Terminology Web Services†

Ceri Binding* and Douglas Tudhope**

Hypermedia Research Unit, University of Glamorgan, Pontypridd, CF37 1DL, UK,
*<cbinding@glam.ac.uk>, ** <dstudhope@glam.ac.uk>

Ceri Binding is a Research Associate in the Hypermedia Research Unit, Faculty of Advanced Technology, University of Glamorgan. Ceri graduated with a BSc in Computer Studies in 1997 whilst working as an Analyst Designer / Programmer for Hyder IT, before joining Glamorgan in 2000. He had responsibility for development work on the FACET project and implemented various standalone and web systems for the project. He is currently conducting research and development work for the STAR project, involving use of SKOS and CRM data. Related research interests include Knowledge Organisation Systems, intelligent web-based retrieval and interface design.

Douglas Tudhope is Professor in the Faculty of Advanced Technology, University of Glamorgan and leads the Hypermedia Research Unit. His area of research is knowledge organization systems and services. He was principal investigator on the FACET and STAR projects. Since 1997, he has been editor of *New Review of Hypermedia and Multimedia*. He is a member of the Networked Knowledge Organisation Systems/Services (NKOS) network. He co-authored the 2006 *JISC State of the Art Review on Terminology Services and Technology* and the *JISC Terminology Registry Scoping Study*.

† This research was supported by the Arts and Humanities Research Council (AHRC). The authors would like to thank Keith May, Phil Carlisle and other staff from English Heritage for their assistance, as well as Andy Priest, Janine Rigby and Caroline Williams, from MIMAS.

Binding, Ceri and Tudhope, Douglas. **Terminology Web Services.** *Knowledge Organization, 37(4)*, 287-298. 16 references.

ABSTRACT: Controlled terminologies such as classification schemes, name authorities, and thesauri have long been the domain of the library and information science community. Although historically there have been initiatives towards library style classification of web resources, there remain significant problems with searching and quality judgement of online content. Terminology services can play a key role in opening up access to these valuable resources. By exposing controlled terminologies via a web service, organisations maintain data integrity and version control, whilst motivating external users to design innovative ways to present and utilise their data. We introduce terminology web services and review work in the area. We describe the approaches taken in establishing application programming interfaces (API) and discuss the comparative benefits of a dedicated terminology web service versus general purpose programming languages. We discuss experiences at Glamorgan in creating terminology web services and associated client interface components, in particular for the archaeology domain in the STAR (Semantic Technologies for Archaeological Resources) Project. We go on to consider the case for more specialised terminology services for different kinds of controlled vocabulary.

1.0 Introduction

Conventional web search involves users manually resolving any ambiguity post search, by choosing relevant documents from a sea of textual matches. Users eventually learn to use term co-occurrence coupled with unusual or less ambiguous terms. Term suggestion tends to be based on transient popularity metrics. Keyword search and manual disambiguation of a vast and diverse range of resources is still disappointing. Certain search engine features originate from library science and historically there have been initiatives towards categorisation of online resources, but there remains a chasm between library content and online content.

Controlled vocabularies are frequently cited as beneficial resources in this area, providing a useful mediating interface for search operations. Controlled vocabularies consist of terms considered useful for retrieval purposes, which are used to represent concepts. This vocabulary can be used by Knowledge Organization Systems (KOS), which structure their concepts via various forms of semantic relationships. Exposing access in the form of terminology services enables programmatic integration of these useful resources into other applications.

2.0 What are terminology services?

A JISC (Joint Information Systems Committee, UK) review of terminology services and technology (Tudhope, Koch, and Heery 2006, 7) describes terminology services as:

> a set of services that present and apply vocabularies, both controlled and uncontrolled, including their member terms, concepts and relationships.… They can be applied as immediate elements of the end-user interface (e.g. pick lists, browsers or navigation menus, search options) or can underpin services behind the scenes.

We are referring in this paper specifically to terminology web services-distributed data service functionality, opening up programmatic access to controlled terminologies for other organisations to base applications on. The services ideally expose open, freely accessible data.

Web Services generally have been applied for some time in a variety of applications and with different underlying bindings. Gardner (2001) gives an introduction in a digital library context. Terminology web services are a more recent development although we can trace one line of descent to earlier work on protocols for programmatic access to networked (distributed) KOS, see for example, Davies (1996). In 1998, the second NKOS workshop had as one of its themes a "functional model of the process of using a KOS over a network." Johnson (2004) outlined a theoretical proposed network of thesaurus access and navigation services. Binding and Tudhope (2004) detailed some early approaches at defining coherent service protocols, notably the CERES (California Environmental Resources Evaluation System), Zthes, and the ADL (Alexandria Digital Library) thesaurus protocols.

Simple Knowledge Organization Systems (SKOS) is about developing specifications and standards to support the use of knowledge organization systems (KOS) within the framework of the Semantic Web. SKOS allows Knowledge Organization Systems to be represented in the Resource Description Framework (RDF) for purposes of interoperability. SKOS is an effort by the W3C Semantic Web Deployment Working Group (SWDWG). In an earlier project leading up to this effort, the Semantic Web Advanced Development (SWAD) Europe project defined the SKOS API and implemented the DREFT (Demo of RDF Thesaurus) server demo.

In common with other APIs, terminology services offer developers the major advantage of not implementing all functionality from scratch. Basic programmatic patterns can be invoked by calling on already existing program libraries. If the patterns correspond to commonly agreed or widely applicable use cases then development proceeds faster by building on previous work. There are a number of advantages of terminology services over other forms of distribution. The terminology provider can maintain version control and the user automatically always has access to the most up to date version of their work. Services are platform/location agnostic; the calling application does not have to be implemented using the same programming language and operating system as used for the service. Furthermore, service providers do not have to be the KOS creators/owners but may offer services based on KOS developed elsewhere. One possible downside is that applications become reliant on constant network availability (assuming the service is located externally) and external server infrastructure, but in general the positives appear to outweigh the negatives.

