

Repositories and Preservation Proposal Cover Sheet

Cover Sheet for Proposals (All sections must be completed)		<i>JISC Capital Programme</i>	
Name of Capital Programme: Repositories and Preservation Programme			
Bid for Call Area : (Please tick ONE BOX ONLY, as appropriate)			
Tools and Innovation (Strand B)			
<input type="checkbox"/>	Call Area I – Tools and Innovation Projects	Please specify area of proposed project eg <i>'metadata generation and validation'</i>	
Discovery to Delivery (Strand C)			
<input type="checkbox"/>	Call Area II – Discovery to Delivery Projects	<input type="checkbox"/> a) Version identification framework <input type="checkbox"/> b) Persistent identifier interoperability demonstrator <input type="checkbox"/> c) Federated access management and repositories <input type="checkbox"/> d) Semantic interoperability demonstrator	
Repository Start-Up and Enhancement (Strand D)			
<input type="checkbox"/>	Call Area III – Repository Start-Up and Enhancement Projects	<input type="checkbox"/> a) Repository start-up projects <input checked="" type="checkbox"/> b) Repository enhancement projects	
Digital Preservation and Records Management (Strand H)			
<input type="checkbox"/>	Call Area IV – Digital Preservation and Records Management Projects	<input type="checkbox"/> a) Digital preservation across the lifecycle <input type="checkbox"/> b) Models and implementation of preservation services <input type="checkbox"/> c) Preservation tools development	
Shared Infrastructure Services (Strand I)			
<input type="checkbox"/>	Call Area V – Shared Infrastructure Services Projects	<input type="checkbox"/> a) Pilot implementation of licence registry <input type="checkbox"/> b) Pilot national name and factual authority service <input type="checkbox"/> c) Scoping an architecture to support digital policy management <input type="checkbox"/> d) Scoping a terminology registry	
Name of Lead Institution: University of Oxford			
Name of Proposed Project: Bridging the Interoperability Divide (BID)			
Name(s) of Project Partner(s): Monash University			
Full Contact Details for Primary Contact:			
Name: Neil Jefferies			
Position: Acting Development Manager			
Email: neil.jefferies@sers.ox.ac.uk			
Address: Oxford University Library Services, Osney One Building, Osney Mead, Oxford, England. OX2 0EW.			
Tel: 01865 280588			
Fax: 01865 204937			
Length of Project: 18 months			
Project Start Date: 1/4/2007		Project End Date: 1/10/2008	
Total Funding Requested from JISC:			
Funding Broken Down over Financial Years (April – March):			

Apr06 – Mar07	Apr07 – Mar08	Apr08 – Mar09
0	120202	63676
Total Institutional Contributions:		
Percentage Contributions over the Life of the Project:	JISC 50%	PARTNERS 50%
Outline Project Description		
<p>The Bridging the Interoperability Divide (BID) project will use the principles underpinning the JISC e-framework to build interoperability between 3 repository systems: SRB, Fedora and ASK. The implementation will help join e-science, academic publishing and learning/ teaching practice communities by creating a joined-up set of repository services. The project will focus on demonstrating interoperability across the federation for the following services: harvesting (OAI), federated search (SRW), authentication (Shibboleth), metadata management (MODS/ METS), identifiers, and discovery (OpenURL). The project will also create a client for authenticated bulk upload (ingest service) into an institutional repository.</p>		
I have looked at the example FOI form at Appendix A and included an FOI form in the attached bid (Tick Box)	YES ✓	NO
I have read the Circular and associated Terms and Conditions of Grant at Appendix B (Tick Box)	YES ✓	NO

1. Introduction

The Bridging the Interoperability Divide (BID) project will use interoperability specifications and service-oriented design principles to enhance 3 repository systems that are currently implemented at Oxford and Monash. In this section we describe each of these systems, the communities they currently serve and the use cases that are driving development to bring the system in line with the JISC e-framework.

