DD 8723-5:2008

DRAFT FOR DEVELOPMENT

Structured vocabularies for information retrieval – Guide –

Part 5: Exchange formats and protocols for interoperability

ICS 35.040

This publication is not to be regarded as a British Standard.

It is being issued in the Draft for Development series of publications and is of a provisional nature because the model and schema in Clauses **5** and **6** are considered to need additional user testing with full-scale applications before they can be regarded as stable. No existing format capable of conveying thesauri that conform to BS 8723-2 was identified in widespread use, therefore original work was undertaken to develop the format recommended herein. DD 8723-5 should be applied on this provisional basis so that information and experience of its practical application can be obtained.

A review of this Draft for Development will be carried out not later than 2 years after its publication.

Notification of the start of the review period, with a request for the submission of comments from users of this Draft for Development, will be made in an announcement in the appropriate issue of Update Standards.

According to the replies received, the responsible BSI Committee will judge whether the Draft for Development can be converted into a British Standard or what other action should be taken.

Observations which it is felt should receive attention before the official call for comments will be welcomed.

These should be sent to the Secretary of BSI Subcommittee IDT/2/2 at British Standards House, 389 Chiswick High Road, London W4 4AL.

Publishing and copyright information

ISBN 978 0 580 53864 3

The following BSI references relate to the work on this standard:
Committee reference IDT/2/2

Publication history

First published July 2008

Amendments issued since publication

Amd. no.	Date	Text affected

Contents

List of tables

Summary of pages

This document comprises a front cover, an inside front cover, pages i to iv, pages 1 to 56, an inside back cover and a back cover.

Foreword

Publishing information

This part of BS 8723 is published by BSI and came into effect on 31 July 2008. It was prepared by Subcommittee IDT/2/2, *Information description, source identification and indexing*, under the authority of Technical Committee IDT/2, *Information and documentation*. A list of organizations represented on these committees can be obtained on request to their secretary.

Relationship with other publications

BS 8723 comprises five parts:

- Part 1: *Definitions, symbols and abbreviations*;
- Part 2: *Thesauri*;
- Part 3: *Vocabularies other than thesauri*;
- Part 4: *Interoperability between vocabularies*;
- Part 5: *Exchange formats and protocols for interoperability.*

Part 1 provides definitions, symbols, abbreviations and other conventions applying to all the parts.

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, especially those used in selecting terms or codes for use in subject metadata.

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 deals with the protocols and formats needed for the exchange of vocabulary data, and is being issued as a Draft for Development in the first instance.

Use of this document

As a guide, this part of BS 8723 takes the form of guidance and recommendations. It should not be quoted as if it were a specification and particular care should be taken to ensure that claims of compliance are not misleading.

Any user claiming compliance with this part of BS 8723 is expected to be able to justify any course of action that deviates from its recommendations.

Presentational conventions

The provisions in this standard are presented in roman (i.e. upright) type. Its recommendations are expressed in sentences in which the principal auxiliary verb is “should”.

Commentary, explanation and general informative material is presented in smaller italic type, and does not constitute a normative element.

Contractual and legal considerations

Liability. The British Standards Institution ("BSI") shall not be liable, to the fullest extent permitted by law, for loss suffered by any party as a result of the use or application of the Draft for Development including loss of profits, loss of business, loss of contract, loss or corruption of data or information, any special, indirect, consequential or pure economic loss, costs, damages, charges or expenses. This exclusion shall not extend to death or personal injury caused by negligence. **Users of the Draft for Development should ensure that its application is entrusted to those who are appropriately qualified and experienced.**

Intellectual property rights. The British Standards Institution ("BSI") gives no warranty that the contents of this Draft for Development do not breach the intellectual property rights, including but not limited to the copyright, design rights and/or patent rights of any party and does not accept any liability, to the fullest extent permitted by law, for any claim that the content of the Draft for Development breaches such third party rights. **Users of the Draft for Development should satisfy themselves with regard to any such rights.**

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

Compliance with a Draft for Development cannot confer immunity from legal obligations.

Introduction

Widespread electronic networks and the almost universal use of computers in information management present opportunities for sharing data and information on a scale not previously imaginable: from a single computer connected to the internet one can simultaneously search different collections of books, articles, photographs, data sets, or museum artefacts residing on remote computers all over the world. The application of standards and facilities for interoperability makes it easier to search any of those collections by subject, where the subject terms used have been drawn from a structured vocabulary.

In the round, interoperability is encouraged by the following components:

- network infrastructure;
- formats for data exchange;
- protocols, including basic computer communication protocols as well as those applying specifically to vocabularies;
- registries of vocabularies, metadata and other common tools;
- robust data architecture.

While all are needed in the context of information retrieval, this part of BS 8723 addresses only the formats and protocols specific to exchange of vocabulary information.

Sometimes the functions of managing a structured vocabulary and making information retrievable using that vocabulary are integrated into a single system. However, information services are often provided by different computer applications running on different computers that cooperate to meet the needs of a user. For example, in order to allow a user to perform a federated search of several different databases with a single search strategy, an information retrieval application might need to translate search terms from a thesaurus used in one database into the equivalent terms used in another. In order to perform this translation, the search application usually needs to consult a structured vocabulary application running on another computer at a different location on the network; the two applications have to communicate and cooperate in order to provide this federated search service to the user.

The goal of accessing information quickly, easily and effectively therefore depends not just on common approaches to defining thesauri and to managing them within a single computer system, as described in earlier parts of BS 8723; it also depends on common approaches that promote interoperability of different computer applications when sharing the responsibility for management and use of structured vocabularies. Included among these common approaches are the way in which the data can be requested by one computer application and delivered by another (i.e. service protocols), and the presentation of the data requested (i.e. exchange formats). This part of BS 8723 addresses these two aspects of interoperability.

At the time of publication of this Draft for Development, networking technology is advancing rapidly. Future editions are likely to provide for enhanced interoperability functions.

1 Scope

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

It recommends a standard format for exchange of whole thesauri or subsets thereof. The format is applicable to thesauri that conform to BS 8723-2, utilizing any or all of the optional features described therein. Equally it supports multilingual thesauri displaying structural unity as described in BS 8723-4:2007, **4.2**. It does not address the additional needs of the vocabulary types described in BS 8723-3, although in some cases it may be adapted to serve such vocabularies. It does not address the exchange of data providing mappings between vocabularies that do not share the same structure.

This part of BS 8723 also considers 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 considers only the highest level of protocol, 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 part of BS 8723 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, and screen displays (although some guidance on these issues is provided in earlier parts of BS 8723).

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

BS 8723-1:2005, *Structured vocabularies for information retrieval – Guide – Part 1: General*

BS 8723-2:2005, *Structured vocabularies for information retrieval – Guide – Part 2: Thesauri*

BS 8723-3:2007, *Structured vocabularies for information retrieval – Guide – Part 3: Vocabularies other than thesauri*

BS 8723-4:2007, *Structured vocabularies for information retrieval – Guide – Part 4: Interoperability between thesauri*

BS ISO 639-1, *Codes for the representation of names of languages – Part 1: Alpha-2 code*

ISO 8601:2004, *Data elements and interchange formats – Information interchange – Representation of dates and times*

RFC 4646, *Tags for identifying languages*. The Internet Society. Best Current Practice recommendation, 2005.[1)]

[1)] Available at http://www.rfc-editor.org/rfc/rfc4646.txt.

3 Definitions

For the purposes of this part of BS 8723, the definitions in BS 8723-1, BS 8723-3, BS 8723-4 and the following apply. Where a term used in a definition is defined here or in the other parts, that term is italicized.

3.1 client
initiating *computer application* in a network interaction

NOTE The client application is the application that usually interacts with the human user.

3.2 computer application
computer program or set of programs that provides high-level processing related to a specific user need

NOTE In DD 8723-5, a computer application is sometimes referred to as an application.

3.3 data model
abstract model that describes how data is represented and used

3.4 exchange format
convention concerning the structure and semantics of data to be transferred between one *computer application* and another

NOTE For example, an exchange format might specify that a broader term should be preceded by a label BT and a specified punctuation symbol, and followed by an end of line marker. The layout and sequence of elements is the structure, and the label BT indicates the semantics of the information being exchanged.

3.5 federated search
simultaneous search of multiple information resources using a single search strategy and a common user interface

NOTE Also known as metasearch or cross-database search.

3.6 markup
annotations or other type of encoding embedded in text, in conformity with a *markup language*

3.7 markup language
set of encoding conventions that can be used to provide instructions for the interpretation of a text, by the use of annotations embedded in the text itself

NOTE The interpretation may concern issues such as content, structure or rendering of the text. Widely used examples include HTML (Hypertext Markup Language) [1], which is largely concerned with presentation, and XML (eXtensible Markup Language) [2–4], which addresses the structure of text.

3.8 protocol
convention that defines the syntax, semantics and synchronization of the communication process between two computers in order to enable a particular service

3.9 server
responding *computer application* in a network interaction

NOTE The server of a structured vocabulary is a vocabulary server.

4 Functions involving vocabularies that need interoperability

All of the functions described in earlier parts of BS 8723 as applying within a single system should be supported in a networked environment. (These are described variously in BS 8723-2:2005, Clause **11**, **13.5** and Clause **14**; BS 8723-3:2007, **6.8** and **6.9**; and BS 8723-4:2007, Clause **11**.) More specifically, there can be a need to:

- publish the whole vocabulary or part of a vocabulary;
- publish vocabulary updates;
- search a collection with direct human interaction;
- search a collection without direct human interaction;
- browse a collection with direct human intervention;
- cross-search multiple collections using different vocabularies;
- search a collection of metadata harvested from different sources each using a different vocabulary;
- manually index or classify documents in a collection;
- automatically index or classify documents in a collection.

In a networked environment, a vocabulary server commonly has to send data to a variety of clients, such as a search engine, an indexing system or a vocabulary management system. To support the range of functions enumerated in this clause, the design of an exchange format and protocol should take account of the fact that different combinations of client and vocabulary server applications are liable to apportion the processing burden differently between them. At one extreme, all processing functions are provided by the server prior to transmission of the data, and the client merely displays the result. At the other, the server merely serves the raw data and all further processing is undertaken by the client.

For example, when a thesaurus is published on the web for display via a standard web browser, the server commonly sends the data with presentational markup in Hypertext Markup Language (HTML) [1] and the client passively displays the result. In this case, what the user sees (including any hyperlinks) is dependent entirely upon the HTML generated by the server; the client browser needs only to render the HTML it receives and often performs no additional processing. The server-generated HTML is needed simply to ensure that the thesaurus data are formatted for visual presentation in accordance with the guidelines provided elsewhere in BS 8723, particularly BS 8723-2:2005, Clause **10** and BS 8723-4:2007, Clause **10**, and structural markup is largely redundant. Optionally, some client applications add value by utilizing both presentational and structural markup.