2.1 Users and uses of terminology services

End users might wish for some ready made "widgets" to slot into their systems, so service users may be systems developers looking to incorporate vocabulary data into their own applications. They may be cataloguers seeking to annotate their repository content with established terminology (see, for example, Vizine Goetz et al. 2006), or web searchers wishing to improve search performance via various forms of vocabulary based query expansion (Binding and Tudhope, 2004).

Terminology services could find usage in a number of complementary areas. Improved search facilities involving term suggestion are already being implemented within commercial search interfaces (e.g., Google Suggest, Flickr). Tag suggestion systems are used to improve search engine rankings by manipulation of metadata indexing for competitive advantage (deriving popular synonyms describing core competencies for an organisation). In the digital library area, suggestion systems can be used to catalogue/index/annotate repository content with controlled vocabulary terms.

Social tagging systems could also benefit from alignment with established common indexing terminology (Golub et al. 2009). The growth of social bookmarking sites indicates a desire for the personal organisation and structuring of web resources. Social tagging produces some interesting results, but also produces ambiguous vocabularies mixing index terms with opinions. Intuitive tools incorporating established controlled terminologies in fields other than libraries remain sparse, yet there are clearly potential gains in facilitating their use in this area.

3.0 Existing terminology services

We review a selection of terminology web services to illustrate some interesting contemporary projects and the breadth of applications in this area (this is not intended as an exhaustive list). Some general definitions are given first.

XML (eXtensible Markup Language) is a standard markup language for Web documents. RDF (Resource Description Framework) is a standard conceptual modelling language for the Semantic Web, based on subject-predicate-object triples. SOAP (Simple Object Access Protocol) is a protocol specification for exchanging structured information using Web services, while REST (Representational State Transfer) is a lighter weight HTTP protocol. JSON (JavaScript Object Notation) is a lightweight computer data interchange format used for serializing and transmitting structured data over a network connection. SRU (Search/Retrieval via URL) is a REST based protocol for Internet search queries. SparQL (Simple Protocol and RDF Query Language) is a standard RDF query language. The concept of Linked Data forms part of the vision of a 'web of data'; content is made available in RDF, addressed via virtual but persistent URIs that allow HTTP clients to "negotiate" their preferred representation of the content.

i. The German National Library of Economics (ZBW) has published an experimental REST (Representational State Transfer) web service interface to the STW Thesaurus for Economics. The service offers both XML and JSON output formats.
ii. OCLC have produced a set of services accessible via the SRU (Search/Retrieval via URL) query language CQL. Concept details can be retrieved in a variety of formats – HTML, MARC XML, SKOS, and Zthes, from a number of controlled vocabulary resources.
iii. The CATCH (Continuous Access To Cultural Heritage, NL) programme, in the context of the STITCH (Semantic Interoperability to Access Cultural Heritage) and TELplus projects, has developed a SKOS-based Vocabulary and Alignment service prototype. The core of the service is SOAP-based, with a REST-like access layer, returning RDF/SKOS data and JSON output for concepts.
iv. The European Environment Information and Observation Network (EIONET) GEMET thesaurus has a REST interface, derived from the SKOS API definition.
v. The Library of Congress Authorities and Vocabularies service is a groundbreaking demonstrator of a REST Linked Data service exposing LCSH SKOS data.
vi. The HILT (High Level Thesaurus) Phase IV project has produced a SRU/W (Search and Retrieve Web) Service operating against a number of common vocabulary resources.
vii. The Food and Agriculture Organisation of the United Nations (FAO), under their Agriculture Information Management Standards (AIMS) initiative, have produced the Agrovoc Concept Server with a set of terminology web services.
viii. The Getty Vocabularies Web Services offer retrieval and update of Getty vocabularies to licensees of the vocabularies in real time.
ix. The American National Biological Information Infrastructure (NBII) Biocomplexity Thesaurus is exposed as a terminology web service based on SKOS API.
x. The Finnish Semantic Computing Research Group (SeCo) have implemented ONKI SKOS – a server for lightweight vocabularies in SKOS and ontologies in RDFS/OWL (RDF Schema/Web Ontology Language) format with web service support (Tuominen et al. 2009)
xi. The UK Becta Vocabulary Bank provides an SRU web services interface to its educational vocabularies via the Zthes profile, with some additional indexes.
xii. The British Oceanographic Data Centre (BODC) Data Grid's Vocabulary Server provides web service access to its vocabularies represented in SKOS. A mapping service is based on the SKOS mapping relationships.

xiii. As part of the Explicator project, Gray et al. (2009) have implemented a vocabulary search web service, applied to SKOS astronomy related vocabularies, which focuses on identifying the best vocabulary concept for a given query string.

Despite the clear success of early terminology service implementations there are still some hurdles to overcome to facilitate greater adoption and use. Some existing large scale "standard" vocabularies have licensing restrictions on their usage. In order to offer terminology as a persistent service, there is first the need to resolve licensing and copyright issues. Perhaps this would be an opportune moment to suggest that (part of) UDC (Universal Decimal Classification) could be released for public use and for incorporation into some of the existing terminology services?