1.1. Computational experimentation

The Oxford e-research centre is committed to developing and maintaining the campus grid named OxGrid. This has been designed so that in the future there will be a single point of entry to university researchers using computational and storage infrastructure. Currently the system has ~30 users in many fields from Computational Chemistry, Materials Science and Theoretical Physics and overall have submitted many thousands of individual computational tasks. The process by which a user runs on the campus grid is as follows: (1) Generate overarching experimental configuration file (2) Submit parent process to the grid which launches a large number of child processes each with their own configuration data. There is an n-times increase in input configuration data at this point. (3) Each child process then runs to completion after which all output data is returned to the parent process. The data volume increase at this point is dependant on the individual research. It can be as much as 1000x. (4) Any post processing is done. This can lead to a data reduction of another increase if visualisation for example is used.

The use cases driving the integration of the SRB system into a wider repository framework are:

Publishing datasets to institutional repository: Each user may repeat the process just outlined many tens of times the amount of individual data generated can be massive. Of this though only a small amount would be of the type that you would want to put into the institutional repository, i.e. the processed output data and possibly any post processing. It is important though that the link is maintained between the input data and output data as well as a full metadata description of the experiment to ensure no wasted time or computational resource. Therefore it is essential that there is a clear link between the campus grid data storage system, where the input and raw output data sets can stay and the institutional repository where the processed output can be held to be able to be directly referenced by uploaded material from the researcher

such as papers to ensure also in the longer term open access is granted to the generated research data for other studies not yet possible but which could be massively assisted by the availability of large amounts of relevant and catalogued raw data.

Access control: As one of the partner nodes of the National Grid Service (NGS) we also provide access to a very large proportion of the registered users on the system by allowing 'zoning' between the Oxford campus grid Storage Resource Broker¹ (SRB) instance and the NGS instance we can allow our users to index the data they have within the NGS instance from within the institutional repository. This will increase the usefulness of the National Grid's data nodes and increase data participation with that project. To read more on this subject please see the DART report².

The SRB software is accessed through a command line (see appendix 5.1 for a screen shot). Please contact David Wallom for a demonstration: david.wallom@oerc.ox.ac.uk

1.2. Learning and teaching

At Oxford as with many other educational organisations teachers and learners are able to store and share resources using a virtual learning environment (Bodington). VLEs do not typically provide comprehensive and interoperable repository services so to address this Oxford is being supported by the JISC in developing the ASK system which will provide a full set of repository services to the Bodington. In brief the project has further developed and integrated a number of open source software components to provide a comprehensive repository system (federated search, Shibboleth authentication integrated with fine-grained authorisation and group management, harvesting, metadata management, discovery (OpenURL) and direct access APIs for web service communication with external systems such as learning design applications, portals and VLEs. To bring the repository into a wider federation the following use cases need to drive further enhancement to ASK repository service:

Desktop to online repository: A student has 100 image files on her desktop that were taken to annotate recent field work and wishes to share them by emailing a link to all the files stored together as a collection.

Controlling access: A teacher needs an easy way to share tutorial resources just with other teachers in his department.

Searching: A teacher wants to create a collection that includes teaching resources, data sets, past exam papers and research papers for a seminar series.

Documentation and access to the demonstration services are available on the ASK web site³, a screen shot is in the appendix 5.2 and please contact Howard Noble with any questions: howard.noble@oucs.ox.ac.uk

1.3. Academic publication

Oxford Research Archive (ORA), the Oxford University Repository for research output, is Fedora-based, and includes the VITAL/VALET user-friendly interface for selected workflows. The repository includes ePrints (pre- and post-prints) and eTheses. These are obtained by a mixture of author deposit (or their representative), mediated deposit and bulk transfer from a variety of sources. Not all objects need to be held in the repository: metadata records for remote objects can form part of the collection, particularly for items such as large datasets. There are three main areas which are priority for the repository: long-term preservation (in line with the collection policy), open access (in compliance with copyright) and increased visibility, enabled by use of rich metadata produced to standard schema. To be useful to the communities which it serves, ORA will provide a means of efficient discovery of research output items, and include the means to link to related objects and provide a consistent metadata interface for harvesters and other third parties. ORA will be a point of discovery, storage and retrieval for objects across a diverse and complex institution. Fedora enables the creation of complex digital objects comprising multiple data-streams together with the creation of relationships between objects. The use cases that are driving the development of ORA to use interoperability specifications and web service integration to join with the proposed federation are:

