



Project Information			
Project Acronym	RDInfokit		
Project Title	Resource Discovery Information Kit		
Start Date	2 July 2007	End Date	30 April 2008
Lead Institution	Centre for Digital Library Research, University of Strathclyde		
Project Director	Gordon Dunsire		
Project Manager & contact details	Gordon Dunsire (e: g.dunsire@strath.ac.uk)		
Partner Institutions	None		
Project Web URL	http://cdlr.strath.ac.uk/rdinfokit/index.htm		
Programme Name (and number)	Not applicable		
Programme Manager	Balviar Notay		

Document Name			
Document Title	Final report		
Reporting Period			
Author(s) & project role	Gordon Dunsire (Project director), Tony Ross (Project researcher)		
Date	8 July 2008	Filename	RDInfokitFinalReport.doc
URL	http://cdlr.strath.ac.uk/rdinfokit/docs/RDInfokitFinalReport.pdf		
Access	<input type="checkbox"/> Project and JISC internal		<input type="checkbox"/> General dissemination

Document History		
Version	Date	Comments
1.0	8 Jul 2008	



Resource Discovery Information Kit

Final Report

Authors: Tony Ross and Gordon Dunsire

Contact: Gordon Dunsire
Depute Director
Centre for Digital Library Research
University of Strathclyde
12.12 Livingstone Tower
26 Richmond Street
Glasgow G1 1XH
g.dunsire@strath.ac.uk
0141 548 4680

Date: 08 July 2008

Table of Contents

Table of Contents.....	1
Acknowledgements.....	3
Executive Summary.....	4
Background.....	5
The Centre for Digital Library Research.....	5
JISC.....	5
Infokits.....	5
Aims and Objectives.....	7
Methodology.....	8
Overview.....	8
Creation of taxonomic framework.....	8
Creation of database structure.....	9
Population of content.....	9
Construction of Web interface.....	10
Exit and sustainability strategies.....	10
Implementation.....	11
Overview.....	11
Creation of taxonomic framework.....	11
Creation of database structure.....	12
Population of content.....	12
Development of the Web service interface.....	14
Exit and sustainability strategies.....	19
Outputs and Results.....	20
Outcomes.....	22
Conclusions.....	23
Implications.....	24
Recommendations.....	26
References.....	27
Appendices.....	28
Appendix A: Case-study of the methodology and implementation of the iKit taxonomic framework.....	29
Introduction: the value of controlled vocabularies.....	29
Literature review and methodology.....	29
Implementation and results.....	31
Conclusion.....	34
References.....	34
Appendix B: Sources of terms for the JISC Information Environment thesaurus.....	35
Appendix C: SKOS case study.....	36
Investigation of the declaration of small-scale thesaurus in RDF/SKOS.....	36
Background.....	36
JISC and the semantic Web.....	36
The JISC Information Environment thesaurus.....	37
Conclusion.....	37
References.....	38
Appendix D: Inventory briefing paper.....	39
JISC inventory.....	39
Purpose.....	39
Methodology.....	39
Inventory database.....	39
Inventory items.....	39
Sources of information.....	39
Identifying inventory items.....	39
Recording inventory items.....	40
Monitoring and feedback.....	40
Further work.....	41
Appendix E: Resource Discovery iKit tutorial.....	42
Tutorial.....	42
Appendix F: Example of IE vocabulary preferred terms (from the interim report).....	47

Appendix G: Example of IE vocabulary non-preferred terms (from the interim report).....	52
Appendix H: Example of Infokit items (from the interim report).....	57
Appendix I: XML manifestation of RDF/SKOS expression of IE vocabulary.....	61
Appendix J: Comments received during beta-testing.....	68

Illustrations

Figure 1: Illustration of project methodology	8
Figure 2: Relationship structure diagram of the database	12
Figure 3: iKit IE thesaurus browse page (top) with CSS and images (default)	16
Figure 4: iKit IE thesaurus browse page (top) without CSS or images	17
Figure 5: iKit IE thesaurus browse page (bottom) without CSS or images	18
Figure 6: Display of tools for developing service registries, using a parameterised URL (http://cdlr.strath.ac.uk/rdinfokit/service/IEItemList.cfm?uEntryID=51)	24
Appendix A. Figure 1: JISC Information Environment concept map	33
Appendix E. Figure 1: iKit main page, including IE thesaurus concept-map	42
Appendix E. Figure 2: Intermediate result for keyword search "Edinburgh" within Contributing institutions facet	43
Appendix E. Figure 3: Full results display for all items associated with "EDINA at Edinburgh, University of"	44
Appendix E. Figure 4: Full record view example	45
Appendix E. Figure 5: Full term display for "terminology services"	46

Acknowledgements

The Resource Discovery iKit was funded by the Joint Information Systems Committee. It is not part of a specific JISC programme.

The project was carried out by staff of the Centre for Digital Library Research at the University of Strathclyde.

The project wishes to thank Jane Barton, Alan Dawson, Emma McCulloch, Keith McGinty, Balviar Notay, and all who responded to requests for feedback during the beta-testing stage.

Executive Summary

The Resource Discovery Information Kit project aimed to create a service to make it easier for JISC projects and those implementing resource discovery services within UK HE institutions to find information on resource discovery solutions.

The service was required to offer views of its content, consisting of reports, software and other relevant outputs from JISC projects, based on a taxonomic analysis of those components of the JISC Information Environment relevant to resource discovery. This entailed the project creating and documenting re-usable controlled vocabularies, including a thesaurus of Information Environment terms, to provide contextual search and retrieval functions. The project also investigated the use of Semantic Web standards (specifically Simple Knowledge Organisation System) to document such vocabularies.

The project was also required to implement the service in an integrated and sustainable way, and develop a plan for its future maintenance.

The overall approach used to create the service was iterative, with different strands of the methodology allowed to interact at all stages. The infrastructure chosen for the service uses a relational database management system to store metadata for resource discovery solutions, as well as the structure and content of the controlled vocabularies. The database is used to drive the content of dynamic web pages which provide the search and retrieve interfaces for the service. Appropriate standards for database design, data extraction, and web page design were used where relevant.

The RDInfokit project has produced an operational service (Resource Discovery iKit) at <http://cdlr.strath.ac.uk/rdinfokit/service/index.cfm>, established a database schema, and investigated exposure of relevant metadata vocabularies to the Semantic Web.

The service would benefit from further development to share metadata and interoperate with the PIMS project, establish a maintenance regime, and register relevant vocabularies in SKOS/RDF. Specific recommendations to JISC are:

- In the short term, JISC should resource the ongoing maintenance of the service, using the methodology developed for populating the inventory. CDLR would need to be kept informed of the completion of projects in the area of resource discovery so that it can identify additions to the inventory.
- In the longer term, JISC should consider other approaches to maintenance, including those identified in this report. Alternatives include:
 1. Development of a maintenance interface which can be used by third parties such as JISC managers and contributing institutions.
 2. Migration and integration of the iKit with another suitable content management infrastructure.
 3. Development of metadata interoperability between the iKit and related services such as PIMS.
- JISC should consider resourcing further development of the iKit to improve functionality, for example to meet the requirements of PIMS and other JISC services.
- JISC should consider the registration of iKit and other controlled vocabularies in RDF/SKOS and other Semantic Web technologies, to encourage re-use, reduce duplication, and improve metadata consistency and coherency.

Background

The Centre for Digital Library Research

The University of Strathclyde's Centre for Digital Library Research¹ was established in 2000, with two key aims, to combine theory with practice in innovative ways and to be a centre of research excellence in the area of distributed digital information and learning environments. The CDLR has a clear guiding philosophy of combining research and development, such that its research is well grounded and readily applicable, and its development work is based on sound principles, demonstrates best practice and wherever possible consolidates and extends skills, experience and knowledge. This philosophy is reflected in the Centre's position within the University, with connections to both the Department of Computer and Information Sciences and the Information Resources Directorate. The CDLR has a wealth of experience in the development of interoperable vocabulary control services (e.g. HILT²), and in the development of Web-based information retrieval systems (e.g. SCONE³; HAIRST⁴).

JISC

The Joint Information Systems Committee (JISC) is a committee of the UK further and higher education funding bodies that, through the funding of services and the management of development programmes, aims to help and facilitate institutions in their use of the Internet and other Information and Communications Technology applications.

As part of this mission, JISC seeks to develop ways to support seamless, flexible and comprehensive resource discovery. JISC's Information Environment⁵ activity aims to provide seamless access to well managed resources for learning and teaching, seeking to address the requirements of universities, colleges, learners and researchers via more flexible delivery. Through this activity, JISC-funded services and research programmes produce a great deal of genuinely innovative material. Such JISC Programmes include:

- the Portals Programme;
- the Presentation Programme;
- the Virtual Research Environments (VRE) Programme;
- the e-Learning Frameworks and Tools Programme;
- the Distributed National Electronic Resource (DNER 5/99) Learning and Teaching Programme;
- the Shared Infrastructure Services Programme;
- the User and Innovation Programme;
- the Semantic Grid and Autonomic Computing Programme.

Infokits

Although there is no formal definition of the term "infokit", it is in wide-spread use. The term generally refers to a set of information resources and tools presented in a contextualised framework to support a specific task or community, and is a contraction of the phrase "information kit".

The framework may be a linear narrative which refers to individual resources at appropriate points; this approach is often used when a specific task is to be supported. Similar frameworks are a simple check-list of resources or a decision-tree or flow-diagram leading to relevant resources.

¹ Centre for Digital Library Research, University of Strathclyde. <http://cdlr.strath.ac.uk/>

² High-Level Thesaurus Project (HILT). <http://hilt.cdlr.strath.ac.uk/>

³ Scottish Collections Network (SCONE). <http://scone.strath.ac.uk/service/>

⁴ Harvesting Institutional Resources in Scotland Testbed (HaIRST). <http://hairst.cdlr.strath.ac.uk/>

⁵ JISC Information Environment. http://www.jisc.ac.uk/whatwedo/themes/information_environment.aspx/

A more flexible approach, suitable for supporting a community carrying out a set of related tasks, is to use an inventory, registry or database of information resources and tools.

Many of JISC's projects and programmes produce resources and tools which support the development of resource discovery services in general. The aim of the Resource Discovery Infokit is to make it easier for institutions and resource discovery service providers to find information on resource discovery solutions, in particular those that support the JISC Information Environment and that have been developed via JISC Development programmes.

The main audiences for the Infokit are:

- Computing, library, learning technologists and those implementing resource discovery services in UK HE
- JISC projects

The Infokit is intended to offer views of information according to themes explaining clearly how to implement solutions and pointing to useful resources and projects.

Aims and Objectives

The main aim of the original project⁶ was to make it easier for JISC projects and those implementing resource discovery services within UK HE institutions to find information on resource discovery solutions. The original objectives were focused on the infokit as a functional product, explicitly to:

- create a Resource Discovery Information Kit (iKit) on the JISC website;
- offer views of the content of the iKit according to key questions and themes;
- explain clearly how to implement resource discovery solutions;
- provide pointers to useful resources and projects;
- develop a plan for the future maintenance of the iKit.

The project produced an interim report⁷ with two main recommendations:

1. to consider the feasibility of integrating the iKit completely with the JISC website;
2. to choose an approach to implementation from three options: a quick-fix by creating static web pages with narrative content; a compromise by creating static web pages from a structured inventory of appropriate content; or an innovative approach by creating dynamic, database-driven web pages (the project's preferred approach).

After consultation with JISC, the project agreed to host the iKit locally, with appropriate links to the JISC website, and to adopt the innovative approach. JISC granted the project an extension in order to support these outcomes⁸ (ref doc on website).

The revised aims and objectives, consolidated with the project extension, were:

- create a Resource Discovery Information Kit (iKit) on the CDLR website;
- offer views of the content of the iKit based on a taxonomic analysis of resource discovery components of the JISC Information Environment;
- to create and document re-usable, controlled vocabularies based on the taxonomic analysis;
- provide contextual pointers to useful resources and projects in the area of resource discovery solutions;
- to investigate, and where possible implement, the use of SKOS within the iKit and document this process as a case study;
- to implement the iKit in an integrated and sustainable way and develop a plan for its future maintenance;
- to document the process of creating the iKit as an exemplar of best practice.

⁶ Resource discovery infokit project. Project plan.
<http://cdlr.strath.ac.uk/rdinfokit/docs/RDInfokitProjectPlanPublic.pdf>

⁷ Resource discovery infokit project. Interim report.
<http://cdlr.strath.ac.uk/rdinfokit/docs/RDInfokitInterimReport.pdf>

⁸ Resource discovery infokit project. Project extension plan.
<http://cdlr.strath.ac.uk/rdinfokit/docs/RDInfokitExtensionPublic.pdf>

Methodology

Overview

- The methodology adopted in the first instance was to develop a faceted taxonomic framework based on the iKit requirements to inform decisions on implementation options.
- At the same time, the project used a combination of desk research and discussion with stakeholders to scope the best solution to hosting and maintaining the iKit service.
- An interim report, based on this work, including a technical appraisal, was produced for consideration by JISC.
- The chosen implementation option was to develop a set of dynamic, database-driven web pages for the iKit in order to improve scalability and support future maintenance of the service.
- The applicability of SKOS to the iKit was investigated through a combination of desk research and use of available authoring tools.

The project methodology was finalised following the interim report, and the programme of work divided into the following four work packages:

- WP1: Creation of taxonomic framework (including investigation of SKOS)
- WP2: Construction of a suitable database structure
- WP3: Population of iKit content
- WP4: Construction of a suitable web interface

A formal approach to the construction of use-cases and the design of system architecture was not thought necessary because of the CDLR's wealth of experience in the creation of Web based information retrieval systems and the fact that project staff (as developers of resource discovery solutions) were themselves representative of the iKit's target audience. This allowed the limited resources of the project to be most effectively allocated.

Instead, a holistic approach was taken where each of the four strands of the methodology were allowed to interact, using iteration so that the iKit as a whole was able to evolve flexibly, as illustrated below.

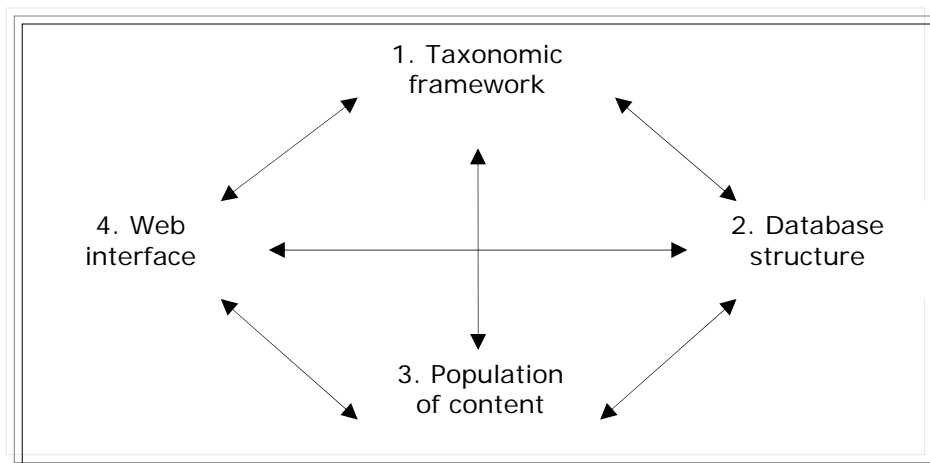


Figure 1: Illustration of project methodology

Creation of taxonomic framework

The use of a taxonomic framework to underlie retrieval of iKit items was preferred over less elegant options (e.g. full-text retrieval) due to its potential for interoperability, re-usability and sustainability. Following initial scoping work, the framework would include controlled lists of: contributing institutions,

JISC projects and JISC programmes and thesauri of controlled terms for both the functional and technical components of the IE. The creation of the taxonomic framework used a methodology adapted from Shearer⁹, which involved seven key stages:

1. Define taxonomic facets
2. Collect raw terms from literature and other sources
3. Identify relationships between terms
4. Add further terms to the thesaurus to fill significant gaps
5. Apply standard rules for the use of terms
6. Define term and add notation
7. Create alphabetical index and develop presentation

Standards applied to the framework include BS 5723¹⁰ (identical to ISO 2788) and BS 8723-1¹¹.

Appendix A of this report is a full case-study of the creation of the iKit taxonomic framework, including its methodology.

Investigation of declaration in SKOS

A secondary investigative aim of the project was to examine the use of the emergent Semantic Web standard Simple Knowledge Organisation System (SKOS)¹² to declare one aspect of the taxonomy, the Information Environment thesaurus, to allow it to be reusable in a standardised way. This task was determined to require the identification of relevant authoring and publishing tools, and the necessary infrastructural commitments required to support such a declaration.

Appendix C of this report is a full case-study of the investigation.

Creation of database structure

A relational database management system (RDBMS) model was adopted for the iKit database structure to support requirements of scalability and flexibility. Due to the need for rapid prototyping, and to use existing skill-sets to best expend project resources, MS Access software was used. It is readily available, in widespread use, is compatible with database-to-web publishing software, and provides sufficient functionality to support the many-to-many relational structure needed to link items to the taxonomic framework of the iKit.

The structure was developed using standard relational database methods, including normalisation to avoid unnecessary duplication of data. Third normal form¹³ was used throughout, with separate data table structures for distinct, logical sets of metadata fields. All tables have primary keys, and tables were related using only these keys. Many-to-many table relationships were instantiated using linking tables containing primary key values as secondary keys.