4.0 Programmatic API approaches / protocols

Currently various approaches are taken to exposing programmatic access to vocabulary data via a network:

a. Linked Data
b. SKOS API, SRU/W
c. SPARQL Endpoints
d. Combinations of the above

The distinction between SKOS API (say) and Linked Data is not necessarily entirely mutually exclusive. SKOS API is an abstract interface so could be implemented via a RESTful approach. While current Linked Data implementations tend to involve more "atomic" implementations, exposing data at the level of individual resources (e.g., concepts), a terminology service could offer various forms of search functionality over associated linked data. This may be necessary for some use cases, where following individual links in turn may be impractical.

However, a discussion on the relative merits of SOAP vs. REST vs. XML-RPC (XML Remote Procedure Call), etc., would risk missing the point; a service API is abstract, specifying what you are able to ask for and what you can expect to get back. The value of an established API can get lost in occasionally zealous discussions about what is essentially a low level delivery mechanism. The issue then is more between using a specific API (linked data, SKOS API, SRU) versus a more flexible query interface (SPARQL).

The specific API approach has a number of attractive features:

i. Abstracts and hides underlying architecture and implementation details.
ii. Predefined functionality – limited defined set of function calls. User does not need to know anything about the underlying data schema, just the expected syntax for calls and responses.
iii. Can implement efficient methods with server side optimisation.
iv. Can take advantage of browser cache for more efficient use of services.

SPARQL endpoints, on the other hand, are a slightly different proposition. Whilst SPARQL undoubtedly offers very powerful server side facilities with advantages of flexibility there are also some not insignificant associated disadvantages, which may serve to limit their viability or attractiveness for use as a reliable outward facing terminology service mechanism.

4.1.1 SPARQL advantages

i. Flexibility – end user decides nature of query and data to be returned
ii. Standardisation – query compatible with any SPARQL enabled system
iii. Native implementations within some platforms, no need to deploy any specific server application.

4.1.2 SPARQL disadvantages

i. To construct a SPARQL query the end user needs to have detailed knowledge of the underlying data schema. It also delegates optimization of queries to the end user.
ii. Use of SPARQL as the API rather dictates the underlying implementation.
iii. Does not easily support implementation of concept expansion and other algorithm based / probabilistic functionality.
iv. Publicly available SPARQL endpoints are elegant but in practice not necessarily an appropriate solution. The same arguments apply as to exposing a public SQL(Structured Query Language) interface – they may expose the server to excessive / malicious activity.
v. SPARQL queries incorporating full text querying can be inefficient, as they involve regular expression filtering. As a consequence, performance may not be sufficient for real-time applications. In fact, we worked around this limitation in the STAR project by supplementing the underlying triple store database with a full-text index. Alistair Miles

also reported encouraging experience of using the Lucene full text search engine in concert with LARQ (a Jena bridge between ARQ and Lucene) to work around the same issues (See SKOS list, February 2009).

To some degree, the appropriate choice depends on the particular circumstances and development context, along with user requirements. This is also currently a fairly quickly moving field.

5.0 Use of terminology services at Glamorgan

A series of projects has explored the use of terminology services and Glamorgan and developed various service and client implementations.

5.1 Pilot Client for SKOS API

In 2003, a use case driven low level SKOS API was developed by ILRT (Institute for Learning & Research Technology, Bristol) for the SWAD Europe project. Although the demonstrator implementation (DREFT) took the form of a set of SOAP based web services, the API was intended as an abstract definition of the standard functionality that a SKOS thesaurus service might typically offer at the API level, independent of whether machine access was via a web service. Development and maintenance of the DREFT software effectively ended when that project ended in 2004, but there has been continuing interest in exposing vocabulary resources to programmatic access and a number of practical approaches have come to the fore.

In 2005, University of Glamorgan created a Windows based client application as a research prototype (Tudhope and Binding 2006) working against this existing SKOS API DREFT service (running but unsupported) at ILRT Bristol. The application was a 'rich client' browser displaying concept details and facilitating browsing via semantic links, as shown in Figure 1 (accessing the GEMET thesaurus).

Due to limitations imposed by the remote server configuration, the application utilized only a small subset (two) of the possible SKOS API calls: 'getConcept' and 'getAllConceptRelatives'. At the time these calls did not return sufficient relationship information, so the browser could only display immediate semantically related terms, without indicating the specific nature of the relationship. The application did, however, provide a fast enough response for satisfactory real-time interaction, and a further enhancement involving the caching of previously retrieved data sig-

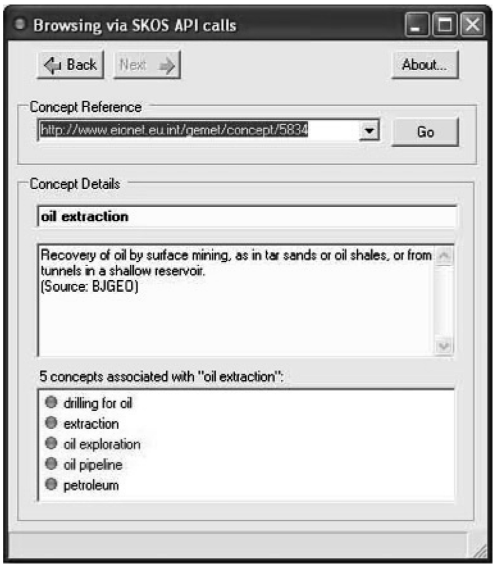

Figure 1. Initial SKOS API client application

nificantly improved the user experience. The exercise provided initial empirical evidence that the SKOS API in the form of a web service could be used to support real-time client applications, and this motivated the development of further services and applications within the scope of our later projects.