In other cases, the client application (whether browser-based or otherwise) needs to add considerable functionality of its own to the data received from the server. For instance, a search application might offer query expansion facilities which require it to interrogate either the cached data on the client or the original data on the server, or both, in order to determine and display broader, narrower and related terms or any preferred term corresponding to the user's entry term. Client applications such as automatic indexing systems need not only to interrogate the server data, but also to undertake the text processing needed to select index terms for particular documents. A description of the latter type of processing is outside the scope of this part of BS 8723, but any presentational markup in the server data is largely irrelevant in this context.

In addition, some of the interactions in a network can be highly interactive (e.g. switching terms at the time of a search) and some are primarily a batch transaction (e.g. sending a complete thesaurus in electronic form to another institution for loading into their system). Where batch functionality is required, the exchange format is important but protocols are likely to be very simple and based on existing communication and application protocols. For example, a network request is sometimes made via File Transfer Protocol [5] for a file that contains a complete copy of the thesaurus. If exchange does not go through an electronic network, a protocol might not be required at all: the request could be made by telephone and a copy of a thesaurus sent on CD-ROM through the postal system, then loaded on another computer system. A standard format should be used for the data.

In order to support these uses and achieve interoperability with a wide range of thesauri and classification schemes, the lower level functions listed in items a) to c) below are required. Not all of these functions apply to all contexts. For example:

- the ability to specify a language is not required to provide access to a monolingual thesaurus;
- in the absence of mappings between one structured vocabulary and another, a request to switch a term is irrelevant.

However, an exchange format or protocol which supports only a subset of these functions is less likely to be widely implemented and hence is less likely to encourage interoperability.

In summary, as many as possible of the following functions should be supported.

a) **Publish or exchange a vocabulary**

- Export/import a whole vocabulary, or a subset which could include some selection of concepts, classes or categories and could include some or all of the information associated with each selected concept, class or category.

b) **Update a vocabulary publication**

- Export all changes made since the last edition of the vocabulary for (ideally) automatic import into another vocabulary system.

c) **Access a vocabulary for indexing or searching**

- Search for a particular concept in a structured vocabulary based on:
  - the full preferred term or non-preferred term in a specified language;
  - words from the preferred term or non-preferred term in a specified language;
  - character strings from the preferred term or non-preferred term in a specified language;
  - the identifier of the term, concept, class or category;
  - the classification code or notation.
- Retrieve for a given term, concept, class, etc., information directly related to that term, concept or class:
  - all information concerning the term, concept or class;
  - all information concerning the term, concept or class except internal housekeeping information;
  - all information related to a particular language (for a multilingual vocabulary);
  - equivalent terms in other languages;
  - other combinations of data elements, when needed to support anticipated use patterns;
  - mappings into other structured vocabularies;
  - alternative forms of a term (to support augmentation or expansion of a search term, possibly to take in hierarchically and/or associatively related terms as well as equivalents).
- Retrieve for a given term or class the context of that term or class:
  - the alphabetical context of a term in a particular language, including a specifiable number of terms that follow that term in the alphabetical sequence of terms and preferably a specifiable number of terms that precede that term in the alphabetical sequence;
  - the hierarchical (and possibly associative) context of a term or class, including terms or classes higher and lower in the hierarchy by a specifiable number of levels;
  - the classified context of a term, including a specified number of terms before or after the term in the classified sequence and, where applicable, any node labels pertaining to the terms displayed.
- Validate a particular term in a particular language.

5 Data model

5.1 General

This clause models the underlying structure of the data which may be conveyed in a thesaurus conforming to BS 8723-2. The XML schema in **6.2** is derived from this model and should be used when sending or receiving thesaurus data in electronic form. It is not intended for use internally within a computer system, although it might be found useful, in whole or part, in that situation too.

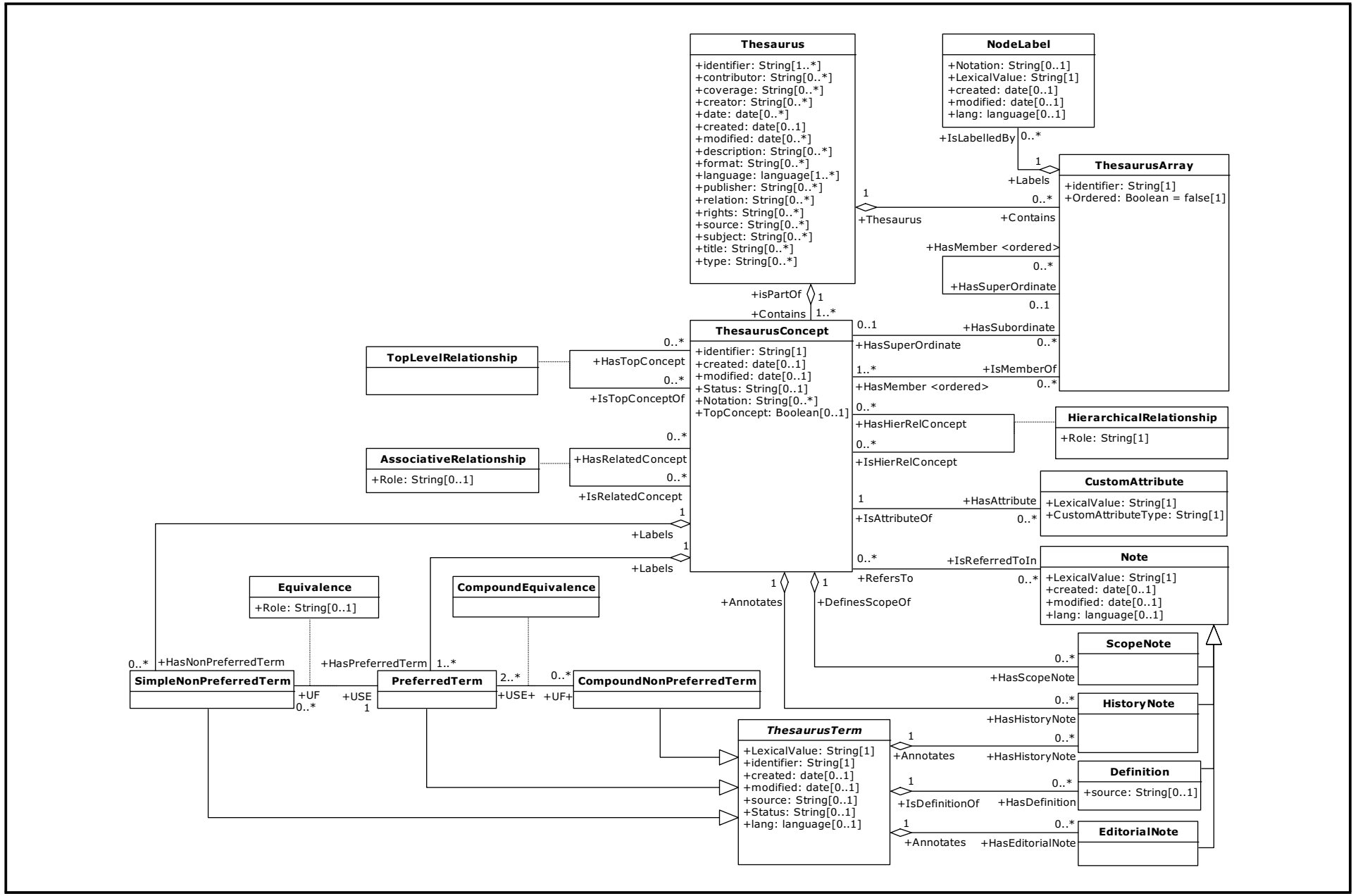
The basic recommendations of BS 8723-2 are quite simple, needing only a simple model. But BS 8723-2 also allows and sometimes encourages a number of optional features that require much greater complexity in the model. For thesaurus exchange to be successful in a wide range of situations, it is important to have a format hospitable to sophisticated as well as simple thesauri.

The model presented in the present part of BS 8723 accommodates the full range of options described in BS 8723-2. Some elements of this model are shown as optional, and need not be used when exchanging thesauri that lack the optional features. Software systems importing data may ignore optional elements that they do not support, and assume appropriate default values (which may be null) for optional elements that they do support but which are not present in the imported data, while interpreting the remaining elements consistently and correctly.

The model is presented first as a UML (Unified Modelling Language [6–8]) class diagram (Figure 1). UML is highly expressive and discourages ambiguity, but as some users might be unfamiliar with its conventions, an alternative tabular presentation is added at **5.3**. The data elements and attributes in the UML diagram and the data tables are identical in substance though some details of relationships are not made explicit in the tabular presentation. In case of doubt, the diagram should be regarded as the authority.

NOTE UML defines its own special conventions and terminology for modelling purposes. In the present clause, the term "class" is used in the UML sense, as a construct that is used to group related instance variables. Similarly, UML gives special meanings to the expressions "String", "Boolean" and "aggregation". More information about UML can be found from references in the Bibliography.

Figure 1 **UML model**



5.2 Notes on the model

*COMMENTARY ON **5.2***
*Conventions for naming and symbols are described first. Since the classes **Thesaurus**, **ThesaurusConcept** and **ThesaurusTerm** are fundamental to every thesaurus, their attributes and other features are described next. Some of the same attributes are found also in the classes **Note** and **NodeLabel**. (For definitions of "thesaurus", "concept", "term", "node label", etc., see BS 8723-1. For a fuller discussion of all of them, see BS 8723-2.) After these come brief notes on **ThesaurusArray** and **NodeLabel**. Attributes and associated features of the relationship classes are described last.*

*These notes apply to the tabular presentation in **5.3**, as well as to the diagram in Figure 1.*

5.2.1 Naming conventions

Some Dublin Core metadata elements [9] are shown in the model and in the derived XML schema. The names of these elements follow the Dublin Core lowerCamelCase convention. Other elements use UpperCamelCase in line with the standard naming convention for UK public sector schema development [10].

5.2.2 Symbols for associations

The diamond symbol in the diagram represents an aggregation. This symbol is used to indicate that in the derived XML schema the definition of an object will be nested in its parent object.

Normal associations (without a diamond) indicate a relationship of the type key/keyref in XML as described in XML documentation at http://www.w3.org/TR/xmlschema-1/#cIdentity-constraint_Definitions and http://www.w3.org/TR/xmlschema-0/#specifyingKeysAndtheirRefs (see [2–4] for the complete reference). Note that the XML schema shown in Clause **6**, Figure 2 does not actually enforce the key/keyref constraints. The reason for this relaxation of the constraints is that enforcement would have the effect of invalidating XML output files that contain only an extracted portion of the thesaurus rather than a whole thesaurus.

5.2.3 Identifier

This is an attribute of ***Thesaurus***, ***ThesaurusConcept***, ***ThesaurusTerm*** and ***ThesaurusArray***. Identifiers for terms and concepts are described in BS 8723-2:2005, **12.3.1.2** and **12.3.1.3**. For trouble-free use of the format described in Clause **6**, the application of both types of identifier is strongly recommended. A separate identifier for each array (see **5.2.5**) may also be applied.