All primary keys are automatically generated by the RDBMS as "autonumber" values; that is, running numbers which are never duplicated or reused within any specific table.

Population of content

In the absence of any existing, definitive lists of either project outputs or indeed, lists of projects, programmes or institutions, the gathering of items for inclusion proceeded according to a thorough

⁹ Shearer, J.R. A practical exercise in building a thesaurus. In: Roe, K.S. and Thomas, A.R. (eds.). *The thesaurus: review, renaissance, and revision*. New York: Haworth Press, 2004. pp.35-36

¹⁰ *Guide to establishment and development of monolingual thesauri*. (BS 5723). British Standards Institution, 1987.

¹¹ *Structured vocabularies for information retrieval - Guide - Part 1: Definitions, symbols and abbreviations*. (BS 8723-1). British Standards Institution, 2005.

¹² Simple Knowledge Organization System (SKOS) - home page. <http://www.w3.org/2004/02/skos/>

¹³ Codd, E.F. Further normalization of the data base relational model. IBM research report RJ909. 1971.

inventory of the JISC website and attendant project websites. This was conducted in line with normal quality assured content and metadata creation processes, via the implementation of a simple workflow, including the application of controlled vocabularies and checks for internal consistency, etc. In particular, this quality assurance included necessary definitions of:

- the scope and criteria against which resources can be judged for their inclusion;
- a systematic process whereby a complete inventory of JISC programmes and services could be carried out;
- clear rules for the naming, description and categorisation of iKit items.

To assist these processes an inventory briefing paper was developed and is included here as Appendix D.

Construction of Web interface

The development of web pages necessarily requires compliance with various standards and accessibility legislation. This compliance was verified in conjunction with the JISC Communications and Outreach Team towards the end of the project and ahead of the publication of the website.

The web interface is fully conformant with XHTML 1.0 Transitional and CSS Level 2 standards. The W3C markup validation service¹⁴ and W3C CSS validation service¹⁵ were used to check conformance.

Reflecting the iterative nature of the iKit methodology, interface functionality was evolved alongside the other project work-packages. As part of quality assurance procedures, the iKit service was subjected to user-testing prior to launch to inform final adjustments to the ergonomics of the interface.

Exit and sustainability strategies

Overarching this methodology was the recognition that the future sustainability and flexibility of the iKit was of central importance; thus it must both easily maintainable and extensible. Two differing triggers for possible future changes to the iKit, namely *content* and *context* were identified, each offering distinct challenges:

Content: JISC programmes and projects will continue to generate resources which fall within the scope of the iKit. The Information Environment is not a paradigm; rather it is an evolving conceptual framework which can only be defined by the services and technologies active within it. Hence the content of the iKit will need to be refreshed periodically in order to ensure it continues to point users to the most up to date resources. This content will not only mean the addition of new items, but also new terms to the taxonomic framework.

Context: The database structure and Web interface of the iKit (i.e., the context within which the iKit service is delivered) must be considered separately from its content. These were identified to require separate consideration of exit and sustainability. The project has delivered a functioning and stable infrastructure, but changes to the context within which the iKit exists might require adaptation or extension of this infrastructure.

¹⁴ W3C markup validation service. <http://validator.w3.org/>

¹⁵ W3C CSS validation service. <http://jigsaw.w3.org/css-validator/>

Implementation

Overview

In light of the need to consider sustainability, longevity, accessibility, interoperability and reusability, an approach based on faceted information retrieval via an extensible taxonomic framework was adopted. This not only enables users to discover and access resources quickly and easily, but also enables JISC to maintain and sustain the iKit in the longer term, and indeed to extend and add value to it in the future.

This approach to organising and displaying information within web pages is broadly similar to that used to develop the Scottish Cultural Portal within the SPEIR project¹⁶, and builds upon knowledge gained in that endeavour. It allows the user to discover and access resources in a number of ways – navigational browsing via vocabularies and taxonomies, contextual browsing, narrative browsing and free-text searching. The vocabularies of JISC programmes, projects, contributing institutions and the technical/functional taxonomic analysis of the JISC IE are broadly equivalent to, but more formally expressed than, the "key questions and themes and FAQs" mentioned in the initial terms of reference for the project.

The interim analysis of technological solutions provided an affirmation of the value of this methodology, and extended the project resources such that an enhanced more structured approach could be taken. Work on the project proceeded generally smoothly with only the unforeseen difficulty in obtaining reusable data-sets of JISC projects, programmes and institutions (which were known to exist), with its attendant need to divert project resources to duplicate these data-sets from openly available sources, requiring significant amendment of the project methodology.

Creation of taxonomic framework

Appendix A provides a full case-study of the methodology and implementation of the iKit taxonomic framework.

External sources of content for the framework include HERO¹⁷ (for institutional names), the JISC standards catalogue¹⁸ (for definitions of technical terms), and Wikipedia¹⁹ (for definitions of technical terms)

Appendices F and G give examples of early versions of preferred and non-preferred terms for the Information Environment thesaurus.

Investigation of SKOS

The NSDL Metadata Registry Sandbox²⁰ was recognised to be an excellent tool for the authoring and publication of a test version of the SKOS declaration. It allows concepts and their properties to be input via an intuitive interface and for that input then to be easily output as valid RDF/SKOS via a simple automated process. Using this tool, the project was able to produce a proof-of-concept version of the IE thesaurus in RDF/SKOS. It is important to note that this script is not usable, as concepts are identified by pseudo-URIs which are used only as placeholders. If it was later decided to assign stable URIs to these concepts, it would be necessary to amend the iKit infrastructure in order to substitute

¹⁶ SPEIR final report. <http://speir.cdfr.strath.ac.uk/documents.htm>

¹⁷ HERO: The official online gateway to Higher Education and Research in the UK. <http://www.hero.ac.uk/uk/home/index.cfm>

¹⁸ JISC standards catalogue. <http://standards.jisc.ac.uk/catalogue/Home.phtml>

¹⁹ Wikipedia. English language version. http://en.wikipedia.org/wiki/Main_Page

²⁰ NSDL Metadata Registry sandbox. <http://sandbox.metadataregistry.org/>

the internal look-up table of IE thesaurus terms for a link to the relevant URI for the externally registered SKOS concept.

Appendix C of this report is a full case-study of the investigation.

The IE thesaurus is seen as a major deliverable of this project, and its potential for reuse within the JISC community is foreseen to be a great strength. Interoperability in this area would be greatly aided by the declaration within RDF/SKOS.

Creation of database structure

A relational database was created in MS Access, with the item data table forming the central element in the database. As per Rowley and Farrow²¹, the traditional approach to the use of thesauri in retrieval systems was adopted. Thus the taxonomic framework was used both to index items within the database and to allow users to search and retrieve the items. To facilitate the assignment of many-to-many relationships between items and taxonomic facets, each facet was located in an individual data table and linked to the item using linking tables. Figure 2, below, shows the data tables, and the relationships between them, which form the stable database structure.

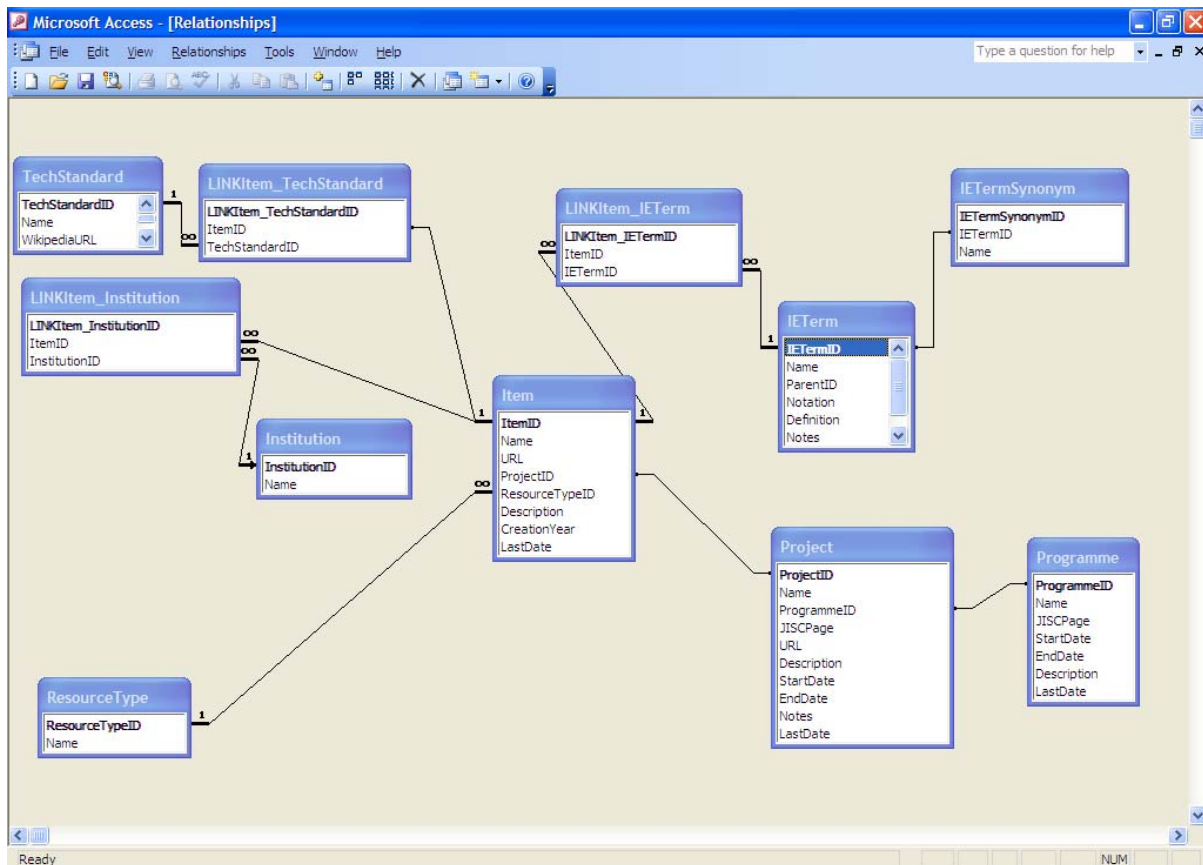


Figure 2: Relationship structure diagram of the database

Population of content

Define inclusion criteria

In scoping items for inclusion within the iKit, it was seen to be possible to apply either a hard or a soft definition of resource discovery, with the latter having the potential to produce a very large number of

²¹ Rowley, J. and Farrow, J. Organizing knowledge: an introduction to managing access to information. 3rd ed. London: Ashgate, 2000. p.161

resources. While some of those resources might prove invaluable, their sheer number might also serve to dilute and devalue the iKit as a whole. Following the initial scoping work, with its opportunity to perform an informal appraisal of the quantity and quality of possible items for inclusion, it was determined that while the types of item to be included should be as broad as possible, those resources must be substantive and authoritative to warrant inclusion. The following digital resources associated with JISC projects and services were identified to be within scope for inclusion within the iKit:

- Final report of the project
- Miscellaneous project reports
- Case studies
- Demonstrators
- Software tools or packages created by the project
- Manuals associated with demonstrators software tools/packages

Identify items

Early assumptions that existing data tables of JISC programmes and projects would be readily obtainable proved incorrect. It was therefore necessary to implement a methodology based on a systematic inventory of the JISC website. Initial scoping identified core JISC programmes which were highly likely to have produced relevant items. It was acknowledged that outputs from many JISC programmes whose core aims were not related to resource discovery might also fall within the scope of the iKit. Thus a thorough inventory of the research outputs of every JISC programme was required. Appendix D describes the methodology used.

The decision was taken to initially only include projects listed as starting earlier than May 2007, given that it was thought unlikely that any relevant inventory items would have been created after this point. A final pass of all project sites towards the end of the project would then be conducted to identify any items which had been created since the initial inventory.

The first-pass identification of items for possible inclusion was deliberately designed to err on the side of inclusiveness, and was assigned to a member of staff unfamiliar with JISC programmes, projects, or typical project outputs. This enabled an early overview of likely items for the inventory, and allowed refinement of the scope of materials to be included. It also identified a granularity issue, since many items were themselves composites of individual project outputs. Appendix H gives some examples. The issue was resolved by determining whether an item was likely to be useful on its own, without the context of a parent item, or whether components needed to be used as a package.

Items then deemed out-of-scope were removed from the first-pass inventory.

A second-pass identification was then carried out, using the refined scope, to ensure that all relevant items were included in the inventory.

Name, describe and categorise the items

Once an inventory of items had been accrued, it was necessary to apply conventions for their naming, description and categorisation:

- **Naming:** It was recognised that the diversity of items required some measure of standardisation to allow outputs to be easily identified. The inventory briefing paper given in Appendix D includes rules for the construction of iKit item names.
- **Description:** Where possible existing descriptive text, either from the resource itself or from other sources, was used and repurposed as necessary (i.e., transposing into the past tense; truncating verbose text, etc.). As part of the iterative methodology of the project, during the construction of the interface in WP4 it was decided that the item description would be displayed with the item name on the search results page, to support early identification of specific items. This in turn informed a decision to keep the content brief, so that each description did not exceed three lines in length.
- **Categorisation:** Finally, terms from the controlled vocabularies were to be applied. It was identified that it made good practice to assign first the JISC Project and Programme, then the

creating institution or organisation, then the technologies employed or described by the resource, then finally adding terms from the IE functional thesaurus. Thus, as each layer of terms is added so the person cataloguing the item gains familiarity with the item, thus being better able to make the more complex interpretative decisions required for the assignment of the IE functional thesaurus terms.

Development of the Web service interface

Complications over the possibility of incorporation of the iKit into the JISC website content management system led to the decision that the iKit would take the form of a dynamic, database-driven website, to be delivered via the CDLR's existing technical infrastructure. The expectation was that the iKit will eventually be handed over to the JISC, and dependencies on the current technical infrastructure were minimised during the development of the service.

The creation of a specific web interface assumed a larger proportion of project resources than had originally been anticipated.

CDLR has considerable expertise and experience in developing web services with database-driven content. Its developmental and operational infrastructure uses Adobe DreamWeaver to create XHTML web pages with embedded ColdFusion Markup Language (CFML)²² scripts. The scripts extract content from the database using Structured Query Language and generate dynamic XHTML to format it for display. DreamWeaver is also used to create Cascading Style Sheets to manage the look and feel of the web pages, and provides offline and online testing facilities. The CFML scripts are processed in real time by ColdFusion Server, and the resulting XHTML page is delivered to the client browser by MS Internet Information Server.

CDLR has its own domain name, which allowed the web pages to be published for testing purposes without any delay. CDLR also has protected areas within the domain name, requiring username and password for access, which allowed publication of early versions of the pages without them being available to the public or search engines.

A major design consideration was the use of the JISC Information Environment functional architecture to provide contextual resource discovery. The JISC Information Environment architecture diagram²³ was used to prototype a visual interface, by turning it into a "clickable map". For the operational service, the diagram was replaced by the concept map developed by the project. A textual interface was developed which allows the thesaurus to be browsed hierarchically. Both interfaces support retrieval of resources associated with the Information Environment at the level and all narrower levels of the term selected by the user; context is preserved by presenting results under the hierarchy of terms.

The other taxonomic facets used for resource discovery, including JISC projects, programmes, contributing institutions, and resource discovery technologies, have flat structures suitable for display as alphabetical browse lists. The number of values or instances of each facet was not large enough to warrant any structured browsing such as initial letter indexes.

The interface therefore supports users who do not know the facet values in advance, for example the names of JISC projects or Information Environment terms, by displaying a complete list in alphabetic or hierarchical order.

A keyword search facility was developed for users who do have some knowledge of the facet values. This allows a pre- and post-truncated search of a user-input term against the values in the chosen facet. Full truncation alleviates some of the problems caused by spelling variations between the user input and the controlled vocabularies.

²² CFML reference.

http://livedocs.adobe.com/coldfusion/8/htmldocs/help.html?content=Part_3_CFML_Ref_1.html

²³ JISC Information Environment architecture. <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/>

The interface was designed to display the results of a search as a list of resource titles with brief descriptions, to give an easily-browsable list. The full details of a resource are displayed using a labelled link. This approach is based on the design of most online library catalogues, and was chosen for its utility and familiarity. For the same reasons, a "more like this" or related items search facility was developed by exploiting the controlled vocabularies and relational database structure.

Ease of use was an important factor in designing the interface. DreamWeaver web page templates were used to provide access to all search types and internal navigation on every page without having to use a hierarchical menu. CSS is used to provide visual distinction between internal hyperlinks, which will not open a new browser window, and external hyperlinks to resources, which are displayed in a new browser window.

CSS is also used to maximise the accessibility of the interface. All font sizes and page element spacing is relative, to preserve page layout if a user increases or decreases the browse font size.

The interface is also designed to be used by screen-reading software. The XHTML follows W3C accessibility guidelines²⁴ (including substantial parts of version 2.0²⁵) by avoiding the use of tables for layout, following strict hierarchy in heading tags, and avoiding deprecated tags. The interface is designed to be displayed and remain easy to use if the CSS is turned off; this was tested using the Firefox²⁶ browser which has this capability. The following screen-shots (Figures 3, 4 and 5) show the JISC Information Environment thesaurus browse web page in its default display, and then with CSS and images turned-off so that it can be used by screen-reader software.