5.2 STAR Project services and clients based on SKOS API

The STAR project subsequently developed a pilot set of web services based on a subset of the SWAD-Europe SKOS API, with extensions for concept expansion. Our implementations typically concentrated on providing the functionality necessary for our own purposes, rather than a complete (re)implementation of the original SKOS API DREFT server. The service currently consists of 7 function calls (see Figure 1). The services provide string matching across the associated thesauri, which are represented in SKOS, along with browsing and semantic concept expansion within a chosen thesaurus. Figure 2 summarises the services. The STAR website provides more details under Semantic Terminology Services, including a WSDL (Web Services Description Language) file and service description and an example client that can be downloaded.

- **GetConceptSchemes** Returns an array of all supported ConceptSchemes in the triple store.
- **GetConceptScheme** Given the URI of a particular ConceptScheme, returns a data structure representing that ConceptScheme.

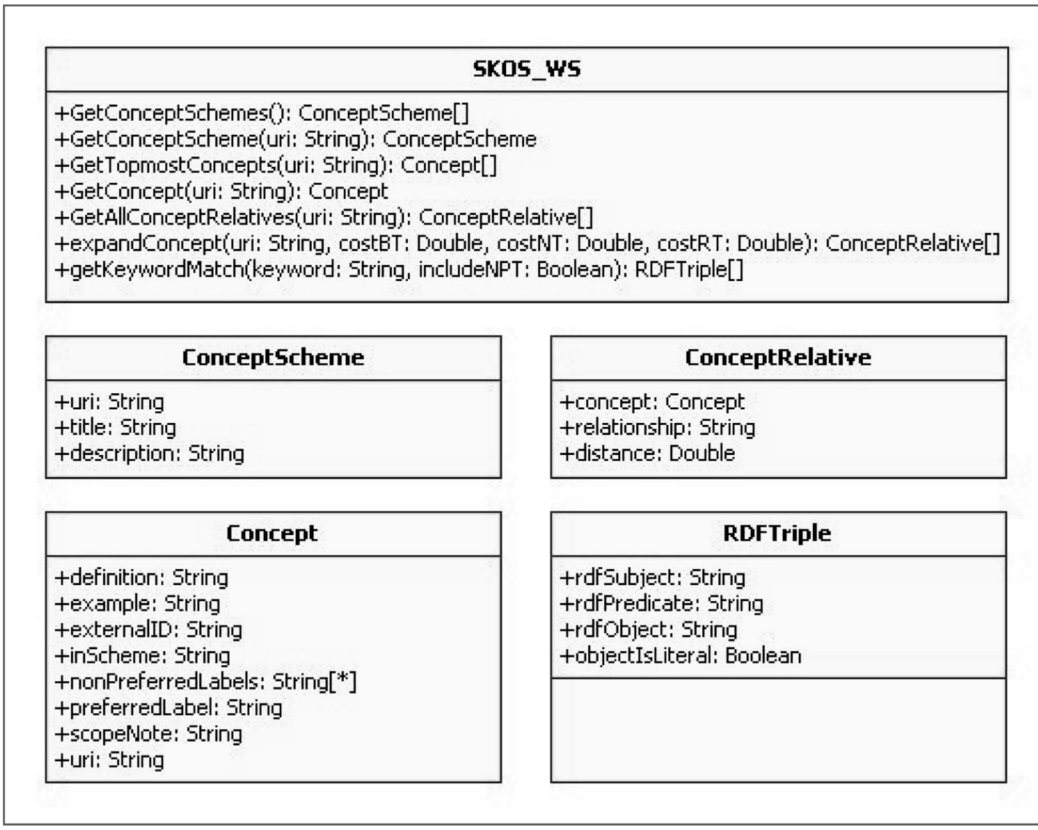

Figure 2. STAR SKOS_WS Service Interface

- **GetTopmostConcepts** Given the URI of a particular ConceptScheme, returns an array of Concepts that are positioned at the top of the hierarchical structure.
- **GetConcept** Given the URI of a particular Concept, returns a data structure representing that Concept.
- **GetAllConceptRelatives** Given the URI of a particular Concept, returns an array of ConceptRelative – consisting of all directly related Concepts and their associated relationship.
- **expandConcept** Given the URI of a particular Concept, performs a spreading expansion of that Concept, using supplied weighting parameters for core thesaurus relationships. Returns an array of ConceptRelative which includes a distance metric representing the semantic distance of each Concept from the originating Concept.
- **getKeywordMatch** General free text search against the preferredLabel (and optionally the nonPreferredLabels) of all Concepts in the triple store. Returns an array of RDFTriple indicating the individual triples where the match occurred.

The thesauri used for the STAR project were SKOS conversions of thesaurus data received from English Heritage. The services were used in conjunction with applications for cross-search of archaeological datasets, allowing searching to be augmented by SKOS-based vocabulary resources. A series of demonstrator client applications were developed (Figure 3) extending the functionality of the initial SKOS API client application.

Queries are often expressed at a different level of generalization from document content or metadata, or may employ a slightly different semantic perspective. In combination with the search system, the services allowed queries to be expanded by synonyms or by concept expansion over the SKOS semantic relationships. Concept expansion was based on a measure of "semantic closeness" (Binding and Tudhope 2004). Subsequently a number of web browser based "widget" controls were developed (Figure 4), working against the same underlying services. These were developed to be incorporated within online STAR demonstrators and other applications.

