HILT IV: Subject Interoperability through Building and Embedding Pilot Terminology Web Services

Dennis Nicholson, Anu Joseph, and Emma McCulloch, **affiliation**

World Digital Libraries 2(1): 1–22

Abstract

This paper presents an expert evaluation of a collaborative annotation system named DiLAS (Digital Library Annotation Service). It can be globally accessed by individuals as well as different user communities, and knowledge is created and shared within English-speaking communities. It contains a collection of textual documents on information science and software science, and it gives access to all kinds of related material such as authors' home pages, photos, articles, and so on. An analytical evaluation was conducted as a participatory group evaluation, which involved presentation beyond the objectives and rationale papers and development of the prototype. The empirical evaluation of DiLAS consisted of two experiments. The first evaluation was a bottom up evaluation that began at the elementary level with an evaluation of the usability of the interface. A Cognitive Walkthrough approach was chosen using a qualitative approach. The next evaluation moved towards the broader work context with a user- and work-centred evaluation involving an entire collaborative task situation, which required knowledge sharing on a common real life work task. This evaluation resulted in a large number of user requirements for the next prototype.

Introduction and overview

This paper reports on work carried out within Phase IV of the JISC-funded HILT[1] project, considers its context against the background of other recent and ongoing terminologies work, describes its outcome and conclusions, and considers future research and development directions, some of which are envisaged as being collaborative. As with HILT Phase III (Nicholson and McCulloch, 2006 a; 2006b), the project utilized the expertise of participants at CDLR[2] and EDINA,[3] and of the HILT terminology advisors,[4] together with some ongoing liaison with MIMAS[5] who run IESR[6] and Intute.[7] It also had ongoing support from OCLC[8] in respect of allowing the use of the electronic files of Dewey Decimal Classification (DDC[9]) and of LCSH[10] mappings to DDC[11] and of a range of other participants on the HILT Steering group.[12] The project is working towards setting up a JISC Shared Infrastructure Service in respect of terminologies and interoperability. Shared Infrastructure Services are an important building block of the JISC Information Environment (IE) architecture.[13] They constitute a common set of infrastructural services, that provide this information through machine-to-machine (M2M) interfaces, on which other elements of the environment, such as portals, brokers and aggregators, can draw. In the case of HILT, the focus is on the provision of M2M access to terminology and terminology interoperability data for other services to interact with and use.

Ensuring that HE (Higher Education) and FE (Further Education) users of the JISC IE can find appropriate learning, research and information resources by subject search and browse in an environment where most national and institutional service providers - usually for very good 'local' reasons - use different subject schemes to describe their resources is a major challenge facing the JISC domain (and, indeed, other domains beyond JISC). Encouraging the use of standard terminologies in some services (institutional repositories, for example) is a related challenge. Under the auspices of the HILT project, JISC has been investigating mechanisms to assist the community with this problem and thereby optimise the value obtained from expenditure on content and services by facilitating resource sharing to benefit users in the learning and research communities. The project has been through a number of phases, with work from earlier phases reported, both in published work elsewhere,[14] and in project reports (see the project website). The focus of this paper is on Phase IV which began in May 2007 and is due to complete a few months from time of writing at the end of May 2009.

HILT Phase IV had two elements - the core project, whose focus was 'to research, investigate and develop pilot solutions for problems pertaining to cross-searching multi-subject scheme information environments, as

[1] http://hilt.cdlr.strath.ac.uk/
[2] Centre for Digital Library Research (CDLR): http://cdlr.strath.ac.uk/
[3] EDINA: http://edina.ac.uk/
[4] Alan Gilchrist and Leonard Will
[5] Manchester Information & Associated Services (MIMAS): http://www.mimas.ac.uk/
[6] Information Environment Services Registry (IESR): http://iesr.ac.uk/
[7] intute: http://www.intute.ac.uk/
[8] OCLC Online Computer Library Center: http://www.oclc.org/
[9] Dewey Decimal Classification (DDC): http://www.oclc.org/dewey/
[10] Library of Congress Subject Headings (LCSH): http://authorities.loc.gov/
[11] LCSH to DDC mappings: http://www.oclc.org/asiapacific/zhcn/dewey/updates/numbers/default.htm
[12] http://hilt.cdlr.strath.ac.uk/hilt4/groups.html
[13] http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/
[14] See http://hilt.cdlr.strath.ac.uk/dissemination.html for a list key HILT publications

well as providing a variety of other terminological searching aids', and a short extension to encompass the pilot embedding of routines to interact with HILT M2M services in the user interfaces of various information services serving the JISC community. Only the outcomes of the core HILT IV work are reported here. These are summarised in section 3, following a short section on related work that sets HILT in the context of wider work in the area. Section 4 describes the current position in the project in respect of available working facilities and project perspectives, adding detail to the section 3 summary outcomes, and section 5 covers the aims of the embedding extension project and addresses future research and development directions beyond the end of HILT IV. Section 6 provides a brief conclusion to the paper.

Related work

There is a good deal of ongoing and recent work in the general area of terminologies, terminological interoperability in cross-searching and other scenarios, and terminology and terminology interoperability services (Day, Koch, and Neuroth 2004, Landry 2004, Tudhope, Koch, and Heery 2006, Zeng, and Chan2006, van Gendt, *et al.* 2006, Vizine-Goetz, Diane, Andrew Houghton, and Eric Childress, *et al.* 2006, Macgregor and McCulloch 2006, Agosti, *et al.* 2007, Mayr and Petras 2007, Binding, Tudhope, May, 2008, Geser 2008, Levergood, *et al.* 2008). Much of it is of relevance, or potential relevance, to HILT – because HILT's intended role as a JISC shared infrastructure service aiming to support the work of end-user focused information services facing a myriad of different problem situations and their various facets means its aim has largely been to address the interoperability problem in general, rather than some specific aspect of it since its primary focus is serving researchers, teachers, and learners in universities, and since this community as a whole is likely to need to find and access subject information located almost anywhere in the world and to do so using almost any of the available knowledge organisation systems (KOS[15]), and, potentially, any language, it is impossible to envisage a future service being a stand-alone enterprise. The need for HILT, and for any similar general service, will be to provide, not just direct access to the KOS and KOS interoperability data it provides, but also to offer services using it facilities to permit discovery of, and integrated interaction with, other sources of such data, as well as means of intelligently and transparently handling this M2M data in user interfaces to best meet the needs of individual user groups tackling particular types of task.