²⁴ W3C Web content accessibility guidelines 1.0. <http://www.w3.org/TR/WCAG10/>

²⁵ W3C Web content accessibility guidelines 2.0. <http://www.w3.org/TR/WCAG20/>

²⁶ Firefox web browser. www.mozilla.com/firefox/



Figure 3: iKit IE thesaurus browse page (top) with CSS and images (default)

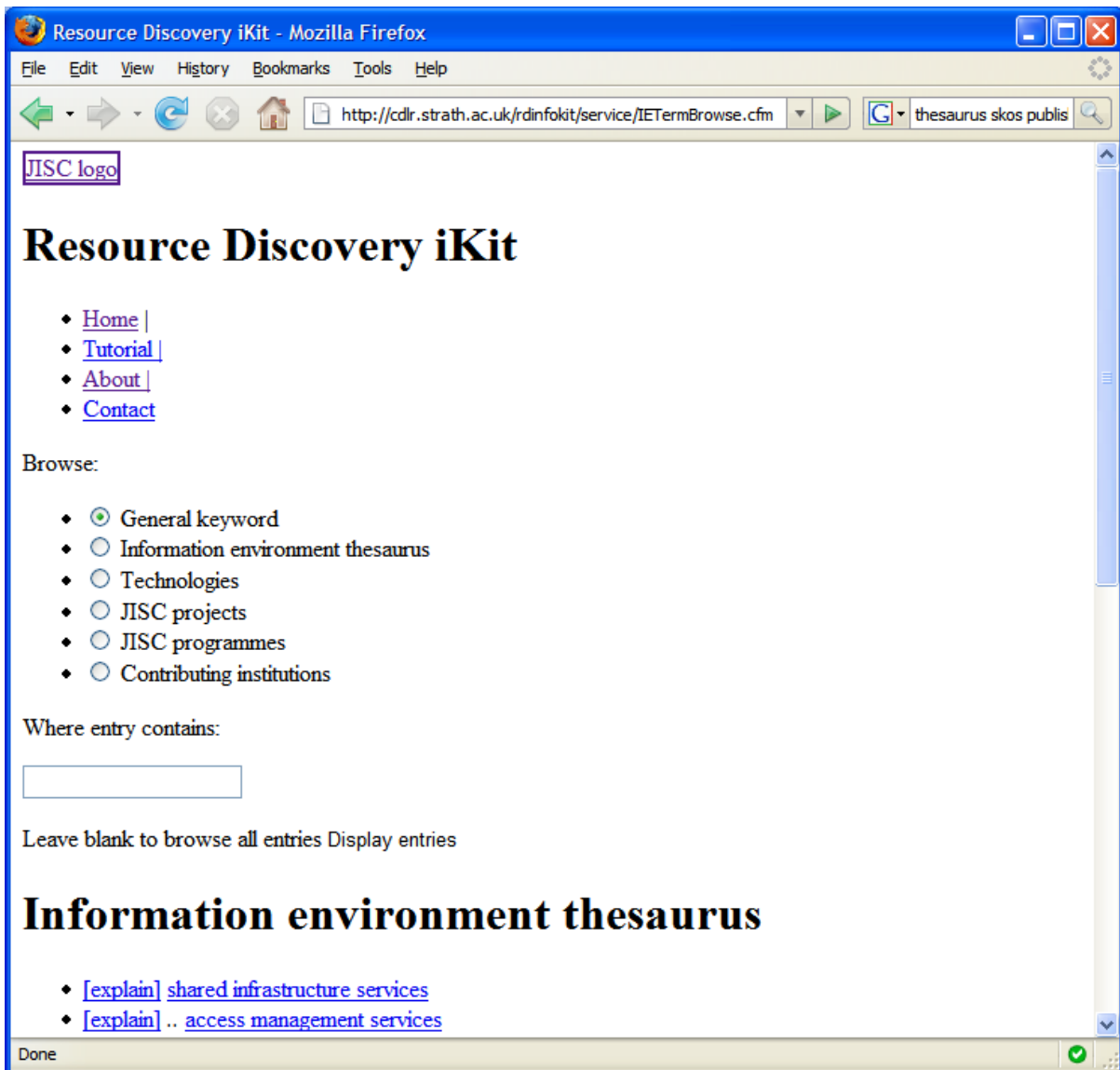


Figure 4: iKit IE thesaurus browse page (top) without CSS or images

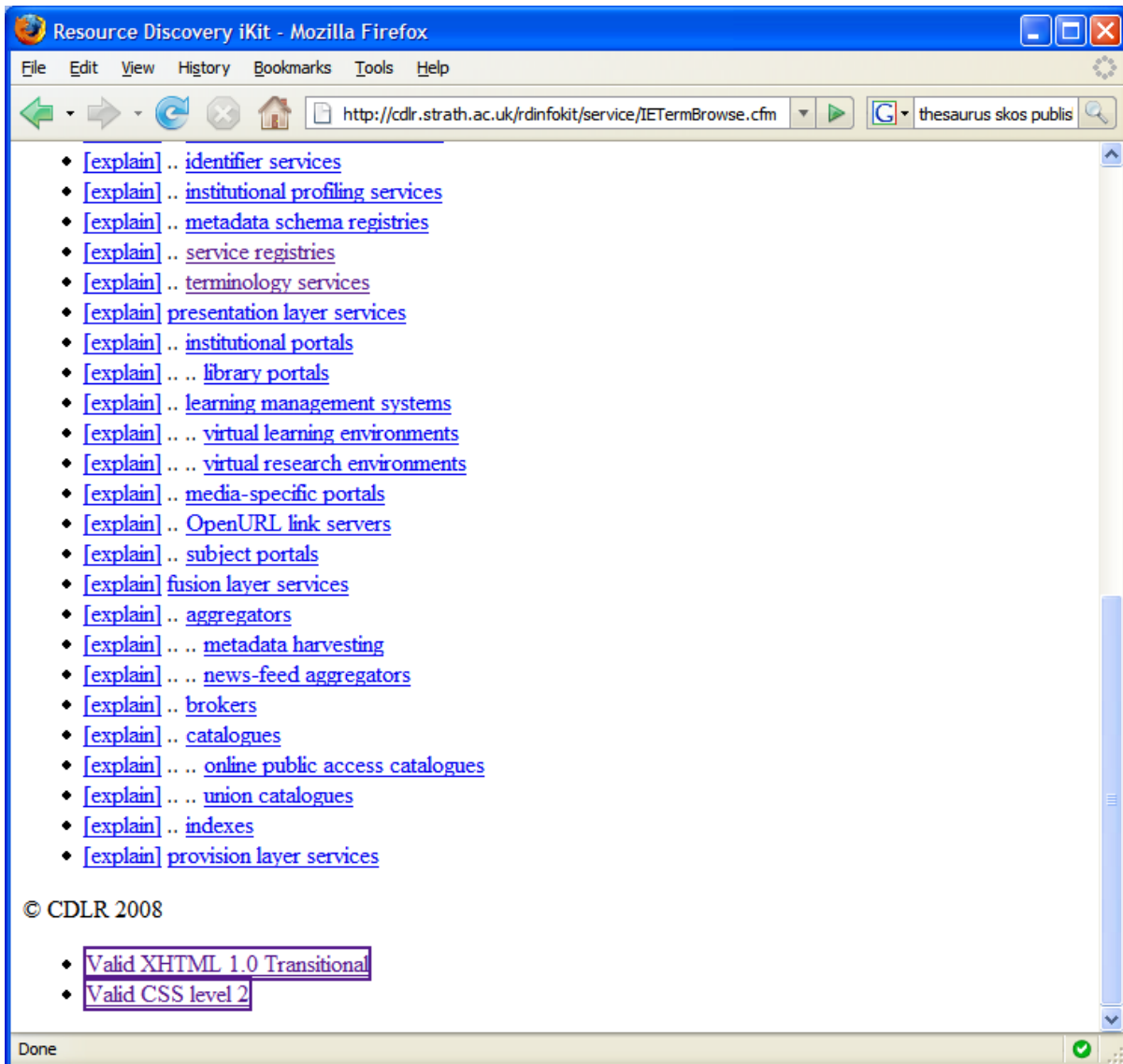


Figure 5: iKit IE thesaurus browse page (bottom) without CSS or images

The general approach adopted for developing the interface was rapid prototyping with iterative refinement. This allowed the early development of basic page structures without needing to wait for the full database structure to be finalised. As data tables, fields and relationships were developed, corresponding amendments to the web pages were made. In turn, the display of dynamic content in the interface enabled further refinement of the database structure.

Global variables containing literal values were used in the CFML scripts to allow rapid testing of display constants. For example, a global variable contains the punctuation indicating at which level of the Information Environment thesaurus a particular term lies. Feedback during beta testing of the interface showed that this should be changed to a more commonly-accepted value, and the amendment was applied once, to the variable, instead of several times in various scripts.

The application of CSS also allowed rapid changes to be made during interface development.

Towards the end of the project, the service was launched as a beta version, and feedback regarding the ergonomics of the interface was requested from key CDLR contacts and relevant JISCmail lists. Response was positive, with suggestions for minor amendments being incorporated where feasible/desirable. One example was the suggestion of a respondent that the creation date of a resource should be displayed on the intermediate results screen to assist in user assessment of relevance; this was implemented immediately.

The construction of the interface involved consultation with the JISC Communications and Outreach Team and conformance to JISC style guide documents. The website is designed to conform to current World Wide Web Consortium standards for both XHTML and CSS, and to adhere to accessibility legislation.

Exit and sustainability strategies

Consideration of the future of the iKit informed its creation at each step. The iKit infrastructure, that is a database structure and Web interface, was designed to be stable so that content could be easily updated, yet extensible to meet future requirements for additional functionality. The taxonomic framework allows for easy adaptation and for further layers of granularity to be added as needed.

Maintenance of the currency of iKit content was identified to require not only the cataloguing and inclusion of new resource discovery outputs from JISC projects and services, but also the updating of taxonomic facets as and when new terms come into use. Therefore, the IE functional thesaurus also requires a method for synchronising with future changes in structure, labelling of concepts, and focus within JISC.

Adapting to changes in context will require more fundamental work to change the database structure and/or Web interface. Possible changes to context might include the extension of the scope of the iKit to include resources other than outputs from JISC projects; work to integrate the iKit with other JISC services to enhance interoperability, etc. At this stage the following three iKit content maintenance options were identified:

1. *Assuming some commitment from JISC, the iKit could easily, and at low cost (in the short-term) be maintained using the same methods used to create it (i.e. centrally, by the developers, using direct methods). This option would have lower costs in the short term, but these costs would be ongoing such that in the longer term this might not constitute a best-value option.*
2. *Develop a standard maintenance interface which can be used by 3rd parties (e.g. JISC programme and project managers) and incorporate standard features such as the maintenance of referential integrity, enabling the production of appropriate management reports, etc.*
3. *Export the iKit content, data structure and Web interface functionality to a different software/hardware platform, in anticipation of its relation to another CMS (e.g. JISC's Programme Information Management Systems (PIMS)) via a M2M interface.*

CDLR will keep the service operational for at least one year.

The service can be ported to any Web server running MS Access and ColdFusion server.

The database can be ported to any relational database management system, using standard or system-specific methods.

The ColdFusion scripts for extracting dynamic web page content from the database use simple programming techniques and Structured Query Language; their functionality can be readily implemented using other programming languages.

Outputs and Results

The main output of the project is an operational service, Resource Discovery iKit²⁷, for the discovery of tools likely to be of interest to the developers of resource discovery services; in other words, a resource discovery service for resource discovery services. The iKit provides users with an intuitive Web interface which assists in the discovery of the outputs of JISC-funded research in the area of resource discovery by allowing them to dynamically search a database of around 230 individually-catalogued quality-assessed resources via a faceted taxonomic framework. These taxonomic facets include: JISC programme; JISC project; contributing institution; technologies used or described by the resource; and a thesaurus of Information Environment functional terms.

The service provides functionality found in resource discovery services for general information resources.

The service provides a general keyword search of resource titles and descriptions.

The home page provides a clickable conceptual map for the JISC Information Environment which provides high-level entry to the thesaurus.

The taxonomic facets can be browsed, or terms can be found by a full-truncation (anywhere-in-term) search. The thesaurus includes synonyms (non-preferred terms). An "explain" facility provides a definition or description for each term in all facets, with the exception of contributing institution. Where appropriate, this information is provided by external links to websites, the JISC standards catalogue²⁸, and Wikipedia²⁹.

When a facet term is selected, the service displays a list of resources, each with a brief description and direct link. A link to a full record for the resource is also provided.

The full record for the resource displays the direct resource link and brief description along with resource type, date of creation, and all associated facet terms. Each facet term can be used to display a list of related resources, a "more like this" or "related works" facility.

All searches are available on all pages, so there is no need to return to the home page.

The service also provides a tutorial in its use, and contact details to encourage feedback. A print version of the tutorial containing more detailed information is given as Appendix E of this report.

The service is designed for accessibility. It can be used by screen-reader software by turning off the style-sheet, and browser controls can increase or decrease font size without distorting the relative positions of display elements.

Another output of the project is the normalised relational database used for the catalogue of resources. This encompasses several datasets which can be regarded as outputs in their own right, including machine-readable, structured data for JISC projects and programmes and institutions investigating resource discovery solutions, technologies used in resource discovery services, and the JISC Information Environment thesaurus. These data can readily be made available for use by other applications. These data can also be used for management information purposes.

The JISC Information Environment thesaurus is available in RDF/SKOS format.

Another significant output is the thesaurus of terms used to describe the JISC Information Environment and the accompanying conceptual map. This updates and refines the functional architecture diagram produced in 2005.

²⁷ Resource Discovery iKit. <http://cdlr.strath.ac.uk/rdinfokit/service/index.cfm>

²⁸ JISC standards catalogue. <http://standards.jisc.ac.uk/catalogue/Home.phtml>

²⁹ Wikipedia. English language version. http://en.wikipedia.org/wiki/Main_Page

The inventory of tools for developing resource discovery services is also an output. The catalogue describes outputs from other JISC projects which have been assessed for their utility for developers. It does not include outputs which are irrelevant, incomplete or redundant.

The technical infrastructure of the iKit service can also be regarded as an output of the project. It includes the structure of the relational database, CSS (cascading style sheet) file controlling the look-and-feel of the interface, DreamWeaver templates controlling the layout of the interface, and Coldfusion scripts controlling dynamic interaction between the interface and database. The infrastructure meets W3C standards for XHTML 1.0 and CSS level 2.

The final output of the project is this final report.

Outcomes

The project has been successful in meeting most of its aims and objectives, including the creation of a Resource Discovery Information Kit (iKit) on the CDLR website which offers views of its content based on a taxonomic analysis of resource discovery components of the JISC Information Environment and provides contextual pointers to useful resources and projects in the area of resource discovery solutions. The service has been implemented in an integrated and sustainable way using a scalable relational database, and a plan for its future maintenance has been developed. The process of creating the service has been documented as an exemplar of best practice. The project has created and documented re-usable, controlled vocabularies based on the taxonomic analysis and has investigated the use of SKOS within the service.

It was not possible to implement the use of SKOS within the service because of requirements outside the scope and resources of the project. The investigation of SKOS has been documented as a case study.

The main outcome of the project is an operational service which is open, easy to use, continuously available, and highly accessible. The target user community is developers of resource discovery services, and the primary impact will be on them. The secondary impact will be on the teaching, learning and research communities which use such services, and on service funders and managers. Both groups should benefit from improved effectiveness and efficiency resulting from use of the iKit.

A related outcome will be better take-up of the outputs of JISC projects, increasing the value for money of JISC's investment.

JISC will also benefit from the review of JISC Information Environment functional architecture and the creation of a concept map and hierarchical thesaurus which includes synonymous terms. These will support further development of the Information Environment, with secondary beneficial impact on teaching, learning and research communities.

The other controlled vocabularies for JISC programmes, projects and contributing institutions will assist JISC in any future development of project management and dissemination services. The SKOS case study will encourage JISC and research communities to further develop and adopt Semantic Web approaches to resource description and access.

The methodology used to manage the development of the iKit service, using rapid prototyping and iterative refinement, was successful and is applicable to the development of similar services using small-scale databases driving dynamic content in Web interfaces.

The methodology used to develop the controlled vocabularies was also successful, and reaffirms the utility of using a standard approach.

The methodology used to investigate SKOS employed openly-accessible, easy to use tools, and is applicable to any project involving controlled vocabularies in the Semantic Web.

Conclusions

The project has demonstrated the feasibility and utility of the use of relational database techniques to:

- Store an inventory of web resources.
- Store controlled vocabularies to access items in the inventory.
- Store the thesaurus structure of the JISC Information Environment.
- Provide data for dynamic web pages for searching and retrieving items in the inventory.

This approach supports:

- Scalability.
- Extensibility.

The application of thesaurus structures to controlled vocabularies for access supports:

- Contextual search and retrieval.
- Hierarchical browsing.

Implications

The approach taken by the iKit, using a database to drive dynamic web pages, can be adopted by infokits and similar services in other areas of JISC's activities, and by other organisations engendering web-based outputs that can be of use to a wider community.

The architecture of the iKit is scalable, and could be expanded readily to accommodate the outputs of all JISC projects, and not just those concerning resource discovery solutions. This would involve the development and implementation of controlled vocabularies or thesauri to describe activities beyond the current Information Environment.

The structured thesaurus of Information Environment terms and the other controlled vocabularies developed for the iKit can be re-used by JISC and other organisations. There is a particular implication for the development of JISC's Programme Information Management System (PIMS) which is likely to contain terms for JISC programmes, projects, contributing institutions, and standards, all of which overlap and duplicate those used in the iKit. Interaction between the iKit and PIMS is highly desirable, to avoid duplication of effort and to ensure a consistent and coherent approach to preferred and non-preferred terms.

