

Structured vocabularies for information retrieval — Guide —

Part 1: Definitions, symbols and abbreviations

ICS 01.140.20

Committees responsible for this British Standard

The preparation of this British Standard was entrusted by Technical Committee, IDT/2, Information and documentation, to Subcommittee IDT/2/2, Indexes, filing and thesauri, upon which the following bodies were represented:

- British Association for Information and Library Education and Research
- CAB International
- CILIP — Chartered Institute of Library and Information Professionals
- INSPEC
- Society of Indexers
- Stationery Office
- Co-opted members

This British Standard was published under the authority of the Standards Policy and Strategy Committee on 1 November 2005

© BSI 1 November 2005

First published November 2005

Amendments issued since publication

Amd. No.	Date	Comments

The following BSI references relate to the work on this British Standard:
Committee reference IDT/2/2
Draft for comment 04/30086620

ISBN 0 580 46798 8

Contents

	Page
Committees responsible	Inside front cover
Foreword	ii
Introduction	1
1 Scope	1
2 Definitions	1
3 Symbols, abbreviations and other conventions	6
Bibliography	9
Table 1 — Symbols and abbreviations	7

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 Together with BS 8723-2 it supersedes BS 5723:1987, which is withdrawn. When published, BS 8723-4 will revise and supersede BS 6723:1985 which will then be withdrawn.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*¹⁾;
- Part 4: *Interoperation between vocabularies*¹⁾;
- Part 5: *Interoperation between vocabularies and other components of information storage and retrieval systems*¹⁾.

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.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

Compliance with a British Standard does not of itself confer immunity from legal obligations.

¹⁾ In preparation.

Summary of pages

This document comprises a front cover, an inside front cover, pages i and ii, pages 1 to 9 and a back cover.

The BSI copyright notice displayed in this document indicates when the document was last issued.

Introduction

Today's information retrieval systems are able to provide many facilities. As well as thesaural guidance they may offer free text searching, filtering by date or by form, automatic clustering and categorization, automatic summarization, identification of names of people and organizations, etc. While aiming to mesh with that diversified environment, this standard focuses on retrieval by subject. Within that still broad scope, it is particularly concerned with the controlled vocabulary approach to indexing and retrieval.

The precursor standards BS 5723 and BS 6723 were concerned only with thesauri. Furthermore, they were developed in a largely paper-based era. BS 8723 acknowledges that the environment has changed and that new needs have emerged. Today's thesauri are mostly electronic tools; they are built and maintained with the support of software and need to integrate with other software, such as search engines and content management systems. Whereas in the past thesauri were designed for information professionals trained in indexing and searching, today there is a demand for vocabularies that untrained users will find to be intuitive. There is also a need for search aids in contexts where "full text" is not available, such as museum collections and image databases. As the Internet and other networks allow simultaneous searching across resource collections that have been indexed using different vocabularies, there is a need to have the means of "translating" search queries across boundaries.

BS 8723 still gives most attention to thesauri (part 2), including their use within computerized systems. It describes, in less detail (part 3), some of the other types of vocabulary with which thesauri may need to interface. Part 4 deals with interoperability between different vocabularies — either different language versions of one multilingual thesaurus, or multiple vocabularies in one natural language. Part 5 presents the formats and protocols needed for data exchange in support of thesauri and other structured vocabularies.

1 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.

2 Definitions

For the purposes of this British Standard, the following definitions apply. Where a term used in a definition is also defined in this clause, the term is italicized.

2.1

array

group of *sibling terms*

EXAMPLE In the following, **outerwear** and **underwear** are *sibling terms* in the same array.

clothing
 outerwear
 overcoats
 underwear