Interoperable Subject Retrieval in a Distributed Multi-Scheme Environment: New Developments in the HILT Project.

Authors: Dennis Nicholson and Emma McCulloch.

1.   Introduction

Overview

The HILT (HIgh-Level Thesaurus) project (http://hilt.cdlr.strath.ac.uk/), based primarily at the Centre for Digital Library Research (CDLR) (http://cdlr.strath.ac.uk/) at Strathclyde University in Glasgow is entering its fourth stage following the completion of Phases I (http://hilt.cdlr.strath.ac.uk/index1.html) and II (http://hilt.cdlr.strath.ac.uk/index2.html) and the Machine to Machine (M2M) Feasibility Study (http://hilt.cdlr.strath.ac.uk/hiltm2mfs/). HILT is funded by the Joint Information Systems Committee (JISC) in the United Kingdom (UK) to examine an issue of global significance – facilitating interoperability of subject descriptions in a distributed, cross-service retrieval environment where different services use different subject and classification schemes to describe content, making cross-searching by subject difficult. HILT Phase I determined that there was a community consensus in the UK in favour of using inter-scheme mapping to achieve interoperability between services using different schemes, an approach followed by several recent projects (Heery et al, 2001; Koch et al, 2001; MACS, 2005; Saeed and Chaudhury 2002). HILT Phase II chose a spine-based approach to mapping and chose the Dewey Decimal Classification (DDC) as the central scheme to which all other schemes would be mapped. It also built an illustrative pilot mapping service, based on an adaptation of the Wordmap (http://www.wordmap.com/) terminology-handling software and made a range of recommendations on issues requiring further research and ongoing development requirements.

There are a range of issues to resolve before a fully operational (as opposed to pilot) JISC terminologies service can become a reality. Of these, one of the most important is the provision of the facilities offered in the Phase II pilot via web-services protocols so as to enable M2M interaction with other components of the JISC Information Environment architecture (JISC, 2005) - services such as Go Geo! (http://www.gogeo.ac.uk/), or the Resource Discovery Network (RDN) (http://www.rdn.ac.uk/) . The HILT Phase II pilot encompassed a range of facilities but only made them available through a direct user interface. An M2M interface is required to allow other machines to query the pilot server in the same way that end users can now, thereby permitting the various JISC services to provide terminology mapping services to their users in a transparent way. Since it proved difficult to specify and cost a project that would develop an M2M based version of the Phase II pilot, an M2M feasibility study was funded to clarify the detail. This, and the resultant follow-up Phase III project is the primary focus of the present paper. Further information is available on Phases I and II in the project reports (HILT, 2002; 2003), and in a range of other papers (McCulloch et al, 2004, 2005; Nicholson, 2003; Nicholson and Shiri, 2003; Nicholson and Wake, 2001; Nicholson et al 2001, 2002, 2005; Shiri et al, 2004).

Brief Description of the Phase II Pilot

The illustrative but functional Wordmap-based HILT Phase II pilot can be accessed and used online at http://hiltpilot.cdlr.strath.ac.uk/pilot/top.php and is demonstrated to some extent in the screen shots provided in Appendix A of this paper. Worked examples can be found in Appendix I of the HILT Phase II Final report (HILT, 2003).

The pilot is based on a DDC (kindly provided in machine-readable form by OCLC)
spine and encompasses:

- Access to the whole of DDC 21, indexed on the DDC captions, standard sub-divisions, relative index, and the other schemes mentioned below
- Mappings of DDC to LCSH as provided by OCLC
- Illustrative mappings to UNESCO and MeSH
- An illustrative staff interface to the system based on standard Wordmap windows 'drag and drop' style interface but utilised in a HILT-specific way
- A user query interface
- A user query disambiguator (e.g. by lotus, do you mean the flower, the car, the software etc)

- A 'find collections appropriate to disambiguated query' function, based on a DDC truncation algorithm and (simulated) interaction with the proposed JISC Information Environment Services Registry (IESR) shared service
- A 'determine subject scheme used by retrieved collection' function
- A 'determine specific term from that scheme that maps to user's query' function
- A 'find hits in retrieved collection' function using this term
- Minimal on-screen user help

More detail on what it does and how it functions can be found in the literature (Nicholson et al, 2005). The M2M feasibility study looked at offering this kind of functionality via an M2M interface, rather than the direct user interface seen in the pilot – although, as we shall see, not necessarily via the Wordmap-based pilot.