All parameters for search and retrieval of iKit tools are specified as URL parameters. Currently, this involves using local database primary keys to identify vocabulary terms and inventory tools, as shown in the example below (Figure 6).

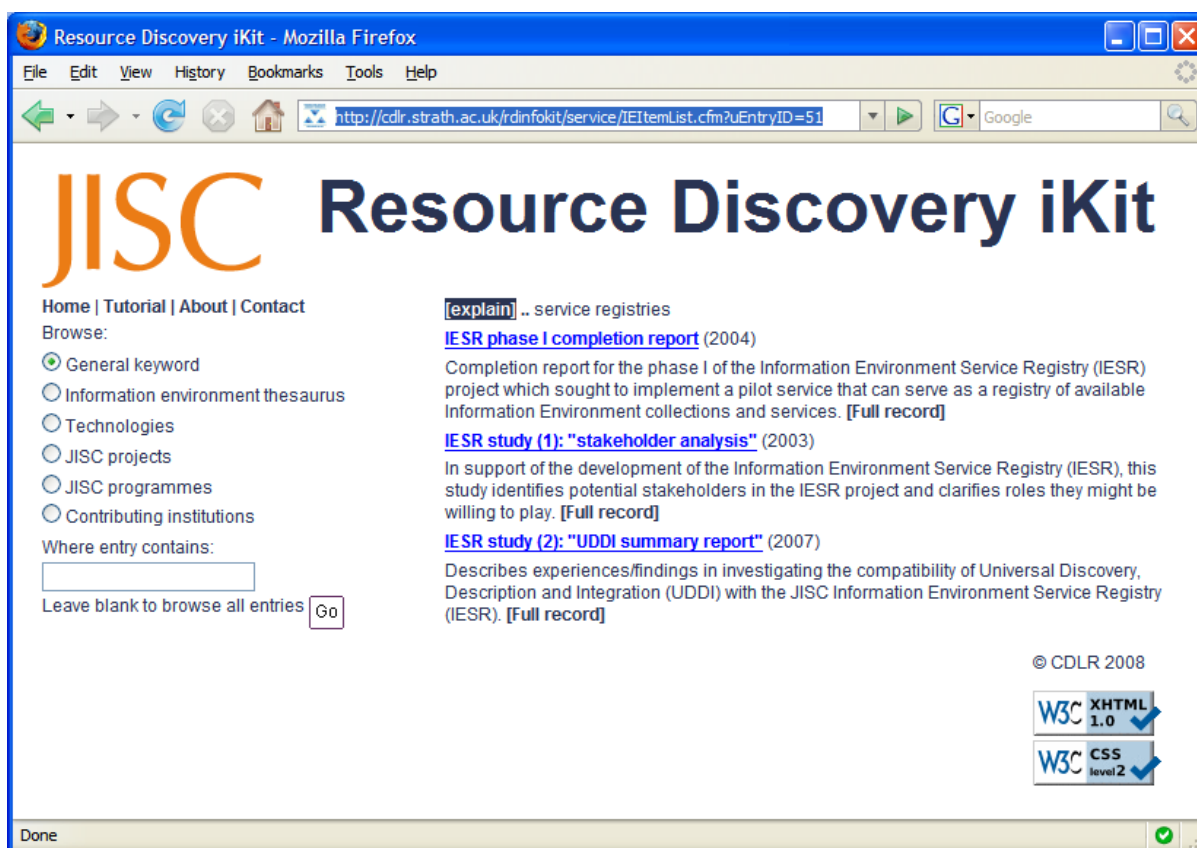


Figure 6: Display of tools for developing service registries, using a parameterised URL (<http://cdlr.strath.ac.uk/rdinfokit/service/IEItemList.cfm?uEntryID=51>)

In this example, the Information Environment term "service registries" is specified by the URL name/value parameter "uEntryID=51". This URL can be used by any web page external to the iKit, including JISC's website, to display the list of tools for developing service registries, for example embedded in a text narrative or list. Creators of such web pages only need to know the ID number for the Information Environment term they wish to use, which can be published easily as part of the iKit

website (with further development). Another approach to fostering re-use of iKit content is to develop the programme scripts to accept the actual term as a URL parameter; for example, "uETerm=service%20registries" ("%20" is the URL encoding for a space). Ultimately, the best method of developing this approach is to register the terms in RDF/SKOS and use the resulting URIs to specify external access to iKit data.

All aspects of the iKit taxonomic framework are amenable to these potential developments. For example, the current URL name/value parameter for "Dublin Core" is "uEntryID=7", and that for the "technologies" vocabulary from which the term is taken is "uEntryType=1"; the iKit display of tools using Dublin Core therefore has the URL
<http://cdlr.strath.ac.uk/rdinfokit/service/ItemList.cfm?uEntryType=1&uEntryID=7>

A related approach to encouraging re-use of iKit content is to develop the service to export data in a suitable XML format which can then be reformatted by the external service.

Registration of iKit terms in RDF/SKOS requires JISC to take ownership of the vocabularies, guarantee their maintenance, identify a domain to be the basis of term URIs, and publish the RDF/SKOS XML file. Further information about these processes is available from W3C³⁰.

A maintenance regime is required to keep the iKit service current. New tools for the inventory will continue to be created by existing and future JISC programmes and projects.

³⁰ Quick guide to publishing a thesaurus on the Semantic Web. <http://www.w3.org/TR/swbp-thesaurus-pubguide/>

Recommendations

The following recommendations are aimed at JISC, rather than any teaching, learning or research community.

- In the short term, JISC should resource the ongoing maintenance of the service, using the methodology developed for populating the inventory. CDLR would need to be kept informed of the completion of projects in the area of resource discovery so that it can identify additions to the inventory.
- In the longer term, JISC should consider other approaches to maintenance, including those identified in this report. Alternatives include:
 1. Development of a maintenance interface which can be used by third parties such as JISC managers and contributing institutions.
 2. Migration and integration of the iKit with another suitable content management infrastructure.
 3. Development of metadata interoperability between the iKit and related services such as PIMS.
- JISC should consider resourcing further development of the iKit to improve functionality, for example to meet the requirements of PIMS and other JISC services. See Appendix J for comments received during beta-testing which may inform further development.
- JISC should consider the registration of iKit and other controlled vocabularies in RDF/SKOS and other Semantic Web technologies, to encourage re-use, reduce duplication, and improve metadata consistency and coherency.

References

- CFML reference. Available at: http://livedocs.adobe.com/coldfusion/8/htmldocs/help.html?content=Part_3_CFML_Ref_1.html
- Codd, E.F. Further normalization of the data base relational model. IBM research report RJ909. 1971.
- Guide to establishment and development of monolingual thesauri. (BS 5723). British Standards Institution, 1987.
- JISC Information Environment. http://www.jisc.ac.uk/whatwedo/themes/information_environment.aspx/
- JISC Information Environment architecture. Available at: <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/>
- JISC standards catalogue. Available at: <http://standards.jisc.ac.uk/catalogue/Home.phtml>
- HERO: The official online gateway to Higher Education and Research in the UK. Available at: <http://www.hero.ac.uk/uk/home/index.cfm>
- Quick guide to publishing a thesaurus on the Semantic Web. Available at: <http://www.w3.org/TR/swbp-thesaurus-pubguide/>
- Resource Discovery iKit. Available at: <http://cdlr.strath.ac.uk/rdinfokit/service/index.cfm>
- Resource discovery infokit project. Interim report. Available at: <http://cdlr.strath.ac.uk/rdinfokit/docs/RDInfokitInterimReport.pdf>
- Resource discovery infokit project. Project extension plan. Available at: <http://cdlr.strath.ac.uk/rdinfokit/docs/RDInfokitExtensionPublic.pdf>
- Resource discovery infokit project. Project plan. Available at: <http://cdlr.strath.ac.uk/rdinfokit/docs/RDInfokitProjectPlanPublic.pdf>
- Rowley, J. and Farrow, J. Organizing knowledge: an introduction to managing access to information. 3rd ed. London: Ashgate, 2000.
- Shearer, J.R. A practical exercise in building a thesaurus. In: Roe, K.S. and Thomas, A.R. (eds.). The thesaurus: review, renaissance, and revision. New York: Haworth Press, 2004.
- Simple Knowledge Organization System (SKOS) - home page. Available at: <http://www.w3.org/2004/02/skos/>
- Structured vocabularies for information retrieval - Guide - Part 1: Definitions, symbols and abbreviations. (BS 8723-1). British Standards Institution, 2005.
- W3C CSS validation service. Available at: <http://jigsaw.w3.org/css-validator/>
- W3C markup validation service. Available at: <http://validator.w3.org/>
- W3C Web content accessibility guidelines 1.0. Available at: <http://www.w3.org/TR/WCAG10/>
- W3C Web content accessibility guidelines 2.0. Available at: <http://www.w3.org/TR/WCAG20/>
- Wikipedia. English language version. Available at: http://en.wikipedia.org/wiki/Main_Page

Appendices

The following appendices give more detailed information about the project's methodology, implementation, and outputs.

Appendix	Title	Page
A	Case-study of the methodology and implementation of the iKit taxonomic framework	29
B	Sources of terms for the JISC Information Environment thesaurus	35
C	SKOS case study	36
D	Inventory briefing paper	39
E	Resource Discovery iKit tutorial	42
F	Example of IE vocabulary preferred terms (from the interim report)	47
G	Example of IE vocabulary non-preferred terms (from the interim report)	52
H	Example of Infokit items (from the interim report)	57
I	XML manifestation of RDF/SKOS expression of IE vocabulary	61
J	Comments received during beta-testing	68

Appendix A: Case-study of the methodology and implementation of the iKit taxonomic framework

Introduction: the value of controlled vocabularies

In the current technological environment, with full-text searching of documents prevalent and inexpensive, the question might be asked whether there is still a need for the creation and use of controlled vocabularies in information retrieval systems. This attitude assumes, however, that a document is only about what it mentions; that it is no more than the sum of the words it uses. Fairthorne (in: Rowley and Farrow, 2000, p125) gives the lie to this argument:

Moby Dick is about a whale, Othello is about a handkerchief, and about other things. The difficulties are to identify which of the things mentioned refer to relevant topics, and how to deal with topics of the document that are not mentioned explicitly ... Parts of the document are not always what the entire document is about, nor is a document usually about the sum of the things it mentions.

The meaning of a document is necessarily a subjective matter; however the assignment of controlled vocabulary terms via intellectual effort can assist in the standardisation of description of the content of a document. Asghar Shiri and Revie (2000, p274) advise that in the online environment, users need quality conceptual and semantic tools to assist in finding resources on the Web. They continue: "The semantic structures provided by thesauri can play a part in both organising and retrieving Web information and knowledge resources."

Rowley and Farrow (2000, p128) list the advantages to be gained from the use of controlled vocabularies in document description, viz. that they facilitate user searching via: Control of synonyms; qualification of homographs; provision of scope notes; make explicit the relationships between terms, and; can express concepts which might remain elusive in free-text searching. In addition they enable us to map an area of knowledge. In the case of the JISC Information Environment, this latter point is particularly beneficial; the iKit taxonomy will contribute to the IE research effort by clarifying concepts and making explicit functional relationships.

Literature review and methodology

The following seven steps, which draw from Shearer (2004), identify the major methodological stages in the creation of a thesaurus.

1. Define taxonomic facets
2. Collect raw terms from literature and other sources
3. Identify relationships between terms
4. Add further terms to the thesaurus to fill significant gaps
5. Apply standard rules for the use of terms
6. Define term and add notation
7. Create alphabetical index and develop presentation

We here first describe the major elements of each of these steps and then later go on to link each of these steps to the real-world implementation of the iKit taxonomic framework.

1. Define taxonomic facets

The construction of facets within a thesaurus can proceed in either a top-down or bottom-up manner. In the top-down approach, we would first identify top-level headings or sections within the thesaurus and then assign terms as they are gathered; in the bottom-up approach, we gather together all our terms before deciding an appropriate schema of subject headings and hierarchies. McCulloch (2005, p404) advises that the best approach will very much be decided by the particular instance. Thus, this step might, in many cases actually come after the initial collection of terms.

Rowley and Farrow (2000, p153) describe facet analysis as the means of creating a systemic display of the terms in a controlled vocabulary according to fundamental, and mutually exclusive, categories.

This is to be used as a more user-friendly approach to thesauri display than merely listing all terms in a single alphabetical sequence. Examples of such categories are: entities, actions, place, time.

2. Collect raw terms from literature and other sources

Terms should be gathered from existing, authoritative resources within the field. McCulloch (2005, p405) advises that, in the lack of any already existent thesauri or taxonomies, terms can be gathered via a variety of literature, citing academic papers, conference proceedings, official publications, organisational websites and mailing list archives as examples.

3. Identify relationships between terms

This stage involves the assignment of thesaural relationships. Chowdhury and Chowdhury (2007, p118) identify three main classes of fundamental thesaural relationship: equivalence, hierarchy and association.

- a) **Equivalence:** refers to the coordinate, or sibling, status of a term in relation to another term in the thesaurus, thus including synonyms and antonyms. Equivalence can occur: where terms have differing linguistic origins; have popular and scientific names; have variant spellings; vary between differing cultures; or where they have abbreviations in common usage.

Shearer advises that one term should be designated as the "preferred form" and the other as an alternative. He says: "Which term to prefer will depend upon the nature of the user group and its needs. Thus arachnids and spiders are synonyms, but arachnids would be preferred in a list for technical users, spiders in one for use by the general public." (Shearer, 2004, p37). Given that the target audience for the iKit is "computing and library staff, learning technologists and others researching resource discovery solutions", an assumption of technical awareness can clearly be made.

- b) **Hierarchy:** refers to the super- or sub-ordinate status of a term in relation to other thesauri terms. This distinction is also known as the parent/child or broader/narrower relationship. The three types of hierarchical relationships are:

- *Generic:* the relationship between a thing and its kinds or types. Also known as genus/species relationship, this relationship exists where *some* As are B, but *all* Bs are A. E.g. some brass instruments (genus) are trombones (species); while all trombones are brass instruments.
- *Partitive:* the relationship between a thing and its components or constituent parts. Also known as the whole/part relationship. The part must be completely subordinate to the whole and hence this relationship applies only if the part is unique to the whole.
- *Poly-hierarchical:* where a term can be located in more than one hierarchy, most often in the case of compound terms. (Broughton, 2006, p122-128)

- c) **Association:** refers to terms that are linked in a manner which is not strictly hierarchical. This will usually reflect a causal connection and be assigned in an ad hoc manner. Broughton (2006, p129-130) gives as examples: an action and its necessary equipment (e.g. digging and spade), an activity and the active agent (e.g. running and athlete), or a process and an outcome (e.g. baking and cake).

4. Add further terms to the thesaurus to fill significant gaps

At this stage, as the structure becomes more complete, Shearer advises it will become more apparent that there are obvious omissions or that further clarification is needed for the way the terms have been divided (Shearer, 2004, p41).

5. Apply standard rules for the use of terms

Grammatical conventions should be adopted to standardise the use and display of terms. Here the formal standards for vocabulary control are the identical International Standard ISO 2788:1986 and British Standard BS 5723:1987³¹.

In order to control the form of terms, e.g. grammatical form and spelling, rules should be established and applied. As an example, Rowley and Farrow (2000, p144) advise that a decision should be made over whether to admit proper name. In this instance, then, a naming convention must also be applied to standardise the use of names. This will be particularly pertinent for the naming of technologies and institutions. Further, Lykke Nielsen (2004, p68) advises that choices should be made over the use of spelling, singular or plural form and abbreviations.

Rowley and Farrow further advise that terms should be either "concrete entities" (i.e. materials or things and their physical parts) or "abstract entities" (i.e. actions or events; abstract entities and properties; disciplines or sciences; units of measurement). They advise that these rules are strict and that if a term does not conform then it should be modified (Rowley and Farrow, 2000, p144).

Of particular issue is the use of compound terms. Aitchison, Gilchrist and Bawden (2000, p38) advise that ISO 2788 states that "it is a general rule that terms in a thesaurus should represent simple or unitary concepts as far as possible, and compound terms should be factored (i.e. split) into simple elements, except when this is likely to affect users' understanding." They acknowledge, though that this rule is difficult to follow in practice, especially in "specialist fields" where unitary concepts might necessarily be best expressed as compound terms.

We must note here that Lykke Nielsen (2004, p69) points out that "these general principles and guidelines must be applied according to the situation, the purpose of the thesaurus, and the characteristics of the information environment." Thus it is important to recognise that while striving to conform to these standards, general guidelines must be subordinate to local requirements in the construction of the vocabulary.

6. Define term and add notation

Where terms are unclear, a definition should be added. In the case of a specialist thesaurus, such definitions shall be particularly necessary. Shearer further advises (2004, p42) that an appropriate notation should then be added to the terms in order to mechanise their order. This can take the form of numbers, letters, or a combination of both.

7. Create alphabetical index and develop presentation

This application of notation will then facilitate the clear presentation of the resultant thesaurus either alphabetically or according to the hierarchical relationships of terms.

