

**Knowledge Organization System Services:
brief review of NKOS activities
and possibility of KOS registries**

Doug Tudhope
Hypermedia Research Unit
University of Glamorgan

NKOS: Networked Knowledge Organization Systems/Services

Informal network for enabling knowledge organization systems (KOS), such as classification systems, thesauri, gazetteers, and ontologies, as networked interactive information services to support the description and retrieval of diverse information resources through the Internet

- Two ongoing series of NKOS workshops
 - JCDL Conferences in the US
 - ECDL Conferences in Europe
 - DC NKOS workshop 2005
- Listserv hosted by OCLC
- Special issues in JoDI (2001, 2004), NRHM (2006)
- See <http://nkos.slis.kent.edu/>

NKOS Workshops

• **US DL/JCDL**

- 1997 - Philadelphia
- 1998 – Pittsburgh (1st)
- 1999 – Berkeley
- 2000 – San Antonio
- 2001 – Roanoke
- 2002 – Portland
- 2003 – Houston
- 2004 – no workshop
- 2005 – Denver (7th)

ECDL

2000 – Lisbon

2003 – Trondheim

2004 – Bath

2005 – Vienna

2006 – Alicante

2007 – Budapest, Sept 21

<http://www.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2007/programme.html>

Dublin Core NKOS Session

2005 – Madrid, 2005

Agenda: KOS integration into DL services

from Hill02 Research Agenda KOS/DL

Taxonomy of KOS - KOS types linked to DL service protocols

Registries of KOS and KOS-level metadata to represent them

XML/RDF KOS representations - customisable

Core set of relationship types across all KOS

General KOS service protocol

from which protocols for specific types of KOS can be derived

Robust linking model in which DL entities (collections, objects, and services) can refer to KOS entities (concepts, labels, and relationships)

Visualization tools that fully use and display the rich semantics embedded in KOS

Terminology Services

Searching for concepts

- schemes in registries

- concepts/terms in taxonomy servers

Search support for queries

- collection finding

- cross-searching, cross-browsing, mapping services

- KOS browsing and user interface/visualisation

- query expansion, disambiguation

- automatic indexing and classification

- extraction/mining of terms

- translation support using vocabularies

JISC review on Terminology Services and Technologies

<http://www.ukoln.ac.uk/terminology/JISC-review2006.html>

NKOS: some example activity

- ADL Gazetteer, protocols <http://www.alexandria.ucsb.edu/research/kos/index.htm>
- AGROVOC (FAO) thesaurus → ontology, mapping
http://www.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2006/presentations/AOS-CS_OWL-model.ppt
- Ecoterm (Gail Hodge) http://www.umweltbundesamt.at/en/ecoterm_vienna
- Glamorgan FACET, STAR projects <http://hypermedia.research.glam.ac.uk>
- HILT mapping via DDC (web services) <http://hilt.cdlr.strath.ac.uk/>
- KnowLib automatic classification <http://www.it.lth.se/knowlib/>
- NKOS website and Marcia Zeng (mapping) <http://nkos.slis.kent.edu/>
- OCLC terminology (mapping) services via DDC, automatic classification <http://www.oclc.org/research/projects/termservices/>
- SKOS RDF representation <http://www.w3.org/2004/02/skos/>

See

JISC review on Terminology Services and Technologies

http://www.jisc.ac.uk/Terminology_Services_and_Technology_Review_Sep_06



FACET - Faceted Access to Cultural hEritage Terminology

***FACET - a collaborative project investigating the potential of
semantic expansion in retrieval***

Aims:

- Integration of thesaurus into search process / interface
- Semantic query expansion
taking advantage of facet structure

<http://www.comp.glam.ac.uk/~FACET/>

Semantic Query Expansion

Times
Agents
Processes
Materials
Objects

Find in
Thesaurus

View
Hierarchy

View
Expansion

Add to
Query

Run Query

Objects...
Furnishings...
furnishings...
 <furnishings by form or function>...
 furniture...
 <furniture by form or function>...
 seating furniture...
 <single seating furniture>...
 chairs...
 <chairs by form>...
 armchairs...
 bergeres
 great chairs...
 elbow chairs...
 ..