Resource discovery through metadata records. Each digital object, collection and community will have a rich metadata record assigned to it. This will be used by, for example, researchers seeking objects

¹ Storage Resource Broker: http://en.wikipedia.org/wiki/Storage_resource_broker

² DART report: <http://ausweb.scu.edu.au/aw06/papers/refereed/treloar/paper.html>

³ ASK web site: <http://ask.oucs.ox.ac.uk/>

authored themselves and by colleagues, by external users wishing to obtain research publications by authors affiliated to Oxford University and for research-led teaching. Metadata will be created in MODS and METS as well as simple DC in order to allow for harvesting of rich metadata via OAI. This facility will benefit those wishing to find out about others working in similar or related fields, possibly for the purpose of cross-disciplinary research.

Linking to data/objects from within papers/theses. For example, an article may include a link to a dataset held in the eScience SRB system. Inclusion of metadata within ORA for the datasets will allow users to discover the dataset and link externally from the archive to access data.

Link to learning objects which reference a resource. A user of a learning object will be able to link directly to the digital object held within ORA in line with the eBank model for re-use of research⁴. This will include those using a VLE, both students studying and staff delivering courses. Likewise there are plans to investigate links and data sharing with VREs or research databases: initially work needs to be undertaken to identify possible targets (in addition to the Medical Sciences Division database) within the institution and then investigate methods of linking.

Provide a consistent metadata interface. ORA will comply with OAI-PMH for harvesting by third parties. As stated above, the metadata will be rich and will be available using multiple schema if harvesters wish to obtain more than simply Dublin Core. Currently it is feasible to ingest items to Fedora from other repositories/systems using OAI-PMH. In order to ingest metadata and/or objects from the ASK and eScience repositories a generic ingest tool is required. In order to offer consistent search and indexing within ORA, it will be necessary to obtain metadata for ingested objects and records from the federated repositories and map to the schema used in ORA.

The ORA system is currently part of a large advocacy drive at Oxford. Please see appendix 5.3 for a screen shot and contact Benjamin O'Steen (Benjamin.OSteen@sers.ox.ac.uk) with any questions.

2. Project description

2.1. Project management

The project will be coordinated by a steering group that is highly experienced in designing and delivering repository services. All project tasks will be agreed by the steering group with the project manager. All tasks will be structured so that they are properly informed by a detailed understanding on user cases and the technical design. Software will be created in an Agile development environment that emphasises open communication. A range of tools will be used to coordinate a potentially large number of developers in creating open source code: a project wiki for technical documentation; a mailing list for frequent communication; messenger software for group discussion; bug/ feature tracking for rapid small scale enhancement and test case tracking; Sourceforge for code distribution. The steering group will monitor task milestones and report to the JISC when milestones are met or need changing.

2.2. Technical design

The BID project will build interoperability between 3 distinct repository systems to deliver a common set of services to the 3 practices communities outlined in the introduction. The project steering group will coordinate the development of an institution-wide technical design that mandates standardisation in a number of key areas so that the benefits of interoperability can be gained. Here we outline the services the project development effort will focus on:

Authentication, authorisation and group management: Shibboleth allows for cross-institutional authentication between organisations but for it to be used effectively it needs to interoperate with authorisation and group management services. With this in place users will be able to control exactly which individuals and groups can See, Edit, Delete and Remove resource from a repository. Such fined-grained access management allows a person to store their resources online in a private briefcase (accessible only to themselves) and then gradually open access to the resource as it becomes more complete (for instance).

