



NEW WORK ITEM PROPOSAL 25964	
Date of presentation 2007-4-30	Reference number (to be given by the Secretariat)
Proposer 2007-07-30	ISO/TC 46 / SC 9 N 458
Secretariat SCC (Canada)	

Title of proposal (in the case of an amendment, revision or a new part of an existing document, show the reference number and current title)	
English title	Information and documentation - Guidelines for the establishment and development of thesauri [revision of ISO 2788 and ISO 5964]
French title (if available)	Information et documentation - Guide pour l'établissement et développement de thesaurus
Scope of proposed project The proposal is to update and extend the scope of these two intimately linked standards, to reflect the information environment of the 21 st century. All of their existing scope would be retained and revised, and the following additional subjects would be added: <ul style="list-style-type: none">- guidance on electronic functions and displays;- functional specification for software to manage thesauri;- guidelines for certain additional types of vocabulary (e.g. classification schemes, taxonomies);- interoperability (mapping) between vocabularies;- formats for exchange of thesaurus data. Additional details concerning the intended scope of NP 25964 are given in Annex 1 of this proposal.	
Concerns known patented items (see ISO/IEC Directives Part 1 for important guidance) <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No If "Yes", provide full information as annex	
Envisaged publication type (indicate one of the following, if possible) <input checked="" type="checkbox"/> International Standard <input type="checkbox"/> Technical Specification <input type="checkbox"/> Publicly Available Specification <input type="checkbox"/> Technical Report	
Purpose and justification (attach a separate page as annex, if necessary) The current versions of ISO 2788 and ISO 5964 are dated 1986 and 1985, respectively. Their content is very out of date. Since these two standards were developed in the era of only printed publications, they do not address the needs of a networked, electronic environment. Furthermore, they do not cover other forms of controlled vocabulary such as taxonomies which are becoming increasingly popular and sometimes supplant thesauri, but for which an international standard does not yet exist. Access to the immense bodies of information interconnected by the Internet and corporate Intranets often demand interoperability between diverse forms of controlled vocabulary, and this can be achieved only by the establishment and widespread adoption of standards. The current drive towards establishing a "Semantic Web" renders all the more urgent this proposal to develop the means of interoperability between vocabularies. The revision and updating of ISO 2788 and ISO 5964 as new standard ISO 25964 should be of benefit to all sectors of society that make use of large collections of information in the private, public and voluntary sectors.	
Target date for availability (date by which publication is considered to be necessary) 2009-07-30 or earlier	
Proposed development track <input type="checkbox"/> 1 (24 months) <input checked="" type="checkbox"/> 2 (36 months - default) <input type="checkbox"/> 3 (48 months)	

Relevant documents to be considered <p>It is proposed to base this work on BS 8723, a five-part standard of which Part 1 (<i>Guide</i>) and Part 2 (<i>Thesauri</i>) are already published. Drafts for Public Comment for Part 3 (<i>Vocabularies other than thesauri</i>) and Part 4 (<i>Interoperability between vocabularies</i>) are scheduled for release in May 2007. A draft of Part 5 (<i>Exchange formats and protocols for interoperability</i>) is currently under committee discussion and may be circulated for public comment later in 2007.</p> <p>Extracts from the Foreword and Scope of these documents are presented in the attached Annex to this project proposal.</p> <p>Although all parts of BS 8723 have been prepared in the UK, the BSI committee welcomes input from other countries, which will enhance international applicability.</p>		
Relationship of project to activities of other international bodies		
Liaison organizations	Need for coordination with: <input type="checkbox"/> IEC <input type="checkbox"/> CEN <input type="checkbox"/> Other (please specify)	
Preparatory work (at a minimum an outline should be included with the proposal) <input type="checkbox"/> A draft is attached <input checked="" type="checkbox"/> An outline is attached. It is possible to supply a draft [of all Parts of BS 8723, except Part 5], immediately The proposer or the proposer's organization is prepared to undertake the preparatory work required <input checked="" type="checkbox"/> Yes <input type="checkbox"/> No		
Proposed Project Leader (name and address) Stella Dextre Clarke Luke House, West Hendred, Wantage, Oxon OX12 8RR United Kingdom	Name and signature of the Proposer (include contact information) Sheeba Mukadam, BSI	
Comments of the TC or SC Secretariat Supplementary information relating to the proposal <input type="checkbox"/> This proposal relates to a new ISO document; <input checked="" type="checkbox"/> This proposal relates to the amendment/revision of two existing ISO documents (ISO 2788 and ISO 5964); <input type="checkbox"/> This proposal relates to the adoption as an active project of an item currently registered as a Preliminary Work Item; <input type="checkbox"/> This proposal relates to the re-establishment of a cancelled project as an active project. Other:		
Voting information The ballot associated with this proposal comprises a vote on: <input checked="" type="checkbox"/> Adoption of the proposal as a new project <input type="checkbox"/> Adoption of the associated draft as a committee draft (CD) (see ISO Form 5, question 2.3.1) <input type="checkbox"/> Adoption of the associated draft for submission for the enquiry vote (DIS or equivalent) (see ISO Form 5, question 2.3.2) Other:		
Annex(es) are included with this proposal (give details) <input checked="" type="checkbox"/> Extracts from British Standard BS 8723 are attached, describing the proposed scope of NP 25964 in greater detail. NOTE - Some parts of BS 8723 are already published; others are still in development.		
Date of circulation 2007-04-30	Closing date for voting 2007-07-30	Signature of the TC or SC Secretary [Original signed by Jane Thacker]