armchairs
Term applied to a wide variety of chairs with arms, to distinguish them from side chairs which have no arms.
(Variations: armchair; arm chairs; armed chairs; arming chairs; chairs, arm; chairs, armed;

See also *curridles (chairs), porters' chairs, sleeping chairs, lolling chairs, side chairs, Sleepy Hollow chairs, hunting chairs, student chairs, dining chairs, reading chairs, Morris chairs*

brocading, wft patterning, brocade, brocatelle
Remove

Term Expansion
Min Max

mahogany, white mahogany, Borneo mahogany, Swietenia, afrkan mahogany, hardwood.
Remove

Term Expansion
Min Max

armchairs, elbow chairs, great chairs, bergeres, Sleepy Hollow chairs, student chairs, noters'
Remove

Term Expansion
Min Max

25 matching items found.

Match	Reference	Collection	Index Terms	Description
<input type="checkbox"/>	292708	NRM - Railway Furniture	Victorian, oak, crests, ovals, brocade, Carver chairs	Carver chair, oak with oval brocade seat, Prince of Wales crest on back from Royal Saloon of 1876.
<input type="checkbox"/>	292763	NRM - Railway Furniture	Edwardian, mahogany, upholstery, floral	Carver Chair, Great Northern Railway, Mahogany, upholstered in cream, pink &

FACET Web Demonstrator

- Illustrates thesaurus based expansion and faceted search
- Intended as an exploration of FACET research outcomes via dynamically generated Web components rather than a complete final interface
- Based on custom API for thesaurus programmatic access
- Browser-based interface (ASP application), using a combination of server-side scripting and compiled components

<http://www.comp.glam.ac.uk/~FACET/webdemo/>

<http://jodi.tamu.edu/Articles/v04/i04/Binding/>

Some lessons learned

- Results show potential of faceted KOS for
 - Query expansion with semantically ranked results
 - Realtime implementation multi-concept matching function
 - Semantic expansion as a browsing tool
 - Potential combine with statistical and linguistic techniques

How to generalise?

→ *need for*

- Common KOS representations and APIs
- Terminology Registries? (JISC scoping study by Glamorgan and UKOLN just starting)

SKOS API

- SKOS Core (RDF/XML) Schema and SKOS API deliverables of SWAD-Europe Thesaurus Activity - <http://www.w3.org/2001/sw/Europe/reports/thes>
- SKOS API designed to provide programmatic access to thesauri and related KOS in SKOS Core – builds on previous NKOS work on KOS protocols
- Example SKOS API calls
 - `getConcept (uri)`
 - `getConceptsMatchingKeyword/Regex (string)`
 - `getAllConceptRelatives (concept)`
 - `getSupportedSemanticRelations`
 - `getAllConceptRelatives (concept, relation)`
 - `getAllConceptsByPath (concept, relation, distance)`

Pilot KOS Browser Client Web Service

- Developed pilot to work with a remote server as an initial experiment with the SKOS API, a 'rich client' browser displaying details for thesaurus concepts via web service calls
- Uses GEMET - GEneral Multilingual Environmental Thesaurus
- DREFT demonstration web services server based on SKOS API developed at ILRT, Bristol University <http://www.w3.org/2001/sw/Europe/reports/thes/dreft/>
- Only a subset of SKOS API calls were available at time of work due to local requirements