This means the project has an interest in a wide variety of work in the field. The work of significant or recent projects in the area (CACAO,[16] MACS,[17] LIMBER, RENARDUS,[18] KoMoHe,[19] and STAR[20] are examples) is important – not just because the issues they face are similar, but because, in future, comprehensiveness, and hence the need to interoperate with initiatives of this kind (Nicholson 2008), will be an issue, and their work helps highlight some of the requirements of that. So the work on different approaches to providing cross-works or cross-walk services (Zeng and Chan 2004), whether it be on differing approaches to intellectual

[15] Taxonomies, classification schemes, thesauri, subject heading lists and similar – see Zeng and Chan, 2004 for a full description
[16] http://www.cacaoproject.eu/
[17] MACS Project: https://macs.vub.ac.be/pub/
[18] RENARDUS Project : http://renardus.sub.uni-goettingen.de/
[19] KoMoHe Project http://www.gesis.org/en/research/information_technology/komohe.htm
[20] STAR Project http://hypermedia.research.glam.ac.uk/kos/star/

mapping, on machine processing or automation based approaches (see for example Doerr 2001, Chan and Zeng 2002, Koch and Vizine-Goetz 1998, Godby et al. 1999, Ardo 2004), or on the advantages and disadvantages of various different strategies or approaches. For historical (Nicholson and McCulloch 2006a), and to some extent, resourcing reasons, HILT's approach has been based on intellectual mapping, but this does not mean it is seen as the only viable approach in every circumstance. For obvious reasons, work on the advantages of cross-walk work (Mayr and Petras 2008), and on problems encountered (Liang and Sini 2006, McCulloch and Macgregor 2008) is of direct benefit, whether for current work or future planning. Multi-lingual efforts in the area (Cacao 2008, Clavel-Merrin 2004, Landry 2004, Agosti et al. 2007) are significant because various groups in the academic community that is the project's focus have research and teaching and learning interests that cross language boundaries. Since there is an assumption that a future service will have to integrate other (perhaps unknown) terminology services into its operations for the benefit of those it serves, and provide mechanisms to aid the discovery of simple information services covering particular subject areas, work on terminologies and services registries is of interest (Apps 2007, Hillman *et al.* 2006, Golub and Tudhope 2008). And since the aim is to aid information services in the provision of enhanced (and transparent) subject retrieval and browse facilities for their users and staff via M2M interactions with HILT, and the subsequent intelligent and user-adaptive handling within their user interfaces of returned terminological data encompassing different KOS and KOS cross-walk approaches with differing characteristics, work on user interfaces (Shiri 2008, Stafford et al. 2008, Shiri et al 2004), work on user behaviour and what motivates it (Blair 1980, Brown 1995, Buckland et al. 1999, McCray et al. 1999, McCulloch et al. 2004, Teevan et al. 2004), and similar, and KOS typologies (Tudhope 2006) is also relevant (for example, when designing the embryonic toolkit described below).

It would not be difficult to add to this list of related work, but it is unnecessary here. A flavour of the complexity of possible relationships is all that is required to provide context and a feel for the landscape of potential interconnections. A detailed investigation of the relevance and implications of the work referenced above for future services of the kind envisaged by HILT would be relevant and illuminating, but is neither appropriate nor feasible in the present paper which is concerned primarily with the work of HILT IV itself. What is perhaps worth mentioning is that the one thing missing from the plethora of ongoing work in the area from the HILT perspective (and, arguably, that of similar enterprises such as the OCLC terminology services suite[21]) is co-ordination of R&D work within an agreed common architecture or architecture set, an issue covered further in section 4.

HILT IV: Project Outcomes

Prior to its extension to encompass a short project to embed the pilot use of HILT M2M services in a variety of service interfaces, HILT IV was scheduled to complete in February 2009. By March 2009, therefore this core work was complete. The project had

- Built and tested a range pilot M2M (SRU/W,[22] SOAP,[23] and SKOS[24]) based web services to deliver terminologies and

[21] http://www.oclc.org/research/projects/termservices/resources/termservices-overview.pdf
[22] Search/Retrieve Web Service (SRU/W): http://www.loc.gov/standards/sru/
[23] SOAP: http://www.w3.org/TR/soap/
[24] Simple Knowledge Organization System (SKOS) Core: http://www.w3.org/2004/02/skos/

terminology mappings to JISC and institutional information services, supporting the transparent enhancement of subject search facilities.
- Built and tested a database of terminologies (DDC,[25] UNESCO,[26] HASSET,[27] IPSV,[28] JITA,[29] MeSH,[30] CAB, GCMD,[31] NMR,[32] AAT,[33] SCAS, JACS,[34]), high level mappings to DDC (HASSET, IPSV, UNESCO, JACS), and limited sample 'deeper' mappings to MeSH and HASSET.[35]
- Built and tested an embedding toolkit to offer information services the core software needed to begin building subject interoperability services for their users by interacting at M2M level with the above web services and database through routines embedded transparently in their service user interfaces - a programmer's toolkit to help build improved subject browse and retrieve facilities.
- Conducted various practical experiments to successfully embed terminology service interaction into JISC community services to create operational pilot subject browse, retrieve, and deposit enhancements for service users (a smaller piece of work preceding the embedding extension project mentioned above).
- Developed a generic terminology services architecture that will (over a period of time) permit the HILT services to grow and improve by incorporating terminology services being developed elsewhere.
- Determined that a terminology services registry is a key part of this architecture and that the core functionality required to build and run such a registry is already inherent in HILT pilot services.
- Developed staff skill sets and experience associated with the problems of subject searching and subject interoperability within and across information services using different subject terminologies, with the best approaches to mapping new schemes into the database, and with an associated distributed architecture to permit the ready integration of new services into the JISC and global subject interoperability landscape.
- Determined in conjunction with JISC that the best general option for a sustainable operational Shared Infrastructure Service based on HILT project outcomes is a Terminologies Interoperability Centre offering a mix of standard 'plug and play' type M2M and toolkit facilities free at the point of use, including a training portal and an associated terminology services registry, more flexible, charged-for, specially-scoped versions of these, tailored to the needs of individual services and institutions, and ongoing development via a mix of collaboration and externally funded R&D, as well as through JISC support.

In consequence, it had developed to the point where it was in a position to tackle the more advanced embedding work required by the extension project (described in outline in

[25] Dewey Decimal Classification (DDC): http://www.oclc.org/dewey/
[26] UNESCO Thesaurus: http://www2.ulcc.ac.uk/unesco/
[27] Humanities and Social Science Electronic Thesaurus (HASSET): http://www.data-archive.ac.uk/search/hassetSearch.asp
[28] Integrated Public Sector Vocabulary (IPSV): http://www.esd.org.uk/standards/ipsv/
[29] http://eprints.rclis.org/jita/
[30] Medical Subject Headings (MeSH): http://www.nlm.nih.gov/mesh/
[31] Global Change Master Directory (GCMD): http://gcmd.nasa.gov/Resources/valids/keyword_list.html
[32] National Monuments Record Thesauri (NMR): http://thesaurus.english-heritage.org.uk/
[33] Art & Architecture Thesaurus (AAT): http://www.getty.edu/research/conducting_research/vocabularies/aat/
[34] Joint Academic Coding System (JACS): http://www.ucas.ac.uk/figures/ucasdata/subject/
[35] Also included are LCSH mappings included with the DDC file provided by OCLC.

section 5), and put forward to JISC proposals for future development towards the set up of an operational shared infrastructure service. It had in place a core of tested pilot facilities and terminology sets and an embryonic mappings set that could be the basis of a future service, an understanding of the terminological, technical, architectural, and staffing requirements for future development, and a clear if outline vision of what was required to move towards a useful, robust, and sustainable service aligned with JISC's strategic aims.[36]

HILT IV: working facilities and project perspectives—the current position

The current position in the project is best described under four headings: Service functions, database, and embedding toolkit; The HILT Architecture and its Implications; Project understanding of service and end-user needs; and The Requirements of a robust, useful, and sustainable service.