2. The HILT M2M Feasibility Study

Background: M2M and HILT Phase II, including UKOLN Recommendations

The HILT Phase II proposal indicated that it would be 'difficult in such a relatively small, relatively low-cost project to fully investigate M2M use of the pilot facility in an operational sense'. It therefore proposed to focus primarily on the use of the demonstrator service by end users and cover the M2M needs by 'examining the requirement for this on an ongoing basis at a mainly theoretical level'.

UKOLN (http://www.ukoln.ac.uk/) undertook the examination of the M2M (UKOLN, 2003) requirement and made the following 'concluding recommendations':

- Provide M2M demonstrator services based on controlled vocabularies mapped within Wordmap. Develop SOAP (http://www.w3.org/TR/soap12-part1/) based interfaces between JISC IE components and Wordmap APIs (Application Programmers Interfaces). Use these services in the short term as an aid to firm up use cases, in the longer term as a basis for pilot service if this approach is still appropriate at that stage.
- Carry out investigative implementation of Zthes (http://www.loc.gov/z3950/agency/profiles/zthes-04.html) based solution, whether data is exchanged using Z39.50 or OAI-PMH, with a view to taking advantage of standards based structured controlled vocabularies (particularly faceted vocabularies) as they become available from third party agencies.
- Track developments within the Semantic Web and eScience activities to ensure decisions made now concerning both syntax for structuring vocabularies and data exchange protocols take account of forward compatibility.

Aims and Objectives of the Feasibility Study

After some discussion, JISC agreed to fund a feasibility study that would design a pilot M2M service and assess the resources required to create it. Taking into account the possibility of a future Zthes-based solution using Z39.50 or OAI-PMH and syntax and data-exchange protocol implications of eScience and Semantic Web developments, the project would, it was agreed:

- Investigate the feasibility of developing SOAP-based interfaces between JISC IE components and Wordmap APIs or a non-Wordmap alternative based on storing terminology mappings in an SQL compliant database.
- Determine the scope and cost of the provision of an actual demonstrator based on each of these approaches.

The 'non-Wordmap alternative pilot' mentioned is a second version of the pilot built using SQL Server and Cold Fusion after the end of HILT Phase II to simplify some aspects of the way it operated. This pilot – available at http://hiltpilot.cdlr.strath.ac.uk/pilot/examples/ - is very similar to the first from the user's point of view.

After discussions with the main project partners, and with UKOLN, it was agreed that the primary concerns of the study should be an assessment of the feasibility, scope, and cost of a follow-up M2M pilot that considered the best options in respect of:

- Query protocols (SOAP, Z39.50, SRW, OAI) and associated data profiles (e.g. Zthes for Z39.50 and for SRW)
- Standards for structuring thesauri and thesauri-type information (e.g. the Zthes XML DTD and SRW version of it and SKOS-Core (Miles et al, 2005).

Overall Approach

The project took the HILT Phase II pilot demonstrator and the subject schemes mapped, or partially mapped within it (DDC, LCSH, UNESCO and MeSH), as its starting point and focused only on the requirements and feasibility of replacing its current direct user interface with an M2M interface. No significant new terminologies mapping work was to be undertaken, although it was agreed that the possibility of looking at a small number of additional MeSH mappings could be considered should this prove helpful in examining the requirements of BIOME (http://biome.ac.uk/) in the RDN (MeSH being one of the schemes in use at BIOME).

Research Plan

The following processes were specified and implemented:

- Construct a model of the whole BIOME to HILT transaction set for a simple interaction involving all steps from subject query to retrieval from remote service using an actual subject search example likely to arise within BIOME.
- Design a SOAP-based version of this model, identify syntax, data exchange protocol, and API interfacing issues that arise, and specify how the project will deal with them.
- Based on the resulting research, determine the feasibility of developing a SOAP-based version of the simple BIOME to HILT interaction. Conduct for both a Wordmap-based pilot and an SQL RDBMS-based pilot.
- Agree a representative set of use scenarios between BIOME and HILT, identify any new issues that arise from these, determine the feasibility of developing a SOAP-based version of BIOME to HILT interactions covering all use scenarios. Conduct for both a Wordmap-based pilot and an SQL RDBMS-based pilot.
- Examine the possible additional implications for delivering a SOAP-based version of BIOME to HILT interactions covering all use scenarios of a possible future need for a service also offering (1) a Zthes-based solution using SRW or OAI-PMH (2) a solution that takes account of the syntax and data-exchange protocol implications of eScience and semantic-web developments (see SWAD-Europe project). Determine whether changes to the design of the SOAP-based interface are required to ensure harmonisation with these possible future needs and whether such changes affect the feasibility of building a SOAP-based version of BIOME to HILT interactions covering all use scenarios. Conduct for both a Wordmap-based pilot and an SQL RDBMS-based pilot.
- Assuming that an adequate SOAP-based interface is feasible, using either Wordmap or an SQL system or both:
    - Agree the scope of a project for creating an operational BIOME to HILT M2M pilot based on the agreed use scenarios identified earlier.
    - Determine the cost of such a project using one or other or both Wordmap and SQL based solutions.
    - If both options are feasible, compare costs and benefits of each.
    - Make recommendations about a possible future project.

Study Outcomes: Overview

The primary outcomes of the feasibility study were:

1. A minimal functionality SOAP demonstrator, an outline description of which is given and illustrated below.
2. A report detailing five agreed 'use cases' that an M2M pilot should address, and assessing the suitability of the various protocols and mark-up approaches considered from handling the issues they raised. This is available in full as Appendix D of the final report of the study (HILT, 2005). Its conclusions are summarised below.

3. Outline proposals for an actual demonstrator based on the findings of the work. This is described in outline below. It was used to agree the final project design described in section 3 of this paper.

Study Outcomes: SOAP Demonstrator

EDINA (http://edina.ac.uk/), working with CDLR, built a simple SOAP demonstrator as part of the project to show that this was feasable. It is available at: http://nevis.ed.ac.uk:8080/asp-misc/public/hilt.asp.

An illustration of a HILT server response via the SOAP demonstrator is shown below. In this example shown, the term 'cakes' is input and a response is sent back in XML showing details of the appropriate DDC caption and (bottom of screen), a mapping of the Scots term 'bannock' from a terminology set used in the SPEIR project (see http://cdlr.strath.ac.uk/projects/speir.htm).

HiLT SOAP Demonstrator

Enter a HiLT search term

search for: cakes   submit

```
<?xml version="1.0" encoding="UTF-8"?>
<hilt:xml xmlns:hilt="http://hilt.cdlr.strath.ac.uk/">
 <inputterm>cakes</inputterm>

 <hilt:term>
  <hilt:termsource>DDC</hilt:termsource>    <hilt:termid>641.8653</hilt:termid>
  <hilt:termtext>Cakes</hilt:termtext>
  <hilt:tree>
   <hilt:branch>Technology</hilt:branch>
   <hilt:branch>Home &amp; family management</hilt:branch>
   <hilt:branch>Food and drink</hilt:branch>
   <hilt:branch>Cooking specific kinds of composite dishes</hilt:branch>
   <hilt:branch>Desserts</hilt:branch>
   <hilt:branch>Pastries</hilt:branch>
   <hilt:branch>Cakes</hilt:branch>
  </hilt:tree>
 </hilt:term>

 <hilt:term>
  <hilt:termsource>SPEIR</hilt:termsource>    <hilt:termid>641.8653</hilt:termid>
  <hilt:termtext>Bannock</hilt:termtext>
 </hilt:term>

</hilt:xml>
```

Note that this illustrates a simplified version of the real situation and is intended only to show the feasibility of the M2M interaction. Inputting 'cakes' as a term in Dewey for Windows returns five possible numbers with distinct captions: 641.815 Breads and bread-like foods; 641.8653 Cakes; 641.8659 Danish, French, related pastries; 664.7525 Pastries; 664.768 Formula feeds. The notes make it clear that some kinds of cakes occur at each of these places. This is a non-trivial problem and will introduce complications in the practical implementation, requiring human intervention. Also, the identical coding of each element as <hilt:branch> would not be appropriate in a full pilot which would have to encode the hierarchical relationships entailed.