5.2.4 Dates

The date created and date modified are optional attributes of ***Thesaurus***, ***ThesaurusConcept***, ***ThesaurusTerm***, ***Note*** and ***NodeLabel***. They should be given in YYYY-MM-DD format, in line with ISO 8601. In the case of ***Thesaurus***, another option is to use the simple attribute "date" without specifying whether it applies to the first edition or a subsequent version. The same format should be used.

5.2.5 Status

This is an optional attribute of ***ThesaurusConcept*** and ***ThesaurusTerm***. See BS 8723-2:2005, **12.3.1.2**b).

5.2.6 Notation

This is an optional attribute of ***ThesaurusConcept*** and ***NodeLabel***. See BS 8723-2:2005, **10.1.2** and **10.5.2**.

5.2.7 Source

This is an optional attribute of ***Thesaurus*** and ***ThesaurusTerm***. See BS 8723-2:2005, **12.3.1.2**a).

5.2.8 Language codes

Language is an optional attribute of ***ThesaurusTerm***, ***Note*** and ***NodeLabel*** and a mandatory attribute of ***Thesaurus***. It should be given as a code selected from BS ISO 639-1, extended where necessary with the additional codes in RFC 4646.

5.2.9 Dublin Core elements

The ***Thesaurus*** class has a number of attributes intended for assigning Dublin Core metadata [9] to a file using the BS 8723 format. As well as identifier, date, source and language already described, these include contributor, coverage, creator, description, format, publisher, relation, rights, subject, title and type. All except language and identifier are optional.

5.2.10 Multiplicity of *PreferredTerm*

The model allows more than one preferred term per concept, to provide for multilingual thesauri, where one is permitted per language. For monolingual thesauri the multiplicity is reduced to one.

5.2.11 Top level concepts

BS 8723-2:2005, **10.2**d) makes brief mention of "Top Term" as an optional feature of alphabetical displays and single term displays, linking a concept to the top concept of the hierarchy in which it occurs. The class ***TopLevelRelationship*** enables this.

BS 8723-2:2005, **14.8**b) recommends a browse capability that starts by presenting editors with a list of top terms. The attribute ***TopConcept*** addresses this, specifying that a concept is at the top of a hierarchy, i.e. it has no broader concepts. This makes it easier, on importing a thesaurus, to pick out all the top level concepts from which to build a navigational tree.

5.2.12 Note references

The association between ***Note*** and ***ThesaurusConcept*** enables any note for one concept to refer (using marked up text) to any other concept in the thesaurus. This capability is particularly useful for scope notes. See BS 8723-2:2005, **6.1.2.4**.

5.2.13 Custom attributes

The model includes a class ***CustomAttribute*** for custom attributes of the concept. This class enables recording of custom information about the concept such as the inclusion of the concept within a subset or micro-thesaurus, or a mapping to a related concept in another vocabulary, etc.

It is included as a separate class rather than as a normal attribute so that the administrator of the thesaurus management system can specify the values of custom attributes that can be assigned to a concept. The class has an attribute ***CustomAttributeType***, allowing the administrator to specify which type of attribute is being used.

5.2.14 Arrays and node labels

Although all thesauri have arrays of sibling terms, it is not always necessary to model them explicitly and so this class is optional. The ***ThesaurusArray*** and ***NodeLabel*** classes of the model are needed particularly to support the generation of displays incorporating node labels and/or non-alphabetic sequences of sibling terms. (See examples in BS 8723-2:2005, Figures 4, 6, 7 and 8; also BS 8723-4:2007, Figure 7 and **10.5.3**.) The attribute ***Ordered*** enables non-alphabetic sequences when necessary.

5.2.15 Relationship roles

The classes ***Equivalence***, ***HierarchicalRelationship*** and ***AssociativeRelationship*** each have a "role" attribute. This can be used to convey the common relationships USE/UF, BT/NT and RT/RT. It can also be used for the optional relationship types described in BS 8723-2:2005, **8.3**, as well as custom relationship types as suggested in BS 8723-2:2005, **8.5**. Each occurrence of a relationship is expressed in one direction only, between a "source" class and a "target" class. For example, "Concept A has a broader concept, namely Concept B", where the role is "broader". If the complementary relationship "Concept B has a narrower concept, namely Concept A" needs to be expressed, then this is shown as a separate occurrence with the role "narrower".

For custom relationship types, the text given in the "role" attribute should be composed of (a) the name of the parent relationship type followed by (b) the symbol forward slash "/" and finally (c) the name of the custom relationship type.

For example the text in the "Role" attribute of ***HierarchicalRelationship*** may be one of the following:

```
NT
NT/NTP
NT/NTI
NT/NTG
NT/NTI/NTX
BT
BT/BTP
BT/BTI
BT/BTG
BT/BTI/BTX
etc.
```

5.2.16 Compound non-preferred terms

The classes ***CompoundNonPreferredTerm*** and ***CompoundEquivalence*** enable representation of complex concepts by a combination of terms, as described in BS 8723-2:2005, **8.2.5**.

5.2.17 Role of relationship tags

Figure 1 shows a relationship between ***SimpleNonPreferredTerm*** and ***PreferredTerm***, with the tags USE and UF, which are described in BS 8723-2:2005, Clause **4** and **8.2** in the context of conventions and displays for human readers. In the context of data exchange, it could be argued that this relationship is redundant, since all the preferred and non-preferred terms for a given concept are linked to that concept. However, some redundancy is retained, to facilitate exchange of data for simpler thesauri and to make it easier to exchange update information and other specialized subsets of the data.

Another detail concerns USE…+. In displays for human readers, there is always a gap between "USE" and "+", to accommodate the term in between (e.g. Coal mining USE coal + mining). But if the tag is to be used behind the scenes by computers, it needs to be compressed, in the style coal mining USE+ coal; coal mining USE+ mining. So in the diagram, the gap is omitted.

5.3 Tabular presentation

The elements of the model shown in Figure 1 are listed in Tables 1 to 5, based on the five classes ***Thesaurus***, ***ThesaurusArray***, ***ThesaurusConcept***, ***ThesaurusTerm*** and ***Note***. Other subsidiary classes have been shown within these tables. The brief text in the description columns is complemented by the more extensive notes in **5.2**. See also definitions for many of the terms in BS 8723-1.

The Type column indicates the data type, using the UML data types "String" and "Boolean" where applicable, the XML data types "date" and "language" where no UML type is applicable, and "class" in the case of an initial or associated class in the model.

The occurrence columns describe whether the elements are mandatory and/or repeatable as shown in Table 6.

Table 1 **Attributes and associations of *Thesaurus***

Element	Type	Occurrence	Description
Thesaurus	class	1	The thesaurus as a whole
Attributes of Thesaurus			
identifier	String	1..*	Identifier for the thesaurus as a whole
contributor	String	0..*	Person or organization who contributed to the thesaurus
coverage	String	0..*	Spatial or temporal coverage of the thesaurus
creator	String	0..*	Person or organization primarily responsible for making the resource
date	date	0..*	Any date associated with the thesaurus
created	date	0..1	The date the thesaurus was created
modified	date	0..*	A date when this version was modified
description	String	0..*	An account of the thesaurus
format	String	0..*	The file format or physical medium of the thesaurus
language	language	1..*	Codes showing languages supported by the thesaurus
publisher	String	0..*	Entity responsible for publication
relation	String	0..*	A related publication
rights	String	0..*	Copyright or other rights information
source	String	0..*	Resource from which the thesaurus was derived
subject	String	0..*	Index terms indicating the subject content
title	String	0..*	Title of the thesaurus
type	String	0..*	The genre of the vocabulary, e.g. “thesaurus”
Associated classes of Thesaurus			
ThesaurusArray	class	0..*	An array of sibling concepts forming part of the thesaurus
ThesaurusConcept	class	1..*	A concept forming part of the thesaurus

Table 2 **Attributes and associations of *ThesaurusArray***

Element	Type	Occurrence	Description
ThesaurusArray	class	0..*	An array of sibling concepts forming part of the thesaurus
Attributes of ThesaurusArray			
identifier	String	1	A unique identifier for the array
Ordered	Boolean	1	A true/false label indicating whether the order of concepts in the array has to be maintained. Default is "false"
Associated classes of ThesaurusArray			
HasSuperOrdinate array	class	0..1	A higher-level array of which this array is a member
HasMember array	class	0..*	A lower-level array which is a member of this array *An array needs to have at least one member, but this can be a concept or another array.*
HasSuperOrdinate concept	class	0..1	A higher-level concept to which this array is subordinated
HasMember concept	class	0..*	A concept which is a member of this array *An array needs to have at least one member, but this can be a concept or another array.*
NodeLabel	class	0..*	A label showing the basis on which the concepts in the array have been chosen and ordered
♦ ***Attributes of NodeLabel***			
♦ Notation	String	0..1	A symbolic notation attached to the node label, which may be used for sorting and display
♦ LexicalValue	String	1	The wording of the node label
♦ created	date	0..1	The date when the node label was created
♦ modified	date	0..1	The date when the node label was last modified
♦ lang	language	0..1	A code showing the language of the node label. This should be included if the thesaurus supports more than one language

Table 3 **Attributes and associations of *ThesaurusConcept***

Element	Type	Occurrence	Description
ThesaurusConcept	class	1..*	A concept which is a member of the thesaurus
Attributes of ThesaurusConcept			
identifier	String	1	A unique identifier for the concept
created	date	0..1	The date when the concept was created
modified	date	0..1	The date when the concept was last modified
Status	String	0..1	Indication of whether the concept is candidate, approved, superseded, etc.
Notation	String	0..*	Symbolic notation attached to the concept, which may be used for sorting and display
TopConcept	Boolean	0..1	A true/false label indicating whether the concept is at the top of a hierarchy, i.e. has no broader concepts

Table 3 **Attributes and associations of *ThesaurusConcept*** (*continued*)