So we investigated possibilities with just 2 API calls

Pilot SKOS API Web Service Browser

getConcept

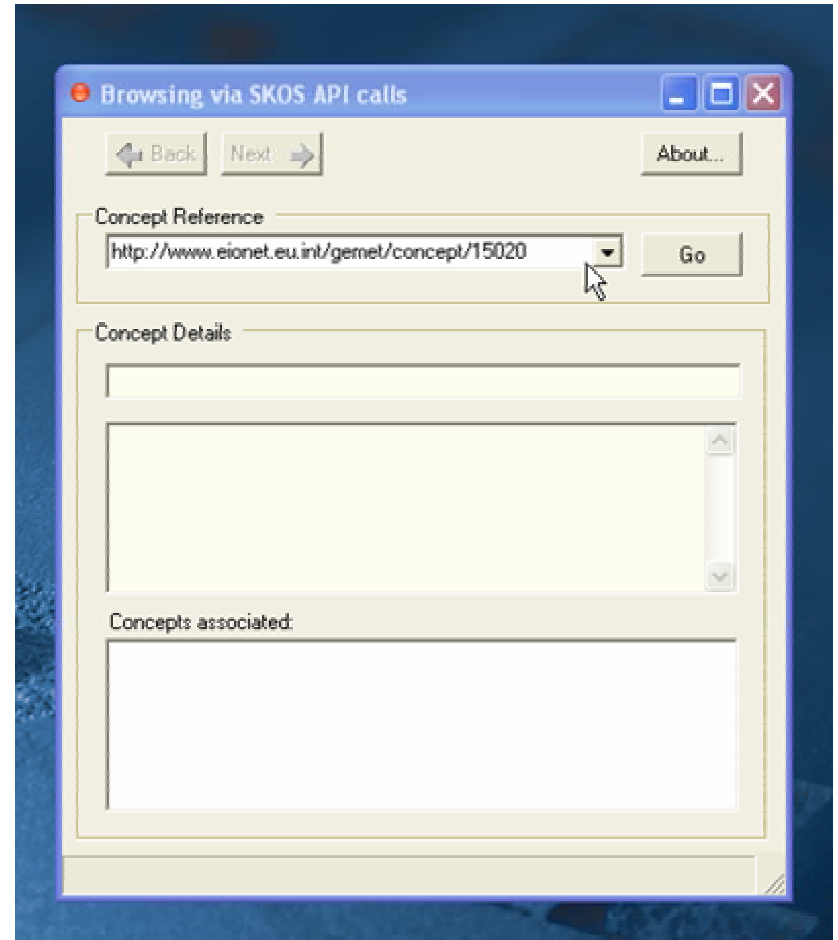
getAllConceptRelatives

show semantically connected
concepts but not relationships

Navigation history and
local cache of retrieved concepts
implemented

API needs more work
but is a possible basis for web services

See http://www.asist.org/Bulletin/Jun-06/tudhope_binding.html



How to encourage use of existing KOS and development of KOS-based tools?

- Various lists of existing KOS available with some form of metadata and links to content as an intellectual resource

Eg

- HILT: Vocabulary resources. <http://hilt.cdlr.strath.ac.uk/Sources/vocabulary.html>
- SPECTRUM Terminology Bank. <http://www.mda.org.uk/spectrum-terminology/termbank.htm>
- Synapse (now: Factiva) Taxonomy Warehouse <http://www.taxonomywarehouse.com/>

Terminology (KOS) Registries ?

- Currently move to registries of collections with defined services
Eg JISC Information Environment Service Registry (IESR):
<http://iesr.ac.uk/>
- Can we also move towards *registries of machine accessible KOS* making KOS content available in different kinds of tools via terminology (web) services?
- Need standards for KOS representation and access protocols
- Need standards for metadata on KOS
- Need to consider business case and cost effectiveness
- Do we need to think more about typologies of KOS, what distinguishes the different types of KOS?