Implementation and results

1. Define taxonomic facets

The construction of facets within a thesaurus can proceed in either a top-down or bottom-up manner. In the top-down approach, we would first identify top-level headings or sections within the thesaurus and then assign terms as they are gathered; in the bottom-up approach, we gather together all our terms before deciding an appropriate schema of subject headings and hierarchies. The main facets of the iKit taxonomy had, in fact, already been defined at the project planning stage, and initial scoping work conducted to confirm their suitability. Therefore, it made methodological sense to assign terms to facets as they were collected. The facets were defined as follows:

- **JISC programme and project:** would allow the searching and browsing of the resultant research discovery outputs of a given JISC programme or project.

³¹ Guide to establishment and development of monolingual thesauri. (BS 5723). British Standards Institution, 1987.

- **Institutions and organisations:** would allow the searching and browsing of an institution's output as it pertains to JISC-funded resource discovery research.
- **Information Environment technical taxonomy:** would allow the specification of the technologies, standards, protocols and specific applications in use within the IE. iKit items will then be able to be described according to the specific technologies they employ or investigate.
- **Information Environment functional taxonomy:** would allow the specification of the functional aspects of the services which constitute the architecture of the IE.

It was noted that some relationships and dependencies might best be handled by the creation of additional or supplementary vocabularies, for example institution type, user type, resource types, etc. The above, however, form the core constituent facets of the iKit taxonomic framework.

2. Collect raw terms from literature and other sources

It was initially assumed that data-tables of JISC programmes, projects and contributing institutions would be obtainable from JISC, and easily re-purposed as necessary to then fit into the database structure of the iKit. Following the initial scoping work, it became apparent this would not be the case and so this data was collected manually from openly available sources, including the JISC website, project websites, etc.

Thus for each facet of the taxonomy, except that of the Information Environment functional taxonomy, the collection of terms would be conducted parallel to the WP3 goal of populating the iKit content (i.e. finding items for inclusion), since the inclusion of terms or names in the vocabularies for Programmes, Projects, Institutions and technologies would be driven by the items themselves. The purpose of the iKit was not to create data tables of all existent JISC programmes or projects, or of all institutions conducting JISC-funded research, or of all technologies being employed in that research. Rather, the aim was to create data-tables of those entities involved in resource discovery research, to enable the use of these terms as descriptors to allow the retrieval of items within the iKit.

The collection of terms for the IE functional taxonomy, as stated, was the only facet where the collection of terms was conducted prior to inventory of iKit resources. Thus, this facet differs from the others in that it is an attempt to provide a complete view of the IE in controlled language. Here, collection of terms was, in the main, conducted via close reading of the IE functional and technical architectural documents produced by UKOLN, and specifically, Andy Powell. Appendix B of the project final report is a full bibliography of the main documents used to gather terms for this facet.

3. Identify relationships between terms

McCulloch (2005, p405) warns that when imposing a hierarchical structure on terms by the assignment of relationships that the level of specificity must be appropriate to the size of the thesaurus in question. She says, "too few levels of granularity is insufficient; too many is unnecessary." It was thus anticipated that the degree of detail required to be applied within each facet would be determined, to a greater or lesser extent, by the number of resources to be described.

In practice, this means that where two terms within the IE functional taxonomy might actually have a hierarchical relationship to each other, if there are not many resources about those terms, then to make the hierarchical distinction is perhaps going to confuse more than assist the user, who is searching for documents in that area. Thus it may be the case that although two terms are, strictly speaking, related hierarchically, they are actually related within the taxonomy as equivalent terms. The approach taken is, however, fully extensible such that further degrees of detail or granularity can be applied as and when required.

The illustration below is a digital photograph of an embryonic conceptual representation of the terms for inclusion in the IE functional taxonomy.

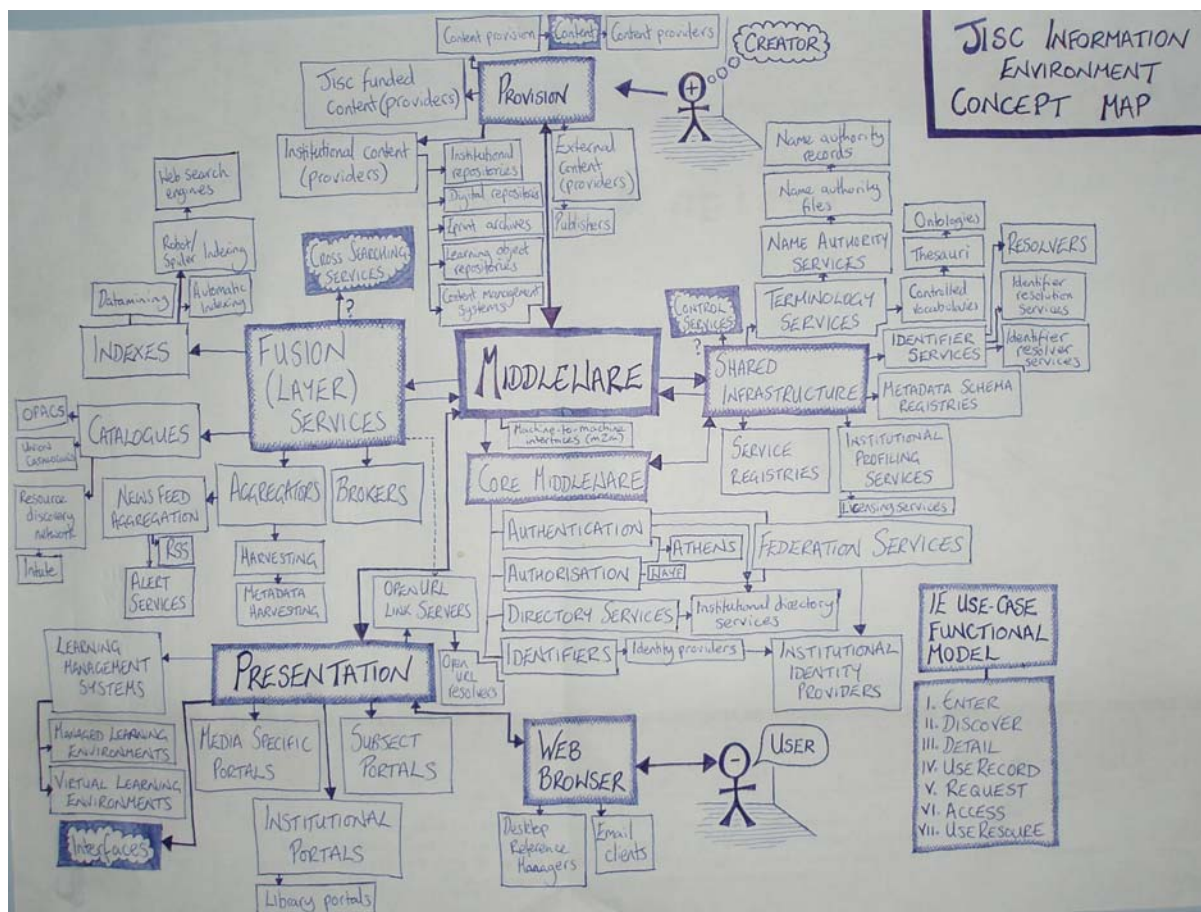


Figure 1: JISC Information Environment concept map

4. Add further terms to the thesaurus to fill significant gaps

As stated, Shearer advises that at this point major gaps in the thesaurus will become clear. In the case of the iKit it was perhaps the opposite which was the case; i.e. that due to its small-scale nature, the taxonomies (especially for the IE functional thesaurus) might be too involved, and with too much granularity in terms for the amount of items to be described (which by this point had been discerned to be between 200-300).

5. Apply standard rules for the use of terms

As per Lykke Nielsen (2004, p69), the imposition of rules must always be subordinate to the needs of the particular case. In the case of the iKit, it is the end-user functionality which a faceted taxonomy delivers which is central; all decisions shall ultimately be informed by this recognition. In any case, the IE taxonomy conforms to the British Standard for controlled vocabularies (BS 8723-1:2005)³²; terms are presented in their plural form and in lower case. Given the technical nature of the area being described, compound terms feature heavily in the IE taxonomy.

The naming of JISC projects and programmes were as they are used on the JISC website. The list of Higher and Further Education institutions found on the HERO website³³ was used as the authority list for the naming of institutions. Where a particular research group within a University was prominent enough in the area of JISC research to warrant separate inclusion, this layer of granularity was added, as for example UKOLN at the University of Bath.

³² Structured vocabularies for information retrieval - Guide - Part 1: Definitions, symbols and abbreviations. BS (8723-1). British Standards Institution, 2005.

³³ HERO: The official online gateway to Higher Education and Research in the UK. <http://www.hero.ac.uk/uk/home/index.cfm>

6. Define term and add notation

Definitions were applied to terms in the IE functional thesaurus, along with scope notes to assist in their use. For the technologies used, it was identified that JISC already maintained, via UKOLN, the JISC Standards Catalogue³⁴. Thus it was not felt necessary to define these terms, but rather a link to the relevant entry in the Standards Catalogue was given, along with a link to the homepage for the relevant standard. Where available, a link to the corresponding entry in Wikipedia³⁵ was also added to enhance the amount of information available.

In the iKit a simple numeral notation was applied, with a full stop indicating a step down in the hierarchy. Thus, term A with notation "1" is a top level term, which has the narrower term B with notation "1.1". Term C, with notation "1.1.1", is thus a narrower term of term B, and so on.

7. Create alphabetical index and develop presentation

In keeping with the iterative nature of the four work packages required to create the iKit methodology, this step was very much informed by the structures of both the database and Web interface. Display of the taxonomies is alphabetical, except in the case of the IE thesaurus, where the inherent hierarchical structure is thought to be more useful to the user as a means of browsing. The standard textual notation for thesauri (e.g. BT to mean "broader term") has been eschewed in favour of more intuitively grasped plain English descriptors.

Conclusion

The taxonomic framework whose creation is detailed here has been successfully used to describe iKit items and so assist users to retrieve information in a structured and intuitive way.

References

- Aitchison, J., Gilchrist, A. and Bawden, D. (2000). *Thesaurus construction and use: a practical manual* (4th ed.). London: Aslib IMI.
- Asghar Shiri, A. and Revie, C. (2000). Thesauri on the Web: current developments and trends. *Online Information Review*. 24(4), pp. 273-279. Available: <http://www.emeraldinsight.com>. Accessed 21 Apr 2008.
- Chowdhury, G.G and Chowdhury, S. (2007). *Organizing information: from the shelf to the web*. London: Facet.
- Ibekwe-SanJuan, F. (2006). Constructing and maintaining knowledge organization tools: a symbolic approach. *Journal of Documentation*. 62(2), pp.229-250. Available: www.emeraldinsight.com. Accessed 21 Apr 2008.
- Jones, K.P. (1981). Problems associated with the use of compound words in thesauri, with special reference to BS 5723:1979. *Journal of Documentation*. 37(2), pp.53-69. Available: www.emeraldinsight.com. Accessed 21 Apr 2008.
- Lykke Nielsen, M. (2004). Thesaurus construction: key issues and selected readings. In: Roe, K. S. and Thomas, A. R. (eds.). *The thesaurus: review, renaissance, and revision*. New York: Haworth Press. pp.57-74.
- McCulloch, E. (2005). Thesauri: practical guidance for construction. *Library Review*. 54(7), pp.403-409. Available: <http://www.emeraldinsight.com>. Accessed 21 Apr 2008.
- Rowley, J. and Farrow, J. (2000). *Organizing knowledge: an introduction to managing access to information* (3rd ed.). London: Ashgate.
- Shearer, J.R. (2004). A practical exercise in building a thesaurus. In: Roe, K. S. and Thomas, A. R. (eds.). *The thesaurus: review, renaissance, and revision*. New York: Haworth Press. pp.35-56.

³⁴ JISC standards catalogue. <http://standards.jisc.ac.uk/catalogue/Home.phtml>

³⁵ Wikipedia. English language version. http://en.wikipedia.org/wiki/Main_Page

Appendix B: Sources of terms for the JISC Information Environment thesaurus

Presented here is an alphabetical list of key resources from which Information Environment terms were gathered, in addition to the JISC website and individual project websites.

- Brophy, P. (2005). The formative evaluation of the 5/99 Programme and its broader environment: the EDNER Project. *VINE*. 35(1/2), pp. 105-112. Available: <http://www.emeraldinsight.com/>. Accessed 21 November 2007.
- Grout, C. (2001). *Information Environment Development Strategy 2001-2005*. Available: http://www.jisc.ac.uk/media/documents/themes/infoenvironment/ie_strategy.pdf Accessed 4 December 2007.
- Harris, N. (2002). *ANGEL: Authenticated Networked Guided Environment for Learning vocabulary*. Available: <http://www.angel.ac.uk/public-files/htmldoc004.htm>. Accessed 7 Apr 2008.
- Powell, A. (2003). *JISC Information Environment Architecture Portal FAQ*. Bath: UKOLN, University of Bath. Available: <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/faq/portal/>. Accessed: 7 Apr 2008.
- Powell, A. (2003). Mapping the JISC IE service landscape. *Ariadne*. Issue 36. Available: <http://www.ariadne.ac.uk/issue36/powell/>. Accessed: 7 Apr 2008.
- Powell, A. (2004). *The JISC Information Environment and Google: a discussion paper*. UKOLN, University of Bath. Available: <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/ie-google/ie-google.pdf>. Accessed 7 Apr 2008.
- Powell, A. (2005). *A service oriented view of the JISC Information Environment*. Bath: UKOLN, University of Bath. Available: <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/soa/jisc-ie-soa.pdf>. Accessed 7 Apr 2008.
- Powell, A. (2005). *JISC Information Environment Architecture Glossary*. Bath: UKOLN, University of Bath. Available: <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/glossary/>. Accessed 7 Apr 2008.
- Powell, A. (2005). *The JISC Resource Discovery Landscape - a personal reflection on the JISC Information Environment and related activities*. UKOLN, University of Bath. Available: <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/resource-discovery-review/>. Accessed 7 Apr 2008.
- Powell, A. (2006). *JISC Information Environment Technical Standards (Version 1.1)*. Bath: UKOLN, University of Bath. Available: <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/standards/>. Accessed 7 Apr 2008.
- Powell, A. and Lyon, L. (2001). *The DNER Technical Architecture: scoping the information environment*. Bath: UKOLN, University of Bath. Available: <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/dner-arch.html>. Accessed 7 Apr 2008.
- Powell, A. and Lyon, L. (2002). The JISC Information Environment and Web services. *Ariadne*. Issue 31. Available: <http://www.ariadne.ac.uk/issue31/information-environments/>. Accessed: 7 Apr 2008.
- Powell, A. and Lyon, L. (2005). *JISC Information Environment Architecture Functional Model*. Bath: UKOLN, University of Bath. Available: <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/functional-model/>. Accessed 7 Apr 2008

Appendix C: SKOS case study

Investigation of the declaration of small-scale thesaurus in RDF/SKOS

Background

Research into the semantic Web seeks to strengthen the underlying semantic structure of the Web, in order to facilitate greater interoperability and machine processability of information. New standards and frameworks are now gradually taking root which aim to allow automated web services to process discrete pieces of semantic information, to retrieve them from heterogeneous sources and dynamically reorganise them according to user need. This is the semantic Web: "*a common framework that allows data to be shared and reused across application, enterprise, and community boundaries*" (Miles, 2005). The semantic Web is an envisioned extension of Web functionality "*in which information is given well-defined meaning, better enabling computers and people to work in co-operation.*" (Berners-Lee et al., 2001)

Within the framework of the semantic Web, individual vocabularies (ontologies) must be specified in order to control the meaning of terms. One such vocabulary is Simple Knowledge Organisation System (SKOS), which aims to express "the basic structure and content of concept schemes such as thesauri, classification schemes, subject heading lists, taxonomies, 'folksonomies', and other types of controlled vocabularies" (Miles and Brickley, 2005). As Cantara (2006, p113) says: "The implications of SKOS Core and other Web ontology languages for taking online searching to new levels of semantic interoperability are immeasurable". The full potential applicability of these technologies shall be discovered only via research and development activity such as that conducted by the the Joint Information Systems Committee (JISC).

JISC and the semantic Web

JISC aims to help and facilitate UK Higher and Further Education institutions in their use of the Internet and other Information and Communications Technology applications. One of JISC's key objectives is to build an online Information Environment (IE) that provides secure and convenient access to comprehensive collection of scholarly and educational material. Vocabulary control services form a key part of the shared infrastructure services being developed to support the development of the IE.

The value of controlled vocabularies (library classifications, thesauri, etc.) is manifest: they facilitate consistent description of documents and control of what is an essentially unstable unit of exchange, viz. language. They are not, of course, a panacea to the vagaries of language. The subjective nature of language dictates that no list of subject terms can ever be exhaustive, and similarly they can never exactly mirror the vocabulary of users. By definition their usage precludes users from searching for documents using their own natural language. However, by making explicit the controlled vocabularies in use in different retrieval systems, via the infrastructure of the semantic Web, it is possible then to allow machine-readable mappings between these vocabularies to increase their applicability. This is the aim of the JISC-funded High-Level Thesaurus (HILT) project at the CDLR, for example.

Using RDF/SKOS it is possible to mark-up not just monolithic classification schemes and thesauri like LCSH, DDC, AAT, etc. but also smaller taxonomies which are used to organise information entities at a more local level. Usually organisation or domain specific, produced by the needs and interests of a specific community and lacking the broader scope of a well-known classification scheme, it is possible to register such taxonomies and folksonomies so that a document automatically indicates its 'web-cultural' references in a machine-processable and interoperable manner (Nasir Uddin and Janecek, 2007, p219). The need for such interoperability was pointed out by Asghar Shiri and Revie (2000, p278), who argue that a lack of standards for the publication of thesauri on the Web produces a lack of "interoperability, reusability and shareability". This is the rationale behind the investigation of the feasibility of the declaration in RDF/SKOS of the small-scale *JISC Information Environment* thesaurus, created for the JISC Resource Discovery iKit by the CDLR.