Search and discovery: One of the main reasons for using a repository is to make resources as accessible as possible. By describing resources with structured metadata complex searches can be executed to retrieve material most relevant to the user. The federation has already chosen the MODS and METS specifications

⁴ Scholarly Knowledge Cycle diagram (Lyon): <http://www.ariadne.ac.uk/issue36/lyon/>

but agreement is still incomplete on the detail of the exact vocabulary to use in populating these XML files. The quality of metadata will affect the way metadata is shared between systems. OAI-PMH and the federated search interoperability specifications Z39.50/ SRW/ SRU all depend on the mappings being consistent between systems. In the case of federated search it is not feasible to search a very large number of targets in any one user request because the speed of service will be too slow. UDDI or the IESR⁵ service can be used to create a registry of repositories that can be queried to configure a federated search service on the fly. Such a service can be used for instance to configure only the searching of scientific repositories when a scientist wants to find material by an author with a common name. Once a resource has been found through searching an OAI compliant repository or using a federated search service the user may want to gain full-text access. For this to be seamless an OpenURL Resolver can be used to construct 'deep links' to a resource in a range of repositories that an institution has access to (find the so-called 'appropriate copy'). For a Resolver to construct deep link a globally unique id is required. For books and journals this is typically achieved through using ISBN and ISSN numbers or, for individual articles, DOI. The federation aims to support deep linking and will research an appropriate solution.

Ingest: Users need to be deposit resources (or relevant metadata in the case of remote objects) into repositories through a variety of channels (e.g. through sending an email, submission through a web form, using a dedicated desktop application). It is also important to enable users to upload a large number of resources into a repository system. Bulk upload of resources through a variety of channels is one of the most important services in allowing repositories to gain a critical mass of material for their value to be realised. There is currently no well recognised 'standard' to implement in this area. The Deposit API initiative⁶ initiated by JISC has made a start at researching this area. Shibboleth-based authentication adds another dimension to any solution.

In **figure 1** below we outline our initial ideas on how the federation of repository services will be joined by implementing interoperability specifications and web services. (Please see corresponding numbers annotated in diagram):

- [1] and [2] denote *the harvestable interface* and authenticated access to the SRB system,
- [3], [4] and [5] denote *the harvestable interface*, federated search and ingest (bulk upload from desktop client) to the ASK repository system,
- [6], [7], [8] and [9] denote *metadata ingest, indexing, federated search/OAI-PMH interfaces and authenticated access to the fedora system respectively*.

⁵ Proposed IESR service: <http://iesr.ac.uk/use/use-cases/>

⁶ Deposit API work undertaken for JISC: http://www.ukoln.ac.uk/repositories/digirep/index/Deposit_API

while we understand the scenarios that require such a tool we are aware of a large number of possible designs and the difficulties posed in achieving the desired functionality and compliance with Shibboleth. We are also aware of the ease of use that emerging so-called Web 2.0 technologies like Flickr are developing for enabling the population of these repositories; we would look to learn from these technologies).

- Demonstrable federated search service working across Fedora, SRB and ASK metadata and with as many systems as possible outside of the federation. Interoperability will be dependent on compliance with Z39.50, SRW, SRU specifications.
- Demonstrable harvesting service across the federation and with systems outside of the federation. Interoperability will be dependent on compliance with the OAI-PMH specifications.
- Demonstrable implementation of a common approach to revision control through implementing an interoperable object identifier 'service'. The identifier service will enable a wide range of other services such as discover, search and harvest. This work will build on Oxford's work on the JISC the JISC River project⁷
- Demonstrable implementation of a common approach to metadata across the federation.
- Demonstrable implementation of a federation-wide discovery service compliant with the OpenURL specification.
- Online collection of scenarios and use cases and how they map to technical design decisions. This work will be conducted by the Client Relations Team (CRT), part of Oxford University Computing Services and will be made freely available via the CRT wiki⁸:

2.5. Value to the JISC

- The project will demonstrate a large-scale implementation of the principles underlying the JISC e-framework.
- This proposal will allow diverse user communities (teaching, learning, scientific and research) to shape the implementation of these and innovative new services.
- The project will produce open source software freely available to the wider community
- The project will support the eBank scholarly communication cycle model
- The project technical design will support the use and re-use of digital objects
- The project will straddle the boundaries of eResearch, teaching & learning and scholarly communications and demonstrate 'joined-up thinking' between these often overly specialised research areas.