Element	Type	Occurrence	Description
Associated classes of ThesaurusConcept			
PreferredTerm	class	1..*	The term used as a label for this concept. There should be one preferred term per language
SimpleNonPreferredTerm	class	0..*	Alternative terms by which this concept could be sought
ScopeNote	class	0..*	A note defining or clarifying the scope of the concept within this thesaurus
HistoryNote	class	0..*	A note recording changes to this concept within this thesaurus
Note	class	0..*	A note of any kind which contains a reference to this concept although it applies more directly to another concept
IsMemberOf array	class	0..*	An array of which this concept is a member
HasSubordinate array	class	0..*	An array which is subordinate to this concept
CustomAttribute	class	0..1	An additional attribute of a concept
♦ ***Attributes of CustomAttribute***			
♦ LexicalValue	String	1	The wording of the custom attribute
♦ CustomAttributeType	String	1	Specification of particular kinds of attribute
TopLevelRelationship	association class		
♦ ***Classes linked by TopLevelRelationship***			
♦ HasTopConcept	class	0..*	The concept at the top of a hierarchy containing this concept
♦ IsTopConceptOf	class	0..*	Concept in the hierarchy of which this concept is top
HierarchicalRelationship	association class		
♦ ***Attributes of HierarchicalRelationship***			
♦ Role	String	1	Specification of the kind of hierarchical relationship that the target concept has to the source concept, e.g. BT, BTI, NT, NTP, etc.
♦ ***Classes linked by HierarchicalRelationship***			
♦ HasHierRelConcept	class	0..*	Target concept, linked to this source concept by the relationship specified in "Role", e.g. if Role is "BT", this is a link to a broader concept of this concept
♦ IsHierRelConcept	class	0..*	Source concept to which this target concept is linked by the relationship specified in Role, e.g. if Role is "BT", this is a link to the concept of which this is a broader concept
AssociativeRelationship	association class		
♦ ***Attributes of AssociativeRelationship***			
♦ Role	String	0..1	Specification of the kind of associative relationship that the target concept has to the source concept. This will usually be "RT", but the relationship could be asymmetric, e.g. "CAUSE" as discussed in BS 8723-2:2005, **8.5**
♦ ***Classes linked by AssociativeRelationship***			
♦ HasRelatedConcept	class	0..*	Target concept, linked to this source concept by the relationship specified in Role, e.g. if Role is "CAUSE", this is a link to a concept which is a cause of this concept
♦ IsRelatedConcept	class	0..*	Source concept to which this target concept is linked by the relationship specified in Role, e.g. if Role is "CAUSE", this is a link to the concept of which this concept is a cause

Table 4 **Attributes and associations of *ThesaurusTerm***

Element	Type	Occurrence	Description
ThesaurusTerm	class	1	A term in the thesaurus by which a concept could be sought
Attributes of ThesaurusTerm			
LexicalValue	String	1	The wording of the term
identifier	String	1	A unique identifier for the term
created	date	0..1	The date when the term was created
modified	date	0..1	The date when the term was last modified
source	String	0..1	The person(s) or document(s) from which the term was taken
Status	String	0..1	Indication of whether the term is candidate, approved, superseded, etc.
lang	language	0..1	A code showing the language of the term. This should be included if the thesaurus supports more than one language
Associated classes of ThesaurusTerm			
HistoryNote	class	0..*	A note recording changes to this term within this thesaurus
EditorialNote	class	0..*	A note for use by the thesaurus editors during the editing process
Definition	class	0..*	A note giving definitions of a term, not necessarily limited to the scope of the concept labelled by the term in this thesaurus
♦ ***Attributes of Definition***			
♦ source	String	0..1	The person(s) or document(s) from which the definition was taken
Specializations of ThesaurusTerm			
PreferredTerm	class		A term that has been chosen to label a concept
SimpleNonPreferredTerm	class		A term that serves as an alternative label for a concept, but which is not the preferred term
CompoundNonPreferredTerm	class		A term which represents a combination of two or more thesaurus concepts
♦ ***Associations of PreferredTerm and SimpleNonPreferredTerm***			
♦ Equivalence	association class		
♦ ***Attributes of Equivalence***			
♦ Role	class	0..1	Specification of kind of equivalence relationship. This will normally be USE, linking the source SimpleNonPreferredTerm to the target PreferredTerm
♦ ***Classes linked by Equivalence***			
♦ PreferredTerm	class	1	Target term, linked to this source term by the relationship specified in Role, e.g. if Role is "USE", this is a link to a PreferredTerm which is to be used in place of the current SimpleNonPreferredTerm
♦ SimpleNonPreferredTerm	class	0..*	Source term to which this target term is linked by the relationship specified in Role, e.g. if Role is "USE", this is a link to the SimpleNonPreferredTerm for which the current PreferredTerm should be used instead
♦ ***Associations of PreferredTerm and CompoundNonPreferredTerm***			
♦ CompoundEquivalence	association class		
♦ ***Classes linked by CompoundEquivalence***			
♦ CompoundNonPreferredTerm	class	0..*	A non-preferred term naming a complex concept which is represented by a combination of preferred terms. See BS 8723-2:2005, **8.2.5**
♦ PreferredTerm	class	2..*	One of the two or more PreferredTerms which together represent the CompoundNonPreferredTerm

Table 5 **Attributes and associations of *Note***

Element	Type	Occurrence	Description
Note	class	0..*	A piece of text giving additional information about the term or concept to which it is linked
Attributes of Note			
LexicalValue	String	1	The wording of the note
created	date	0..1	The date when the note was created
modified	date	0..1	The date when the note was last modified
lang	language	0..1	A code showing the language of the note. This should be included if the thesaurus supports more than one language
Associated classes of Note			
ThesaurusConcept	class	0..*	A concept referred to in the note, other than the one to which it is directly attached
Specializations of Note, for particular types			
ScopeNote	class		A note defining or clarifying the scope of the concept within this thesaurus
HistoryNote	class		A note recording changes to a concept or a term within this thesaurus
Definition	class		A note giving a definition of a term, not necessarily limited to the scope of the concept labelled by the term in this thesaurus
EditorialNote	class		A note for use by the thesaurus editors during the editing process

Table 6 **Definitions of occurrence as used in Tables 1 to 5**

Occurrence	Meaning
1	mandatory, not repeatable (one occurrence only)
1..*	mandatory, repeatable (one or more occurrences)
0..1	optional, not repeatable (zero or one occurrence)
0..*	optional, repeatable (zero or more occurrences)

6 Exchange formats

6.1 XML schema for thesaurus exchange

This clause presents a single format that is recommended for exchange of thesauri, independent of what sort of system has produced them or needs to import them. It is expressed as an XML (eXtensible Markup Language) [2, 3, 11, 12] schema, since this language is widely used and hence promotes interoperability. Furthermore it is system- and platform-independent, and through adoption of UNICODE [13] it supports most of the scripts and languages likely to be required for multilingual vocabularies. The format corresponds to the model described in Clause **5** and is designed to convey thesauri that comply with BS 8723-2, utilizing any or all of the optional features described therein. Equally it supports multilingual thesauri as described in the relevant clauses of BS 8723-4.

The format is also capable of conveying many thesauri that do not conform fully to BS 8723-2, but are broadly compatible with it. The extensibility options may be used where necessary to accommodate special features.

The following guidelines apply to its design.

- A complex type (see XML schema Parts 0 and 1 [2, 3]) has been defined for each UML class of the model.
- In line with guidelines (see [10]) for UK public sector schemas, the name of a complex type always terminates with "Structure".
- In line with guidelines for UK public sector schemas, the name of a simple type (see XML schema Parts 0 and 1 [2, 3]) always terminates with "Type".
- An element in the complex type has been defined for each attribute of a UML class.
- When relevant, the Dublin Core naming convention was followed for the elements.

In accordance with XML guidelines, a namespace for the schema has been established, at http://www.bsigroup.com/resources/standards/bs8723/.

Annex A shows some examples of thesaurus data encoded using the schema.

6.2 Exchanging a whole thesaurus

The XML schema shown in Figure 2 should be used. As explained in **5.2.2**, the schema as provided does not validate key/keyref links. Users may add this validation if it is deemed helpful.

6.3 Updating a thesaurus

An update may be supplied as a whole updated version, or as a subset comprising just the sections that are new or have changed in some way. In the event that the latter approach is used, it is often necessary to provide separate information about any terms or concepts that have been deleted. Data delivery should preferably be supplemented by a text explanation of the rationale for the update.

6.4 Exchanging thesaurus subsets

In principle, exchanges of at least four different types of thesaurus subset are often found to be needed. The format is applicable for exchanges of any of these types of subset:

a) single records, i.e. all the data associated with a single concept or term. Such subsets are trivially simple and may be conveyed using the appropriate elements from the recommended format, without the validations required for a whole thesaurus;

b) updates (see **6.3**);

c) responses in the course of live indexing/searching applications (see Clause **8**);

d) selections of data large and coherent enough to serve as an independent thesaurus.

Figure 2 **XML schema**

```
<xs:schema targetNamespace="http://www.bsigroup.com/resources/standards/bs8723/"
elementFormDefault="qualified" attributeFormDefault="unqualified">
  <xs:annotation>
    <xs:documentation>
LIABILITY
The British Standards Institution ("BSI") shall not be liable, to the fullest extent
permitted by law, for loss suffered by any party as a result of the use or application of
DD 8723-5 including loss of profits, loss of business, loss of contract, loss or corruption
of data or information, any special, indirect, consequential or pure economic loss, costs,
damages, charges or expenses. This exclusion shall not extend to death or personal injury
caused by negligence. Users of the Draft for Development should ensure that its application
is entrusted to those who are appropriately qualified and experienced.

IPR
The British Standards Institution ("BSI") gives no warranty that the contents of DD 8723-5
do not breach the intellectual property rights, including but not limited to the copyright,
design rights and/or patent rights of any party and does not accept any liability, to the
fullest extent permitted by law, for any claim that the content of the Draft for
Development breaches such third party rights. Users of the Draft for Development should
satisfy themselves with regard to any such rights.
    </xs:documentation>
  </xs:annotation>
  <xs:import namespace="http://www.govtalk.gov.uk/CM/gms" schemaLocation="DD8723-5-
  eGms.xsd"/>
  <xs:import namespace="http://purl.org/dc/terms/" schemaLocation="DD8723-5-
  DCTerms.xsd"/>
  <xs:import namespace="http://purl.org/dc/elements/1.1/" schemaLocation="DD8723-5-
  DC.xsd"/>
  <xs:import namespace="http://www.w3.org/XML/1998/namespace"
  schemaLocation="xml.xsd"/>
  <xs:element name="Thesaurus" type="ThesaurusStructure">
    <xs:annotation>
      <xs:documentation>Root element of the XML file containing the Thesaurus
      data.</xs:documentation>
    </xs:annotation>
    <xs:unique name="CustomAttributeIdentifier">
      <xs:annotation>
        <xs:documentation>Makes the identifier of the custom attributes
        unique.</xs:documentation>
      </xs:annotation>
      <xs:selector xpath="bs8723:CustomAttribute"/>
      <xs:field xpath="@dc:identifier"/>
    </xs:unique>
    <xs:unique name="ThesaurusConceptArrayIdentifier">
      <xs:annotation>
        <xs:documentation>Makes the identifier of the concepts and arrays
        unique.</xs:documentation>
      </xs:annotation>
      <xs:selector xpath="bs8723:ThesaurusConcept|bs8723:ThesaurusArray"/>
      <xs:field xpath="@dc:identifier"/>
    </xs:unique>
    <xs:unique name="ThesaurusConceptIdentifier">
      <xs:annotation>
        <xs:documentation>Makes the identifier of the concepts
        unique.</xs:documentation>
      </xs:annotation>
      <xs:selector xpath="bs8723:ThesaurusConcept"/>
      <xs:field xpath="@dc:identifier"/>
    </xs:unique>
```

