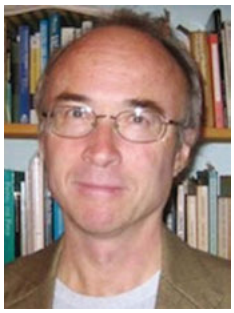
Still Quite Popular After all Those Years— The Continued Relevance of the Information Retrieval Thesaurus†

Douglas Tudhope* and Ceri Binding**

University of South Wales, Hypermedia Research Group, Pontypridd, UK, CF37 1DL,
*<douglas.tudhope@southwales.ac.uk>, **<ceri.binding@southwales.ac.uk>



Douglas Tudhope is Professor in the Faculty of Computing, Engineering and Science, University of South Wales and leads the Hypermedia Research Group. He was PI on the AHRC funded STAR, STELLAR and SENESCHAL projects (Semantic Tools for Archaeological Resources), in collaboration with English Heritage and the preceding EPSRC funded FACET project, in collaboration with the Science Museum, investigating thesaurus-based query expansion. He leads the Linking Archaeological Data Work Package for the ARIADNE FP7 Infrastructures Project. Since 1977, he has been Editor of the journal *New Review of Hypermedia and Multimedia.*

Ceri Binding is a research fellow in the Hypermedia Research Group, Faculty of Computing, Engineering and Science, University of South Wales. Ceri graduated with a BSc in Computer Studies in 1997 whilst working as an analyst designer-programmer for Hyder, before joining the University in 2000. He had responsibility for development work on various research projects including FACET, STAR, STELLAR, SENESCHAL and Digging Into Metadata. He is currently conducting research and development work for the ARIADNE FP7 project.

Tudhope, Douglas and Ceri Binding. 2016. "Still Quite Popular After All Those Years: The Continued Relevance of the Information Retrieval Thesaurus." *Knowledge Organization* 43 no. 3: 174-179. 40 references.

Abstract: The recent ISKO-UK conference considered the question of whether the traditional thesaurus has any place in modern information retrieval. This note is intended to continue in the spirit of that good-natured debate, arguing that there is indeed a role today and highlighting some recent work showing the continued relevance of the thesaurus, particularly in the linked data area. Key functionality that a thesaurus makes possible is discussed. A brief outline is provided of prominent work that employs thesauri in three key areas of infrastructure underpinning advanced retrieval functionality today: metadata enrichment, vocabulary mapping and web services.

Received: 27 January 2016; Accepted: 30 January 2016

Keywords: thesaurus, thesauri, linked data

† The ARIADNE project is funded by the European Commission's 7th Framework Programme (FP7-INFRASTRUCTURES-2012-1-313193). An early version of the ARIADNE work was presented at the 14th European Networked Knowledge Organization Systems (NKOS) Workshop in association with the Digital Libraries 2015 conference. *Still Quite Popular After All Those Years* is the title of a jazz album by Peter Brötzmann and Han Bennink released in 2005 on Brötzmann's BRO label.

1.0 Introduction

The recent debate held by the United Kingdom chapter (hereafter ISKO-UK) of the International Society for Knowledge Organization (ISKO) (ISKO-UK 2015), considered the question of whether the traditional thesaurus has any place in modern information retrieval. Dextre Clarke (2016) gives the historical background to the "long thesaurus debate" and an overview of the origins of the information retrieval thesaurus. This note is intended to continue in the spirit of that good-natured debate, arguing that there is indeed a role today and highlighting some recent work showing the continued relevance of the thesaurus, particularly in the linked data area.

The question of course is not new. As far back as 2000, the introduction to the 4th edition of a widely used text on thesaurus development and use (Aitchison, Gilchrist and Bawden 2000) began by considering the continuing need for thesauri with the advent of full text retrieval and disintermediation with regard to end-users. At

that time, those authors discussed how they were encouraged by the recent use of thesauri in the organisation of large corporate and other Internet resources and by Milstead's (1995, 94) suggestion, that in the future, the thesaurus might be quite widely used to assist search and be "almost invisible to most users." Today, the thesaurus is often pitted against contrasting alternatives based either on statistical full-text retrieval methods or the more recent application of formal logic in ontologies associated with semantic web research.

The thesaurus is one of the most common forms of a knowledge organization system (KOS). Combining an entry vocabulary of synonyms for member concepts, which are linked by a restricted set of semantic relationships, the thesaurus is designed with search and browsing use cases in mind rather than formal logical reasoning.