Terminology Registries – some work to date

- NKOS Reference document for data elements. Vazine Goetz 2001
http://staff.oclc.org/~vazine/NKOS/Thesaurus_Registry_version3_rev.htm
- (Metadata and) Vocabulary Registries. Rachel Heery, NKOS Special Session at DC-2005.
<http://www.ukoln.ac.uk/terminology/events/NKOSatDC2005/NKOSatDC2005/heery-nkos-v2.ppt>
- Bargmeyer, Bruce (2005). eXtended Metadata Registries (XMDR). Presentation at the 7th NKOS Workshop at JCDL 2005.
<http://nkos.slis.kent.edu/2005workshop/Bargmeyer.ppt>
- XMDR <http://www.xmdr.org/>
- NSDL Registry. <http://metadataregistry.org/>

JISC Scoping Study on Terminology Registries

UKOLN-Glamorgan, 6 months, just starting

- The study will analyse issues related to the **potential delivery of a Terminology Registry as a shared infrastructure service within the JISC Information Environment**. The study will consider how a Registry might support development of terminology and other services within the context of a services oriented environment.
- The role of a terminology registry will be considered in relation to other components of the information landscape, in particular with regard to **other JISC IE shared infrastructure services**, such as the JISC IE metadata registry (IEMSR) and the JISC IE service (and collection) registry (IESR).
- The study will describe **usage scenarios and use cases, investigate requirements and sustainability, study costs and benefits**. It will look at organisational questions such as **who is to create, maintain and host the content of the registry** and at cooperation with similar registries.
- Architectural issues will be explored, in particular the potential for co-ordination of registry efforts within the JISC IE and across domains.
- The scoping study will make recommendations on which JISC can base decisions on future provision of a terminology registry.

Terminology Registries?

All input very welcome!

towards a KOS of KOS?

Idea of Typology of KOS discussed at ECDL06 NKOS workshop:

- “That discussion tended towards the view of a need for a terminology registry of individual instances of KOS systems, rather than, or in parallel to, a general typology. This was because of extreme wide variation in characteristics of individual vocabularies regardless of what their originators termed them. Eg the ‘types’ of KOS don’t say much in practice about what the KOS is likely to contain. There was also the view, however, that a typology would be required for an effective large scale registry of vocabularies. “
- A tentative typology of KOS: towards a KOS of KOS? Doug Tudhope 5th European NKOS Workshop. <http://www.ukoln.ac.uk/nkos/nkos2006/presentations/tudhope.ppt>

Dagobert Soergel 2001a

Underlying characteristics for defining elements in a Taxonomy of KOS

Potential Facets in Classification of KOS?

- Entities covered
- Information given
- Arrangement
- Purpose for which designed

Dagobert Soergel 2001b

Characteristics for describing and evaluating KOS

- Purpose
- Coverage of concepts and terms. Sources, quality of usage analysis
- Conceptual analysis and conceptual structure. Terminological analysis
- Use of precombination in the index language
- Access and display. Format of presentation of the vocabulary
- Updating

How are different types of KOS used?

- Important to consider intended purpose/application of a KOS
- How are KOS concepts applied to objects they refer to?
- Distinction between classification and indexing
 - classification groups similar items together
 - indexing brings out differences to help distinguish in search
- (AI) Ontologies Vs Search/Discovery oriented KOS

Rationale for (speculative) draft template of (some) KOS characteristics

- Not exhaustive/complete - for exploration
 - other characteristics to be included
 - Some characteristics to be omitted
- for types of KOS, rather than a specific instance
- Orienting particularly to search/discovery purposes
- Tentative facets (a subset)
Partly chosen to help make distinctions
between some common types of KOS
- Begin to consider KOS purposes and contexts of use
 - how we might describe *purpose*?

Factors governing types of KOS Template (draft)

Entities

Concepts, terms, strings,

Atomic - Composite (attributes)

Enumerative - Synthetic

Low – medium - high degree precombination (coordination in KOS itself)

Size: small – large

Depth: small – medium - large

Relationships (internal)

Types / expressivity of relationships:

low (core set) – medium – high (definable)

concept-concept, concept-term, term-term

monohierarchies - polyhierarchies

Formality: low – medium – high

Typical application to objects in domain of interest

Metadata element: subject, various elements, general

Granularity of application objects: unstructured - complex

Relationship applying concepts to objects in domain

about (fuzzy), instance

Exhaustivity: low - high

Specificity: low - high

Coordination: low - high

expressivity and formality of relationships in coordination (synthesis rules)