Outline of New Work Item Proposal 25964 concerning the revision of ISO 2788 and ISO 5964 on Guidelines for the establishment and development of thesauri

The proposal is to base the revision of ISO 2788:1986 and ISO 5964:1985 on British Standard BS 8723, *Structured vocabularies for information retrieval*.

BS 8723 is a new five-part standard intended to revise and extend BS 5723 and BS 6723. These British Standards are/were identical to ISO 2788 and ISO 5964 respectively, and BS 5723 has already been withdrawn.

To clarify the scope of the revision and extension, this Annex presents extracts from the parts of BS 8723 already published, and from drafts of the parts still under development.

BS 8723-1, *Structured vocabularies for information retrieval – Guide*

Foreword

This part of BS 8723 has been prepared by Subcommittee IDT/2/2. Together with BS 8723-2 it supersedes BS 5723:1987, which is withdrawn. When published, BS 8723-4 will revise and supersede BS 6723:1985 which will then be withdrawn. The other parts of BS 8723 will cover new scope, not previously published in a British Standard.

The other parts of BS 8723 are as follows:

- *Part 2: Thesauri;*
- *Part 3: Vocabularies other than thesauri;*
- *Part 4: Interoperation between vocabularies* [renamed *Interoperability between vocabularies*];
- *Part 5: Interoperation between vocabularies and other components of information storage and retrieval systems* [now renamed *Exchange formats and protocols for interoperability*].

A single index is provided, published with BS 8723-2.

Part 2 covers thesauri, designed for situations in which human indexers analyse documents and express their subjects using thesaurus terms, before searchers retrieve the documents with the same vocabulary.

Part 3 covers other types of structured vocabulary, including those for applications such as automatic categorization schemes and free-text search aids.

Part 4 applies to situations in which more than one language or vocabulary is in use, but access to all resources is needed through the one vocabulary chosen by the user.

Part 5 sets out the protocols and formats needed for the exchange of vocabulary data.

Part 1 together with part 2 of BS 8723 correspond broadly to ISO 2788-1986, since this is equivalent to the withdrawn BS 5723. Part 1 together with part 4 will correspond to ISO 5964:1985 (equivalent to BS 6723:1985), while extending the scope of the latter considerably. ISO 5964:1985 will remain current as an international Standard.

BS 8723-1 covers the development and maintenance of structured vocabularies rather than how to use them in indexing, which is covered by BS ISO 999.

Scope

This part of BS 8723 sets out the definitions, symbols and abbreviations and other conventions for information retrieval applications of structured vocabularies.

The BS 8723 series gives recommendations for the development and maintenance of structured vocabularies of subject terms and concepts for information retrieval. It also gives guidance on the exchange of data between multiple vocabularies and between a vocabulary and the software with which it interacts.

BS 8723 applies to vocabularies used for retrieving information from text resources, such as knowledge bases, bibliographic databases or collections of full text documents, and also catalogues of images or artefacts, such as those in a museum collection. Structured vocabularies are used in post-coordinate retrieval systems, hierarchical directories, pre-coordinate indexes and classification systems.