Typically, early thesauri were developed to assist with indexing a specific collection and to aid users in searching that collection, often mediated by an information professional. With the proliferation of online resources today, a thesaurus may be employed in a variety of ways by users who are unknown to its original designers. Contrariwise, just to discover whether a suitable thesaurus already exists can be an issue for developers of a new information resource. Various projects have constructed registries of thesauri and other KOSs to assist users to locate thesauri on the Web that may serve their particular purposes, with proposals for a core set of KOS metadata to assist discovery a current concern (Golub et al. 2014).

This note continues with a discussion of recent applications of thesauri. The aim is to provide a picture of the trends rather than attempting an exhaustive review (though some pointers to the literature are included) and thus some aspects are described in outline rather than in detail. The next section starts by outlining key functionality that a thesaurus makes possible and a consideration of the role that it plays in linked data and related developments, including the contrasting contributions of thesauri (and related vocabularies) versus logic-based ontologies. This is followed by a brief outline of prominent work that employs thesauri in three key areas of infrastructure underpinning advanced retrieval functionality today: metadata enrichment, vocabulary mapping and web services.

2.0 Modern applications of thesauri

Thesauri and related vocabularies have been seen as necessary for discovery and retrieval of online information resources by various users since the advent of networked online access systems (see, e.g., NKOS). In addition to vocabulary control and disambiguation, traversing a concept structure allows humans and also computers to find and connect information resources. Thus a thesaurus can enhance an interactive user interface, offering pathways for a user to improve and refine a search, with (a view of) the concept graph made explicit. Alternatively, the thesaurus can be available in the background to assist an indexing or retrieval system, make suggestions or perform query expansion. The thesaurus is based on semantic relationships suitable for retrieval applications, where the relationship between information resource and subject metadata is a loose relevance rather than a logical connection. The relationships can be formally defined and made accessible to semi-automated rule-based tools but are not usually appropriate for logical inferencing applications.

The last decade has seen a widespread adoption of the faceted browsing paradigm proposed and demonstrated many years earlier in the thesaurus world (Pollitt 1997; Yee et al. 2003 amongst others) including many commercial web applications and prominent heritage systems, such as Europeana and ArchSearch. Not all facets may be based on thesauri but corporate taxonomies that include thesaurus-like elements drive many systems. Echoing Milstead's point above, a thesaurus can also be used behind the scenes to expand a user search with synonyms or related concepts.

2.1 Thesauri and linked data

The early semantic web literature tended to present one single dimension for comparing different types of KOSs and ontologies based solely on the capability to afford logical reasoning (Souza et al. 2012). It was often presented as a linear spectrum with a progression to formal ontologies as the ultimate form. This is appropriate if the purpose of the comparison is to identify support for logical inferencing. It is misleading, however, if intended as a general comparison. There are various use cases for the different types of KOSs and many potential bases for comparison. Focusing only on one dimension blurs the key similarities and differences and makes it harder to choose an appropriate (combination of) KOS for a particular application's purpose. The thesaurus appears to be particularly suited to typical linked data applications.

Undoubtedly, part of the current uptake of thesauri is due to the growth of linked data and its applications and the prominent role played by the W3C standard, simple knowledge organization system (SKOS) in this context, although the close connection between SKOS and thesauri is not always recognised explicitly. Original formulations of the semantic web project placed a great emphasis on formal ontologies and logic driven applications. This has been overtaken to some extent by the growth of linked data approaches (Bizer, Heath and Berners-Lee 2009) with more flexible KOSs and a renewed focus on interactive applications and retrieval-based functionality.

In particular, SKOS is widely used today in linked data and semantic applications. The SKOS data model is a slightly simplified representation of the thesaurus standard, for use in semantic web applications. The SKOS working group ensured that the SKOS standard was compatible with ISO thesaurus standards and conformed to standard thesaurus design principles (Baker et al. 2013). SKOS is based on the Resource Description Framework (RDF), and this permits thesauri to be published as linked data, in a format which is machine readable and accessible to tools and applications designed to work with the RDF standard. Thus the SKOS standard allows thesauri to play a part in the web of (linked) data and in the various application program interface (API) services offered for programmatic access to online resources indexed by thesauri (Binding and Tudhope 2010).