The JISC Information Environment thesaurus

This thesaurus of the functional service components of the JISC Information Environment is seen as a major deliverable of the project to create the iKit, and its potential for reuse within a range of contexts is seen as a key strength. Machine interoperability is recognised to be a major determining factor in the reuse of thesauri, and thus as a secondary task, the iKit project undertook to investigate the declaration of the IE thesaurus in RDF/SKOS.

The NSDL Metadata Registry (henceforth Registry) was recognised to be a useful tool for this investigation. The Registry is designed to assist in the maintenance and upkeep of vocabularies. The NSDL Registry performs a similar function to that of JISC's own Information Environment Metadata Schema Registry (IEMSR). Such metadata registries will form core infrastructural components of any truly functioning semantic Web. Alongside tools to register, archive and find metadata schema, the registry also offers simple, intuitive self-authoring tools for collection holders or communities who wish to generate metadata schema. These authoring tools make use of the evolving SKOS vocabulary schema to state the properties of terms and the relationships between them. Further, in order to allow testing before full-scale declaration, the Registry has a parallel registry, the *Sandbox*. This was identified as an ideal environment for the investigation of the declaration of the IE thesaurus in SKOS.

Thus, via the use of the Registry's simple web-form input it was possible to input IE thesaurus terms, including their relationships, scope notes, definitions, etc. and then to output the resultant file as RDF/SKOS file. This output is included as part II of this document

It must be made clear, however, that this file has been created only as a demonstration; it does not constitute an interoperable RDF/SKOS file. This is due to the fact that each concept declared in SKOS requires the assignment of a stable URI and no platform was immediately identifiable as a suitable host domain to provide the permanence required for such URIs. Hence, the SKOS:Concept URIs to which the file refers are merely pseudo-URIs, input as placeholders to demonstrate the markup potential of SKOS.

The use of pseudo-URIs means the RDF/SKOS version of the IE thesaurus can merely act to demonstrate the potential of SKOS as a thesaurus markup language. If JISC later wishes to implement the real declaration of the IE thesaurus via the provision of a stable domain to support the URIs, it would be a simple process to substitute the pseudo-URIs in the RDF/SKOS file.

This is not a minor point, however; it would require a commitment from JISC towards the long-term maintenance and upkeep of the thesaurus. The iKit currently resides on the CDLR Web infrastructure, but this is temporary. It is hence not possible for the thesaurus to be hosted on the CDLR website, neither was it thought desirable to have an official JISC output, the IE thesaurus, hosted external to JISC (i.e. by NSDL). The IEMSR is identified as one possible way of registering the RDF/SKOS file; it may then be possible to use the IEMSR as the base domain for the registry of stable URIs to the IE thesaurus concepts.

If JISC later wishes to declare the full thesaurus, they might also consider the possibility of declaring the names of contributing institutions in FOAF, and the extension of the JISC Standards Catalogue to include declaration of this in RDF/SKOS, so as to promote the use of machine-readable, interoperable controlled vocabularies throughout the JISC landscape. Such declarations, while offering an excellent possibility for the promotion of interoperability, are clearly beyond the scope of the iKit project however.

If JISC were to adopt full declaration of the IE thesaurus, it would be a relatively simple matter to amend the iKit infrastructure in order to substitute the internal look-up table of IE thesaurus terms for a link to the relevant URI for the externally registered SKOS concept. This is a proof of concept for the declaration of the taxonomic framework.

Conclusion

Thus, the investigation of the feasibility of the implementation of SKOS for the IE thesaurus has proved successful. It has been demonstrated to be simple to author such tools in SKOS using freely available web-tools. Equally, the usability of such a thesaurus for practical applications such as the iKit is also proven. Only one barrier to implementation remains, the need to assign a permanent URI

to concepts declared in SKOS. Such permanence cannot be guaranteed by a small-scale project such as this; it is hence recommended that JISC designate a service to host such declarations (the IEMSR, once operational, being the obvious choice).

References

- Aud Franklin, R. (2003). Re-inventing subject access for the semantic web. *Online Information Review*. 27(2), 94-101.
- Asghar Shiri, A. and Revie, C. (2000). Thesauri on the Web: current developments and trends. *Online Information Review*. 24(4), pp. 273-279. Available: www.emeraldinsight.com. Accessed 21 Apr 2008.
- Broughton, V. and Slavic, A. (2007). Building a faceted classification for the humanities: principles and procedures. *Journal of Documentation*. 63(5), pp.727-754.
- Cantara, L. (2006). Encoding controlled vocabularies for the Semantic Web using SKOS Core. *International digital library perspectives*. 22(2), pp.111-114.
- Greaves, M. and Mika, P. (2008). Semantic Web and Web 2.0. *Web Semantics: Science, Services and Agents on the World Wide Web*.
- Gu, T., Pung, H.K. and Zhang, D. (2007). Information retrieval in schema-based P2P systems using one-dimensional semantic space. *Computer Networks*. 51, pp. 4543-4560.
- Hillmann, D., Phipps, J., Sutton, S.A. and Laundry, R. (????). *A metadata registry from the vocabularies up: the NSDL Registry project*.
- Miles, A. (ed.) (2005). *Quick Guide to Publishing a Thesaurus on the Semantic Web*. W3C Working Draft 10 May 2005. Available: <http://www.w3.org/TR/2005/WD-swbp-thesaurus-pubguide-20050510>. Accessed 19 Jun 2008.
- Miles, A. and Brickley, D. (eds.) (2005). *SKOS Core Guide*. W3C Working Draft 2 November 2005. Available: www.w3.org/TR/swbp-skos-core-guide. Accessed 19 Jun 2008.
- Mukundan, S. (2004). Spinning the Semantic Web. *On The Horizon*. 12(2), pp.74-78.
- Nasir Uddin, M. and Janecek, P. (2007). Faceted classification in web information architecture: a framework for using semantic web tools. *The Electronic Library*. 25(2), pp.219-233.
- Sadeh T. and Walker, J. (2003). Library portals: toward the semantic web. *New Library World*. 104(1184/1185), pp. 11-19.
- Sicilia, M. (2006). Metadata and semantics research. *Online Information Review*. 30(3), pp.213-216.
- Sure, Y. and Studer, R. (2005). Semantic Web technologies for digital libraries. *Library Management*. 26(4/5), pp.190-195.

Appendix D: Inventory briefing paper

JISC inventory

Purpose

To provide a single point of access to information, advice, tools, standards and technologies to support the development of resource discovery (information retrieval) services.

Methodology

- Create an inventory database
- Identify relevant items from JISC projects
- Record items in the database

Inventory database

1. Create a database in MS Access named JISCInventory
2. Create a table named Item
3. Add a field named ItemID (autonumber)
4. Add a field named Name (char, 120)
5. Add a field named URL (char, 120)
6. Add a field named Summary (memo)
7. Add a field named Standards (char, 120)
8. Add a field named Notes (memo)
9. Add a field named LastDate (small date) – autofill with current system date
10. Create a record add/edit/delete form

Inventory items

The following digital resources associated with JISC projects are relevant for the inventory:

- Final report of the project
- Case studies (web pages and/or separate documents such as DOC and PDF)
- Software packages created by the project
- Software tools created by the project
- Demonstrators
- Manuals associated with software packages, tools and demonstrators (web pages and/or separate documents)

If in doubt, assume the resource is relevant; it will be easier to delete non-relevant resources at a later stage.

Sources of information

The main source of information is the "What we do section" of the JISC website:

<http://www.jisc.ac.uk/whatwedo.aspx>

Other sources of information are the websites of individual projects and organisations, linked from the JISC web pages.

Identifying inventory items

Relevant items are more likely to be associated with the projects belonging the following JISC programmes:

- e-Learning Frameworks and Tools Programme
- Portals Programme

Presentation Programme

Shared Infrastructure Services Programme;

Users and Innovation Programme: personalising technologies

Virtual Research Environments (VRE) Programme

* Distributed National Electronic Resource (DNER 5/99) Learning and Teaching Programme

* Semantic Grid and Autonomic Computing Programme

(* These programmes are not listed on the JISC website, so ignore for the time being.)

To identify relevant items in these programmes:

1. Identify the programme in the list at: <http://www.jisc.ac.uk/whatwedo/programmes.aspx>
2. Follow the link to the programme's summary page.
3. If the start date of the programme is later than April 2007, do not continue (it is unlikely that any relevant inventory items will have been created so soon).
4. If the start date is before April 2007:
5. Identify any relevant inventory items listed on the programme summary page; these may be called background papers or something similar.
6. At the bottom of the summary page is a list of the JISC projects within the programme. Follow the link to each project's summary page.
7. If the start date of the project is later than April 2007, do not continue.
8. If the start date is before April 2007:
9. Identify any relevant inventory items listed on the project summary page.
10. Follow the link on the summary page to the project website.
11. Identify any relevant inventory items available on the project website. This may involve checking every web page; there is no common design or layout for project websites.

Recording inventory items

1. Enter a name for the item in the Name field. The name should be short, but long enough identify the item. Use the name supplied by the project or on the JISC website. In some cases, a name may have to be made-up; if so, use the following guide:
 - Name a project final report as Project name + "Final"; e.g. Mandate Final
 - Name a case study as Project name + "Case"; e.g. Mandate Case
 - If there is more than one case study requiring a made-up name, add a running number; e.g. Mandate Case 2.
 - Use the same pattern for demonstrators ("Demonstrator"), tools ("Tool"), etc.
2. Enter the full URL for the item in the URL field.
3. Enter a single-paragraph summary or description of the item in the Summary field. Where appropriate, use text from the website and edit it as required.
4. Enter acronyms for standards and technologies used by the item in the Standards field, as a delimited list. Use space-plus-space as the delimiter (" + "); e.g. XML + OAI-PMH + DC. Only include standards used by the item itself, and ignore standards such as PDF used to publish or access the item. The field will be parsed later to create a look-up table.
5. Enter any personal notes about the item, including reminders to check further, or ambiguities, etc. as a single paragraph in the Notes field. These will be used for further refinement of the inventory in due course, and the Notes field will eventually be deleted from the inventory.

Monitoring and feedback

- Send a copy of the Access mdb to the project manager once a week (say Monday).
- Feel free to ask the project manager for help whenever they are available.

Further work

We expect to get tables of programme names, project names, and associated institution names from JISC. These will be added to the Access database as look-up tables. The inventory items will be revisited to add the project name(s) and institution(s) at a later date.

Gordon Dunsire
8 Aug 2007

Appendix E: Resource Discovery iKit tutorial

Tutorial

The iKit provides users with an intuitive Web interface which assists in the discovery of the outputs of JISC-funded research in the area of resource discovery by allowing them to dynamically search a database of around 230 individually-catalogued, quality-assessed resources via a faceted taxonomic framework. These taxonomic facets include: JISC programme; JISC project; contributing institution; technologies used or described by the resource; and the thesaurus of Information Environment functional terms.

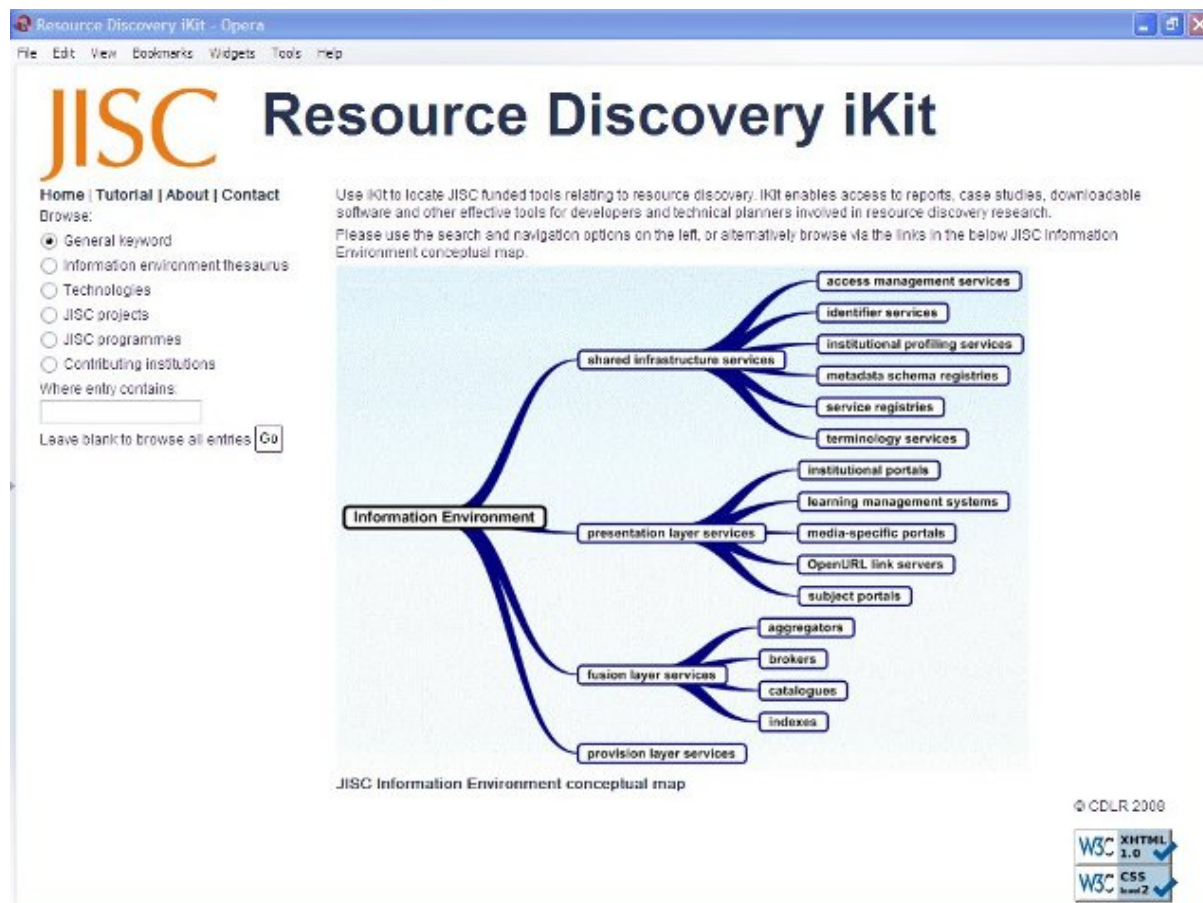


Figure 1: iKit main page, including IE thesaurus concept-map

Figure 1 shows the main page of the iKit. To the right of the screenshot can be seen the clickable image which displays the top two levels of the IE thesaurus as a conceptual map in order to offer an alternate browsing feature. On the left can be seen the search and browse interface, which forms part of the iKit Web template and appears on every page of the iKit. The search and browse interface prompts the user to select the facet of the iKit taxonomic framework they wish to search within (via choice of radio button) and then enter their search term. The default search is set to "general keyword". If no search term is entered the iKit will return all entries listed under the taxonomic facet selected, thus enabling the same interface to be used as both the searching and browsing of iKit content.

If the user chooses to search within one facet of the taxonomic framework they will be presented with an intermediate results screen showing all the controlled vocabulary terms which correspond to their search criteria. For example, the term "Edinburgh" entered within the "contributing institutions" facet will return the screen shown in figure 2, allowing the user then to further define their criteria.



Figure 2: Intermediate result for keyword search "Edinburgh" within Contributing institutions facet

Choosing one of these terms (e.g. "EDINA at Edinburgh, University of") then prompts the iKit to dynamically query the database via a Cold Fusion server script and return all items catalogued using this term. The results are then displayed in an alphabetical list which gives the item title, description and year of creation. This screen is shown in figure 3. Clicking the hyperlink at this point opens a new browser window to display the item directly (or to open the download dialogue in the case of an item of software).



Figure 3: Full results display for all items associated with "EDINA at Edinburgh, University of"

A keyword search under "general keyword" will immediately return all iKit items where the search term occurs in either the assigned keywords, in the item title or item description. Clicking the hyperlink will open the item, thus making it possible for a user having accessed the iKit and entered their search term, to potentially be able to access a relevant item in just two keystrokes.

The screenshot shows a web browser window titled "Resource Discovery iKit - Opera". The page header features the JISC logo and the title "Resource Discovery iKit". Below the header, there are navigation links: "Home | Tutorial | About | Contact".

The main content area displays the following information:

- HILT M2M feasibility study: final report**
- URL: http://hilt.cdli.strath.ac.uk/hiltm2mfs/0HIL_TM2MFinalReportRepV3_1.pdf
- Resource type: Report
- Created: 2005
- Small-scale study which aimed to determine the feasibility of providing a machine to machine (M2M) demonstrator of the functionality created in the HILT Phase II terminologies server pilot.
- Delivered by the [HILT High Level Thesaurus](#) project, funded under JISC's Repositories and Preservation Programme.
- The HILT project, now in phase IV, aims to research, investigate and develop pilot solutions for problems pertaining to cross-searching multi-subject scheme information environments, as well as providing a variety of other terminological searching aids.
- Further information is available from the JISC [project description](#) and [programme description](#).