BS 8723 does not apply to the preparation of back-of-the-book indexes, although many of its recommendations may be useful for that purpose. It does not cover the structure of authority files of proper names of people, organizations and places, while recognizing that these may sometimes interact with vocabularies of subject terms. [The exclusion of authority files is currently being reconsidered.]

BS 8723-2, Structured vocabularies for information retrieval – Part 2: Thesauri

Scope

This part of BS 8723 gives recommendations for the development and maintenance of thesauri intended for information retrieval applications. It applies to vocabularies used for retrieving information from text resources, such as knowledge bases, bibliographic databases or collections of full text documents, and also catalogues of images or artefacts, such as those in a museum collection.

This part of this British Standard applies to multilingual and monolingual thesauri, but does not cover recommendations specific to multilingual situations, which are given in BS 8723-4.

The BS 8723 series does not apply to the preparation of back-of-the-book indexes, although many of its recommendations may be useful for that purpose. It does not cover

the structure of authority files of proper names of people, organizations and places, while recognizing these may sometimes interact with vocabularies of subject terms.

BS 8723-3, *Structured vocabularies for information retrieval – Part 3: Vocabularies other than thesauri* [Draft, dated 2007-02-26]

Scope

This part of BS 8723 gives recommendations and guidance on some types of vocabulary, other than thesauri, that are used for information retrieval. It does not give comprehensive guidance on vocabulary construction but focuses on the most common features of the vocabulary types. Where the principal application of the vocabulary is information retrieval (as is the case for classification schemes, subject heading schemes and taxonomies) recommendations are given for content, structure and some additional aspects. Where the principal application is for another function, but the vocabulary may also be applied to information retrieval (as is the case with many ontologies), characteristic features of the vocabulary are described, and the only recommendations concern aspects that are useful, or which should be avoided, for purposes of information retrieval.

As authority lists are often applied to information retrieval, but not always to retrieval by subject, this standard does not give comprehensive normative advice but describes some characteristic features.

While information retrieval is the context, this standard is concerned only with the vocabularies themselves and does not include requirements for indexes, catalogues, directories or other information retrieval tools.

NOTE – Thesauri and classification schemes provide two basic models, respectively, for the organization of vocabularies useful in retrieval. As their features are employed in most of the other types of vocabulary, comparisons and contrasts with thesauri and classification schemes are provided among the descriptions of the other types in order to assist with interpretation of BS 8723-4.

BS 8723-4, *Structured vocabularies for information retrieval – Part 4: Interoperability between vocabularies* [Draft, dated 2007-02]

Scope

This part of BS 8723 gives recommendations and describes the requirements for structured vocabularies, when two or more of them are required to interoperate for the purposes of information retrieval. Since any one concept is typically represented by different terms or notations in different vocabularies, the emphasis is on establishing equivalences or other relationships between the various representations of each concept. The case of a multilingual thesaurus, although in one sense this is a single vocabulary, is

included within the scope as a special case of interoperability between vocabularies in different languages. There are no comparable provisions for multilingual classification schemes, since their notation applies equally across all languages and the need for equivalence relationships between terms does not arise.

This standard provides recommendations on the way relationships may be established across vocabularies, between the terms and other elements of those vocabularies, so that information may be retrieved from resources indexed with the different vocabularies. It also gives recommendations for managing and maintaining data about the terms and relationships.

BS 8723-5, *Structured vocabularies for information retrieval – Part 5: Exchange formats and protocols for interoperability* [Draft, dated 2006-09-14. Note: Scope is still subject to discussion and adjustment]

Scope

This part of BS 8723 gives guidance on the exchange of structured vocabulary data between different computer applications involved in information retrieval.

Specifically, it involves the interaction between a computer application acting as a server for a structured vocabulary (the server) and an application applying that structured vocabulary to a particular information retrieval problem (the client). It provides guidance only on the highest level of application protocols, and assumes the existence of lower level network services and protocols which will be required to support computer-to-computer communication over a network.

This discussion does not address the internal design of systems that provide multiple functions and require communication only between one part and another part of the same application (such as an indexing system that incorporates its own internal subject authority system with a structured vocabulary). Nor does it address the interaction of a computer application and a user including user interface issues such as commands, menus, windows, buttons, menus and screen displays (though some guidance on these issues has been provided in earlier parts of this standard).