According to the new thesaurus standard, ISO 25964-1 (International Organization for Standardization 2011, **vii**), the modern thesaurus:

> Retains the assumption that human intellect is usually involved in the selection of indexing terms and in the selection of search terms. If both the indexer and the searcher are guided to choose the same term for the same concept, then relevant documents will be retrieved. This is the main principle underlying thesaurus design, even though a thesaurus may also be applied in situations where computers make the choices.

This stands in contrast to statistical or logic-based approaches, where the tendency has often been to emphasize automated systems where users are positioned mostly out of the loop. Thesauri are designed for use in interactive applications (both indexing and search), where the user plays an active role, perhaps assisted by "smart tools" with semi-automated suggestions.

OWL (Web Ontology Language) ontologies with formal axioms afford the development of automated logical reasoning in applications where that is required. This was essentially the original formulation of the semantic web project (Berners-Lee et al. 2001). Today, however, there is a realisation (Isaac and Baker 2015) that a wide range of interactive applications may be more suited to a less formal approach based on SKOS. Major projects have moved from an initial logic-based implementation to a more flexible approach based on SKOS vocabularies, although with strict attention to standards in the thesaurus hierarchies (Caracciolo et al. 2012; Lappalainen, Frosterus and Nykyri 2014). When designing an application, there is a need to consider carefully which sort of KOS to use depending on the application's requirements (Isaac and Baker 2015):

> The lack of a way to express less formal semantics hindered many early projects that tried to apply Semantic Web technology in the cultural sector by massaging existing knowledge organization systems into formal ontologies. Given the scope of the artifacts considered, this effort required considerable ontological debugging that was ultimately of dubious value. Indeed, most information retrieval scenarios using KOS for searching or browsing collections do not require more than the information that one concept is broader than another.

It is sometimes the case that formal ontologies and thesauri can be employed in a complementary manner, e.g., ISO25964-2 (International Organization for Standardization 2013, section 21). The W3C Library Linked Data Incubator Group (Isaac et al. 2011) sees a similar combination of metadata and "value vocabularies" (including thesauri) describing metadata element sets that can form part of a more formal data schema. The Europeana cultural heritage portal is a prominent example of this approach, combining various thesauri and other KOS with the Europeana Data Model. A HealthFinland prototype is another example that combines SKOS representations of thesauri with ontologies to provide a faceted user interface oriented to user perspectives (Suominen et al. 2009).

2.2 Search and enrichment using thesauri

Enrichment, mapping and web services are research and development areas currently seeing widespread attention in linked data projects. The Europeana implementation is devoting significant effort to tools and techniques for the semi-automatic enrichment of metadata elements, originally held as (ambiguous) text strings, with thesaurus concepts. Recent developments include enrichment with URIs (Uniform Resource Identifiers) from thesauri, such as the Getty *Art and Architecture Thesaurus* (*AAT*) incorporated into the metadata (Charles, Freire and Isaac 2014). A case study (Stiller et al. 2014) of seven datasets in Europeana evaluates results from the different stages of the enrichment process and discusses challenges and best practice. Another case study (de Boer et al. 2012) on the Amsterdam Museum is reported from a long running linked data project that has also contributed to Europeana and faceted retrieval interfaces. An example of enrichment of Europeana data with *AAT* is discussed in a recent Europeana Data Model case study (Charles and Devarenne 2014). Enrichment is considered to bring potential benefits for developing multilingual services. In the example, if the language of the user interface is switched from English to Dutch then the subject metadata "astronomy" displayed can be switched to the corre-

sponding Dutch label "astronomie" since it is one of the languages covered by the *AAT*.

In the archaeology domain, the ARIADNE FP7 collaborative infrastructure project is creating a cross-searchable registry of different kinds of archaeological datasets and reports. These are natively indexed by various thesauri and vocabularies in a range of European languages. The native subject metadata is expressed as literal text strings. The effect of the ARIADNE enrichment service, which forms part of the upload of partner data to the registry, will be to add corresponding linked data identifiers (URIs) to concepts from the *AAT* (recently available in linked data form) as additional derived subject metadata for an ARIADNE resource (Binding and Tudhope 2016). Correspondingly, the search or browsing system should be based on the thesaurus (*AAT*) URI. Thus a search need not depend on employing the exact string used in the original subject indexing but can return results originally indexed with synonym terms, or even closely related concepts, such as hierarchical expansion over narrower concepts, if the search functionality includes that option. This enrichment service makes use of the outcomes of a mapping exercise between the various partner vocabularies and the *AAT*.