Study Outcomes: Protocols and Mark-ups Summary

The key conclusions on protocols and mark-ups were as follows:

•   Since OAI-PMH is a protocol designed for harvesting metadata rather than searching, it does not look appropriate for the task of providing the services required of HILT by the five use cases. In HILT, subject terms input by users are used to search the mappings database for the records matching the term and the records in question are returned. Since there is no way of predicting what the user term will be in any given instance, the process must be dynamic. OAI-PMH is a metadata harvesting protocol, not a retrieval protocol and relies on using SET

membership (details at: http://www.openarchives.org/OAI/openarchivesprotocol.html#Set) to return records selectively in this way. This, however, is not a dynamic process. The harvester can only harvest pre-selected and labelled sets. This is not adequate for HILT. It relies on being able to 'select' records based on any term a user may come up with, so OAI-PMH will not work for HILT.

• SRW and Z39.50 both appear able to handle the issues that arise, although implementing a Z39.50-based M2M pilot service may involve greater complexity than would be entailed in implementing an SRW-based pilot service.

• On mark-up for returned classification, thesaurus, and mappings data, Zthes, SKOS-Core, and MARC all look adaptable to the task, although Zthes appears to be less suited to handling classification data than the other two. MARC has at least one advantage in that some major thesauri are available in that format (Vizine-Goetz et al, 2004). SKOS-Core is more flexible and more suited to the Web Services perspective and the Semantic Web community.

Study Outcomes: Outline Design for a Follow-up Project

The study concluded that there were two sensible options as regards a baseline follow-up M2M pilot project. The simplest would implement SRW, probably with SKOS-Core (but a case could be made for MARC and even ZThes). A more complex (and inevitably more expensive) version would seek to offer both SRW and Z39.50 services (perhaps through an SRW-Z39.50 gateway) and would offer a choice of Zthes, SKOS-Core, and MARC mark-ups. A sensible compromise would be to implement the simplest approach, but ensure that the pilot design provided for later developments encompassing the more complex version. This implied a follow-up pilot that would:

• Use the SRW protocol only, but be designed so that a possible extension offering other protocols such as Z39.50 could be introduced at a later date.

• Use SKOS-Core as the 'mark-up' for sending out terminology and classification set responses, but be designed so that adding other formats such as MARC and Zthes would be later option.

The project also suggested that a two-server distributed version of the pilot was worth considering, perhaps using SKOS-Core concept URIs as the basis for mapping between different schemes on the two servers. This seemed to offer the basis of an approach that might ultimately lead to a matrix of servers being available across the world with mappings between schemes being based on SKOS Core concept URIs and being built up slowly over a long period of time. Such an approach was theoretically attractive in that it might implement the kind of mapping-based solution HILT had envisaged to subject interoperability issues in a way that would spread the cost and effort over many organisations and a longer period of time. Such an approach would not be any cheaper than setting up the kind of service initially envisaged by HILT (shown to be expensive in HILT Phase II), but it would spread the cost over a number of players and the effort over a longer period of time. An alternative that opened up the same possibility was a one server pilot using SKOS-Core concept URIs to identify concepts uniquely, so that a distributed version of the service could be a later option.

In relation to software, it was decided that any follow-up project should use the SQL Server based pilot, but that the Wordmap software approach be kept in mind as a possible future option offering richer functionality in some areas.

3. The HILT M2M Feasibility Study

Overview

After additional consultations with JISC contacts, and further discussions amongst project partners, a bid for a follow up project was presented to JISC and subsequently funded. The project will run for 15 months from November 2005 and will create an M2M version of the HILT Phase II Pilot based on SRW and SKOS-Core, but with facilities extended to take account of the five use cases drawn up under the HILT M2M Feasibility Study. With JISC's agreement, two versions of the pilot were costed – a single server version and a distributed server version. In the event, budgetary levels for the period meant that only the single server version was funded.