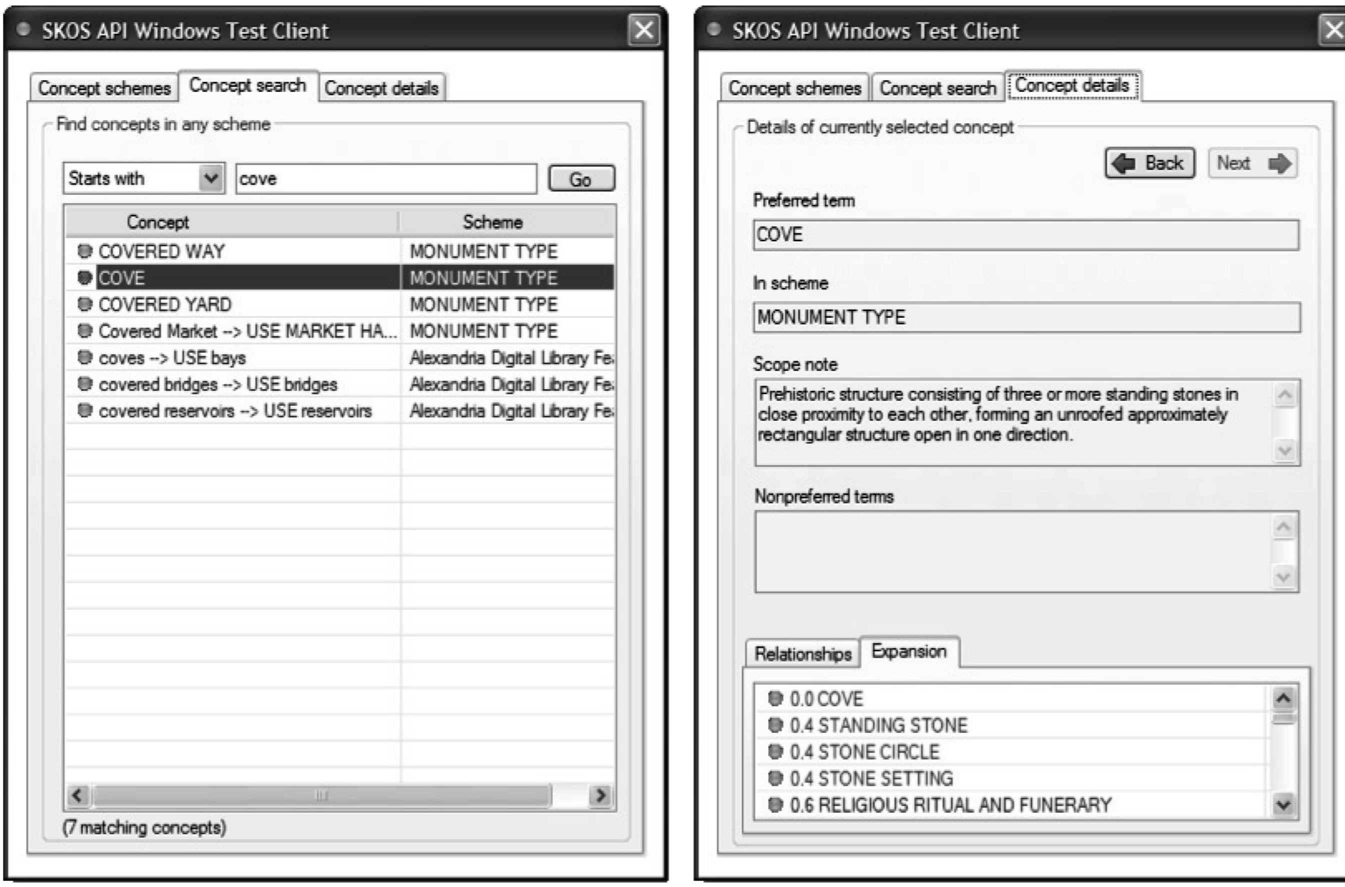

Figure 3. SKOS API client application further developed for the STAR project

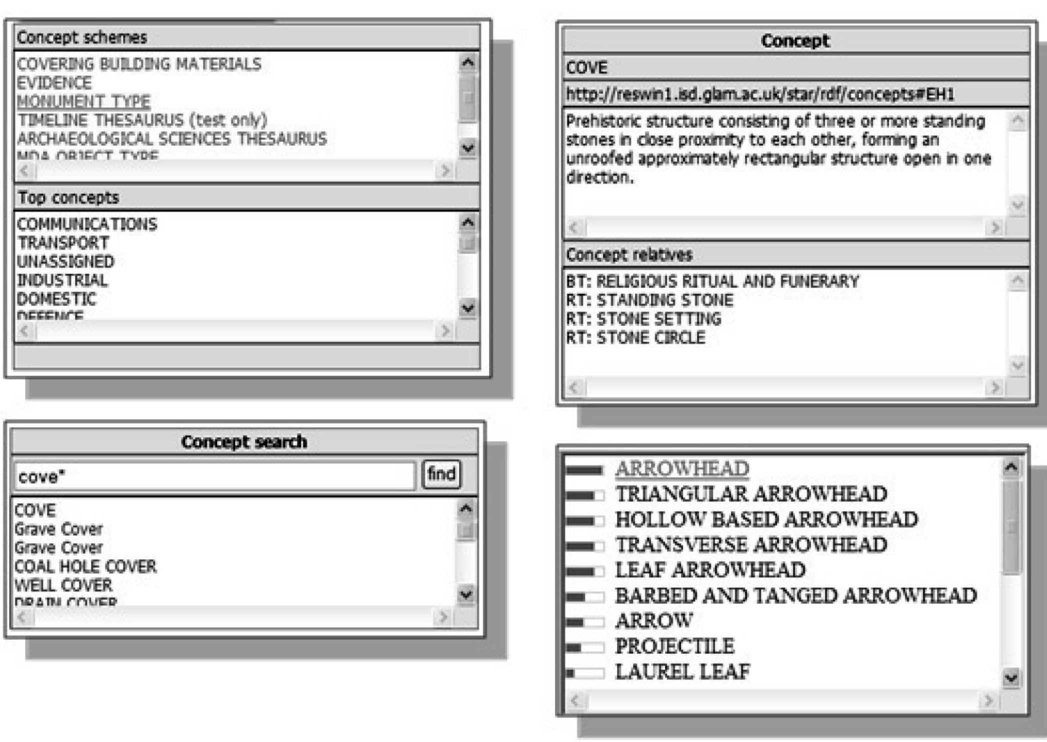

Figure 4. Browser widgets developed for the STAR project

During the course of the project the STAR ser STAR demonstrators and other applications.vices have gone on to be utilized by other projects – notably the ADS (Archaeological Data Service) ArchaeoTools project and a DELOS prototype Digital Library Management System (Binding et al. 2007). They have also been used by undergraduate projects within the University. This demonstrates their utility beyond the particular domain for which they were originally developed.

6.0 A case for more specialised services

Current service implementations tend to conflate different kinds of vocabularies in a common programmatic interface; indeed in areas where there is a degree of commonality it makes sense to provide common service functionality across multiple vocabularies. However, there is also a potential case for more specialist services.

The reference documentation for SKOS refers to a common data model for knowledge organization systems. Short of creating specialized subclasses of skos:ConceptScheme there is currently no way to specify the "type" of a vocabulary in SKOS, so applications accessing the data would potentially treat thesauri and classification schemes (for example) as if they are same. Thus there is a case in general for specialised extensions to SKOS.

Our work to date has primarily involved exposing thesauri for programmatic access. More recently, however, building on core elements of the STAR work, we developed a term suggestion service working against the *Dewey Decimal Classification* (*DDC*), with a URL-based service call interface returning JSON/XML data. This service was developed for a project PERTAINS (PERsonlisation Tagging interface Information in Services), led by MIMAS (University of Manchester) to explore personalization of tag suggestions for users of their COPAC and Intute systems. This initial work surfaced a number of observations concerning the differences between thesauri and other vocabularies. With particular emphasis on major schemes, classification schemes:

i. tend to be more general, covering a wider subject area (i.e., whole library);
ii. tend to have longer, more descriptive captions;