2.3 Mapping between thesauri

Mapping between thesauri is seen as a key element for interoperability in (SKOS) linked data and is particularly important for multilingual capability. We mention just a few examples. In the agricultural domain, the AGROVOC thesaurus and the related VocBench publishing platform is a major extended initiative by the United Nations Food and Agriculture Organization (FAO), which has now evolved into an SKOS-based resource (Caracciolo et al. 2013) that underpins the AIMS (Agricultural Information Management Standards) portal. Connecting AGROVOC to other vocabularies is an ongoing linked data activity, with links established to 13 thesauri and other KOSs, including *LCSH* (*Library of Congress Subject Headings*), GEMET (General Multilingual Environmental Thesaurus) and STW Thesaurus for Economics.

The ARIADNE project is employing the *AAT* as a central hub. An exercise mapping partners' native thesauri to the *AAT* is currently underway, with concepts from various vocabularies in four different languages already mapped. A variety of mapping techniques are being employed, including an interactive mapping tool for vocabularies expressed as linked data developed for the project, which generates SKOS mapping relationships (Binding and Tudhope 2016). The ARIADNE portal will support cross search via *AAT* URI concept identifiers over metadata in multiple languages. Historically, a thesaurus tended to support retrieval within a single collection, although mapping has been a continuing facet of research (see the review in Zeng and Chan 2004). The various standards underlying linked data, SKOS and associated mapping technologies makes it easier to extend the reach of thesauri across multiple collections and thus afford a different style of retrieval. The effects of vocabulary mapping are particularly apparent in multilingual collections. In the case of ARIADNE, the mapping of partner vocabularies to the *AAT* will allow a search in one language to retrieve results originally indexed by terms in a variety of languages (Binding and Tudhope 2016). This has the potential to improve both recall (results can take account of indexing in multiple languages) and precision (a string may have different meanings in different languages and false results may arise from literal string search).

2.4 Terminology services

APIs make it easier for application developers to make use of infrastructure based on thesauri and linked data without requiring detailed knowledge of the underlying data models and low level implementation details. These APIs are often expressed as web services. SKOS based terminology services are reviewed in Binding and Tudhope (2010), which also reports on work by the authors on thesaurus web services and corresponding programmatic interface elements (widgets), including a service for tailorable concept expansion based on a notion of conceptual distance from the originating concept. The latest expression of these thesaurus services forms part of the linked data publication of UK archaeological thesauri (Heritage Data 2016) and has been incorporated into the content management system of the Archaeology Data Service (Charno 2014). Various widgets, such as term suggestion and related concepts, have been developed that are based on these web services. The widgets comprise a suite of configurable interactive user interface controls that can be embedded directly within browser-based applications. The source code is available as open source and a demonstration with examples of their use is explained in a project blog (Binding 2014).

The Finnish Ontology Library Service ONKI, originally a research project development of web services and widgets for third party integration is now being implemented as Finto—an operational service provided by the Library of Finland national vocabulary service (Suominen et al. 2014). Another longstanding implementation of SKOS-based web services, which includes mapping services and a linked data implementation, is provided for the STW Thesaurus for Economics (STW); an STW case study forms one of the contributions to this special issue

(Kempf and Neubert 2016). We now also see commercial offerings of SKOS-based services beginning to appear.

There is more to be said beyond the scope of this note. A discussion of performance gains in retrieval effectiveness of SKOS-based expansion is reported recently by Haslhofer et al. (2013). There is a long tradition in studies of the use of thesauri to assist search. The interdependence of performance, evaluation and indexing and the difficulty in attempting to isolate the contribution of any single element, when evaluating performance is demonstrated by Soergel (1994). For a review of thesaurus-based query expansion, part of an investigation into multi-concept (faceted) query expansion, see Tudhope et al. (2006). The review includes consideration of thesaurus relationships, the query matching function and the balance between interactive and automatic control of query expansion.

3.0 Conclusions

Returning to the original question of the ISKO-UK Great Debate, "this house believes that the traditional thesaurus has no place in modern information retrieval," this paper has discussed several examples of existing use of thesauri in real-world, contemporary information retrieval scenarios, emphasising their continued relevance. The answer in brief is that the thesaurus is still relevant and is seeing an encouraging new wave of use with recent developments in the web of data.