2.6. Risks assessment

Risk	Probability (1-5)	Severity (1-5)	Score (P x S)	Action to Prevent/Manage Risk
Staffing availability	2	2	4	The project can draw from a pool of staff to cover for unexpected loss of staff. Task planning is key to managing this risk.
Organisational	2	3	6	The steering group is well positioned to gain early access to any organisation change and allow the project manager to plan for any alterations required.
Technical	3	2	6	The project is proposing some difficult technical challenges and will allow sufficient time to design a solution before beginning implementation (coding).
External suppliers	2	2	4	All external suppliers will be contractually bound to the University and we will only work with trusted suppliers.
Legal	1	3	3	The project can draw on research services offices and consult with OSSW on open source license issues.

⁷ JISC River project: http://www.jisc.ac.uk/index.cfm?name=jcie_scg

⁸ CRT wiki: http://wiki.oucs.ox.ac.uk/ltg-public/Client_Relations_Team

2.7. IPR Statement

Under the University of Oxford's policy on intellectual property (which covers all University employees and students), the University claims ownership of a range of intellectual property rights with commercial potential. The University does not assert any claim to the ownership of copyright in artistic works, books, articles or lectures, apart from those specifically commissioned by the University. Results arising from projects funded by the JISC at Oxford would therefore usually be owned in the first instance by the University as the employing institution. The University seeks to maximize the commercial potential of its intellectual property through its wholly-owned technology transfer company, ISIS Innovation Ltd. In accordance with the desires of the Repositories Enhancement Programme, however, it is proposed to release project deliverables under either a Creative Commons license or, in the case of software, under an open source software license to maximize the benefit for the wider community.

3. Budget

Directly Incurred Staff	March 07	April 07– March 08	April 08– March 09	TOTAL £
Shivram Rajamanickam (OUCS, Grade 7, 0.5 FTE)	£0	£19598	£10206	£29804
Ben O'Steen (OULS, Grade 7, 0.5 FTE)	£0	£19598	£10206	£40356
Teijun Ma (oeRC, grade 7, 1.0 FTE)	£0	£19598	£10206	£80712
Total Directly Incurred Staff (A)	£0	£58794	£30618	£89412
Non-Staff	March 07	April 07– March 08	April 08– March 09	TOTAL £
Travel and expenses	£0	£6438	£4244	£10682
Consultancy and non-UK partners ⁹	£0	£25750	£15914	£41664
Dissemination	£0	£515	£530	£1045
Total Directly Incurred Non-Staff (B)	£0	£32703	£20688	£53391
Directly Incurred Total (A+B=C) (C)	£	£91497	£51306	£142804
Directly Allocated	March 07	April 07– March 08	April 08–March 09	TOTAL £
Staff ¹⁰	£	£36512	£18727	£55239
Estates	£	£19732	£10063	£29795
Directly Allocated Total (D)	£	£56244	£28790	£85034
Indirect Costs (E)	£	£92663	£47255	£139918
Total Project Cost (C+D+E)	£	£240404	£127351	£367755
Amount Requested from JISC	£	£120202	£63676	£183878
Institutional Contributions	£	£120202	£63676	£183878
Percentage Contributions over the life of the project		JISC 50 %	Partners 50 %	Total 100%

⁹ Participation from Monash University and the consultants detailed on page 10

¹⁰ Neil Jefferies (0.2 FTE), Michael Fraser (0.025 FTE), Sally Rumsey (0.1 FTE), Howard Noble (0.1 FTE), David Wallom (0.1 FTE) See Section 4.2 for justification/details

