technology Dissemination Initiative, and EASEIT-Eng, a TLTP phase 3 project that evaluated computer-based learning resources used in Engineering. Currently he is Learning Technology Adviser to the HE Academy Engineering Subject Centre and coordinator of the CETIS Metadata and Digital Repository Special Interest Group. Dr Barker has extensive experience in evaluating how science and engineering academics use online resources to support their teaching, and in supporting the development of open technical standards for metadata and digital repositories.

Malcolm Moffat. Malcolm Moffat has held positions as EEVL Development Manager, Science Faculty Librarian and Lecturer in Information Management. In his current position at EEVL he has responsibility for; identifying information sources in engineering, metadata and interoperability projects (including RSS and OAI-PMH), SPP usability evaluation, user needs analysis, engineering portal focus groups, and liaison with external parties and organisations. He has extensive experience of resource discovery and federated search technologies and specialist knowledge of information sources (Z39.50 targets etc.) in the area of engineering.

Santiago Chumbe. Dr. Santiago Chumbe works on the technical sides of EEVL projects and he has been a research associate at ICBL since 2000. Santiago worked on the distributed development of JAVA-based resource discovery services while working on the Subject Portals Project, and is currently involved in Heriot-Watt's implementation of VLEs. His software expertise includes J2EE, XSLT, SQL, Z39.50 and OAI.

APPENDIX A: POSSIBLE REPOSITORY CONTENT SOURCES

Potential repository types of interest may include the following;

- Institutional Repositories (e.g. QUEprints (digital repository of Cranfield University research), ePrintsUK)
- Technical Reports (e.g. NASA Technical Reports, NACA, ARC)
- Learning Object Repositories (e.g. JORUM, SMETE, MERLOT, CARERO, NLN)
- Image Repositories (e.g. EMOL, SCRAN)
- Multimedia Repository Sources (e.g. GROW)
- Thesis Repositories (e.g. ETHOS project outcomes)
- Technical Data Repositories (e.g. IHS)
- Publishing Community Repository Sources (e.g. leading engineering publishers; Elsevier, Springer, Thomas Telford, Pearson, Morgan & Claypool, Inderscience)
- Professional Society Repository Sources (e.g. ICE Virtual Library, CISTI)
- Funding Council Sources (e.g. ESPSC funded research collection)