Pilot Software

The terminologies pilot built in HILT Phase II was based on an adaptation of the Wordmap software. In the period between HILT Phase II and the feasibility study an opportunity arose to build a pilot offering similar services based on a more generic solution (SQL Server was used). This alternative pilot (http://nevis.ed.ac.uk:8080/asp-misc/public/hilt.asp) lacks the advanced terminology and mappings management facilities of Wordmap, but is purpose-built to provide the HILT pilot facilities themselves and is easier to work with on this front. In the Feasibility Study, this latter pilot was used to provide the simple SOAP demonstrator service put in place (http://hilt.cdlr.strath.ac.uk/hilt3/top.cfm) to establish the baseline feasibility of an M2M pilot. The proposal for HILT Phase III is to use the SQL- based approach to build the M2M pilot but to continue to liaise with Wordmap as their product develops with a view to ensuring that this development avenue is not closed off as a possible future solution for an operational service. As indicated in HILT Phase II, the continued use of Wordmap could be advantageous in the long-term if, as HILT Phase II concluded, a multi-user interface for maintaining mappings in a distributed fashion is likely to be required. Points against basing the M2M pilot on Wordmap are (1) this interface is not needed for the M2M pilot and (2) whereas it was necessary to adapt the Wordmap database structure to provide the HILT Phase II pilot, the SQL Server version, having been specifically designed for HILT, is easier to work with in the experience of the CDLR HILT team. It does, however, remain true that the Wordmap staff management interface is a possible longer term need.

Two SRW Clients

It is proposed that the project should develop two M2M clients: one for users of Go Geo! at EDINA, and one for the HILT interface itself. Since the purpose of offering a SOAP-based SRW service is specifically platform independence, it has been agreed that adopting a common approach and developing a single client would not be a sensible strategy – that, on the contrary, it would be better to develop (pilot versions of) two different clients. This will serve JISC and its users best by ensuring a more robust service that would be more likely to work with the various new clients that others in the community would need if they wanted to interface with the SOAP/SRW server. Accordingly, the proposal is that two clients be developed separately at Go Geo!, and HILT, but with help and advice from EDINA in the latter case. The needs of the RDN will be taken into account by continuing to liaise with the service as their planned new hardware and software platform is identified and installed.

A Distributed Future?

A distributed version of the service remains an attractive option and, as such, a small study considering the feasibility of using SKOS Core concept URIs to link between schemes on distributed servers will be undertaken as part of the project to examine the credibility of the approach.

References

Heery, R.; et al. (2001). Renardus Project Developments and the Wider Digital Library Context. // D-Lib Magazine. 7:4 (2001). URL: <http://www.dlib.org/dlib/april01/heery/04heery.html>.

HILT (2002). HILT Phase I Final Report. URL: <http://hilt.cdlr.strath.ac.uk/reports/finalreport.html>.

HILT (2003). HILT Phase II Final Report. URL: <http://hilt.cdlr.strath.ac.uk/hilt2web/finalreport.htm>.

HILT (2005), HILT M2M Feasibility Study II Final Report. URL: <http://hilt.cdlr.strath.ac.uk/hiltm2mfs/0HILTM2MFinalReportRepV3.1.doc>.

JISC (2005), JISC Information Environment Architecture. URL: <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/>.

Koch, T.; et al. (2001). Renardus: Cross-browsing European subject gateways via a common classification system (DDC). // Paper presented at IFLA satellite meeting: Subject Retrieval in a Networked Environment, OCLC, Dublin, Ohio, 14-16 August 2001.

McCulloch, E.; et al. (2004). Subject searching requirements: the HILT II experience. // Library Review. 53:8 (2004). 408-414.

McCulloch, E.; et al. (2005). Challenges and issues in terminology mapping: a digital library perspective. // Electronic Library. In press (2005).

MACS, Multilingual access to subjects. URL: <http://laborix.kub.nl/prj/macs/>.

Miles, A., J.; et al. (2005). SKOS-Core 1.0 Guide, An RDF Schema for Thesauri and Related Knowledge Organisation Systems. URL: <http://www.w3.org/2001/sw/Europe/reports/thes/1.0/guide>.

Nicholson, D. (2003). Subject-Based Interoperability: Issues from the High Level Thesaurus (HILT) Project. // International Cataloguing and Bibliographic Control. 32: 1. (2003).

Nicholson, D.; Shiri, A. (2003). Interoperability in subject searching and browsing. // OCLC Systems and Services. 19:2 (2003) 58-61.

Nicholson, D.; Wake, S. (2001). Interoperability in Subject Terminologies: The HILT Project. // New Review of Information Networking. 7 (2001).

Nicholson, D.; Wake, S. (2001). High-level thesaurus project: an overview. // Proceedings of the 2001 IFLA Satellite Conference on Subject Retrieval in a Networked Environment. Dublin, Ohio, August 14-16, 2001.