On the left side, there is a search interface with the following elements:

- Home | Tutorial | About | Contact
- Browse:
- General keyword
- Information environment thesaurus
- Technologies
- JISC projects
- JISC programmes
- Contributing institutions
- Where entry contains:
-
- Leave blank to browse all entries

Below the search interface, there is a box titled "Show all items associated with:" containing the following text:

HILT: High Level Thesaurus Repositories and Preservation Programme
terminology services
MARC (Machine-Readable Cataloguing)
SKOS (Simple Knowledge Organisation System)
SOAP
SRU-SRW (Search/Retrieve via URL and Search/Retrieve Web Service)
Z39.50
CDLR at Strathclyde, University of
EDINA at Edinburgh, University of

On the right side, there are sections for "IE thesaurus terms" and "Standards, protocols and applications:"

IE thesaurus terms:
terminology services

Standards, protocols and applications:
MARC (Machine-Readable Cataloguing)
SKOS (Simple Knowledge Organisation System)
SOAP
SRU-SRW (Search/Retrieve via URL and Search/Retrieve Web Service)
Z39.50

Contributing Institutions:
CDLR at Strathclyde, University of
EDINA at Edinburgh, University of

At the bottom right of the page, there is a copyright notice: "© CDLR 2008".

Figure 4: Full record view example

The user can then either click the hyperlink to go directly to the item or the internal link "[Full record]" to be taken to the full iKit record for that item. This view, shown in figure 4, gives all the information held about that item. Visible to the left of figure 4 is a search interface offering internal links to search the iKit for items associated with all the controlled vocabulary terms included in that item description. In "Full record" view, a box appears on the search interface giving internal links to search the iKit for all the taxonomic terms which appear in that item record.

When searching within either the "IE thesaurus" or "technologies" facets, the user is presented with an option to explain that term. In the case of the IE thesaurus, this takes the user to a screen which gives the term definition, scope notes, and hierarchically and synonymously related terms (see figure 5). In the case of "Technologies", it takes the user to a screen which points the user towards definitions of that technology in (depending on availability) the JISC Standards Catalogue, the official Website for that technology, and/or its corresponding entry in Wikipedia.



Figure 5: Full term display for "terminology services"

Appendix F: Example of IE vocabulary preferred terms (from the interim report)

This is an extract of a structured thesaurus of controlled terms for describing the operational components of the JISC Information Environment, as determined at the interim report stage of the project.

The terms have been identified from the IE architecture diagram, other documents about the IE, and JISC project and programme descriptions.

The thesaurus complies with appropriate international standards, and supports hierarchical and equivalence relationships.

Terms have been assigned a numeric notation which reflects the hierarchy to which they belong.

Up to 6 levels of granularity have been identified for some sections of the thesaurus.

The granularity of the thesaurus is linked to the granularity of descriptions of infokit tools (see Appendix H): if tools are described at the project level, it is unlikely that the lower level thesaurus terms are required; if tools are described at a more specific level, lower level terms would improve retrieval.

The thesaurus is incomplete with respect to related terms (see-also relationship).

IE Vocabulary Preferred Terms (Example – 30 Oct 2007)

Notation	Term	Notes	Related Terms
1	provision layer services	Layer within the JISC Information Environment that encompasses all resources made available by content providers, including JISC-funded data services, and e-journal providers.	
1.1	JISC funded content providers		
1.2	institutional content providers		
1.2.1	institutional repositories		public repositories; open access initiative
1.2.2	digital repositories		public repositories; open access initiative
1.2.3	open access initiative		
1.3	external content providers		
1.3.1	private publishers		
1.3.2	public repositories		public repositories; digital repositories; institutional repositories
2	fusion layer services	Layer within the JISC Information Environment that contains middleware responsible for combining metadata records for one or more content providers, as a result of cross-searching, harvesting, or alerting.	
2.1	brokers	A structured network service that provides (search) access to a range of other, heterogeneous, local or remote structured network services.	
2.2	aggregators	A structured network service that gathers metadata from a range of other, heterogeneous, local or remote structured network services.	
2.2.1	news-feed aggregation		
2.2.2	metadata harvesting		
2.3	catalogues	A network service that provides access to a collection of human-generated metadata records.	
2.3.1	library catalogues		
2.3.2	online union catalogues		

Notation	Term	Notes	Related Terms
2.3.3	Intute		subject portals
2.4	indexes	A network service that provides access to a machine-generated database of information derived from the content of items in a collection.	
2.4.1	automatic indexes		
2.4.1.1	web search engines		
3	shared infrastructure services	A common set of infrastructural services which provide information through machine-to-machine (m2m) interfaces, and on which other elements of the environment (for example: portals, brokers and aggregators) can draw.	
3.1	service registries	A network service that stores and makes available descriptions of (i.e. metadata about) services and the content of collections made available through those services.	
3.2	institutional profiling services	Services which provide machine-readable information about resolution preferences and local library holdings information; intended to support OpenURL and Resolver services.	
3.2.1	licencing services		
3.3	metadata schema registries	Store, manage and make available descriptions of the metadata schemas in use by other service components.	
3.4	identifier services	A network service that takes metadata about a resource and supplies pointers to services on, or related to, that resource, for example delivery services. In the context of the JISC IE, metadata is passed to the resolver using an OpenURL.	
3.5	terminology services	A structured network service that offers terminology-related services, for example mapping a term from one controlled vocabulary to another or expanding terms within a thesaurus.	
3.5.1	name authority services		
3.6	authentication and authorisation services	Authentication is the process whereby a digital ID is determined to relate to the real life individual who has the rights to use it. Authorisation is the process whereby a digital ID is determined to have the necessary rights to access a given resource.	

Notation	Term	Notes	Related Terms
3.6.01	Athens	Athens is an Eduserv service delivering "authentication, authorisation and distributed user administration services" to UK HE/FE.	
3.6.1	core middleware	The central services that are essential to middleware as a whole. These are: authentication, authorisation, directory services, and identifiers.	
3.6.1.1	identifiers		
3.6.1.1.1	identity providers		
3.6.1.1.1.1	institutional identity providers		
3.6.1.1.1.2	single sign-on services		
3.6.1.2	where are you from services		
3.6.1.3	directory services		
3.6.1.3.1	institutional directory services		
4	presentation layer services	Layer within the JISC Information Environment that encompasses all potential interfaces available to users.	
4.1	institutional portals		
4.1.1	library portals	A library portal extends the functionality offered by the library catalogue, typically including cross-searching of local and remote collections.	
4.2	learning management systems		
4.2.1	virtual learning environments		
4.2.2	virtual research environments		
4.3	media specific portals	A media-specific portal provides a single point of access to resources of a particular type (books, images, geo-spatial data, etc.).	
4.4	openURL link servers	A network service that takes metadata about a resource in the form of an OpenURL and supplies links (pointers) to services on, or related to, that resource, for example delivery services.	

Notation	Term	Notes	Related Terms
4.5	subject portals	A subject portal provides personalised discovery services across multiple, heterogeneous content providers (offering books, journals, Web sites, learning objects, images, etc.) within a specific subject area.	Intute;

Appendix G: Example of IE vocabulary non-preferred terms (from the interim report)

This is an extract of a database table of non-preferred terms for the JISC Information Environment, as determined at the interim report stage of the project.

The terms have been identified from IE documentation, JISC project and programme descriptions, and standard methods for thesaurus construction such as inverted entries for phrase terms.

The preferred term from the controlled vocabulary is identified by its notation.

The list of non-preferred terms is incomplete, and may be expanded indefinitely with the addition of spelling variants, acronyms, non-UK Anglophone terms, etc.

IE Vocabulary Non-Preferred Terms (Example – 30 Oct 2007)

Synonym	Term	Notation
access management, federated	core middleware	3.6.1
aggregation	aggregators	2.2
alert services	news-feed aggregation	2.2.1
Athens authentication system	Athens	3.6.01
authentication	authentication and authorisation services	3.6
authentication mechanisms, federated	core middleware	3.6.1
authentication services	authentication and authorisation services	3.6
authorisation	authentication and authorisation services	3.6
authorisation services	authentication and authorisation services	3.6
automatic indexing	automatic indexes	2.4.1
catalogue	catalogues	2.3
cataloguing	catalogues	2.3
content	provision layer services	1
content management	provision layer services	1
content management systems	provision layer services	1
content providers	provision layer services	1
content providers, external	external content providers	1.3
content providers, institutional	institutional content providers	1.2
content providers, JISC	JISC funded content providers	1.1
content provision	provision layer services	1
content provision, external	external content providers	1.3
content provision, institutional	institutional content providers	1.2
content provision, JISC	JISC funded content providers	1.1
content, external	external content providers	1.3
content, institutional	institutional content providers	1.2
content, JISC funded	JISC funded content providers	1.1
control services	shared infrastructure services	3

Synonym	Term	Notation
controlled vocabularies	terminology services	3.5
cross-searching services	fusion layer services	2
e-print archives	institutional repositories	1.2.1
federated access management	core middleware	3.6.1
federated authentication mechanisms	core middleware	3.6.1
federation	core middleware	3.6.1
federation services	core middleware	3.6.1
fusion	fusion layer services	2
fusion services	fusion layer services	2
gateways, subject	subject portals	4.5
Google	web search engines	2.4.1.1
handle	single sign-on services	3.6.1.1.1.2
handle services	single sign-on services	3.6.1.1.1.2
harvesting	metadata harvesting	2.2.2
harvesting (metadata)	metadata harvesting	2.2.2
identifier resolution services	identifier services	3.4
identifier resolver services	identifier services	3.4
identity provider	identity providers	3.6.1.1.1
identity providers, institutional	institutional identity providers	3.6.1.1.1.1
IdP	identity providers	3.6.1.1.1
IESR	service registries	3.1
indexing	indexes	2.4
infrastructural services	shared infrastructure services	3
institutional preference services	institutional profiling services	3.2
interface services	presentation layer services	4
interfaces	presentation layer services	4
JISC IE service registry	service registries	3.1
library catalogue	library catalogues	2.3.1
licensing services	licencing services	3.2.1
managed learning environments	virtual learning environments	4.2.1

Synonym	Term	Notation
name authority files	name authority services	3.5.1
name authority records	name authority services	3.5.1
news feed services	news-feed aggregation	2.2.1
OAI	open access initiative	1.2.3
online public access catalogue	library catalogues	2.3.1
online public access catalogues	library catalogues	2.3.1
online union catalogue	online union catalogues	2.3.2
ontologies	terminology services	3.5
OPAC	library catalogues	2.3.1
OPACs	library catalogues	2.3.1
openURL resolvers	openURL link servers	4.4
portals	presentation layer services	4
portals, institutional	institutional portals	4.1
portals, library	library portals	4.1.1
portals, media specific	media specific portals	4.3
portals, subject	subject portals	4.5
presentation	presentation layer services	4
presentation services	presentation layer services	4
provision	provision layer services	1
provision, content	provision layer services	1
provision, content, external	external content providers	1.3
provision, content, institutional	institutional content providers	1.2
provision, content, JISC	JISC funded content providers	1.1
provision, external	external content providers	1.3
provision, institutional	institutional content providers	1.2
provision, JISC	JISC funded content providers	1.1
publishers, private	private publishers	1.3.1
repositories, digital	digital repositories	1.2.2
repositories, institutional	institutional repositories	1.2.1
repositories, public	public repositories	1.3.2
resolvers	identifier services	3.4

Synonym	Term	Notation
Resource Discovery Network	Intute	2.3.3
RSS	news-feed aggregation	2.2.1
shared infrastructure	shared infrastructure services	3
shared services	shared infrastructure services	3
Shibboleth	core middleware	3.6.1
single sign-on	single sign-on services	3.6.1.1.1.2
subject gateways	subject portals	4.5
thesauri	terminology services	3.5
union catalogue	online union catalogues	2.3.2
union catalogues	online union catalogues	2.3.2
vocabularies, controlled	terminology services	3.5
WAYF	where are you from services	3.6.1.2
WAYF services	where are you from services	3.6.1.2
web crawlers	automatic indexes	2.4.1
web robots	automatic indexes	2.4.1
web spiders	automatic indexes	2.4.1
where are you from	where are you from services	3.6.1.2

Appendix H: Example of Infokit items (from the interim report)

This is an extract from a database table of "tools" (items) identified as suitable for the RD Infokit.

The extract illustrates two main levels of granularity, for the output of a project as a whole, or specific tools generated by a project.

The JISC project and programme name information is taken from related tables.

The extract contains first-pass assignment of technical standards access points. These have not yet been normalised into a separate database table. Further work needs to be done to identify additional access points and assign them to Infokit tools, and to consolidate the list of technical standards.

Item Report (Example – 30 Oct 2007)

Item.Name ANGEL Tool2

URL <http://www.angel.ac.uk/current->

Summary ANGEL - User Manager, handles users of the Angel system.

It allows users to log in to the system, provides authentication of the user and delivers authorisation details to the user thereby allowing them the rights to access the appropriate resources.

Each Angel system is intended to be an individual service to users of a particular institution. The UM enables easier portability to further institutions that wish to adopt it as a solution. The Angel User Manager (UM) queries one or more institutional services to obtain user verification and directory data, presenting it in a unified format. This allows the main Angel software to treat any local user information in the same way, disguising all the technical and organisational idiosyncrasies.

The main advantage to end-users of this approach is that once logged in to the Angel system there is no further need for them to enter username/password combinations to access resources controlled by Angel, such as libraries, content repositories and so on. They don't have to learn (or write down, insecurely) yet another password for each new service that the institution creates or connects to.

Standards

Project.Name ANGEL: Authenticated Networked Guided Environment

Programme.Name Learning and teaching (5/99) programme

Item.Name ANGEL Tool3

URL <http://www.angel.ac.uk/current->

Summary DELIVER Smart Link Finder: The DELIVER Smart Link Finder is a tool for building and maintaining simple reading lists. It includes a Universal Harvesting Tool (UHT) to allow users to simply and quickly add resource to URLs to course pages, and for these URLs to be supported via a central database. In order to allow the UHT tool to use SmartLinks it must be configured to communicate with its own UHT database and an instance of the Angel Resource Manager.

Standards VLE+LMS

Project.Name ANGEL: Authenticated Networked Guided Environment

Programme.Name Learning and teaching (5/99) programme

Item.Name	ARCHES Final
URL	http://www.jisc.ac.uk/uploaded
Summary	Antiquity Related Collections Harnessed for Educationa. ARCHES aims to support and link institutions, departments, courses and modules as they introduce, evaluate and disseminate exemplary, transformative and innovative pedagogy through re-purposing new and existing collections of digital resources pertaining to ancient Greece and Rome.
Standards	
Project.Name	ARCHES - Antiquity Related Collections Harnessed for
Programme.Name	Exchange for learning (X4L) programme
Item.Name	ARCHES Tool
URL	
Summary	Online database
Standards	
Project.Name	ARCHES - Antiquity Related Collections Harnessed for
Programme.Name	Exchange for learning (X4L) programme
Item.Name	ASK: Accessing and Storing Knowledge
URL	http://ask.oucs.ox.ac.uk/
Summary	The ASK project will develop a suite of open source software artefacts that support learners, researchers and teachers in securely accessing and sharing learning objects. The project will implement a pilot design pertaining to a repository reference model based on the JISC E-Learning Framework (ELF) . The project will implement services that fall within the ELF i.e. metadata management, content management, authentication, and authorisation (Shibboleth) components. The overall goal of the project is to build interoperability between user agents (such as VLEs and Activity Management systems) and information systems normally under the management of library departments. The ASK project is a direct response to the project partners institutional needs to integrate library and learning services.
Standards	ELF
Project.Name	ASK: Accessing and Storing Knowledge
Programme.Name	Digital repositories programme 2005-7

Item.Name	CATS Final Report
URL	http://www.jisc.ac.uk/media/do
Summary	The overall aim of the project is to create a toolkit of loosely-coupled web services which support the various tasks inherent to automated assessment construction e.g. searching for, retrieving and aggregating assessment items held in multiple item banks. The outputs of the project will be disseminated to users via the CETIS Assessment SIG (Special Interest Group). In addition, the project will undertake a consultation exercise with several domain practitioners involved in creating assessments from item banks. This task will give an overview of current practice, as well as indicating future functionality which CATS could support.
Standards	SPAID+Discovery Plus
Project.Name	CATS: Constructing Assessments using Tools and Services
Programme.Name	e-Learning framework and tools programme
Item.Name	CC-Interop (Copac and CLUMPS interoperability project) Final
URL	http://ccinterop.cdrl.strath.ac.uk
Summary	<p>Copac and CLUMPS interoperability project. CC-interop was a JISC funded collaborative project, involving the M25 Systems Team, CDLR, MIMAS and RIDING and built on the expertise acquired in recent years in the area of union catalogues, both physical and virtual, aiming to take the JISC vision of the Information Environment ever closer.</p> <p>Beginning May 2002 and finishing in June 2004, CC-interop was a two year project which was split into three work packages: Work Package A was the responsibility of M25 Systems Team based at LSE and MIMAS, which hosts the established JISC service COPAC. It included:</p> <ul style="list-style-type: none"> Bullet working co-operatively to test the feasibility of inter-linking between union catalogues, both physical and virtual Bullet investigating various ways of getting a virtual union catalogue to be co-searchable with a physical union catalogue without having to re-key the search Bullet testing further some of the findings of the UKNUC Feasibility Study Report in relation to comparative speed of searching, de-duplication and ranking of records, and precision of results sets between COPAC and the clumps
Standards	Copac+CLUMPS
Project.Name	Copac and CLUMPS interoperability project (CC-Interop)
Programme.Name	Digital library infrastructure programme

Appendix I: XML manifestation of RDF/SKOS expression of IE vocabulary

This output is courtesy of the NSDL Metadata Registry Sandbox, using the vocabulary "JISC IE Terms" at <http://sandbox.metadataregistry.org/vocabulary/show/id/94.html>.