Figure 2 **XML schema** (*continued*)

```
  <xs:unique name="ThesaurusArrayIdentifier">
    <xs:annotation>
      <xs:documentation>Makes the identifier of the arrays
      unique.</xs:documentation>
    </xs:annotation>
    <xs:selector xpath="bs8723:ThesaurusArray"/>
    <xs:field xpath="@dc:identifier"/>
  </xs:unique>
  <xs:unique name="ThesaurusTermIdentifier">
    <xs:annotation>
      <xs:documentation>Makes the identifier of the terms
      unique.</xs:documentation>
    </xs:annotation>
    <xs:selector
    xpath="bs8723:ThesaurusConcept/bs8723:PreferredTerm|bs8723:ThesaurusConcept/
    bs8723:NonPreferredTerm|bs8723:CompoundEquivalence/bs8723:CompoundNonPreferredTerm"/>
    <xs:field xpath="@dc:identifier"/>
  </xs:unique>
</xs:element>
<xs:complexType name="ThesaurusStructure">
  <xs:annotation>
    <xs:documentation>Data type representing the structure of the
    thesaurus</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element ref="dc:identifier" maxOccurs="unbounded"/>
    <xs:element ref="dc:contributor" minOccurs="0" maxOccurs="unbounded"/>
    <xs:element ref="dc:coverage" minOccurs="0" maxOccurs="unbounded"/>
    <xs:element ref="dc:creator" minOccurs="0" maxOccurs="unbounded"/>
    <xs:element ref="dc:date" minOccurs="0" maxOccurs="unbounded"/>
    <xs:element ref="dcterms:created" minOccurs="0"/>
    <xs:element ref="dcterms:modified" minOccurs="0" maxOccurs="unbounded"/>
    <xs:element ref="dc:description" minOccurs="0" maxOccurs="unbounded"/>
    <xs:element ref="dc:format" minOccurs="0" maxOccurs="unbounded"/>
    <xs:element ref="dc:language" minOccurs="0" maxOccurs="unbounded"/>
    <xs:element ref="dc:publisher" minOccurs="0" maxOccurs="unbounded"/>
    <xs:element ref="dc:relation" minOccurs="0" maxOccurs="unbounded"/>
    <xs:element ref="dc:rights" minOccurs="0" maxOccurs="unbounded"/>
    <xs:element ref="dc:source" minOccurs="0" maxOccurs="unbounded"/>
    <xs:element ref="dc:subject" minOccurs="0" maxOccurs="unbounded"/>
    <xs:element ref="dc:title" minOccurs="0" maxOccurs="unbounded"/>
    <xs:element ref="dc:type" minOccurs="0" maxOccurs="unbounded"/>
    <xs:element name="ThesaurusArray" type="ThesaurusArrayStructure" minOccurs="0"
    maxOccurs="unbounded">
      <xs:annotation>
        <xs:documentation>Array of the thesaurus.</xs:documentation>
      </xs:annotation>
    </xs:element>
    <xs:element name="ThesaurusConcept" type="ThesaurusConceptStructure"
    maxOccurs="unbounded">
      <xs:annotation>
        <xs:documentation>Concept contained in the Thesaurus</xs:documentation>
      </xs:annotation>
      <xs:unique name="TermIdentifier">
        <xs:annotation>
          <xs:documentation>Makes the identifier of the terms
          unique.</xs:documentation>
        </xs:annotation>
```

Figure 2 **XML schema** (*continued*)

```
      <xs:selector
      xpath="bs8723:ThesaurusConcept/bs8723:PreferredTerm|bs8723:ThesaurusConcept
      /bs8723:NonPreferredTerm"/>
      <xs:field xpath="@dc:identifier"/>
    </xs:unique>
    <xs:unique name="PreferredTermIdentifier">
      <xs:annotation>
        <xs:documentation>Makes the identifier of the preferred terms
        unique.</xs:documentation>
      </xs:annotation>
      <xs:selector xpath="bs8723:ThesaurusConcept/bs8723:PreferredTerm"/>
      <xs:field xpath="@dc:identifier"/>
    </xs:unique>
    <xs:unique name="NonPreferredTermIdentifier">
      <xs:annotation>
        <xs:documentation>Makes the identifier of the non-preferred terms
        unique.</xs:documentation>
      </xs:annotation>
      <xs:selector xpath="bs8723:ThesaurusConcept/bs8723:NonPreferredTerm"/>
      <xs:field xpath="@dc:identifier"/>
    </xs:unique>
  </xs:element>
  <xs:element name="CompoundNonPreferredTerm"
  type="CompoundNonPreferredTermStructure" minOccurs="0" maxOccurs="unbounded">
    <xs:annotation>
      <xs:documentation>Details of the compound non-preferred
      term</xs:documentation>
    </xs:annotation>
  </xs:element>
 </xs:sequence>
</xs:complexType>
<xs:complexType name="ThesaurusConceptStructure">
 <xs:annotation>
  <xs:documentation>Data type representing the structure of a concept in a
  Thesaurus.</xs:documentation>
 </xs:annotation>
 <xs:sequence>
  <xs:element ref="dcterms:created" minOccurs="0"/>
  <xs:element ref="dcterms:modified" minOccurs="0"/>
  <xs:element ref="Notation" minOccurs="0" maxOccurs="unbounded"/>
  <xs:element ref="eGMS:Status" minOccurs="0"/>
  <xs:element name="TopConcept" type="xs:boolean" minOccurs="0">
    <xs:annotation>
      <xs:documentation>Indicates whether this concept is a top
      concept</xs:documentation>
    </xs:annotation>
  </xs:element>
  <xs:element name="ScopeNote" type="ScopeNoteStructure" minOccurs="0"
  maxOccurs="unbounded">
    <xs:annotation>
      <xs:documentation>Scope note of the concept.</xs:documentation>
    </xs:annotation>
  </xs:element>
  <xs:element name="HistoryNote" type="HistoryNoteStructure" minOccurs="0"
  maxOccurs="unbounded">
    <xs:annotation>
      <xs:documentation>History note of the concept.</xs:documentation>
    </xs:annotation>
  </xs:element>
```

Figure 2 **XML schema** (*continued*)

```
    <xs:element name="PreferredTerm" maxOccurs="unbounded">
      <xs:annotation>
        <xs:documentation>Preferred term of the concept. A concept may have only
        one preferred term per language.</xs:documentation>
      </xs:annotation>
      <xs:complexType>
        <xs:complexContent>
          <xs:extension base="PreferredTermStructure"/>
        </xs:complexContent>
      </xs:complexType>
    </xs:element>
    <xs:element name="NonPreferredTerm" minOccurs="0" maxOccurs="unbounded">
      <xs:annotation>
        <xs:documentation>Non-preferred term of the concept.</xs:documentation>
      </xs:annotation>
      <xs:complexType>
        <xs:complexContent>
          <xs:extension base="SimpleNonPreferredTermStructure"/>
        </xs:complexContent>
      </xs:complexType>
    </xs:element>
    <xs:element name="HasAttribute" type="CustomAttributeStructure" minOccurs="0"
    maxOccurs="unbounded">
      <xs:annotation>
        <xs:documentation>Reference to an attribute defined by the
        thesaurus.</xs:documentation>
      </xs:annotation>
    </xs:element>
    <xs:element name="HasHierRelConcept" type="HierRelStructure" minOccurs="0"
    maxOccurs="unbounded">
      <xs:annotation>
        <xs:documentation>Reference to a broader or narrower concept of this
        concept.</xs:documentation>
      </xs:annotation>
    </xs:element>
    <xs:element name="HasRelatedConcept" type="AssociativeRelationshipStructure"
    minOccurs="0" maxOccurs="unbounded">
      <xs:annotation>
        <xs:documentation>Reference to the related concepts of this
        concept.</xs:documentation>
      </xs:annotation>
    </xs:element>
    <xs:element name="HasTopConcept" type="xs:string" minOccurs="0" maxOccurs="unbounded">
      <xs:annotation>
        <xs:documentation>Contains the identifier of the top
        concept.</xs:documentation>
      </xs:annotation>
    </xs:element>
  </xs:sequence>
  <xs:attribute ref="dc:identifier" use="required"/>
</xs:complexType>
<xs:complexType name="ThesaurusTermStructure" abstract="true">
  <xs:annotation>
    <xs:documentation>Data type defining the structure of a term in a
    Thesaurus.</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element ref="LexicalValue"/>
    <xs:element ref="dcterms:created" minOccurs="0"/>
    <xs:element ref="dcterms:modified" minOccurs="0"/>
    <xs:element ref="dc:source" minOccurs="0"/>
    <xs:element ref="eGMS:Status" minOccurs="0"/>
```

Figure 2 **XML schema** (*continued*)

```
    <xs:element name="HistoryNote" type="HistoryNoteStructure" minOccurs="0">
      <xs:annotation>
        <xs:documentation>History note of the term.</xs:documentation>
      </xs:annotation>
    </xs:element>
    <xs:element name="Definition" type="DefinitionStructure" minOccurs="0"
    maxOccurs="unbounded">
      <xs:annotation>
        <xs:documentation>Definition of the term.</xs:documentation>
      </xs:annotation>
    </xs:element>
    <xs:element name="EditorialNote" type="EditorialNoteStructure" minOccurs="0">
      <xs:annotation>
        <xs:documentation>Editorial note about the term.</xs:documentation>
      </xs:annotation>
    </xs:element>
  </xs:sequence>
  <xs:attribute ref="xml:lang"/>
  <xs:attribute ref="dc:identifier"/>
</xs:complexType>
<xs:complexType name="NoteStructure">
  <xs:annotation>
    <xs:documentation>Data type representing the structure of notes (scope notes,
    history notes, etc.).</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="LexicalValue">
      <xs:annotation>
        <xs:documentation>The notes may include references to other
        concepts.</xs:documentation>
      </xs:annotation>
      <xs:complexType mixed="true">
        <xs:sequence>
          <xs:element name="RefersTo" type="xs:string" minOccurs="0" maxOccurs="unbounded"/
          >
        </xs:sequence>
      </xs:complexType>
    </xs:element>
    <xs:element ref="dcterms:created" minOccurs="0"/>
    <xs:element ref="dcterms:modified" minOccurs="0"/>
  </xs:sequence>
  <xs:attribute ref="xml:lang"/>
</xs:complexType>
<xs:complexType name="ScopeNoteStructure">
  <xs:annotation>
    <xs:documentation>Data type representing the structure of scope notes. This
    complex type extends the generic structure of
    NoteStructure.</xs:documentation>
  </xs:annotation>
  <xs:complexContent>
    <xs:extension base="NoteStructure"/>
  </xs:complexContent>
</xs:complexType>
<xs:complexType name="HistoryNoteStructure">
  <xs:annotation>
    <xs:documentation>Data type representing the structure of history notes. This
    complex type extends the generic structure of
    NoteStructure.</xs:documentation>
  </xs:annotation>
  <xs:complexContent>
    <xs:extension base="NoteStructure"/>
  </xs:complexContent>
</xs:complexType>
```