4. Key personnel

4.1. Project Leads

Michael Fraser is head of the Research Technologies Service at Oxford University Computing Services, an Associate Director of the Oxford e-Research Centre, and Director of Institute Arts and Humanities (previously the Humbul Humanities Hub). He is a co-investigator for three VRE projects (Integrative Biology; Sakai VRE Demonstrator; and Building a VRE for the Humanities). He has also been principal investigator for a number of other projects relating to access management, institutional repositories, and humanities computing. He is a member of various advisory committees and is currently Technical Secretary for Oxford's ICT Strategy. He has an academic background in the humanities with a PhD in theology from Durham University.

Neil Jefferies (Project Manager) is Acting Development Manager & Strategy Coordinator for Oxford University Library Services, responsible for the development and delivery of new services. He has recently managed projects to (1) use of Citrix farms to deliver legacy data resources to browser clients; (2) the implementation a Content Management System for use by Oxford's Libraries; (3) the deployment of Thin Clients as library terminals and the Eprints and Fedora Repositories (with corresponding involvement in the SHERPA and PRESERV projects). Previously he has worked in a broad range of computer-related fields from chip designs and parallel algorithm development for Nortel, through writing anti-virus software for Dr Solomon's and developing corporate systems for several major blue-chips. He has an MA in Natural Sciences from Cambridge and an MBA from Warwick Business School.

4.2. Steering Group

As this project involves the integration of three repositories which use different technologies and have differing design objectives, it is felt that the Managers of the three repositories need to have significant involvement in order to ensure success. As with any such systems integration effort, organisational issues need to be managed closely. To this end, in addition to an overall Project Manager, a hands-on steering group is proposed, comprising:

Howard Noble (OUCS) manages the Client Relations Team at Oxford University and is PM of the JISC Accessing and Storing Knowledge (ASK) project. While at Oxford Howard has managed the JISC TIP and MDC projects and been a project member of the JISC RIVER study. Before joining Oxford Howard was a project officer on the JISC OLIVE project at the Royal Holloway University. Before working in the educational sector Howard was an IT consultant for IBM Global Services and Extraprise (a Client Relationship Management consultancy). Howard has a BSc. in Genetics from Manchester University and a MSc. in Cognitive Psychology from Sussex University.

Sally Rumsey (OULS) is the Project Manager leading the development and implementation of a sustainable repository for research output (initially ePrints and eTheses) of Oxford University. She previously worked at the London School of Economics where she was eServices Librarian. While there, she was the Manager of the LSE research repository (LSE Research Online), led the team which implemented an electronic resource management system and was involved with most aspects of electronic resources including licences, authentication and digitisation. Her book, "How to find information: a guide for researchers" is published by the Open University Press. She is a member of the JISC eBooks Working Group. Sally has a BMus and an MA in Information Studies.

Dr. David Wallom (OeRC) is the Technical Manager of the Oxford e-Research Centre. In his present role he is expected to engage the wider community within the university with e-Science tools and technologies and especially with the e-Infrastructure within the university including Campus grid, virtual research environments and the institutional repository. He has advised the [ShibGrid](#) project on usage models within campus grids. He has a degree in Applied Physics from Coventry University with a year in industry and a PhD in Experimental Particle Physics from the University of Bristol. He is also currently chair of the UK e-Science Engineering Task Force and co-chair of the Open Grid Forum Production Grid Services Research Group.

4.3. Development Team

Ben O'Steen (OULS) is the Software Engineer implementing the design and delivery of a sustainable repository for research output (initially ePrints and eTheses) for Oxford University. He has over 10 years experience building bespoke business-critical systems and tools for corporate and private clients, including London Underground Limited, [MacAlpine](#) Helicopters and recently, Oxford Aviation Training. He holds a Bachelor's degree in Chemistry from Oxford.