iii. have an associated notation (often encompassing a specific ordering principle);
iv. tend to be more associated with browsing usage;
v. tend to be intended for classification, not indexing;
vi. tend to encourage pre-coordinated descriptor strings for use in indexing and browsing (as opposed to post-coordinated thesauri) – see, for example, Broughton (2001) and FATKS (Facet Analytical Theory in Managing Knowledge Structures).

Pre-coordinated descriptors and ordering based on notation have been emphasised as important distinctive elements of classification schemes (Broughton 2001, Gnoli and Hong 2006). These differences have potential implications for the service calls to be exposed. Possible specialisation extensions to services for classification schemes would be services to handle pre-coordination of terms informed by facet grammar or synthesis rules, incorporating validity checking constraints and also ranking/ordering services.

Term suggestions in a "type ahead" style interface work well when every term is unique, as is the case in a thesaurus. Term lookup in classifications and subject heading schemes however becomes more complex, since a term can appear in many more places within captions. The context of DDC terms depends on their ancestry for clarity in online display (this issue was observed in another Glamorgan project, EnTag: Enhanced Tagging for Discovery). When offering suggestions starting with the characters typed, even just within the 1000 top level classes of the *DDC* Summaries, the term "Philosophy and theory" occurs over 100 times; only with the associated context of the broader term would each suggestion be useful.

The "reverse order" characteristic of *LCSH* (*Library of Congress Subject Headings*) terms (see Figure 5) would make them less appropriate for interactive type ahead style interfaces, as they often share a common prefix:

Laurence-Moon Syndrome — ultrastructure
Laurence-Moon Syndrome — therapy
Laurence-Moon Syndrome — surgery
Laurence-Moon Syndrome — rehabilitation
Laurence-Moon Syndrome — radiotherapy
(etc.)

Figure 5. LCSH subject headings

In order to reduce the volume of suggestions (due to the nature of the *DDC* captions as described previously), the term suggestion service for the PERTAINS project incorporated an extra parameter allowing the user to specify areas of interest from the higher level categories. In the demonstration applica-