Factors governing types of KOS

Term List

Entities

Concepts, **terms**, strings,

Atomic - Composite (attributes)

Enumerative - Synthetic

Low – medium - high degree precombination (coordination in KOS itself)

Size: **small** – large

Depth: **small** – medium - large

Relationships (internal)

Types / expressivity of relationships:

low (core set) – medium – high (definable)

concept-concept, concept-term, term-term

monohierarchies - polyhierarchies

Formality: low – medium – high

Typical application to objects in domain of interest

Metadata element: subject, various elements, **general**

Granularity of application objects: unstructured - complex

Relationship applying concepts to objects in domain

about (fuzzy), instance

Exhaustivity: low - high

Specificity: low - high

Coordination: low - high

expressivity and formality of relationships in coordination (synthesis rules)

Factors governing types of KOS Taxonomy

Entities

Concepts, **terms**, strings,

Atomic - Composite (attributes)

Enumerative - Synthetic

Low – **medium** - high degree precombination (coordination in KOS itself)

Size: **small** – large

Depth: **small** – medium - large

Relationships (internal)

Types / expressivity of relationships:

low (core set) – medium – high (definable)

concept-concept, concept-term, **term-term**

monohierarchies - polyhierarchies

Formality: **low** – medium – high

Typical application to objects in domain of interest

Metadata element: subject, various elements, **general**

Granularity of application objects: unstructured - complex

Relationship applying concepts to objects in domain

about (fuzzy), instance

Exhaustivity: **low** - high

Specificity: **low** - high

Coordination: **low** - high

expressivity and formality of relationships in coordination (synthesis rules)

Factors governing types of KOS Subject Headings

Entities

Concepts, terms, strings,

Atomic - Composite (attributes)

Enumerative - Synthetic

Low – medium - high degree precombination (coordination in KOS itself)

Size: small – large

Depth: small – medium - large

Relationships (internal)

Types / expressivity of relationships:

low (core set) – medium – high (definable)

concept-concept, concept-term, term-term

monohierarchies - polyhierarchies

Formality: low – medium – high

Typical application to objects in domain of interest

Metadata element: subject, various elements, general

Granularity of application objects: unstructured - complex

Relationship applying concepts to objects in domain

about (fuzzy), instance

Exhaustivity: low - high

Specificity: low - high

Coordination: low - high

expressivity and formality of relationships in coordination (synthesis rules)

Factors governing types of KOS

Classification Scheme

Entities

Concepts, terms, strings,

Atomic - Composite (attributes)

Enumerative - Synthetic

Low – medium - high degree precombination (coordination in KOS itself)

Size: small – large

Depth: small – medium - large

Relationships (internal)

Types / expressivity of relationships:

low (core set) – medium – high (definable)

concept-concept, concept-term, term-term

monohierarchies - polyhierarchies

Formality: low – medium – high

Typical application to objects in domain of interest

Metadata element: subject, various elements, general

Granularity of application objects: unstructured - complex

Relationship applying concepts to objects in domain

about (fuzzy), instance

Exhaustivity: low - high

Specificity: low - high

Coordination: low - high

expressivity and formality of relationships in coordination (synthesis rules)

Factors governing types of KOS

Faceted Classification Scheme

Entities

Concepts, terms, strings,

Atomic - Composite (attributes)

Enumerative - Synthetic

Low – medium - high degree precombination (coordination in KOS itself)

Size: small – large

Depth: small – medium - large

Relationships (internal)

Types / expressivity of relationships:

low (core set) – medium – high (definable)

concept-concept, concept-term, term-term

monohierarchies - polyhierarchies

Formality: low – medium – high

Typical application to objects in domain of interest

Metadata element: subject, various elements, general

Granularity of application objects: unstructured - complex

Relationship applying concepts to objects in domain

about (fuzzy), instance

Exhaustivity: low - high

Specificity: low - high

Coordination: low - high

expressivity and formality of relationships in coordination (synthesis rules)