Tiejun Ma (OeRC) is the Campus Grid Developer within the OeRC. He has special responsibilities for the data system and uptake by non-traditional grid users. He also has general responsibilities for user uptake and code integration. He has a BSc from Dalian University of Technology, China and is submitting his thesis for a PhD in Computer Science from University of Edinburgh with a project on development of failure detection for dependable Grid and Web services.

Shivram Rajamanickam (OUCS) is a Software Developer in Oxford University specialised in the web based software programming and development. He is currently working for the JISC funded ASK project in designing the system using Struts MVC architecture and UML, development of user interface using JSP, CSS, XML and XSLT, enhancement of Federated search feature by performing the search on SRW/SRU repositories using Java / J2EE technologies, web services like SOAP, AXIS, XSLT and eXist database. Shivram has an MSc. in *Embedded systems and Robotics* in University of Essex in 2002.

Peter Crowther (consultant) BSc, MIEE, MBCS brings 25 years experience of software design and development to his role as IT Director of Melandra Limited. His research work spans metadata, repository, e-framework component, VLEs, data and knowledge representation and interchange for the universities of Oxford, Manchester, Southampton, Leeds and UHI. Peter's 5 years as a professional educator have refined his communication skills, an important area sometimes overlooked by his profession. Peter graduated top of his year containing 120 Computer Science students at Manchester University.

Andrew Foster (consultant) is a technical architect specialising in the development of multi-tier and distributed web-based systems. Having been in the industry for over ten years he has worked for IBM, Atos Origin, Sema, Hewlett Packard and most recently Oxford University. Andrew has shown his ability to adapt by working in a wide range of industry sectors including finance, medical, utilities, government and transportation. Technically Andrew is an object-oriented designer with full lifecycle experience using Java and J2EE. He has worked on multiple research projects for JISC along with being a key contributor to the JAFER API updating its Z3950, SRW, CQL and UDDI support. Andrew also has strong web service and XML/XSLT skills and a good understanding of the eXist XML database.

Richard Doe (consultant) is a web developer specialising in database driven website applications, with a front-end focus on web standards, accessibility and usability. He has worked in website development for 5 years, primarily for educational institutions and NGO's, including the University of Oxford, medical research charity Action Medical Research, student campaigning network People & Planet, and the University of Sussex. He has a BA in Philosophy from Sussex.

5. Appendices

5.1. SRB repository system

```
chem0162@oxgrid-vom chem0162]$ Sls -al
/IeRC/home/chem0162.OxGrid/ATTOSEC:
C-/IeRC/home/chem0162.OxGrid/ATTOSEC/0_10_10
C-/IeRC/home/chem0162.OxGrid/ATTOSEC/0_14
C-/IeRC/home/chem0162.OxGrid/ATTOSEC/0_14_trpzd
C-/IeRC/home/chem0162.OxGrid/ATTOSEC/0_14_trpzd_cmprss
C-/IeRC/home/chem0162.OxGrid/ATTOSEC/0_17
C-/IeRC/home/chem0162.OxGrid/ATTOSEC/0_20
C-/IeRC/home/chem0162.OxGrid/ATTOSEC/0_20_10
C-/IeRC/home/chem0162.OxGrid/ATTOSEC/0_20_trpzd
C-/IeRC/home/chem0162.OxGrid/ATTOSEC/0_24
C-/IeRC/home/chem0162.OxGrid/ATTOSEC/0_28
C-/IeRC/home/chem0162.OxGrid/ATTOSEC/0_28_6C
C-/IeRC/home/chem0162.OxGrid/ATTOSEC/0_2_6C
chem0162@oxgrid-vom chem0162]$ Sls -ar| 0_10_10
/IeRC/home/chem0162.OxGrid/ATTOSEC/0_10_10:
C-/IeRC/home/chem0162.OxGrid/ATTOSEC/0_10_10/ATTO-20060616142302
C-/IeRC/home/chem0162.OxGrid/ATTOSEC/0_10_10/ATTO-20060616144327
C-/IeRC/home/chem0162.OxGrid/ATTOSEC/0_10_10/ATTO-20060616145918
C-/IeRC/home/chem0162.OxGrid/ATTOSEC/0_10_10/ATTO-20060618122944
/IeRC/home/chem0162.OxGrid/ATTOSEC/0_10_10/ATTO-20060616142302:
chem0162 0 OxGrid-data01 73 2006-09-07-14.19 % 0-dynn_nd.inp
chem0162 0 OxGrid-data01 73 2006-09-07-14.19 % 1-dynn_nd.inp
chem0162 0 OxGrid-data01 74 2006-09-07-14.19 % 10-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 100-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 101-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 102-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 103-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 104-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 105-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 106-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 107-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 108-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 109-dynn_nd.inp
chem0162 0 OxGrid-data01 74 2006-09-07-14.19 % 11-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 110-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 111-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 112-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 113-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 114-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 115-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 116-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 117-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 118-dynn_nd.inp
chem0162 0 OxGrid-data01 75 2006-09-07-14.20 % 119-dynn_nd.inp
chem0162@oxgrid-vom chem0162]$
```