Nicholson, D.; et al. (2005). HILT: A Pilot Terminology Mapping Service with a DDC Spine. // Cataloguing and Indexing Quarterly. In press (2005).

Nicholson, D.; et al. (2002). HILT: moving towards interoperability in subject terminologies. // Journal of Internet Cataloguing. 5:4 (2002) 97-111.

Nicholson, D.; et al. (2001). High-level Thesaurus Project: Investigating the Problem of Subject Cross-searching and Browsing between Communities. // Global Digital Library Development in the New Millennium: Fertile Ground for Distributed Cross-Disciplinary Collaboration, edited by Chin-chich Chen. Beijing: Tsinghua University Press. 2001 219-226.

Saeed, H.; Chaudhury, A. S. (2002). Using Dewey Decimal Classification scheme (DDC) for building taxonomies for knowledge organisation. // Journal of Documentation. 58:5 (2002) 575-583.

Shiri, A.; et al. (2004). User evaluation of a pilot terminologies server for a distributed multi-scheme environment. // Online Information Review. 28:4 (2004) 273-283.

SWAD-Europe Project. URL: <http://www.w3.org/2001/sw/Europe/>.

UKOLN (2003). HILT Final Report Appendix J. URL: <http://www.ukoln.ac.uk/metadata/hilt/m2m-report/hilt-final-report.pdf>.

Vizine-Goetz, D.; et al. (2005). Vocabulary Mapping for Terminology Services. // Journal of Digital Information. 4:4 (2004) URL: <http://jodi.ecs.soton.ac.uk/Articles/v04/i04/Vizine-Goetz/>.

Glossary

API: Application Programmers Interface

BIOME: BIOME is a collection of gateways providing access to evaluated, quality Internet resources in the health and life sciences, aimed at students, researchers, academics and practitioners.

CQL: Common Query Language

DDC: Dewey Decimal Classification

DTD: Document Type Definition

EDINA: A JISC-funded national datacentre based at Edinburgh University Library, offering the UK tertiary education and research community networked access to a library of data, information and research resources.

e-Science: Research Councils UK (http://www.rcuk.ac.uk/escience/) describe e-Science in the following terms 'In the future, e-Science will refer to the large scale science that will increasingly be carried out through distributed global collaborations enabled by the Internet. Typically, a feature of such collaborative scientific enterprises is that they will require access to very large data collections, very large scale computing resources and high performance visualisation back to the individual user scientists'.

FE: Further Education

Go Geo!: A tool designed to help users find details about geo-spatial datasets and related resources within Great Britain tertiary education and beyond. A trial service provided by EDINA.

HE: Higher Education

HILT: HIgh Level Thesaurus

IESR: JISC Information Environment Service Registry

JISC: Joint Information Systems Committee

JISC IE: Joint Information Systems Committee Information Environment

LCSH: Library of Congress Subject Headings

M2M: Machine to machine interaction

MeSH: Medical Subject Headings

OAI-PMH: The Open Archives Initiative Protocol for Metadata Harvesting

OCLC: Online Computer Library Center

RDBMS: Relational Database Management System

RDN: Resource Discovery Network

Semantic Web: A collaborative initiative led by the W3C, the Semantic Web provides a common framework that facilitates data sharing and reuse across application, enterprise, and community boundaries.

SKOS-Core: SKOS Core supports the RDF description of language-oriented knowledge organisation systems (KOS) such as thesauri, glossaries, controlled vocabularies, taxonomies and classification schemes.

SOAP: Simple Object Access Protocol

SQL: Structured Query Language

SRU: Search & Retrieve URL – Z39.50 Next Generation

SRW: Search/Retrieve Web Service – Z39.50 Next Generation

UKOLN: A centre of expertise in digital information management, providing advice and services to the library, information, education and cultural heritage communities. Based at the University of Bath and formerly known as the UK Office for Library & Information Networking.

UNESCO Thesaurus: United Nations Educational, Scientific and Cultural Organization subject scheme.

Use Case: A Use Case represents a series of interactions between a user (human or machine) and the system, utilising (in the present case) an M2M link. Typically, the interaction starts with an enquiry and leads to a resource that should answer that enquiry.