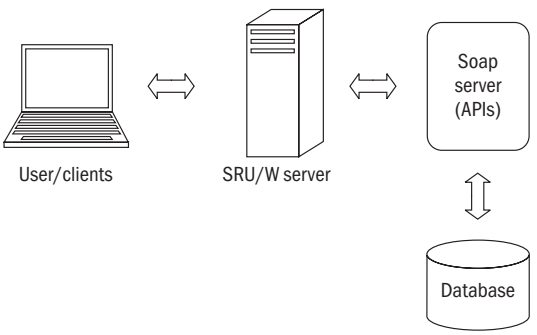
Service functions, database, and embedding toolkit

A full description of these and how these functions in practical circumstances require greater detail than is possible here and is provided in McCulloch, 2008. The present paper deals only with the technical details of the working facilities and with possible ways in which they can be used by information services through M2M interaction with HILT.

Technical Summary

Although the following description omits some of the detail, it gives a useful overview of the technical basis of the HILT services.

The system is designed in a modular fashion and the diagram below shows how each of the different modules interact. This



design facilitates the replacement or modification of the components if a need arises in the future. SQL Server 2005 is a highly scalable, programmable, secure database management system which is used for data storage. The SOAP server enables a distributed web based environment and provides the APIs to interact with the database. A SRW/U server is implemented to support M2M interaction and the APIs can be accessed through the SRW/U protocol.

Further information is provided as follows.

SRU/W server The SRU/W server facilitates M2M interaction between clients and the server. Index Data's SimpleServer (http://www.indexdata.dk/simpleserver/) is used to build the SRU/W server. This is a Perl module intended to develop Z39.50, SRU and SRW servers over any type of database. It is based on the popular YAZ toolkit and is robust, efficient, portable, and interoperable with all Z39.50, SRU and SRW clients. The server involves a SOAP envelop to wrap and transfer data. This implementation allowed us to have a distributed environment.

SOAP server HILT uses a SOAP server to interact with the database and wrap the results in SKOS. REST is a recent alternative to SOAP; the project evaluated both options and decided on SOAP because it is more stable. This implementation is based on simple and easy-to-use NuSOAP (a group of PHP classes)

[36] http://www.jisc.ac.uk/aboutus/strategy/strategy0709/strategy_aims.aspx

which helps to create and consume SOAP web services based on SOAP 1.1, WSDL 1.1 and HTTP 1.0/1.1. It does not require any special PHP extensions. By using SOAP, it was possible to encapsulate database access through a few functions. It is also possible to replace SOAP with REST at a future date if the need arises. In this case, HILT functions would require to be implemented in REST.

Database HILT data is stored using SQL Server 2005 database software and contain 12 Vocabularies[37] (AAT, CAB, GCMD, HASSET, IPSV, JACS, JITA, LCSH, MeSH, NMR, SCAS and UNESCO) separate from DDC. Tables are designed in such a way to accommodate future versions of vocabularies and flat tables are also designed to improve performance. Interaction with the database is through stored procedures, which adds a layer of security and better performance. A full-text index facilitates the search and retrieval of data efficiently and SQL Server's ability to rank result records (based on the frequency of search terms), is useful in multi term queries. Database access is supported through the SOAP server.

APIs As part of the HILT toolkit we have various APIs to offer, which are accessible using the SRW/U server. Results are wrapped in SKOS (Simple Knowledge Organisation System).

The APIs are:

1 Get_DDC_records
Returns DDC captions and numbers related to a subject term. The user can then choose the most appropriate to his/her interest. This function searches within DDC and any schemes that are mapped to DDC for an input term and returns only DDC instances.

2 Get_collections
Returns collections classified under a specified DDC number or its stem, including subject scheme used. A process of truncation is used within this function in order to maximise retrieval potential and avoid a 'no hits' situation. For example, if a user searches for collections with DDC number 336.26, HILT returns collections matching 336.26, 336.2, 336, 330 and 300.

3 Get_non_DDC_records
Returns terms from schemes other than DDC by matching user terms to DDC notations, before identifying mappings to those particular numbers. This function also employs truncation to retrieve relevant mappigs. For an input 336.36, HILT returns mappings corresponding to 336.36, 336.3, 336, 330 and 300. Its up to the client to choose the mappings of their interest.

4 Get_all_records
This function combines the output of get_DDC_records and get_non_DDC_records.

5 Get_filtered_set
Get_filtered_set enables a user to search a particular scheme or schemes directly for a specific term, together with its broader, narrower, related and non-preferred terms, if selected and where applicable. You can also search for the id of a term in a scheme (if known), to build a hierarchy of that scheme where available.

6 Get_SP_suggestions
HILT Spell Checker can suggest a list of words similar to the input word. This implementation is based on David Spencer's code using the n-gram method and the Levenshtein distance. An index (the dictionary) with all the possible words from the HILT database (a lucene index) is created and suggestions are based on this index. The index can be extended by including terms from other sources as well. A single term query returns multiple records whereas a multi term query returns a single record.

[37] Plus the DDC spine

7 Get_wordnet_suggestions
Returns the description about the input term from a large lexical database of English - WordNet. Implementation is based on JwordNet - Java interface to WordNet.

Toolkit A toolkit has been designed to illustrate how these various APIs can be embedded within a service and provide intending implementers with a basic 'starter' package. Further details are available at: http://hilt4.cdlr.strath.ac.uk/toolkit/intro.cgi. A Perl version is available to download and a PHP version may be available at a later date.

Possible Uses

As indicated earlier, a description of how the functions and the toolkit may be used in practice is provided in McCulloch, 2008. At present, these are pilot facilities, but the aim is to move gradually towards operational services as described in the appropriate section below. Operational facilities will enable national, institutional, and other information service providers to access and use terminological and interoperability data to enhance their own services in a variety of ways, including, but not necessarily limited to, the following.

a. Improving recall in a subject search of one or more databases by enriching the set of terms known to a user by providing synonyms and related terms.
b. Providing the best terms for a subject search in a remote service that uses a subject scheme unfamiliar to 'home service' users (or in a cross-search of a group of such services).
c. Taking a user's subject term and using it to identify relevant information services via registries such as IESR (http://iesr.ac.uk/).
d. Generating an interactive browse structure where a scheme is arranged hierarchically.
e. Offering the ability to send a term from a chosen subject scheme and receive back data on broader terms, narrower terms, hierarchy information, preferred and non-preferred terms, and so on.
f. Providing cataloguing staff with information on subject schemes and inter-scheme mappings to assist in metadata creation.
g. Providing a spell-check mechanism to assist user searching.
h. Providing a service to assist user search formulation by providing information on search terms entered (for example, what the term means, alternative meaning, synonyms of the word so that it might be useful in a search, and so on).