Figure 2 **XML schema** (*continued*)

```
<xs:complexType name="DefinitionStructure">
  <xs:annotation>
    <xs:documentation>Data type representing the structure of definitions. This
    complex type extends the generic structure of
    NoteStructure.</xs:documentation>
  </xs:annotation>
  <xs:complexContent>
    <xs:extension base="NoteStructure">
      <xs:sequence>
        <xs:element ref="dc:source"/>
      </xs:sequence>
    </xs:extension>
  </xs:complexContent>
</xs:complexType>
<xs:complexType name="EditorialNoteStructure">
  <xs:annotation>
    <xs:documentation>Data type representing the structure of editorial notes.
    This complex type extends the generic structure of
    NoteStructure.</xs:documentation>
  </xs:annotation>
  <xs:complexContent>
    <xs:extension base="NoteStructure"/>
  </xs:complexContent>
</xs:complexType>
<xs:complexType name="PreferredTermStructure">
  <xs:annotation>
    <xs:documentation>Data type defining the structure of a preferred
    term.</xs:documentation>
  </xs:annotation>
  <xs:complexContent>
    <xs:extension base="ThesaurusTermStructure">
      <xs:sequence>
        <xs:element name="UF" type="UFRelationshipStructure" minOccurs="0"
        maxOccurs="unbounded">
          <xs:annotation>
            <xs:documentation>References the non-preferred terms of the concept or a
            compound non preferred term.</xs:documentation>
          </xs:annotation>
        </xs:element>
      </xs:sequence>
    </xs:extension>
  </xs:complexContent>
</xs:complexType>
<xs:complexType name="SimpleNonPreferredTermStructure">
  <xs:annotation>
    <xs:documentation>Data type defining the structure of non-preferred
    terms.</xs:documentation>
  </xs:annotation>
  <xs:complexContent>
    <xs:extension base="ThesaurusTermStructure">
      <xs:sequence>
        <xs:element name="USE" type="USERelationshipStructure">
          <xs:annotation>
            <xs:documentation>References the preferred term of the
            concept.</xs:documentation>
          </xs:annotation>
        </xs:element>
      </xs:sequence>
    </xs:extension>
  </xs:complexContent>
</xs:complexType>
```

Figure 2 **XML schema** (*continued*)

```
<xs:complexType name="NodeLabelStructure">
  <xs:annotation>
    <xs:documentation>Data type defining the structure of node
    labels</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element ref="Notation" minOccurs="0"/>
    <xs:element ref="LexicalValue"/>
    <xs:element ref="dcterms:created" minOccurs="0"/>
    <xs:element ref="dcterms:modified" minOccurs="0"/>
  </xs:sequence>
  <xs:attribute ref="xml:lang"/>
</xs:complexType>
<xs:complexType name="ThesaurusArrayStructure">
  <xs:annotation>
    <xs:documentation>Data type defining the structure of
    arrays</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="IsLabelledBy" type="NodeLabelStructure" minOccurs="0"
    maxOccurs="unbounded">
      <xs:annotation>
        <xs:documentation>Label given to the array.</xs:documentation>
      </xs:annotation>
    </xs:element>
    <xs:element name="HasSuperOrdinate" type="xs:string" minOccurs="0">
      <xs:annotation>
        <xs:documentation>Identifies the superordinate concepts or
        arrays.</xs:documentation>
      </xs:annotation>
    </xs:element>
    <xs:element name="HasMember" type="xs:string" maxOccurs="unbounded">
      <xs:annotation>
        <xs:documentation>Identifies the concepts or other arrays composing this
        array</xs:documentation>
      </xs:annotation>
    </xs:element>
  </xs:sequence>
  <xs:attribute ref="dc:identifier" use="required"/>
  <xs:attribute name="Ordered" type="xs:boolean" default="false">
    <xs:annotation>
      <xs:documentation>Indicates whether the order of the collection should be
      maintained.</xs:documentation>
    </xs:annotation>
  </xs:attribute>
</xs:complexType>
<xs:complexType name="CompoundNonPreferredTermStructure">
  <xs:annotation>
      <xs:documentation>Data type defining the structure of non-preferred
        terms.</xs:documentation>
  </xs:annotation>
  <xs:complexContent>
    <xs:extension base="ThesaurusTermStructure">
      <xs:sequence>
        <xs:element name="USE" type="USEPlusRelationshipStructure" minOccurs="2"
        maxOccurs="unbounded"/>
      </xs:sequence>
    </xs:extension>
  </xs:complexContent>
</xs:complexType>
```

Figure 2 **XML schema** (*continued*)

```
<xs:element name="LexicalValue">
  <xs:annotation>
    <xs:documentation>Lexical value of a term or node label</xs:documentation>
  </xs:annotation>
</xs:element>
<xs:element name="Notation" type="xs:string">
  <xs:annotation>
    <xs:documentation>Notation of the node label or concept</xs:documentation>
  </xs:annotation>
</xs:element>
<xs:complexType name="HierRelStructure">
  <xs:annotation>
    <xs:documentation>Data type defining the structure of a BT and NT
    relationship</xs:documentation>
  </xs:annotation>
  <xs:simpleContent>
    <xs:extension base="xs:string">
      <xs:attribute name="Role" type="HierRelType"/>
    </xs:extension>
  </xs:simpleContent>
</xs:complexType>
<xs:simpleType name="HierRelType">
  <xs:annotation>
    <xs:documentation>Data type containing the list of common BT and NT
    relationships</xs:documentation>
  </xs:annotation>
  <xs:union>
    <xs:simpleType>
      <xs:restriction base="xs:string">
        <xs:enumeration value="BT"/>
        <xs:enumeration value="BT/BTG"/>
        <xs:enumeration value="BT/BTI"/>
        <xs:enumeration value="BT/BTP"/>
        <xs:enumeration value="NT"/>
        <xs:enumeration value="NT/NTG"/>
        <xs:enumeration value="NT/NTI"/>
        <xs:enumeration value="NT/NTP"/>
      </xs:restriction>
    </xs:simpleType>
    <xs:simpleType>
      <xs:restriction base="xs:string"/>
    </xs:simpleType>
  </xs:union>
</xs:simpleType>
<xs:complexType name="AssociativeRelationshipStructure">
  <xs:annotation>
    <xs:documentation>Data type defining the structure of an associative
    relationship</xs:documentation>
  </xs:annotation>
  <xs:simpleContent>
    <xs:extension base="xs:string">
      <xs:attribute name="Role" type="AssociativeRelationshipType"/>
    </xs:extension>
  </xs:simpleContent>
</xs:complexType>
<xs:simpleType name="AssociativeRelationshipType">
  <xs:annotation>
    <xs:documentation>Data type containing the list of common RT
    relationships</xs:documentation>
  </xs:annotation>
```

Figure 2 **XML schema** (*continued*)

```
  <xs:union>
    <xs:simpleType>
      <xs:restriction base="xs:string">
        <xs:enumeration value="RT"/>
      </xs:restriction>
    </xs:simpleType>
    <xs:simpleType>
      <xs:restriction base="xs:string"/>
    </xs:simpleType>
  </xs:union>
</xs:simpleType>
<xs:complexType name="USERelationshipStructure">
  <xs:annotation>
    <xs:documentation>Data type defining the structure of a USE
    relationship</xs:documentation>
  </xs:annotation>
  <xs:simpleContent>
    <xs:extension base="xs:string">
      <xs:attribute name="Role" type="USERelationshipType"/>
    </xs:extension>
  </xs:simpleContent>
</xs:complexType>
<xs:simpleType name="USERelationshipType">
  <xs:annotation>
    <xs:documentation>Data type containing the list of common USE relationship
    types</xs:documentation>
  </xs:annotation>
  <xs:union>
    <xs:simpleType>
      <xs:restriction base="xs:string">
        <xs:enumeration value="USE"/>
      </xs:restriction>
    </xs:simpleType>
    <xs:simpleType>
      <xs:restriction base="xs:string"/>
    </xs:simpleType>
  </xs:union>
</xs:simpleType>
<xs:complexType name="USEPlusRelationshipStructure">
  <xs:annotation>
    <xs:documentation>Data type defining the structure of a USE +
    relationship</xs:documentation>
  </xs:annotation>
  <xs:simpleContent>
    <xs:extension base="xs:string">
      <xs:attribute name="Role" type="USEPlusRelationshipType"/>
    </xs:extension>
  </xs:simpleContent>
</xs:complexType>
<xs:simpleType name="USEPlusRelationshipType">
  <xs:annotation>
    <xs:documentation>Data type containing the list of common USE + relationship
    types</xs:documentation>
  </xs:annotation>
  <xs:restriction base="xs:string">
    <xs:enumeration value="USE +"/>
  </xs:restriction>
</xs:simpleType>
<xs:complexType name="UFRelationshipStructure">
  <xs:annotation>
    <xs:documentation>Data type defining the structure of a UF
    relationship</xs:documentation>
  </xs:annotation>
```

Figure 2 **XML schema** (*continued*)

```
    <xs:simpleContent>
      <xs:extension base="xs:string">
        <xs:attribute name="Role" type="UFRelationshipType"/>
      </xs:extension>
  </xs:simpleContent>
  </xs:complexType>
  <xs:simpleType name="UFRelationshipType">
    <xs:annotation>
      <xs:documentation>Data type containing the list of common UF relationship
      types</xs:documentation>
    </xs:annotation>
    <xs:union>
      <xs:simpleType>
        <xs:restriction base="xs:string">
          <xs:enumeration value="UF"/>
          <xs:enumeration value="UF +"/>
        </xs:restriction>
      </xs:simpleType>
      <xs:simpleType>
        <xs:restriction base="xs:string"/>
      </xs:simpleType>
    </xs:union>
  </xs:simpleType>
  <xs:complexType name="CustomAttributeStructure">
    <xs:annotation>
      <xs:documentation>Data type representing the structure of a custom
      attribute</xs:documentation>
    </xs:annotation>
    <xs:sequence>
      <xs:element ref="LexicalValue"/>
      <xs:element name="CustomAttributeType" type="xs:string"/>
    </xs:sequence>
  </xs:complexType>
</xs:schema>
```

7 Alternative formats in widespread use

While this part of BS 8723 recommends the format described in Clause **6**, a number of other formats are already in use in particular communities. In order to promote interoperability, it might sometimes be necessary to work alongside these alternatives, and seek ways of minimizing the inconvenience.

Three existing formats are particularly well known:

- **MARC (MAchine-Readable Cataloguing)** [14–16]. This set of standards is widely used for exchange of bibliographic information between libraries. Although most frequently used to exchange cataloguing records, it also offers a linked format for "authority records", which may be used to convey thesaurus data, albeit in a cumbersome way.
- **SKOS.** The SKOS (Simple Knowledge Organization Systems) format [17, 18] is particularly applicable to Semantic Web developments, and is encoded using XML and RDF (Resource Description Framework). As well as thesauri, it is being developed for classification schemes, taxonomies and other vocabulary types.

- **Zthes.** Zthes [19] was originally developed as an application profile of Z39.50 to enable exchange of thesaurus data, but has since evolved, and has a version that can be used with SRW/SRU. (See Clause **8**.)