Factors governing types of KOS Thesaurus

Entities

Concepts, terms, strings,

Atomic - Composite (attributes)

Enumerative - Synthetic

Low – medium - high degree precombination (coordination in KOS itself)

Size: small – large

Depth: small – medium - large

Relationships (internal)

Types / expressivity of relationships:

low (core set) – medium – high (definable)

concept-concept, concept-term, term-term

monohierarchies - polyhierarchies

Formality: low – medium – high

Typical application to objects in domain of interest

Metadata element: subject, various elements, general

Granularity of application objects: unstructured - complex

Relationship applying concepts to objects in domain

about (fuzzy), instance

Exhaustivity: low - high

Specificity: low - high

Coordination: low - high

expressivity and formality of relationships in coordination (synthesis rules)

Factors governing types of KOS

Lexical database

Entities

Concepts, terms, strings,

Atomic - Composite (attributes)

Enumerative - Synthetic

Low – medium - high degree precombination (coordination in KOS itself)

Size: small – large

Depth: small – medium - large

Relationships (internal)

Types / expressivity of relationships:

low (core set) – medium – high (definable)

concept-concept, concept-term, term-term

monohierarchies - polyhierarchies

Formality: low – medium – high

Typical application to objects in domain of interest

Metadata element: subject, various elements, general

Granularity of application objects: unstructured - complex

Relationship applying concepts to objects in domain

about (fuzzy), instance

Exhaustivity: low - high

Specificity: low - high

Coordination: low - high

expressivity and formality of relationships in coordination (synthesis rules)

Factors governing types of KOS (AI) Ontology

Entities

Concepts, terms, strings,

Atomic - Composite (attributes)

Enumerative - Synthetic

Low – medium - high degree precombination (coordination in KOS itself)

Size: small – large

Depth: small – medium - large

Relationships (internal)

Types / expressivity of relationships:

low (core set) – medium – high (definable)

concept-concept, concept-term, term-term

monohierarchies - polyhierarchies

Formality: low – medium – high

Typical application to objects in domain of interest

Metadata element: subject, various elements, general

Granularity of application objects: unstructured - complex

Relationship applying concepts to objects in domain

about (fuzzy), instance

Exhaustivity: low - high

Specificity: low - high

Coordination: low - high

expressivity and formality of relationships in coordination (synthesis rules)

How to apply KOS?

- What is the purpose of a given KOS?
- we need to specify/articulate more clearly
- Cost/benefit issues for KOS applications
in granularity of relationships and degree of formalisation
- Important to take into account how concepts are used
Some KOS informal by design
with relationships at a useful level of generality
for many search/retrieval applications (with some specialisation?)

Contact Information

Doug Tudhope
School of Computing
University of Glamorgan
Pontypridd CF37 1DL
Wales, UK

`dstudhope@glam.ac.uk`

<http://hypermedia.research.glam.ac.uk/>

References

Binding C., Tudhope D. 2004. KOS at your Service: Programmatic Access to Knowledge Organisation Systems. JoDI 4(4), <http://journals.tdl.org/jodi/article/view/jodi-124/109>

NKOS website. <http://nkos.slis.kent.edu/>

Soergel D. 2001a. The representation of Knowledge Organization Structure (KOS) data.: a multiplicity of standards. JCDL 2001 NKOS Workshop, Roanoke. <http://www.clis.umd.edu/faculty/soergel/SoergelNKOS2001KOSStandards.PDF>

Soergel D. 2001b. Evaluation of Knowledge Organization Systems (KOS): Characteristics for describing and evaluating KOS. JCDL 2001 NKOS Workshop, Roanoke. <http://nkos.slis.kent.edu/2001/SoergelCharacteristicsOfKOS.doc>

Tudhope D., Koch T., Heery R. 2006. JISC review on Terminology Services and Technologies http://www.jisc.ac.uk/Terminology_Services_and_Technology_Review_Sep_06