Uses of this kind can, of course, already be tested using the pilot facilities currently in place. Some work in this area has already taken place through two small elements of embedding work undertaken within the core HILT IV project; further work is already underway within the embedding extension project.

The HILT Architecture and its Implications

A key element of the current state of play with the project is a developing understanding of the architecture required to support a rich and user adaptive set of relevant terminology and subject interoperability services, and the implications of this architecture for a distributed and collaborative approach to future service growth and associated research and development.

Arising out of JISC-funded work with UK user communities in the first phase of HILT, the project's original focus as regards solving subject interoperability problems was on mapping between subject schemes via a DDC spine. Mapping is an established and effective approach (for example, Mayr and Petras 2008) and has attracted significant resourcing in

other European countries (for example Agosti et al. 2007; Mayr and Petras 2008). However, HILT now has an architectural approach that allows it (a) to adopt a range of interoperability strategies appropriate to specific use cases and cost-benefit levels (for example, expensive deep mapping where the problem and user area justifies this, or high-level or browse-based retrieval where significant costs would be less justifiable), (b) to incorporate in the model - for the benefit of the JISC user communities - solutions offered and funded by other players, whether they be based on mapping to a spine, scheme to scheme, or some variety of automated approach.

The original assumption underpinning HILT was that it would be, in essence, a stand-alone service that would facilitate subject interoperability in the JISC Information Environment by mapping commonly used schemes to a DDC spine and providing the best terms to use in services using a particular subject scheme via these mappings. All through the progress of the project, it was very clear (a) that, even within the JISC communities this was likely to be a large and probably expensive task (b) that mapping to a DDC spine and, indeed, mapping in itself was only one of many possible approaches to the problem (c) that the variety of subject interoperability problems within JISC would probably be better tackled using a mix of these methods. However, the likely additional costs entailed in encompassing a variety of approaches in a HILT service meant that only the one approach - the aforesaid mapping via DDC spine - could be encompassed in a project with limited resources.

Fortunately, the move towards HILT offering M2M functionality has changed the picture somewhat. In the kind of web services environment now envisaged, it becomes possible to think in terms of a distributed and devolved approach to subject interoperability described in Appendix A (Nicholson 2008, for a more detailed description) - with information services, both within JISC and elsewhere, utilising M2M connections, not just to HILT, but to terminology services worldwide as they come on stream. These would be funded by a range of different players, and would utilise all kinds of different approaches to interoperability. Information services would find and interact with them via data obtained through infrastructural services such as IESR or dedicated terminology services registries. Other data from the same sources would allow the information services to elegantly handle the different kinds of data and different approaches to interoperability served up by these different terminology services.

With this as the backdrop, it becomes necessary to look at the question of research required for subject interoperability service development in the context of the distributed and devolved architecture it implies. The downside of this is that a research for development agenda that is more complex and wider in scope than might otherwise be the case (since it must handle not just interactions with one service of known functionality and scope, but (a) a large variety of different services of unknown functionality and scope, and (b) intermediate interactions with infrastructural services such as registries). The upside is that the distributed and devolved environment means (a) that JISC need only focus on those aspects of research for development needed by specific JISC communities at specific times (b) that other players in the field will likely seek to tackle elements of the research for development agenda required in their own domains, and (c) it is probable that some of the issues faced in the JISC communities will also be faced (and, in some cases, tackled) elsewhere.

This still emerging perspective on the likely technical and networked environment within

which a future service will grow will impact on upcoming considerations relating to the shape and form of a sustainable service and on probable research and development paths pursued.

Project understanding of service and end-user needs

At present, HILT has only a limited understanding of the needs of the services likely to use HILT facilities and an even more limited understanding of the needs of the end users of such information services. Work on embedding HILT facilities transparently into the user interfaces (or clones of those interfaces) of three services has been undertaken - the GoGeo! Service, which used HILT terminological data to enrich user search terms sets when searching distributed external services external to GoGeo!; the Intute service, which used HILT facilities in spell-check and in offering users a drop down list of terms related to the term input (e.g. motherhood and mother and child relationships as alternatives to mothers); and the CAIRNS distributed catalogue service which established a proof-of-concept demonstrator to show how HILT could be used to improve recall. These were small scale exercises, but nevertheless provided useful insights into the kinds of issues likely to be encountered as HILT aimed to offer its services operational to a range of disparate services with different aims and different terminological issues (it is safe to say that almost every information service operates in a unique subject terminologies environment making 'plug and play' solutions to all but a few core problems difficult to provide). They also helped inform the need for the embedding extension project and, to some extent, its shape and form.

HILT IV also carried out some limited work with a number of individuals who joined a HILT email list (HILT-Collaborators@jiscmail.ac.uk) in response to a call for service staff who were potentially interested in using HILT pilot facilities. This work showed that there was a good deal of interest from the community in the kind of uses of HILT listed above. Twenty-nine people joined the list which was set up specifically to look at embedding work of this kind. A questionnaire asking what kinds of uses were of interest was answered by sixteen people. All indicated an interest in at least one of the above potential ways of using HILT, most indicated an interest in three or more ways. However, the work also showed that in many cases - via some practical tests of actual training and attempts to do simple embedding work - it would be difficult for the early stages of embedding work using HILT facilities to be supported remotely and that more work was needed to discover the best way of ensuring that embedding work at service sites was facilitated and well supported by HILT.

This work notwithstanding, however, it is safe to say that further in-depth work with both staff of services interested in using HILT's M2M services to support subject searching and interoperability functions within their service interfaces, and, more particularly, with the end-users of such information services is needed to inform the process of transforming pilot facilities into a robust, useful, and sustainable service. Some of this work is scheduled to take place within the embedding extension to HILT IV, but a good deal of it will have to be encompassed within a later scoping study for a future service, and in a projected subsequent 'soft launch' of an actual service.