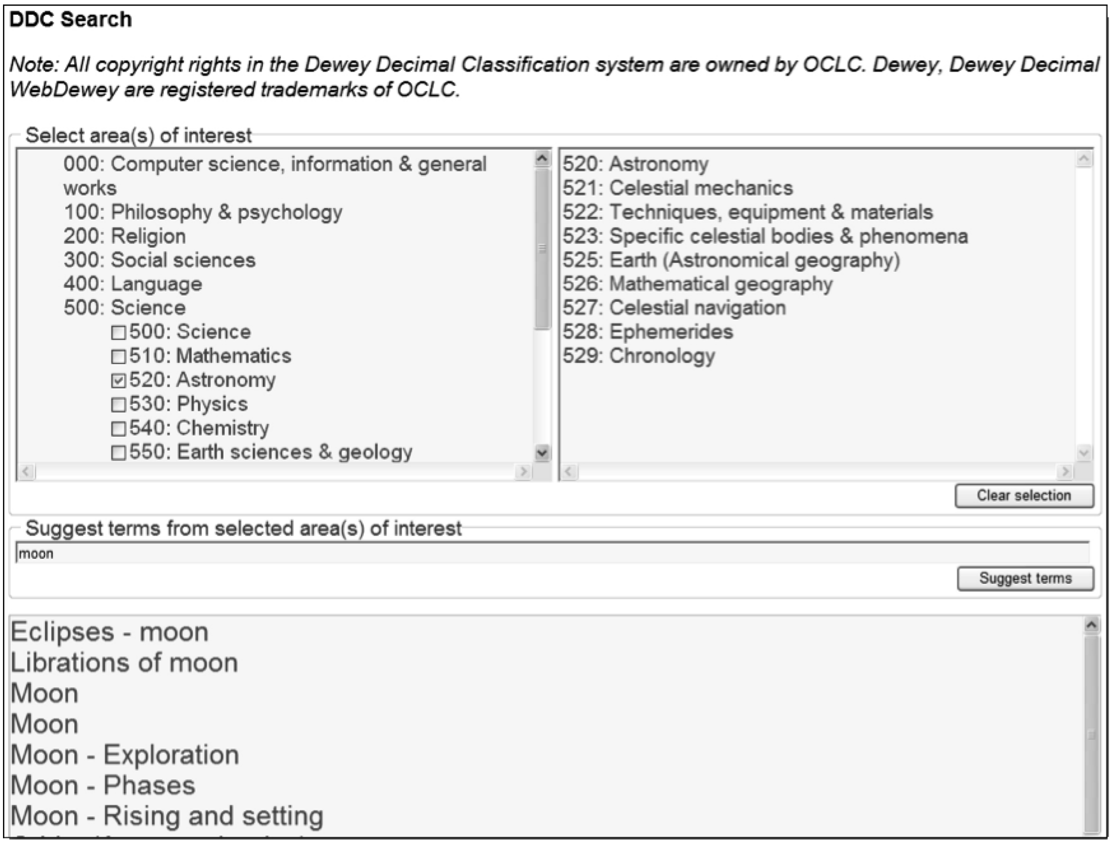

Figure 6. DDC search within specific categories

tion (Figure 6), a search on "moon" is restricted to suggestions from class 520 (Astronomy). This prevents suggestions, e.g., from astrology, author names, place names, from being returned. The problem of qualifying the returned suggestions is however still evident in this particular example.

7.0 Conclusions

This paper has introduced terminology services and reviewed work in the area. Implementations in various projects at Glamorgan have been discussed along with some issues arising. The choice of employing a terminology service over alternative methods of delivering programmatic access to vocabularies depends on the application use cases and the skill set of developers involved. Some situations may involve a combination of (say) terminology services, linked data, general query languages not designed specifically for vocabularies.

Section 4 discusses pros and cons. General purpose languages (such as SPARQL or SRU) may offer flexibility if developers are familiar with the language. Furthermore, terminology services rely on network availability (assuming the service is located externally) and external server infrastructure. On the other hand, the limited set of function calls provided by a terminology service can offer advantages in hiding details of the underlying architecture or representation, while being optimised for common use cases involving online vocabularies. A terminology web service is not restricted to any particular client platform nor development language. This may suit some development situations.

Thus terminology services enjoy a set of distinctive advantages for many contexts and situations. These include:

i. Abstracts and hides underlying architecture and implementation details;
ii. Predefined functionality – limited defined set of function calls. User does not need to know anything about the underlying data schema, just the expected syntax for calls and responses;
iii. Services are platform/location agnostic;
iv. Can implement efficient methods with server side optimisation;
v. Can take advantage of browser cache for more efficient use of services;
vi. Can assist the terminology provider maintain version control.

Strong commonality exists in the abstract API of current terminology services. We have discussed programmatic API approaches and observed how this commonality can sometimes be lost in technical discussions of low level delivery mechanisms such as REST/SOAP/RPC. Current terminology services and associated data models have tended to conflate various types of vocabulary in the interests of common purpose. However, there are compromises inherent in this approach, and we have discussed the case for more specialised services, particularly for major classification schemes.

References

Binding, Ceri et al. 2007. *DelosDLMS: infrastructure and services for future digital library systems, 2nd DELOS Conference, Pisa (2007)*. Available http://www.delos.info/index.php?option=com_content&task=view&id=602&Itemid=334

Binding, Ceri and Tudhope, Douglas. 2004. KOS at your service: Programmatic access to knowledge organization systems. *Journal of Digital Information,* 4 (4). Available http://journals.tdl.org/jodi/article/view/110/109

Broughton, Vanda. 2001. Faceted classification as a basis for knowledge organization in a digital environment: The Bliss Bibliographic Classification as a model for vocabulary management and the creation of multi-dimensional knowledge structures. *New review of hypermedia and multimedia,* 7: 67-102.

Davies, Ron. 1996. Thesaurus-aided searching in search and retrieval protocols. In Rebecca Green ed., *Knowledge organization and change: proceedings of the 4th International ISKO Conference.* Frankfurt: Indeks Verlag, pp. 137-43.

Gardner, Tracy. 2001. An Introduction to Web services. *Ariadne,* 29. Available http://www.ariadne.ac.uk/issue29/gardner/intro.html

Gnoli, Claudio and Hong Mei. 2006. Freely faceted classification for web based information retrieval. *New review of hypermedia and multimedia* 12: 63-81.

Golub, Koraljka, et al. 2009. EnTag: enhancing social tagging for discovery. *Proc. 9th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2009), Austin, TX, June, ACM Press*, pp 163–72.

Gray, Alasdair et al. (2009 in press). Finding the right term: Retrieving and exploring semantic concepts in astronomical vocabularies. *Information processing and management* doi:10.1016/j.ipm.2009.09.004

Johnson, Eric H. 2004. Distributed thesaurus Web services. *Cataloging and classification quarterly* 37 (3-4): 121-53.