5.2. ASK repository system

The screenshot shows the ASK repository system interface. At the top, there is a navigation bar with "ASK Home" and "Briefcase" tabs. Below this, there is a search bar and a "Filter by type" dropdown menu. The main content area displays a table of resources with columns for Title, Author, Date, Type, Open URL, Email, Web URL, and Download. The table lists several resources, including "UML Diagram", "JISC B10 proposal", "Mentalese", "Noble Authors", "UML and JISC Proposal", and "Learning design at Athab".

Title	Author	Date	Type	Open URL	Email	Web URL	Download
<input type="checkbox"/> UML Diagram	Howard Noble	21/11/2006	RESOURCE	Open URL			
<input type="checkbox"/> JISC B10 proposal	Howard Noble	21/11/2006	RESOURCE	Open URL			
<input type="checkbox"/> Mentalese	Howard Noble	21/11/2006	READING_LIST	Open URL			
<input type="checkbox"/> Noble Authors	Howard Noble	11/21/06 11:44 AM	READING_LIST	Open URL			
<input type="checkbox"/> UML and JISC Proposal	Howard Noble	21/11/2006	READING_LIST	Open URL			
<input type="checkbox"/> Noble Authors	Howard Noble	11/21/06 11:48 AM	READING_LIST	Open URL			
<input type="checkbox"/> Learning design at Athab	Scott Wilson	21/11/2006	RESOURCE	Open URL			

5.3. ORA repository system

VITAL AccessPortal 2.1 | Item Display

File Edit View Go Bookmarks Tabs Help

Oxford Research Archiv... http://fenix.ouls.ox.ac.uk/access/detail.php?locale=en-us Go

Oxford Research Archive

OXFORD UNIVERSITY LIBRARY SERVICES

Home Browse Advanced Search Highlights Contribute About Help

Efficient computational appro Search Show All

Home > Search Results > Item Display

Here are some details about the authors, titles/alternate titles, and organization which published the work.

Label	Size	Format	Download
Thesis Descriptive Metadata	3.6 kB	XML Document	
Thesis Abstract	1.2 MB	Adobe PDF Document	
Full Thesis	592.4 kB	Adobe PDF Document	

Title: **Efficient computational approach to identifying overlapping documents in large digital collections**

Creator: **Monostori, Krisztian,1975-**

Description: Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy, School of Computer Science and Software Engineering (Cauffield).

Description: Summary: p. 7-8.

Subject: **Computer access control equipment industry.**

Subject: **Data structures (Computer science)**

Subject: **Data mining.**

Subject: **Combinatorial analysis**

Subject: **Computer algorithms.**

Subject: **Computer network resources.**

Identifier: eprints:32

Resource Type: Manuscript

Original filename: examplesthesisR242.xml

6. FOI Withheld Information Form

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.

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	None	None