The Requirements of a robust, useful, and sustainable service

As previously noted, one of the outcomes of HILT Phase IV was the conclusion that the best general option for a sustainable operational Shared Infrastructure Service

based on HILT project outcomes was a Terminologies Interoperability Centre. More precisely, it was determined - based on an understanding of issues in the area of subject interoperability as developed over a number of phases of HILT, and a discussion with JISC itself on how to best fit future developments into JISC's strategic requirements and also provide a useful, reliable, and sustainable service for JISC-related communities - that the best route forward was to work towards the 'soft launch' of a Terminologies Interoperability Centre. Once launched, this would offer the community a mix of free and charged-for services and would be supported by a mix of JISC funding, externally earned income and R&D funding, and collaboration, and would provide:

- M2M and user-level access to terminology sets, the detail of those terminology sets, and data to facilitate interoperability between them.
- Open source software toolkits that would enable M2M interaction with HILT web services to be transparently embedded in the user interfaces of local, national and project information services.
- A basic architecture for terminology and interoperability services in the JISC Information Environment (and potentially beyond).
- A way of mounting and developing new terminologies and terminologies interoperability data required by the community, including JISC-specified work to facilitate improvements in subject access in and between the various JISC user communities and their external partners based on ongoing assessments of user and service needs.
- Advisory and M2M support services for projects, services, and other initiatives in JISC or JISC institutions where there is a subject description, subject retrieval, or subject interoperability facet.
- A JISC funded free advisory and training service on using the above facilities in local or national services and projects in the percentage of cases where this was relatively straightforward (plug 'n' play, but after a bit of advice and training).
- The development and hosting of a JISC-focused terminologies services registry.
- A charged-for consultancy service where the work and advice required by local and national services, projects, and organisations (both within and out of JISC) was less straightforward or more sophisticated (because of unique client and client service circumstances and terminology sets).
- A portal for tools and training in the areas described above.
- A focus for wider work in the terminologies area, funded through a variety of sources, including non-JISC sources (for example though successful bids for European funding).
- Ongoing work to facilitate JISC involvement and leadership in a strategically important area with significance for both subject-based retrieval needs in research, learning, teaching, and elsewhere and semantic web developments.

The initial plan in respect of this was to move directly from HILT Phase IV to the proposed soft launch. However, initial consideration of this proposal quickly proved that an interim stage - a scoping study for the proposed Centre and its services – was a necessary and sensible first step. If funded, this scoping study will provide a well-researched evidence base that will inform and guide a future 'soft launch' of a Terminologies Interoperability Centre by

- Putting in place service quality infrastructure to support the work of the Centre, including further development and testing of the components from HILT IV

and work on a pilot terminology services registry. This will ensure that the standard services offered at the soft launch will be robust and usable in a range of JISC service and user environments.
- Determining service user and end user needs via iterative feedback from hands-on experience, utilising outcomes in TIC scoping and soft launch plans, creating mechanisms for an ongoing assessment of such needs, and identifying specific players to work with TIC during the soft launch period.
- Scoping in detail what free and charged-for services the Centre should offer and what they would cost.
- Producing a bid for TIC start-up costs, a programme of works, and a well-researched Sustainability Plan.

The proposal is that this study would take place between June 2009 and November 2010, with an actual soft launch of the Centre taking place in December 2010.

The Embedding Extension Project and Future Research and Development Directions

This, then, is the position with the project at the end of the core element of Phase IV. The next steps for HILT are, first of all, to conduct the more complex embedding work planned under the embedding extension project, then - assuming a positive outcome to the proposal for a Terminologies Interoperability Centre scoping study - to begin to determine the best shape and form for a future scalable and sustainable service. Following that - and regardless of the outcome of the scoping study proposal - the logical path to follow is one that will lead towards increasing international collaborative work based on an agreed architecture and common research and development programme.

Futures: The Embedding Extension Project

As indicated earlier, a small extension project was funded towards the end of HILT Phase IV with the aim of embedding the pilot use of HILT M2M terminology and interoperability facilities within the user interfaces of a number of information services in order to enhance functionality for the end-users of these services. The services in question are based at EDINA, Mimas (Intute), and CDLR. A second extension project, funded in parallel at IESR in Manchester, also entailed HILT services embedding plans, and collaboration between the two in respect of HILT embedding is taking place. The project runs from January to May 2009. It is underway at time of writing, but has yet to produce significant results in respect of embedding as such (although some associated mapping work between DDC and JACS and DDC and RAE headings has been completed).

The focus of the combined CDLR and IESR work as presently envisaged is as follows.

IESR

The aim will be to use HILT M2M services to enhance the user browse interface to the IESR collection discovery service in a variety of ways by sending user terms to HILT, and enrich the interface by getting back broader, narrower, and related terms associated with various HILT schemes and using them to improve retrieval.

Intute

Intute is currently being redesigned, with a target date of end of July. This means embedding demos would not be sustained much past the end of April. Accordingly, it was suggested that Intute work on a clone of their service to show what HILT can offer - the likely focus being enhancing user searches in various ways (for example, by offering more specific terms when a user puts in a very broad one (like 'farming').

Edina

Edina will aim to improve their Depot[38] service by using HILT to guide user-provided subject tags via an M2M fed drop down list of standard terms (for example, from JACS). They will also look to improve their VSM service by aiming to provide users with term expansion - taking a user input search term and expanding it via HILT to include terms from multiple schemes.

CDLR

The intention here is to enhance the recently developed Google Maps based Scotland's Information[39] service by using HILT to translate existing subject terms into the same subjects in different languages. Languages being considered are those where HILT has some access to DDC captions in the language concerned and can therefore cross-map - German, French, and Welsh-Gaelic.

In addition to the above, there will also be some generic work on optimizing the HILT spell check pilot mechanism for use in UK HE an FE.

Futures: The TIC Scoping Study

The aim of the Scoping Study is to provide a well-researched evidence base that will inform and guide a future 'soft launch' of a Terminologies Interoperability Centre by scoping out the detail of a sustainable mix of JISC-funded and charged for services designed to help meet JISC strategic aims in respect of serving its stakeholders. The basic outline of what is required in respect of addressing subject interoperability issues is known. What is needed now is a clear, evidence-based, indication of what is required in detail to ensure a robust, sustainable service sensitive to community needs and the strategic aims of the JISC. Amongst other things, this requires intensive work in which specific JISC-related user communities (including both service and end users) are given hands on experience of the terminological tools as they become available as part of an iterative and ongoing approach to understanding, and implementing solutions to, user needs in specific communities.

The intention is to work with four target groups:

- HE and FE libraries (subject librarians, cataloguers/repository managers, systems librarians/repository developers, end-users of various kinds, national service providers.
- E-learning (teachers, learners, librarians, VLE managers, VLE developers, national service providers.
- NHS e-library (and their services)
- A cross-domain group entailing Museums, Archives, Libraries, and so on - not just in HE (for example, public libraries as well as HE libraries and work with the Strategic Content Alliance)

Working with these groups, the project will move gradually and interactively towards a reliable understanding of user requirements via the following threads:

- Four one-day workshops at four UK sites with hands-on work, focus groups, and other activities designed to engage and inform, and set up an ongoing iterative process to allow improved understanding of needs - each to have a mix of the target groups, and be shared with IESR.
- Workshop demonstrations of terminology service embedding possibilities and uses and discussions on approaches, problems and issues aimed at finding out a more about real needs.
- A purpose built social networking site would be set up and populated prior to the workshops and tools showcase added to it - creating a 'virtual lab' for ongoing interaction with user groups.