Wordmap: A commercially available taxonomy management software application that supports management of multiple controlled vocabularies.

XML: Extensible Mark-up Language

Z39.50: An international standard specifying a client/server-based protocol for searching and retrieving information from remote databases.

Zthes: The Zthes profile is an abstract model for representing and searching thesauri and specifies how this model may be implemented using the Z39.50 and SRW protocols.

Appendix A: HILT Pilot Screenshots

Figure 1. Homepage of the HILT Pilot Terminologies Service

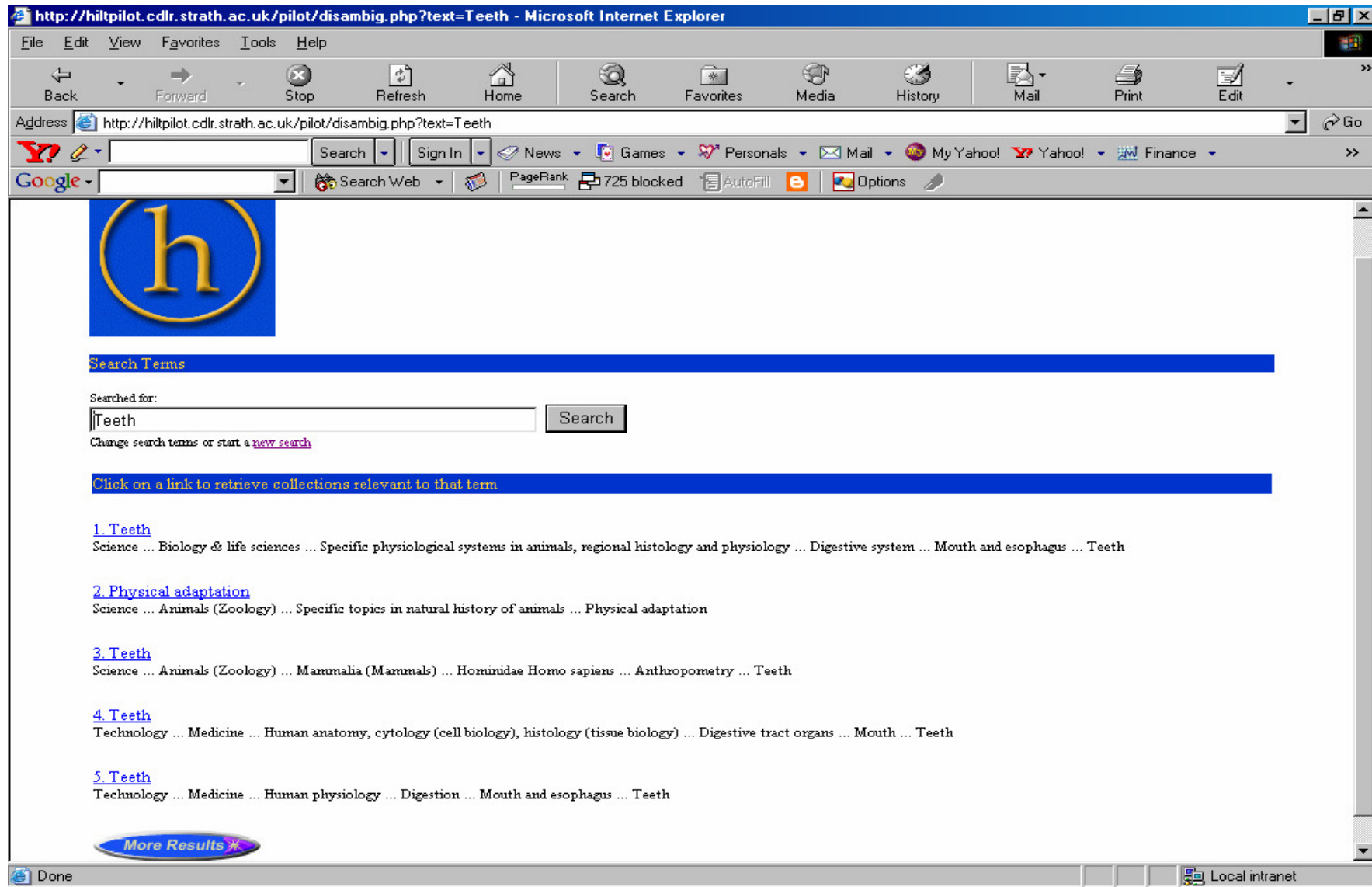Figure 2. Disambiguation page of the HILT Pilot Terminologies Service