SKOS public discussion list: public-swd-wg@w3.org: Mail Archives. Available http://lists.w3.org/Archives/Public/public-swd-wg/

Summers, Ed; Isaac, Antoine; Redding, Clay; Krech, Dan. 2008. *LCSH*, SKOS and linked data. *Proceedings of the International Conference on Dublin Core and Metadata Applications, (DC 2008)*. Available http://arxiv.org/abs/0805.2855

Tudhope, Douglas and Binding, Ceri. 2006. Toward terminology services: experiences with a pilot web service thesaurus browser. *ASIS&T bulletin* 32(5): 6–9, June/July 2006.

Tudhope, Douglas and Binding, Ceri. 2008. Faceted thesauri. *Axiomathes* 18(2): 211–22. Available http://www.springerlink.com/content/m67378t5576g8670/fulltext.pdf

Tudhope, Douglas, Koch Traugott and Heery, Rachel. 2006. *Terminology services and technology: JISC state of the art review.* Available http://www.jisc.ac.uk/media/documents/programmes/capital/terminology_services_and_technology_review_sep_06.pdf

Tuominen, Jouni; Frosterus, Matias; Viljanen, Kim; Hyvönen, Eero. 2009. ONKI SKOS server for publishing and utilizing SKOS vocabularies and ontologies as services. *Proceedings of the 6th European Semantic Web Conference (ESWC 2009), Heraklion, Greece, May 31 - June 4, 2009*. Springer-Verlag. Available http://www.seco.tkk.fi/publications/2009/tuominen-et-al-onki-skos-2009.pdf

Vizine-Goetz, Diane; Houghton, Andrew; Childress, Eric. 2006. Web services for controlled vocabularies. *ASIS&T bulletin* 32 (5-6): 9-12.

All URLs are last checked in February 2010.

Appendix 1: Terminology Web Services

Alexandria Digital Library. *The ADL Thesaurus Protocol.*	http://www.alexandria.ucsb.edu/thesaurus/specification.html
Archeological Data Service. *ADS ArcheoTools project.*	http://ads.ahds.ac.uk/project/archaeotools/
Becta Vocabulary Bank Web Services.	http://bank.vocman.com/bank-webapp/technical
BODC British Oceanographic Data Centre, Natural Environment Research Council, NERC Vocabulary Server.	http://www.bodc.ac.uk/products/web_services/vocab/
CATCH *Vocabulary and alignment repository demonstrator.*	http://www.cs.vu.nl/STITCH/repository/
CERES and National Biological Information Infrastructure (NBII) Biological Resources Division (BRD). *The CERES/NBII Thesaurus Partnership Project.*	http://ceres.ca.gov/thesaurus/
Copac National, Academic, and Special Library catalogue.	http://copac.ac.uk/
EIONET GEMET web services.	http://www.eionet.europa.eu/gemet/webservices?langcode=en
EnTag - Enhanced Tagging for Discovery Project.	http://www.ukoln.ac.uk/projects/enhanced-tagging/
Explicator Project.	http://explicator.dcs.gla.ac.uk/
FAO Agrovoc web services.	http://aims.fao.org/website/Documentation/sub
FATKS - Facet Analytical Theory in Managing Knowledge Structures.	http://www.ucl.ac.uk/fatks/
Getty vocabularies web services.	http://www.getty.edu/research/conducting_research/vocabularies/vocab_web_services.pdf
HILT SRU/W Server.	http://hilt4.cdlr.strath.ac.uk/hilt_sru.cgi
Jena – A Semantic Web Framework for Java.	http://openjena.org/
LARQ - Free Text Indexing for SPARQL.	http://jena.sourceforge.net/ARQ/lucene-arq.html
Library of Congress. *Authorities and Vocabularies service.*	http://id.loc.gov/authorities/
Linked Data : Connect Distributed Data Across the Web.	http://linkeddata.org/
MIMAS. *Centre of Excellence, University of Manchester.*	http://mimas.ac.uk/
CSA/NBII Biocomplexity Thesaurus Web Services.	http://nbii-thesaurus.ornl.gov/thesaurus/
OCLC terminology services project and prototype.	http://www.oclc.org/research/projects/termservices/ [accessed 7/17/09, no longer available
PERTAINS - PERsonlisation Tagging interface INformation in Services presenting tag recommenders in UK national services.	*http://www.jisc.ac.uk/whatwedo/programmes/resourcediscovery/pertains.aspx*
SKOS - Simple Knowledge Organisation Systems - W3C Semantic Web Deployment Working Group.	http://www.w3.org/2004/02/skos/
SKOS API (SWAD Europe).	http://www.w3.org/2001/sw/Europe/reports/thes/skosapi.html
SPARQL endpoint. *Semantic Web.*	http://semanticweb.org/wiki/SPARQL_endpoint

STAR - Semantic Technologies for Archaeological Resources Project. University of Glamorgan Hypermedia Resarch Unit.	http://hypermedia.research.glam.ac.uk/kos/STAR/
STITCH @ CATCH - Semantic Interoperability to access Cultural Heritage.	http://www.cs.vu.nl/STITCH/
STW Web Services (beta)- German National Library of Economics REST web service.	http://zbw.eu/beta/stw-ws
SWAD Europe.	http://www.w3.org/2001/sw/Europe/Overview.html
SWAD-Europe Thesaurus Activity. Deliverable 8.7. RDF Thesaurus Prototype.	http://www.w3.org/2001/sw/Europe/reports/thes/8.7/#sec-demo-server
The Zthes specifications for thesaurus representation, access and navigation.	http://zthes.z3950.org/