[38] http://depot.edina.ac.uk/
[39] http://www.scotlandsinformation.com/

- During and after the workshops the environment would be used to show possibilities and get feedback aimed at pinpointing work needed in a soft launch; we would also use it during the soft launch.
- Discussions with these groups would be mainly on their own experiences with subject searching and terminology issues but would also go beyond terminology issues to discuss issues like the availability of local expertise, and how we might best pitch (a) an 'off the shelf' (and 'free') service (b) how we might cost a paid for consultancy (something that would also be discussed with consultants in the field). There would also be discussions on protocols, mark-up languages (SKOS, MARC, Zthes), and specific mapping and tools functionality needs.
- Work with the groups during and after the workshops to show them illustrative functionality and terminological improvements based on ongoing interaction.
- Work on the detail of how the chosen groups (or individual services or organisations/institutions from these groups) might be involved in a direct practical sense in a soft launch and what level of funding would be required by them for their involvement.

Futures: Towards a Common Collaborative Research and Development Programme

As indicated in above under The HILT Architecture and its Implications, HILT sees a need, both within the JISC community, and in the world at large, for a globally-scoped programme of collaborative research and development based on a common view of an inclusive architecture for subject interoperability service design. Although HILT progress in this area is currently limited, efforts have nevertheless been made to begin work towards agreeing a collaborative approach with other 'players' in the terminologies field. A paper on the idea was presented at an ontologies conference in Helsinki in November 2007, a paper on the topic published in the international version of the Signum journal in November 2008, and, more recently (December 2008), steps have been taken to contact major European projects in the terminologies area to begin the process of talking about collaboration and about applying for FP7 funding to carry the work forward.

Discussions are at a very early stage, but if a proposal were put forward, it would be based on the idea that, despite the plethora of different means of tackling inter-KOS and inter-lingual interoperability currently in play and the size and complexity of the problem itself, it is both desirable and possible to facilitate gradual cumulative progress towards optimizing subject interoperability across the networked world through geographically distributed co-ordinated collaborative effort - that it can be done by agreeing a model set of requirements for interoperability service design and collectively pursuing a common research and development agenda based on it.

A key aim would be to bring together major (but currently stand-alone) European projects working in the area, including HILT, MACS, KoMoHe, CACAO and others. If funded, the project would aim to kick-start and give strategic direction to the major and long term programme of work required to research and develop (1) the creation of a pan-European (and ultimately global) web-services based M2M architecture to support the level of inter-KOS and inter-lingual subject interoperability required to support such things as seamless access to distributed European digital resources for all EU citizens and a richer inter-initiative semantic web; (2) the wide and complex range of user interface handling routines that individual European

information services would require to embed in their interfaces to enable them utilize distributed terminology and terminology cross-walk services elegantly and transparently for the benefit of those of their users seeking to search external information services; (3)The ontology of the key elements of the domain and their inter-relationships that would be required to drive these user interface routines - uniquely categorising such things as KOS types, KOS cross-walk types (e.g.based on intellectual mapping via different spines, direct scheme to scheme mappings, automated statistical mappings, clustering, autoclassification), services characteristics, user characteristics and other elements implied by the architecture shown below; (4) an ongoing infrastructure to support future R&D efforts over the long term; and (5) mechanisms to support person-specific, user group-specific, and task specific customization of system responses to service requests.

It is by no means certain that such a proposal will materialise in the event. However, a number of key players have at least indicated a clear interest in the idea.

Conclusion

The Phase IV project has taken HILT to the point where the launch of an operational support service in the area of subject interoperability is a feasible option and where both investigation of specific needs in this area and practical collaborative work are sensible and feasible next steps. Moving forward requires detailed work, not only on terminology interoperability and associated service delivery issues, but also on service and end user needs and engagement, service sustainablity issues, and the practicalities of interworking with other terminology services and projects in UK, Europe, and global contexts.

Acknowledgement

Acknowledgements are due to George Macgregor who contributed to HILT IV during the early part of its duration, to James Farnhill and Tom Franklin for providing solid and helpful JISC support, Kathleen Menzies for help with references, OCLC for ongoing support and advice, and to JISC for funding HILT.

References

Agosti, et al. 2007
Roadmap for MultiLingual Information Access in the European Library
In *Research and Advanced Technology for Digital Libraries*
Berlin/Heidelberg: Springer
Volume 4675/2007. Abstract online at: http://www.springerlink.com/content/g8126155w7518333/

Apps A. 2007
Using an Application Profile Based Service Registry
In *Proceedings of the International Conference on Dublin Core and Metadata Applications* (*DC 2007: Application Profiles: Theory and Practice*), edited by S A Sutton, S C Chaudry, C Khoo
Singapore, 27–31 August 2007, pp 63–73. Details available at <http://epub.mimas.ac.uk/papers/2007/dc2007/apps-dc2007.html>

Ardo A. 2004
Automatic Subject Classification and Topic Specific Search Engines – Research at KnowLib
Presented at DELOS Regional Awareness Event: Between Knowledge Organization and Semantic Web: Semantic Approaches in Digital Libraries
Lund, Sweden. Details available online at: <http://www.delos.info/eventlist/LUB1/Anders_Ardo/DELOS.PDF>

Binding C, Tudhope D, May K. 2008
Semantic Interoperability in Archaeological Datasets: Data Mapping and Extraction via the CIDOC CRM
Proceedings (ECDL 2008) 12th European Conference on Research and Advanced Technology for Digital Libraries Aarhus, 280–290
Lecture Notes in Computer Science, 5173
Berlin: Springer. final preprint presentation

Blair D C. 1980
Searching biases in large interactive document retrieval systems
Journal of the American Society for Information Science, **31**(4), 217–277

Brown M E. 1995
By any other name: accounting for failure in the naming of subject categories
Library and Information Science Research. **17**(4), pp. 325–429

Buckland M K, Chen A., Gebbie M, Kim Y and Norgard B. 2000
Variation by Subdomain in Indexes to Knowledge Organization Systems
Presented at International Society for Knowledge Organization, ISKO6, *Toronto, 2000*
Details available at <http://metadata.sims.berkeley.edu/papers/iskopaper.html>

Cacao. 2008
Cross-language Access to Catalogues and Online Libraries partners
CACAO Presentation. Details available at <http://www.cacaoproject.eu/fileadmin/media/presentations/CACAO_presentation.pdf>

Chan L, Zeng M. 2002
Ensuring interoperability among subject vocabularies and knowledge organisation schemes: a methodological analysis. 68th IFLA Council and General Conference, 18–24 August 2002
Glasgow: Details available at <http://www.ifla.org/IV/ifla68/papers/008-122e.pdf>

Clavel-Merrin G. 2004
MACS (Multilingual Access to Subjects): A Virtual Authority File Across Languages
In *Cataloguing and Classification Quarterly*, **39**(1/2)
Berlin: Springer