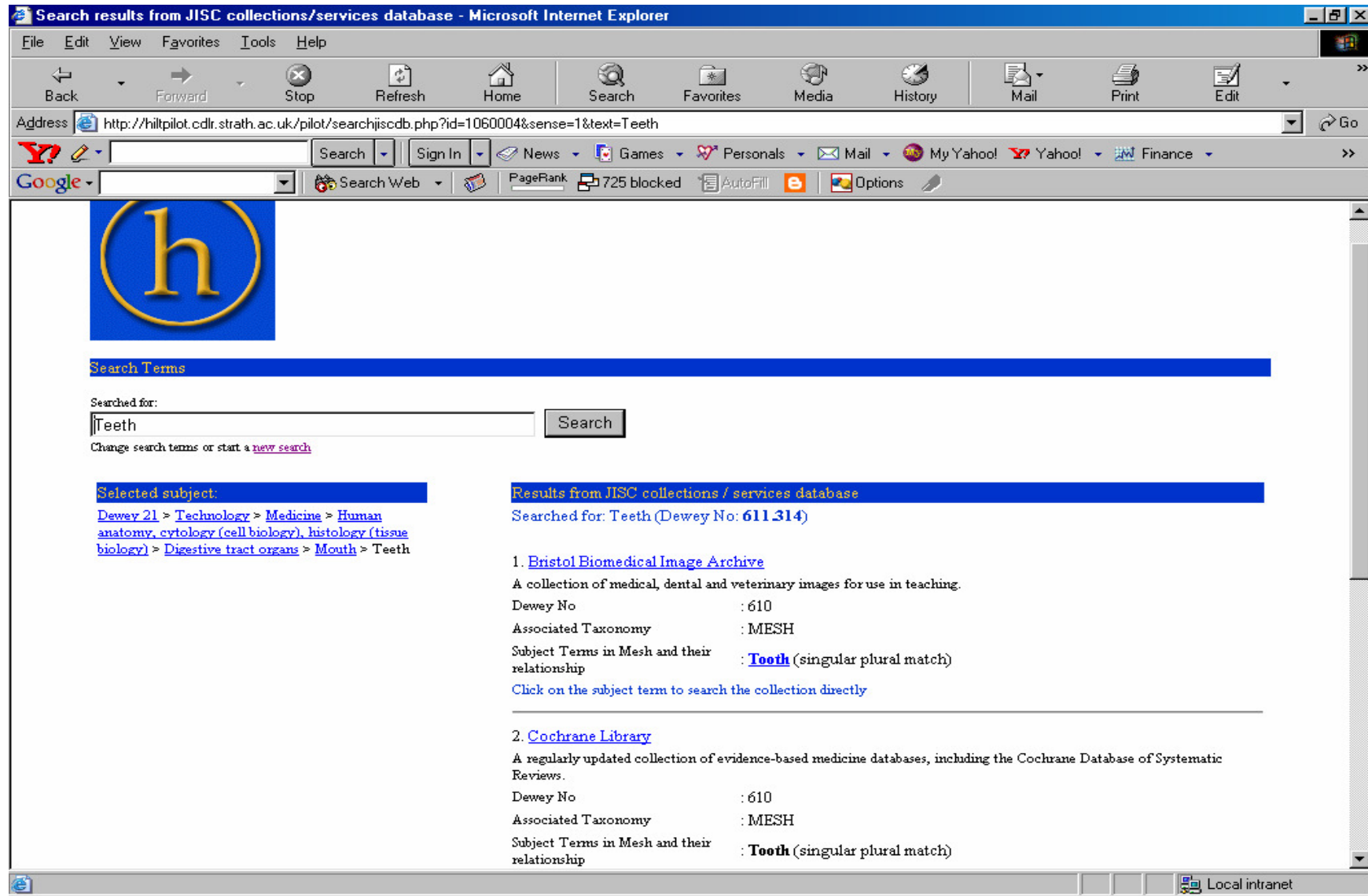Figure 3. Collection selection page of the HILT Pilot Terminologies Service

Figure 4. JISC collection found by the search term "Teeth"