Looking to the future, more work is needed on tools and refinement of thesaurus based services more generally, including thesaurus-based auto-completion, concept expansion and mapping services. These tools should be accompanied by example patterns of use to support developers. A review of the recent developments in vocabulary mapping would be helpful. Reports on evaluation of results and discussion of the relative merits of different techniques and tools for thesaurus mapping work would also be valuable, together with guidelines on their practical use.

References

Aitchison, Jean, Alan Gilchrist and David Bawden. 2000. *Thesaurus Construction and Use: A Practical Manual.* 4th ed. London: Aslib.

Baker, Thomas, Sean Bechhofer, Antoine Isaac, Alistair Miles, Guus Schreiberand Ed Summers. 2013. "Key Choices in the Design of Simple Knowledge Organization System (SKOS)." *Web Semantics: Science, Services and Agents on the World Wide Web* 20:35-49.

Berners-Lee, Tim, James Hendler and Ora Lassila. 2001. "The Semantic Web." *Scientific American* 284, no. 5:34-43.

Binding, Ceri. 2014. "Vocabularies in a Useful Form." Blog posting for SENESCHAL project. http://www.heritagedata.org/blog/vocabularies-in-a-useful-form/

Binding, Ceri and Douglas Tudhope. 2010. "Terminology Web Services." *Knowledge Organization* 37:287-98.

Binding, Ceri and Douglas Tudhope. 2015. "Improving Interoperability using Vocabulary Linked Data." *International Journal on Digital Libraries.* 17 no. 1:5-21. doi:10.1007/s00799-015-0166-y

Bizer, Christian, Tom Heath and Tim Berners-Lee. 2009. "Linked Data—The Story So Far." *International Journal on Semantic Web and Information Systems* 5, no. 3:1–22.

Caracciolo, Caterina, Armando Stellato, Sachit Rajbahndari, Ahsan Morshed, Gudrun Johannsen, Yves Jaques and Johannes Keizer. 2012. "Thesaurus Maintenance, Alignment and Publication as Linked Data: The AGROVOC Use Case." *International Journal of Metadata, Semantics and Ontologies* 7, no. 1:65-75.

Caracciolo, Caterina, Armando Stellato, Ahsan Morshed, Gudrun Johannsen, Sachit Rajbahndari, Yves Jaques and Johannes Keizer. 2013. "The AGROVOC Linked Dataset." *Semantic Web* 4:341-48.

Charles, Valentine and Cécile Devarenne. 2014. "Europeana Enriches its Data with the Art and Architecture Thesaurus." Europeana Data Model Case Study. http://pro.europeana.eu/page/europeana-aat

Charno, Michael. 2014. "SENESCHAL Vocabularies: Implementation." Blog post. http://archaeologydataservice.ac.uk/blog/2014/04/seneschal-vocabularies-implementation/

de Boer, Victor, Jan Wielemaker, Judith van Gent, Michiel Hildebrand, Antoine Isaac, Jacco van Ossenbruggen, and Guus Schreiber. 2012. "Supporting Linked Data Production for Cultural Heritage Institutes: The Amsterdam Museum Case Study." In *The Semantic Web: Research and Applications, 9th Extended Semantic Web Conference: ESWC 2012, Heraklion, Crete, Greece, May 27-31, 2012, Proceedings*, edited by Elena Simperl, Philipp Cimiano, Axel Polleres, Oscar Corcho, Valentina Presutti. Lecture Notes in Computer Science 7295. Berlin: Springer, pp. 733-47.

Dextre Clarke, Stella. 2016. "Origins and Trajectory of the Long Thesaurus Debate." *Knowledge Organization* 43: 138-44.

Golub, Koraljka, Douglas Tudhope, Marcia Zeng and Maja Žumer. 2014. "Terminology Registries for Knowledge Organization Systems: Functionality, Use, and Attributes." *Journal of the Association for Information Science and Technology* 65: 1901-16.

Haslhofer Bernhard, Flávio Martins and João Magalhães. 2013. "Using SKOS Vocabularies for Improving Web Search." *Proceedings of the 22nd international conference on World Wide Web*. New York: ACM, pp. 1253-58.

Heritage Data. 2016. "About Heritage Data." *Linked Data Vocabularies for Cultural Heritage* (blog). http://www.heritagedata.org/blog/

International Organization for Standardization. 2011. ISO 25964-1: *Information and Documentation—Thesauri and Interoperability with other Vocabularies - Part 1: Thesauri for Information Retrieval*. Geneva: International Organization for Standardization.