Day M, Koch T, and Neuroth H. 2004
Searching and browsing multiple subject gateways in the Renardus Service
In *Recent developments and applications in social science methodology: proceedings of the Sixth International Conference on Logic and Methodology* edited by C van Dijkum, J Blasius, Kleijer H, and B Hilten van B. 17–20 August 2004
Amsterdam*:* Amsterdam: SISWO Instituut voor Maatschappijwetenschappen. (CD-ROM). Details available at <http://www.ukoln.ac.uk/metadata/publications/rc33-2004/renardus-paper.pdf>

Doerr M. 2001
Semantic problems of thesaurus mapping
Journal of Digital information, **1**(8). Details available at <http://jodi.ecs.soton.ac.uk/Articles/v01/i08/Doerr/>

Geser G. 2008
STERNA Technology Watch Report. Full Report
Ref: STERNA Del.6.5**,** 10 December 2008. Salzburg Research
Details available at <http://www.sterna-net.eu/index.php/en/downloads>

Godby C J, Miller E J and Reighart R R. 1999
Automatically Generated Topic Maps of World Wide Web Resources
Annual Review of OCLC Research, 1999. Details available at <http://digitalarchive.oclc.org/da/ViewObject.jsp?fileid=0000002655:000000059193&reqid=9300>

Golub K, Tudhope D. 2008
Terminology Registry Scoping Study (TRSS): Excerpt on Metadata
Details available at <http://www.ukoln.ac.uk/projects/trss/dissemination/metadata.pdf>

Hillman D, Stuart A S, Phipps J, Ryan L. 2006
A Metadata Registry from Vocabularies UP: The NSDL Registry Project, submitted to Dublin Core 2006 Conference.
Details available at <http://www.citebase.org/abstract?id=oai:arXiv.org:cs/0605111>

Koch T and Vizine-Goetz D. 1998
Automatic Classification and Content Navigation Support for Web Services: DESIRE II Cooperates with OCLC
Annual Review of OCLC Research, 1998. Details available a <http://www.lub.lu.se/tk/demos/class-ws/automatic.htm>

Landry P. 2004
Multilingual Subject Access: The Linking Approach of MACS
In *Cataloging & Classification Quarterly*, **37**(3/4)

Levergood B, *et al.* 2008
The Specification of the Language of the Field and Interoperability: Cross-Language Access to Catalogues and Online Libraries (CACAO)
International Conference on Dublin Core and Metadata Applications
Berlin. Details available at: <http://dcpapers.dublincore.org/ojs/pubs/article/viewFile/933/929>

Liang A C and Sini M. 2006
Mapping AGROVOC and the Chinese Agricultural Thesaurus: Definitions, tools, procedures
New Review of Hypermedia and Multimedia, **12**(1), pp. 51–62

Macgregor G and McCulloch E. 2006
Collaborative tagging as a knowledge organisation and resource discovery tool
Library Review, **55**(5), pp. 291–300. Pre-print available at <http://strathprints.strath.ac.uk/2335/1/strathprints002335.pdf>

Mayr P and Petras V. 2007
Building a terminology network for search: the KoMoHe project
In *Proceedings of the International Conference on Dublin Core and Metadata Applications 2008*. Details available at <http://arxiv.org/pdf/0808.0518>

Mayr P and Petras V. 2008
Cross-concordances: terminology mapping and its effectiveness for information retrieval
IFLA 2008 Conference Proceedings. Details available at <http://eprints.rclis.org/13828/>

McCulloch E, Shiri A, and Nicholson D. 2004
Subject searching requirements: the HILT II experience
Library Review, **53**(8). pp. 408–414

McCulloch E. 2008
Developing a pilot toolkit to improve subject interoperability between collections and services within the JISC Information Environment (ref needs finishing)

McCulloch E and Macgregor G. 2008
Analysis of equivalence mapping for terminology services
Journal of Information Science, **34**(1) pp.70–92

McCray A T, Loane R F, Browne A C, Bangalore A K. 1999
Terminology Issues in User Access to Web-based Medical Information
Details available at <http://www.amia.org/pubs/symposia/D005626.PDF>

Nicholson D and McCulloch E. 2006a
Investigating the feasibility of a distributed, mapping-based, approach to solving subject interoperability problems in a multi-scheme, cross-service, retrieval environment
International Conference on Digital Libraries, 5–8 December 2006
New Delhi: Details available at <http://strathprints.strath.ac.uk/2875/1/strathprints002875.pdf>

Nicholson D and McCulloch E. 2006b
HILT Phase III: Design requirements of an SRW-compliant Terminologies Mapping Pilot, 5th European Networked Knowledge Organization Systems (NKOS) Workshop, 10th ECDL Conference, 21 September 2006
Alicante, Spain

Nicholson D, 2008
A Common Research and Development Agenda for Subject Interoperability Services?
Signum, **5**

Shiri A, Nicholson D, and McCulloch E. 2004
User Evaluation of a Pilot Terminologies Server for a Distributed Multi-scheme Environment
Online Information Review, **28**(4), 273–283

Shiri A. 2008
Metadata-enhanced Visual Interfaces to Digital Libraries
Journal of Information Science, **34**(6), 763–775

Stafford A, Shiri A, Ruecker S, Bouchard M, Mehta P, Anvik K, Rossello X. 2008
Making Sense of Interface Design: Implementing Searchling, a Visual Thesaurus-Enhanced Interface for Multilingual Digital Libraries.
Chicago Colloquium on Digital Humanities and Computer Science, 1–3 November 2008
Chicago: University of Chicago

Teevan J, Alvarado C, Ackerman M S, Karger D R. 2004
The Perfect Search Engine is not Enough: A Study of Orienteering Behaviour in Directed Search
In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (CHI '04)

Tudhope D. 2006
A tentative typology of KOS: towards a KOS of KOS?
5th European NKOS Workshop, European Conference on Digital Libraries (ECDL06)
Alicante. Details available at: <http://www.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2006/papers/paper-018.doc>

Tudhope D, Koch T, Heery R. 2006
Terminology Services and Technology JISC state of the art review
Details available at: <http://www.jisc.ac.uk/media/documents/programmes/capital/terminology_services_and_technology_review_sep_06.pdf>

van Gendt M, *et al.* 2006
Semantic Web Techniques for Multiple Views on Heterogeneous Collections: a Case Study
Proceedings of the 10th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2006), edited by Julio Gonzalo, Constantino Thanos, M. Felisa Verdejo and Rafael C Carrasco
LNCS, **4172,** pp. 426–437, Alicante: Springer Verlag. Details available at: http://www.cs.vu.nl/~mtvgendt/STITCH-ECDL06.pdf

Vizine-Goetz, Diane, Andrew Houghton, and Eric Childress. 2006
Web Services for Controlled Vocabularies
Bulletin of the American Society for Information Science and Technology, **35**(5). Details available at <http://www.asis.org/Bulletin/Jun-06/vizine-goetz_houghton_childress.html>

Zeng Marcia Lei and Lois Mai Chan. 2004
Trends and issues in establishing interoperability among knowledge organization systems
Journal of American Society for Information Science and Technology **55**(5): pp. 377–395