In addition to the three mentioned, others are sometimes used to meet the needs of specific domains.

When multiple formats are involved, one way to minimize inconvenience is to convert the data from one of these formats to the one recommended in DD 8723-5, or vice versa, by a suitable transformation. Some information is liable to be lost if a thesaurus is transformed into a format that does not support all the features of the source format. Transformations to SKOS and Zthes are currently under development. It is hoped that links to them will be provided on the development website for the BS 8723 XML schema, at http://schemas.bs8723.org/.

8 Protocols

Several different protocols are already in existence, enabling the live querying of thesauri for purposes of indexing and/or retrieval as described in Clause **4**c). The following are fairly widely used:

- **ADL Thesaurus Protocol.** The Alexandria Digital Library (ADL) Protocol [20] is a protocol for accessing monolingual thesauri, adapted to use over the World Wide Web. It supports live querying of, and navigation around, the thesaurus.
- **SKOS API.** This is a Web Service API (Application Programming Interface), designed to provide access to thesauri and other simple knowledge organization systems (SKOS) via the Web [21]. It defines a core set of operations for programmatically accessing and querying a thesaurus. While intended for web service calls, the API itself remains independent of such concrete implementation details. It could also be capable of adaptation to formats other than SKOS Core.
- **SRW/SRU.** The acronym stands for "Search and Retrieve via the Web", for which either the SOAP protocol or URLs (Uniform Resource Locators) can provide an access mechanism, and "Search and Retrieve via URLs" [22], referring to the URL-based version. Unlike Z39.50, it uses XML encoding.
- **Z39.50.** This protocol is more formally known as ANSI/NISO Z39.50, *Information retrieval (Z39.50): Application service definition and protocol specification* [23]. It was designed for querying databases in general, not just vocabularies.

Each protocol has strengths in a different set of circumstances or application environment. The most appropriate one should be selected after evaluating the needs of a particular application.

Annex A (informative) Examples of thesaurus data in DD 8723-5 format

A.1 General

The purpose of this annex is to illustrate how the XML schema shown in Clause **6** should be applied. Each example illustrates different features of the data model. More extensive examples appear on the development website for this standard, at http://schemas.bs8723.org/. The website examples give additional information, for example about transformations between DD 8723-5 and alternative formats.

The tags in all the examples of alphabetical displays are as described in BS 8723-1:2005, Clause **3**. Some of the examples use additional tags, and these are explained where they occur. The colons following the tags are optional, and are used in some but not all of the examples.

None of the examples in this annex shows any Dublin Core metadata, for two reasons. Firstly, such metadata are optional in the schema. Secondly, the examples are all tiny fragments, not real thesauri, and so they do not have an independent existence with realistic metadata. However, the website examples include some real thesauri with Dublin Core metadata.

None of the examples show how to handle specialized relationship types. However, work on this aspect is ongoing, with a view to inclusion in a future edition of this standard. The extension will be made in a way that avoids data incompatibilities. The work in progress can be seen on the development website.

A.2 Example A: A very simple thesaurus

A.2.1 Source data description

This example is drawn from a thesaurus having terms with scope notes and only the most basic relationship types. It has no classified display, but a hierarchical display is shown in Figure A.1, as well as the normal alphabetical one in Figure A.2. This has been used to supply the data that are shown encoded in the DD 8723-5 format in Figure A.3. One of the terms has two broader terms, and so illustrates polyhierarchy. Like many simple thesauri, this one has no concept identifiers, but each term has a term identifier, tagged “TID” in the alphabetical display. (Normally the identifier would not be shown in an alphabetical display for end-users, but it is shown here to facilitate comparison with the XML output.)

Figure A.1 **Example A: Hierarchical display**

```
animals
. birds
. . parrots
. . . pet parrots
. cats
. pets
. . pet parrots

food
. petfood
. . birdseed
```

Figure A.2 **Example A: Alphabetical display**

```
animals
  TID:  1
  NT:   birds
        cats
        pets

birds
  TID:  3
  BT:   animals
  NT:   parrots
  RT:   birdseed

birdseed
  TID:  6
  SN:   Seed selected for
        palatability to most pet
        birds and also some wild
        birds common in gardens
  BT:   petfood
  RT:   birds

cats
  TID:  9
  UF:   pussycats
  BT:   animals

food
  TID:  4
  NT:   petfood

parrots
  TID:  7
  BT:   birds
  NT:   pet parrots

pet parrots
  TID:  8
  BT:   parrots
        pets

petfood
  TID:  5
  BT:   food
  NT:   birdseed
  RT:   pets

pets
  TID:  2
  BT:   animals
  NT:   pet parrots
  RT:   petfood

pussycats
  TID:  10
  USE:  cats
```

Figure A.3 **Example A: Data encoded in DD 8723-5 format**

```
<Thesaurus xsi:schemaLocation="http://www.bsigroup.com/resources/standards/bs8723/
../XmlSchema/DD8723-5.xsd">
  <dc:identifier>http://schemas.bs8723.org/Examples/ExampleSimple/</dc:identifier>
  <dc:language>en</dc:language>
  <ThesaurusConcept dc:identifier="C001">
    <dcterms:created>2007-05-02</dcterms:created>
    <PreferredTerm dc:identifier="1" xml:lang="en">
      <LexicalValue>animals</LexicalValue>
      <dcterms:created>2007-05-02</dcterms:created>
    </PreferredTerm>
    <HasHierRelConcept Role="NT">C002</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C004</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C009</HasHierRelConcept>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C002">
    <dcterms:created>2007-05-02</dcterms:created>
    <PreferredTerm dc:identifier="3" xml:lang="en">
      <LexicalValue>birds</LexicalValue>
      <dcterms:created>2007-05-02</dcterms:created>
    </PreferredTerm>
    <HasHierRelConcept Role="BT">C001</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C006</HasHierRelConcept>
    <HasRelatedConcept Role="RT">C003</HasRelatedConcept>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C003">
    <dcterms:created>2007-05-02</dcterms:created>
    <ScopeNote xml:lang="en">
      <LexicalValue>Seed selected for palatability to most pet birds and also some
      wild birds common in gardens</LexicalValue>
    </ScopeNote>
    <PreferredTerm dc:identifier="6" xml:lang="en">
      <LexicalValue>birdseed</LexicalValue>
      <dcterms:created>2007-05-02</dcterms:created>
    </PreferredTerm>
    <HasHierRelConcept Role="BT">C008</HasHierRelConcept>
    <HasRelatedConcept Role="RT">C002</HasRelatedConcept>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C004">
    <dcterms:created>2007-05-02</dcterms:created>
    <PreferredTerm dc:identifier="9" xml:lang="en">
      <LexicalValue>cats</LexicalValue>
      <dcterms:created>2007-05-02</dcterms:created>
      <UF>10</UF>
    </PreferredTerm>
    <NonPreferredTerm dc:identifier="10" xml:lang="en">
      <LexicalValue>pussycats</LexicalValue>
      <dcterms:created>2007-05-02</dcterms:created>
      <USE>9</USE>
    </NonPreferredTerm>
    <HasHierRelConcept Role="BT">C001</HasHierRelConcept>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C005">
    <dcterms:created>2007-05-02</dcterms:created>
    <PreferredTerm dc:identifier="4" xml:lang="en">
      <LexicalValue>food</LexicalValue>
      <dcterms:created>2007-05-02</dcterms:created>
    </PreferredTerm>
    <HasHierRelConcept Role="NT">C008</HasHierRelConcept>
  </ThesaurusConcept>
```

Figure A.3 **Example A: Data encoded in DD 8723-5 format** (*continued*)

```
  <ThesaurusConcept dc:identifier="C006">
    <dcterms:created>2007-05-02</dcterms:created>
    <PreferredTerm dc:identifier="7" xml:lang="en">
      <LexicalValue>parrots</LexicalValue>
      <dcterms:created>2007-05-02</dcterms:created>
    </PreferredTerm>
    <HasHierRelConcept Role="BT">C002</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C007</HasHierRelConcept>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C007">
    <dcterms:created>2007-05-02</dcterms:created>
    <PreferredTerm dc:identifier="8" xml:lang="en">
      <LexicalValue>pet parrots</LexicalValue>
      <dcterms:created>2007-05-02</dcterms:created>
    </PreferredTerm>
    <HasHierRelConcept Role="BT">C006</HasHierRelConcept>
    <HasHierRelConcept Role="BT">C009</HasHierRelConcept>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C008">
    <dcterms:created>2007-05-02</dcterms:created>
    <PreferredTerm dc:identifier="5" xml:lang="en">
      <LexicalValue>petfood</LexicalValue>
      <dcterms:created>2007-05-02</dcterms:created>
    </PreferredTerm>
    <HasHierRelConcept Role="BT">C005</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C003</HasHierRelConcept>
    <HasRelatedConcept Role="RT">C009</HasRelatedConcept>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C009">
    <dcterms:created>2007-05-02</dcterms:created>
    <PreferredTerm dc:identifier="2" xml:lang="en">
      <LexicalValue>pets</LexicalValue>
      <dcterms:created>2007-05-02</dcterms:created>
    </PreferredTerm>
    <HasHierRelConcept Role="BT">C001</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C007</HasHierRelConcept>
    <HasRelatedConcept Role="RT">C008</HasRelatedConcept>
  </ThesaurusConcept>
</Thesaurus>
```

A.2.2 Commentary

Because this thesaurus is so simple, optional features of the data model (such as ***NodeLabel*** and ***ThesaurusArray***) are not used. However, the lack of concept identifiers in the source data made it necessary to generate concept identifiers prior to encoding for data exchange purposes. As can be seen in the XML output, each concept was given a sequential number in the series C001, C002, etc.

A.3 Example B: Extract from a bilingual thesaurus, with notation and node labels

A.3.1 Source data description

This example is drawn from a bilingual thesaurus, with notation, node labels, history notes, and dates of term creation and modification. Because this thesaurus is much more sophisticated than Example A, the data cannot conveniently be compressed into a single user-friendly display. A combination of three displays is needed, thus Figure A.4 and Figure A.5 show alphabetical displays for each of the languages, while Figure A.6 presents a classified bilingual display, as described in BS 8723-4:2005, **10.3.2**. For ease of checking, the tags have not been translated. In both languages, therefore, they are as shown in Table A.1.