International Organization for Standardization. 2013. ISO 25964-2: *Information and Documentation—Thesauri and Interoperability with other Vocabularies - Part 2: Interoperability with other Vocabularies*. Geneva: International Organization for Standardization.

ISKO-UK. 2015. "The Great Debate: The Traditional Thesaurus vs Modern IR." London: ISKO-UK Annual General Meeting 19 February 2015, 14:00-17:30. http://www.iskouk.org/content/great-debate

Isaac, Antoine, William Waites, Jeff Young and Marcia Zeng, eds. 2011. *Library Linked Data Incubator Group: Datasets, value vocabularies, and metadata element sets* [W3C Incubator Group Report, October 25, 2011]. http://www.w3.org/2005/Incubator/lld/XGR-lld-vocabdataset/

Isaac, Antoine and Thomas Baker. 2015. Linked Data Practice at Different Levels of Semantic Precision: The Perspective of Libraries, Archives and Museums. *Bulletin of the Association for Information Science and Technology* 41 no. 4: 34-39. https://www.asis.org/Bulletin/Apr-15/AprMay15_Isaac_Baker.html

Kempf, Andreas Oskar and Joachim Neubert. 2016. "The Role of Thesauri in the Open Web: A Case Study of the STW Thesaurus for Economics." *Knowledge Organization* 43 no. 3: 160-73.

Lappalainen, Mikko, Matias Frosterus and Susanna Nykyri. 2014. "Reuse of Library Thesaurus Data as Ontologies for the Public Sector." *IFLA WLIC 2014, 16-22 August 2014, Lyon, France: Libraries, Citizens, Societies: Confluence for Knowledge,* Session. http://library.ifla.org/819/1/086-lappalainen-en.pdf

Milstead, Jessica L. 1995. "Invisible Thesauri: The Year 2000." *Online and CD-Rom Review* 19 no. 2: 93-94.

Pollitt, A. Steven. 1997. "Interactive information retrieval based on faceted classification using views." In *Knowledge Organization for Information Retrieval: Proceedings of the Sixth International Study Conference on Classification Research, Held at University College London, June 16-18, 1997*, organized by FID/CR and University College London in association with Aslib, the Classification Research Group and the International Society for Knowledge Organization. FID Publication 716. Würzburg: Ergon Verlag, pp. 51-56.

Soergel, Dagobert. 1994. "Indexing and Retrieval Performance: The Logical Evidence." *Journal of the American Society for Information Science* 45: 589-99.

Souza, Renato, Douglas Tudhope and Maurício Barcellos Almeida. 2012. "Towards a Taxonomy of KOS: Dimensions for Classifying Knowledge Organization Systems." *Knowledge Organization* 39: 179-92.

Stiller, Juliane, Vivien Petras, Maria Gäde and Antoine Isaac. 2014. "Automatic Enrichments with Controlled Vocabularies in Europeana: Challenges and Consequences." In *Digital Heritage: Progress in Cultural Heritage: Documentation, Preservation, and Protection: 5th International Conference, EuroMed 2014, Limassol, Cyprus, November 3-8, 2014. Proceedings*, ed. Marinos Ioannides, Nadia Magnenat-Thalmann, Eleanor Fink, Roko Žarnić, Alex-Yianing Yen and Ewald Quak. Lecture Notes in Computer Science 8740. Cham: Springer, pp. 238-47.

Suominen, Osma, Sini Pessala, Jouni Tuominen, Mikko Lappalainen, Susanna Nykyri, Henri Ylikotila, Matias Frosterus and Eero Hyvönen. 2014. "Deploying National Ontology Services: From ONKI to Finto." In *SEMWEB (2014)*. http://seco.cs.aalto.fi/publications/2014/suominen-et-al-deploying-onki-finto-2014.pdf

Tudhope, Douglas, Ceri Binding, Dorothee Blocksand Daniel Cunliffe. 2006. "Query Expansion via Conceptual Distance in Thesaurus Indexed Collections." *Journal of Documentation* 62: 509-33.

Yee, Ka-Ping, Kirsten Swearingen, Kevin Li and Marti Hearst. 2003. "Faceted Metadata for Image Search and Browsing." In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. New York: ACM, pp. 401-08.

Zeng, Marcia and Lois Chan. 2004. "Trends and Issues in Establishing Interoperability Among Knowledge Organization Systems." *Journal of the American Society for Information Science and Technology* 55:377-95.