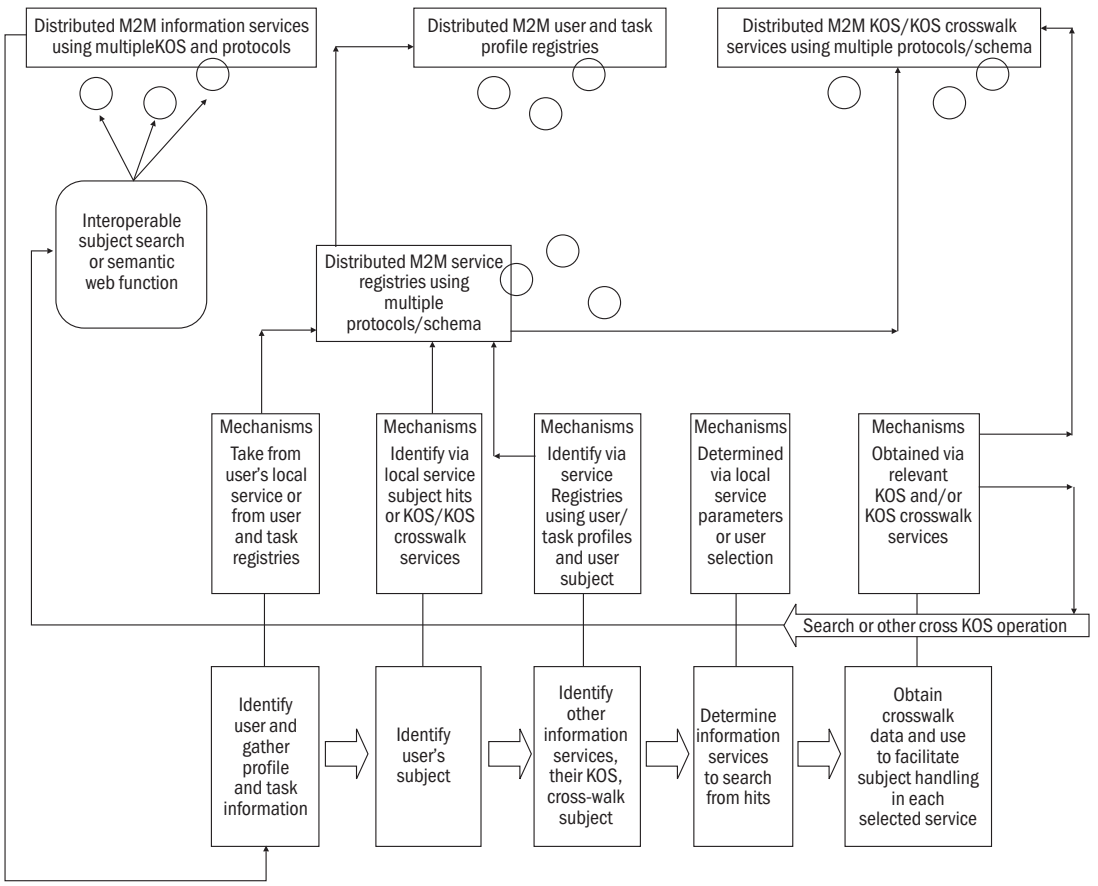
Zeng M L and Chan L M. 2006.
Metadata Interoperability and Standardization – A Study of Methodology Part II: Achieving Interoperability at the Record and Repository Levels
In *D-Lib Magazine*, **12**(6). Details available at <http://www.dlib.org/dlib/june06/zeng/06zeng.html>

Appendix A: the HILT architecture

The diagram below shows the HILT architecture in outline.

In a working model, the following are true.

- There are multiple service registries available that 'know' about each other, so any given information service only needs to know about their own 'home' service registry
- There are multiple information services available, recorded in and discoverable via one or more service registries, and classified according to their subject coverage
- There are multiple user and task profile registries available, also recorded in and discoverable via one or more service registries
- There are multiple KOS and KOS 'crosswalk' services available, also recorded in and discoverable via one or more service registries (or possibly terminology registries)
- There are many information services available using many different KOS and access protocols
- The user's start point is usually (but not necessarily) one of these information services
- The users home information service offers the user a facility to cross search other information services appropriate to her subject needs and uses a subject interoperability service (encompassing the whole of the diagram above) to underpin this service.
- Most of the workings of this service are transparent to the user
- The assumption is that the user starts at one of the information services and needs to search one or more (possibly unknown) others
- The home service goes through the processess shown in the five boxes along the bottom of the diagram, using the services registries, user and task profile registries and KOS and KOS crosswalk services shown in the upper half of the diagram as necessary



- In order to elegantly and intelligently (and transparently) handle often complex interactions as described above, a local information service would need access to complex metadata and sophisticated embedded user interface routines to handle the interactions

Appendix B: Glossary

AAT	Art and Architecture Thesaurus
API	Application Programming Interface
CAB	Commonwealth Agricultural Bureaux
CAIRNS	Co-operative Information Retrieval Network for Scotland
CDLR	Centre for Digital Library Research
DDC	Dewey Decimal Classification
Depot	A UK national open access repository for researchers not yet having an institutional repository in which to deposit their papers, articles, and book chapters (e-prints)

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.
FE	Further Education
GCMD	Global Change Master Directory
Go Geo!	A tool designed to help users find details about geo-spatial datasets and related resources within the UK tertiary education sector and beyond. A trial service is provided by EDINA.
HASSET	Humanities and Social Science Electronic Thesaurus
HE	Higher Education
HILT Project	High-level Thesaurus Project
IESR	JISC Information Environment Service Registry
intute	intute is a free online service providing access to the very best web resources for education and research. Formerly the Resource Discovery Network (RDN).
IPSV	Integrated Public Sector Vocabulary
JACS	Joint Academic Coding System
JISC	Joint Information Systems Committee
JISC IE	Joint Information Systems Committee Information Environment
LCSH	Library of Congress Subject Headings
M2M	Machine to machine interaction
MARC	Machine readable cataloguing
MeSH	Medical Subject Headings
NHS	National Health Service
NMR	National Monuments Records Thesauri
OCLC	Online Computer Library Center
RAE	Research Assessment Exercise
REST	Representational State Transfer
RIN	Research Information Network
SCAS	Standard Classification of Academic Subjects
SKOS	Simple Knowledge Organization System. SKOS Core supports the Resource Description Framework (RDF) description of language-oriented knowledge organisation systems (KOS), such as thesauri, glossaries, controlled vocabularies, taxonomies and classification schemes
SOAP	Originally the Simple Object Access Protocol, but now more simply referred to as SOAP. Used to exchange XML-based messages over computer networks, normally using HTTP
SQL	Structured Query Language
SRU/W	Search/Retrieve Web Service and Search & Retrieve URL - 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
VLE	Virtual Learning Environment
WSDL	Web Services Description Language
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