Note: The domain "namespaces.jiscexample.ac.uk/IETerms" is a place-holder; it does not exist, and needs to be replaced by a real domain to register this vocabulary and publish the XML file on the web.

```
<?xml version="1.0" encoding = "UTF-8"?>
<rdf:RDF
  xmlns="http://www.w3.org/2004/02/skos/core#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">

<!-- Scheme: JISC IE Terms -->
<skos:ConceptScheme rdf:about="namespaces.jiscexample.ac.uk/IETerms/">
  <dc:title>JISC IE Terms</dc:title>
  <skos:hasTopConcept rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1005"/>
  <skos:hasTopConcept rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1018"/>
  <skos:hasTopConcept rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1019"/>
  <skos:hasTopConcept rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1021"/>
</skos:ConceptScheme>

<!-- Concept: access management services -->
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1001">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>access management services</skos:prefLabel>
  <skos:definition>Network services that store, manage and make available information that govern
the authentication and/or authorisation of sign-on credentials to determine a user's right to use a
requested resource.</skos:definition>
  <skos:note>Authentication is the process whereby a digital ID is determined to relate to the real life
individual who has the rights to use it. Authorisation is the process whereby a digital ID is determined
to have the necessary rights to access a given resource. Authentication is often achieved via the use
of a username/password combination or a digital certificate, depending on the degree of assurance
required. A conflation of two conceptually distinct though interrelated network services, these services
have traditionally been married in the JISC IE due to the fact that the Athens Access Management
System has provided a single, shared authentication and authorisation service for the UK higher and
further education community since 1996. Research is currently underway to replace this system with
one based on Shibboleth federated access technology.</skos:note>
  <skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1021"/>
</skos:Concept>

<!-- Concept: aggregators -->
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1002">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>aggregators</skos:prefLabel>
  <skos:definition>Network services that gather metadata records from a range of other,
heterogeneous, local or remote network services and offer them for searching or harvesting by others
or as an aggregated news channel.</skos:definition>
  <skos:altLabel>Aggregators are intended for use by software applications. In the context of the
JISC IE, aggregators interact with indexes, catalogues, content providers and other aggregators using
the OAI-PMH and/or RSS/HTTP.</skos:altLabel>
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1013"/>
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1015"/>
  <skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1005"/>
</skos:Concept>
```

```
<!-- Concept: brokers -->
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1003">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>brokers</skos:prefLabel>
  <skos:definition>Network services that accept a search query from a presentation layer service,
transfer it federally (and in parallel) to one or more provision layer service, before combining the
results and returning them to the requester.</skos:definition>
  <skos:note>Brokers are intended for use by software applications. In the context of the JISC IE,
brokers interact with indexes, catalogues, aggregators, content providers, other brokers and portals
using Z39.50. Results from multiple content providers are usually merged and ranked before being
passed back to the presentation layer service.</skos:note>
  <skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1005" />
</skos:Concept>

<!-- Concept: catalogues -->
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1004">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>catalogues</skos:prefLabel>
  <skos:definition>Network services that provide access to collections of (typically human-generated)
metadata records about content held elsewhere.</skos:definition>
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1016" />
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1024" />
  <skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1005" />
</skos:Concept>

<!-- Concept: fusion layer services -->
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1005">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>fusion layer services</skos:prefLabel>
  <skos:definition>Stratum of the JISC Information Environment encompassing middleware
responsible for combining metadata records for one or more provision layer services via means of
cross-searching, harvesting, alerting and/or manual cataloguing.</skos:definition>
  <skos:note>The positioning of some services within the fusion layer might be seen to be
ambiguous, as it is possible for that activity defined as ""fusion"" activity to be undertaken directly by
services more readily identified as belonging in either the provision or presentation layers. To take an
example, commercial abstracting and indexing services, such as ISI Web of Science, are hard to
place within the JISC IE architecture. While they might be thought an ""index"", based in part on the
generic name of the type of service that they offer, their interfaces are made directly available to the
user, and as such might be thought to be a ""portal"", thus placing them in the presentation layer. In
other cases, stand-alone fusion services may be developed. In the case of cross-searching, such
stand-alone services are typically referred to as brokers. In the case of harvesting and alerting, such
services are referred to as aggregators. Services in the fusion layer may combine the harvesting and
distributed searching approaches. For example, a fusion service may gather metadata records from
content providers using the OAI-PMH, and make the combined database of gathered records
available for searching using Z39.50.</skos:note>
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1002" />
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1003" />
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1004" />
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1007" />
</skos:Concept>

<!-- Concept: identifier services -->
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1006">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>identifier services</skos:prefLabel>
  <skos:definition>Network services that store, manage and make available an association between
an identifier and some metadata about the identified resource, with the intention of supplying pointers
to services that relate to the identified resource.</skos:definition>
```

<skos:note>In the JISC IE, metadata is passed to a resolver service encoded in an OpenURL. Typically, such services take an identifier encoded as a URI and return an HTTP redirect to the current location of the resource being identified. Such a service can provide extended functionality, which may include redirection to document delivery services, redirection to online bookshops, redirection to local library services and further "discovery" activities. Further, it is anticipated that resolver services may make use of institutional profiling services to determine institutional resolution preferences.</skos:note>

```
<skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1021"/>
</skos:Concept>
```

<!-- Concept: indexes -->

```
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1007">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>indexes</skos:prefLabel>
  <skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1005"/>
  <skos:definition>Network services that provide access to machine-generated databases of
information derived, by Web robots or other techniques, from the content of items in a
collection.</skos:definition>
</skos:Concept>
```

<!-- Concept: institutional portals -->

```
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1008">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>institutional portals</skos:prefLabel>
  <skos:definition>Network interface services that provide members of the host institution with a
single, personalised interface to institutionally-selected resources relevant to programmes of
study.</skos:definition>
  <skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1018"/>
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1011"/>
</skos:Concept>
```

<!-- Concept: institutional profiling services -->

```
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1009">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>institutional profiling services</skos:prefLabel>
  <skos:definition>Network services that store, manage and make available machine-readable
information about the preferences and profiles of institutions operating within the JISC
IE.</skos:definition>
  <skos:note>The primary intention of this service is to allow resolver services to automatically
determine institutional preferences and holdings information when resolving OpenURLs by making
available information such as preferred e-content suppliers and existing licensing
agreements.</skos:note>
  <skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1021"/>
</skos:Concept>
```

<!-- Concept: learning management systems -->

```
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1010">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>learning management systems</skos:prefLabel>
  <skos:definition>Network interface services that provide a single, integrated environment for a range
of functionality that supports teaching, learning and research activities.</skos:definition>
  <skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1018"/>
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1025"/>
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1026"/>
</skos:Concept>
```

<!-- Concept: library portals -->

```
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1011">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>library portals</skos:prefLabel>
```

```
<skos:definition> Network interface services that extends the functionality offered by the library catalogue, typically including cross-searching of local and remote collections.</skos:definition>
<skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1008"/>
</skos:Concept>
```

```
<!-- Concept: media-specific portals -->
```

```
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1012">
<skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms"/>
<skos:prefLabel>media-specific portals</skos:prefLabel>
<skos:definition>Network interface services that provide a single point of access to resources of a particular type.</skos:definition>
<skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1018"/>
</skos:Concept>
```

```
<!-- Concept: metadata harvesting -->
```

```
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1013">
<skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms"/>
<skos:prefLabel>metadata harvesting</skos:prefLabel>
<skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1002"/>
<skos:definition>The process whereby the metadata descriptions of records in heterogeneous archives are collected (harvested) to build a central metadata aggregation.</skos:definition>
</skos:Concept>
```

```
<!-- Concept: metadata schema registries -->
```

```
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1014">
<skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms"/>
<skos:prefLabel>metadata schema registries</skos:prefLabel>
<skos:definition>Network services that store, manage and make available machine-readable and human-readable information about the metadata schemas in use by other service components.</skos:definition>
<skos:note>The primary intention of such a service is to allow portals, brokers and aggregators to automatically determine information about appropriate search terms and the structure of metadata records that will be returned to them. Metadata registries, however, also provide a useful human-oriented service, allowing people to see what metadata schemas are in use by which services - providing a basis for metadata schema sharing and re-use. Typically some combination of metadata semantics and syntax will be described.</skos:note>
<skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1021"/>
</skos:Concept>
```

```
<!-- Concept: news-feed aggregation -->
```

```
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1015">
<skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms"/>
<skos:prefLabel>news-feed aggregation</skos:prefLabel>
<skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1002"/>
<skos:definition> The process whereby client software or Web applications collect together syndicated web content such as news headlines, blogs, podcasts, and vlogs in a single location for viewing by the user.</skos:definition>
<skos:note>In addition to a human-oriented Web interface, such services also typically offer machine interfaces in the form of aggregated RSS channels.</skos:note>
</skos:Concept>
```

```
<!-- Concept: online public access catalogues -->
```

```
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1016">
<skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms"/>
<skos:prefLabel>online public access catalogues</skos:prefLabel>
<skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1004"/>
<skos:definition>Computerised collection of bibliographic metadata that describes the holdings of a library and is available, locally or remotely, to users.</skos:definition>
</skos:Concept>
```

```
<!-- Concept: Open-URL link servers -->
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1017">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>Open-URL link servers</skos:prefLabel>
  <skos:definition>Network services that take metadata about an item, encoded in the form of an
OpenURL, and provide the end-user with a set of links to delivery services where the item can be
obtained and to value-added services associated with the item.</skos:definition>
  <skos:note>OpenURL resolver services are usually offered within the institution, typically by the
library, using proprietary software.</skos:note>
  <skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1018"/>
</skos:Concept>

<!-- Concept: presentation layer services -->
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1018">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>presentation layer services</skos:prefLabel>
  <skos:definition>Stratum of the JISC Information Environment encompassing all potential
interfaces with which end-users can interact.</skos:definition>
  <skos:note>Includes all Graphic User Interfaces, whether hosted locally (on a user's
desktop/browser) or remotely (on the web.) In the context of the JISC IE, presentation layer services
interact with fusion and provision layer services via the following protocols and standards: Z39.50,
SRW, the OAI-PMH and RSS/HTTP. Presentation layer services are sometimes synonymised with
portals; portals, however, should be understood more narrowly, as discrete network services that
provide a personalised, single point of access to a range of heterogeneous network services, local
and remote. Thus, portals can be seen to be a type of presentation layer service; to synonymise
portals with all possible human-computer interfaces can broaden the term beyond meaningful
use.</skos:note>
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1008"/>
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1010"/>
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1012"/>
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1017"/>
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1022"/>
</skos:Concept>

<!-- Concept: provision layer services -->
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1019">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>provision layer services</skos:prefLabel>
  <skos:definition>Stratum of the JISC Information Environment encompassing all services
responsible for the storage, management and making available of content.</skos:definition>
  <skos:note>Such content includes all resources made available by content providers, including
JISC-funded data services and e-journal providers. Content is the material that end-users want to
access and use - scholarly journals, monographs, textbooks, learning objects, abstracts, manuscripts,
maps, music scores, still images, geospatial images and other kinds of vector and numeric data, as
well as moving picture and sound collections.</skos:note>
</skos:Concept>

<!-- Concept: service registries -->
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1020">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>service registries</skos:prefLabel>
  <skos:definition>Network services that store, manage and make available machine-readable
descriptive metadata about the services offered within the JISC IE by other service components, and
the content of collections made available via such services.</skos:definition>
  <skos:note>Instantiated as the JISC IE Service Registry (IESR), the term "service registry" actually
refers to two conceptually distinct, though complementary, network services: "collection description"
and "service description". A collection description service provides machine-readable descriptions of
the content of collections that are available as part of the JISC IE. The primary intention of the service
is to allow portals, brokers and aggregators to automatically determine what collections are available
to end-users of the JISC IE. A service description service provides detailed technical information in
```

machine-readable form about JISC IE services. The primary intention of the service is to allow portals, brokers and aggregators to automatically determine how they should interact with content providers and other brokers and aggregators.</skos:note>

```
<skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1021"/>
</skos:Concept>
```

<!-- Concept: shared infrastructure services -->

```
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1021">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>shared infrastructure services</skos:prefLabel>
  <skos:definition>Common set of foundational services that support the activities of all other
  services within the JISC IE by providing standardized, machine-readable information about services,
  content, rights and users via machine-to-machine (m2m) interfaces.</skos:definition>
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1001"/>
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1006"/>
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1009"/>
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1014"/>
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1020"/>
  <skos:narrower rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1023"/>
</skos:Concept>
```

<!-- Concept: subject portals -->

```
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1022">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>subject portals</skos:prefLabel>
  <skos:definition>Network interface services that provide a single point of access to personalised,
  non-media-specific, discovery services across multiple, heterogeneous content providers within a
  specific subject area.</skos:definition>
  <skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1018"/>
</skos:Concept>
```

<!-- Concept: terminology services -->

```
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1023">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>terminology services</skos:prefLabel>
  <skos:definition>Network services that store, manage and make available controlled vocabularies
  used by other services within the JISC IE and provide automated, machine-readable mappings
  between terms in one or more thesauri or classification schema.</skos:definition>
  <skos:note>Includes ontologies, classification schemes, thesauri and any other scheme used to
  manage the use of language. The intention of mapping terms is to allow the user to use one
  terminological set and for their terms to be automatically mapped to alternative or additional terms in
  thesauri or classification schemes used by target services. Terminology is not limited only to subject
  but also includes audience level, resource type and certification.</skos:note>
  <skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1021"/>
</skos:Concept>
```

<!-- Concept: union catalogues -->

```
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1024">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>union catalogues</skos:prefLabel>
  <skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1004"/>
  <skos:definition>Collection of physically or virtually centralised bibliographic metadata records
  sourced from two or more heterogeneous online public access catalogues.</skos:definition>
</skos:Concept>
```

<!-- Concept: virtual learning environments -->

```
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1025">
  <skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
  <skos:prefLabel>virtual learning environments</skos:prefLabel>
```

```
<skos:definition>Network interface services that aggregate integrated tools to provide a single
framework of resources and tools to support online learning and teaching.</skos:definition>
<skos:note>Examples of such resources and tools are: learning delivery mechanisms; student
tracking and assessment; and applications to control access to resources. These integrated tools may
be one product (e.g. BlackBoard, WebCT) or an integrated set of individual, perhaps open-source,
tools.</skos:note>
<skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1010"/>
</skos:Concept>

<!-- Concept: virtual research environments -->
<skos:Concept rdf:about="namespaces.jiscexample.ac.uk/IETerms/1026">
<skos:inScheme rdf:resource="namespaces.jiscexample.ac.uk/IETerms/" />
<skos:prefLabel>virtual research environments</skos:prefLabel>
<skos:definition>Network interface services that aggregate integrated tools to provide a single
framework of resources and tools to support the underlying processes of research.</skos:definition>
<skos:note>The purpose of a virtual research environment (VRE) is to help researchers in all
disciplines manage the increasingly complex range of tasks involved in carrying out research. VREs
add value to the research process across all disciplines by complementing and inter-working with
existing resources and by being flexible and adaptable to changing requirements.</skos:note>
<skos:broader rdf:resource="namespaces.jiscexample.ac.uk/IETerms/1010"/>
</skos:Concept>

</rdf:RDF>
```

Appendix J: Comments received during beta-testing

The following selected comments were received as feedback during beta-testing of the iKit service. They may inform future development of the service.

- This looks like a potentially useful tool, but one that has wider application than the limited area of Resource Discovery. For instance, another area I'm watching is that of Personal Portfolios where there are plenty of similar resources. Hope the data tagging requirements are not too massive and it could be extended to such areas.
- How about a Web2 feature? ie 'Rate the value of this project' and 'Addcomments'. But that might be a bit mischievous (!)
- Any list of project reports should say if the linked document is a pdf or doc or html etc, so the user knows what she is opening up. Maybe there could be a function that converted the pdfs to html? (as happens on Google)
- Will the underlying technology for this be extensible? I would be keen to use it for my JISC programme (Digitisation)
- A future JISC funding programme area beneficial could include possible funding timespan for resource discovery
- I don't know if you have seen it, but you should take a look at the work ... OSS-Watch is doing with Simal. It may be too late to incorporate into your project, but it may be worth sharing the data set with him to see what Simal can do for this use-case.
- I wish I was aware of this before you entered beta. However, discovering you now is better than at the end of your project. Simal is a very early stage project borne out of an internal need here at OSS Watch. It is open source and is intended to draw together info many sources, such as yours ... Having access to your data set (perhaps through an API) would add considerable value to our own application.
- Is it only a matter of time before they are looking for a full-text search of significant project reports as well I wonder?
- I think it likely that people and organisations running projects will want an opportunity to make sure that their project is presented in a clear/accurate/up-to-date light. Not sure how this can be managed. Maybe through an 'update' button that tells users the up to date situation? The text and links that result would have to be fed by the projects themselves.
- There is a similar issue I think with demos. Again, I'm not sure how it can be addressed. I think if JISC plans to pursue this idea, they need to be thinking in terms of having the update button go to an area on project websites specifically designed for the purpose. This would have to be kept up to date during projects and 'finalised' at the end of projects. JISC would have to make it a funding requirement.