Figure A.4 **Example B: Alphabetical display in English**

agricultural chemicals
CID: C002
TID: 51
USE: agrochemicals
DC: 2004-06-25
DTI: 2004-07-09
DTM: 2007-05-21

agriculture
es: agricultura
CID: C001
TID: 58
CC: AA000
DC: 2004-06-25
DTI: 2007-05-21
DTM: 2007-05-21

agrochemicals
es: productos químicos agrícolas
CID: C002
TID: 46
CC: FF100
CHN: Concept added 25 June 2004
UF: agricultural chemicals
NT: fertilizers
pesticides
DC: 2004-06-25
DTI: 2004-06-25
DTM: 2007-05-21

feeds
es: piensos
CID: C008
TID: 22
CC: FM100
DC: 2004-06-18
DTI: 2004-06-18
DTM: 2007-05-21

fertilizers
es: fertilizantes
CID: C003
TID: 21
CC: FF110
BT: agrochemicals
DC: 2004-06-18
DTI: 2004-06-18
DTM: 2007-05-21

fuel
es: combustibles
CID: C009
TID: 23
CC: FP100
DC: 2004-06-18
DTI: 2004-06-18
DTM: 2007-05-21

fumigants
CID: C004
TID: 49
USE: pesticides
DC: 2004-06-18
DTI: 2004-07-09
DTM: 2007-05-21

fungicides
es: fungicidas
CID: C005
TID: 37
CC: FF130
BT: pesticides
DC: 2004-06-18
DTI: 2004-06-18
DTM: 2007-05-21

herbicides
es: herbicidas
CID: C006
TID: 36
CC: FF140
UF: weedkillers
BT: pesticides
DC: 2004-06-18
DTI: 2004-06-18
DTM: 2007-05-21

insecticides
es: insecticidas
CID: C007
TID: 35
CC: FF150
BT: pesticides
DC: 2004-06-18
DTI: 2004-06-18
DTM: 2007-05-21

pesticides
es: plaguicidas
CID: C004
TID: 20
CC: FF120
SN: Excludes growth regulators and repellents
UF: fumigants
BT: agrochemicals
NT: fungicides
herbicides
insecticides
DC: 2004-06-18
DTI: 2004-06-18
DTM: 2007-05-21

weedkillers
CID: C006
TID: 53
USE: herbicides
THN: Synonym added in May 2007
DC: 2004-06-18
DTI: 2007-05-20
DTM: 2007-05-21

Figure A.5 **Example B: Alphabetical display in Spanish**

agricultura
en: agriculture
CID: C001
TID: 58E
CC: AA000
DC: 2004-06-25
DTI: 2007-05-21
DTM: 2007-05-21

alimentos para animales
CID: C008
TID: 56E
USE: piensos
THN: Sinónimo agregado en mayo de 2007
DC: 2004-06-18
DTI: 2007-05-20
DTM: 2007-05-21

combustibles
en: fuel
CID: C009
TID: 23E
CC: FP100
DC: 2004-06-18
DTI: 2004-06-18
DTM: 2007-05-21

fertilizantes
en: fertilizers
CID: C003
TID: 21E
CC: FF110
BT: productos químicos agrícolas
DC: 2004-06-18
DTI: 2004-06-18
DTM: 2007-05-21

fumigantes
CID: C004
TID: 49E
USE: plaguicidas
DC: 2004-06-18
DTI: 2004-07-09
DTM: 2007-05-21

fungicidas
en: fungicides
CID: C005
TID: 37E
CC: FF130
BT: plaguicidas
DC: 2004-06-18
DTI: 2004-06-18
DTM: 2007-05-21

herbicidas
en: herbicides
CID: C006
TID: 36E
CC: FF140
BT: plaguicidas
DC: 2004-06-18
DTI: 2004-06-18
DTM: 2007-05-21

insecticidas
en: insecticides
CID: C007
TID: 35E
CC: FF150
BT: plaguicidas
DC: 2004-06-18
DTI: 2004-06-18
DTM: 2007-05-21

piensos
en: feeds
CID: C008
TID: 22E
CC: FM100
UF: alimentos para animales
DC: 2004-06-18
DTI: 2004-06-18
DTM: 2007-05-21

plaguicidas
en: pesticides
CID: C004
TID: 20E
CC: FF120
SN: Excluye repelentes y reguladores del crecimiento
UF: fumigantes
BT: productos químicos agrícolas
NT: fungicidas
herbicidas
insecticidas
DC: 2004-06-18
DTI: 2004-06-18
DTM: 2007-05-21

productos químicos agrícolas
en: agrochemicals
CID: C002
TID: 46E
CC: FF100
CHN: Concepto agregado 25 de junio de 2004
NT: fertilizantes
plaguicidas
DC: 2004-06-25
DTI: 2004-06-25
DTM: 2007-05-21

Figure A.6 **Example B: Bilingual classified display, with node labels and notation**

agriculture	AA000	agricultura
. <materials>		. <materiales>
. . agrochemicals	FF100	. . productos químicos agrícolas
. . . fertilizers	FF110	. . . fertilizantes
. . . pesticides	FF120	. . . plaguicidas
. . . . fungicides	FF130	. . . . fungicidas
. . . . herbicides	FF140	. . . . herbicidas
. . . . insecticides	FF150	. . . . insecticidas
. . feeds	FM100	. . piensos
. . fuel	FP100	. . combustibles

Table A.1 **Tags used for Example B**

Tag	Meaning
en	English
es	Spanish
CID	Concept identifier
TID	Term identifier
CC	Notation
SN	Scope note
CHN	Concept history note
THN	Term history note
UF	Use for
USE	Use
BT	Broader term
NT	Narrower term
RT	Related term
DC	Date concept was created
DTI	Date term was created
DTM	Date term was modified

A.3.2 Commentary

Figure A.7 show how this example should be encoded in DD 8723-5 format. In contrast with Example A, several optional features of the data model and XML schema are used. In particular, ***NodeLabel*** and ***ThesaurusArray*** are both needed to enable adequate representation of node labels. Use of the language attributes and codes makes it straightforward to convey data in as many languages as are required. Figure A.7 demonstrates how the data from three separate displays for human readers can be carried in a single machine-readable file.

Figure A.7 **Example B: Data encoded in DD 8723-5 format**

```
<Thesaurus xsi:schemaLocation="http://www.bsigroup.com/resources/standards/bs8723/
../XmlSchema/DD8723-5.xsd">
  <dc:identifier>http://schemas.bs8723.org/Examples/ExampleBilingual/</dc:identifier>
  <dc:language>en</dc:language>
  <dc:language>es</dc:language>
  <ThesaurusArray dc:identifier="A002">
    <IsLabelledBy xml:lang="en">
      <LexicalValue>materials</LexicalValue>
    </IsLabelledBy>
    <IsLabelledBy xml:lang="es">
      <LexicalValue>materiales</LexicalValue>
    </IsLabelledBy>
    <HasSuperOrdinate>C001</HasSuperOrdinate>
    <HasMember>C002</HasMember>
    <HasMember>C008</HasMember>
    <HasMember>C009</HasMember>
  </ThesaurusArray>
  <ThesaurusConcept dc:identifier="C001">
    <dcterms:created>2004-06-25</dcterms:created>
    <Notation>AA000</Notation>
    <PreferredTerm xml:lang="en" dc:identifier="58">
      <LexicalValue>agriculture</LexicalValue>
      <dcterms:created>2007-05-21</dcterms:created>
      <dcterms:modified>2007-05-21</dcterms:modified>
    </PreferredTerm>
    <PreferredTerm xml:lang="es" dc:identifier="58E">
      <LexicalValue>agricultura</LexicalValue>
      <dcterms:created>2004-05-21</dcterms:created>
      <dcterms:modified>2007-05-21</dcterms:modified>
    </PreferredTerm>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C002">
    <dcterms:created>2004-06-25</dcterms:created>
    <Notation>FF100</Notation>
    <HistoryNote xml:lang="en">
      <LexicalValue>Concept added 25 June 2004</LexicalValue>
    </HistoryNote>
    <HistoryNote xml:lang="es">
      <LexicalValue>Concepto agregado 25 de junio de 2004</LexicalValue>
    </HistoryNote>
    <PreferredTerm xml:lang="en" dc:identifier="46">
      <LexicalValue>agrochemicals</LexicalValue>
      <dcterms:created>2004-06-25</dcterms:created>
      <dcterms:modified>2007-05-21</dcterms:modified>
      <UF>51</UF>
    </PreferredTerm>
    <PreferredTerm xml:lang="es" dc:identifier="46E">
      <LexicalValue>productos químicos agrícolas</LexicalValue>
      <dcterms:created>2004-06-25</dcterms:created>
      <dcterms:modified>2007-05-21</dcterms:modified>
    </PreferredTerm>
    <NonPreferredTerm dc:identifier="51" xml:lang="en">
      <LexicalValue>agricultural chemicals</LexicalValue>
      <dcterms:created>2004-07-09</dcterms:created>
      <dcterms:modified>2007-05-21</dcterms:modified>
      <USE>46</USE>
    </NonPreferredTerm>
    <HasHierRelConcept Role="NT">C003</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C004</HasHierRelConcept>
  </ThesaurusConcept>
```

Figure A.7 **Example B: Data encoded in DD 8723-5 format** (*continued*)

```
<ThesaurusConcept dc:identifier="C003">
  <dcterms:created>2004-06-18</dcterms:created>
  <Notation>FF110</Notation>
  <PreferredTerm xml:lang="en" dc:identifier="21">
    <LexicalValue>fertilizers</LexicalValue>
    <dcterms:created>2004-06-18</dcterms:created>
    <dcterms:modified>2007-05-21</dcterms:modified>
  </PreferredTerm>
  <PreferredTerm xml:lang="es" dc:identifier="21E">
    <LexicalValue>fertilizantes</LexicalValue>
    <dcterms:created>2004-06-18</dcterms:created>
    <dcterms:modified>2007-05-21</dcterms:modified>
  </PreferredTerm>
  <HasHierRelConcept Role="BT">C002</HasHierRelConcept>
</ThesaurusConcept>
<ThesaurusConcept dc:identifier="C004">
  <dcterms:created>2004-06-18</dcterms:created>
  <Notation>FF120</Notation>
  <ScopeNote xml:lang="en">
    <LexicalValue>Excludes growth regulators and repellents</LexicalValue>
  </ScopeNote>
  <ScopeNote xml:lang="es">
    <LexicalValue>Excluye repelentes y reguladores del crecimiento</LexicalValue>
  </ScopeNote>
  <PreferredTerm xml:lang="en" dc:identifier="20">
    <LexicalValue>pesticides</LexicalValue>
    <dcterms:created>2004-06-18</dcterms:created>
    <dcterms:modified>2007-05-21</dcterms:modified>
    <UF>49</UF>
  </PreferredTerm>
  <PreferredTerm xml:lang="es" dc:identifier="20E">
    <LexicalValue>plaguicidas</LexicalValue>
    <dcterms:created>2004-06-18</dcterms:created>
    <dcterms:modified>2007-05-21</dcterms:modified>
    <UF>49E</UF>
  </PreferredTerm>
  <NonPreferredTerm xml:lang="en" dc:identifier="49">
    <LexicalValue>fumigants</LexicalValue>
    <dcterms:created>2004-07-09</dcterms:created>
    <dcterms:modified>2007-05-21</dcterms:modified>
    <USE>20</USE>
  </NonPreferredTerm>
  <NonPreferredTerm xml:lang="es" dc:identifier="49E">
    <LexicalValue>fumigantes</LexicalValue>
    <dcterms:created>2004-07-09</dcterms:created>
    <dcterms:modified>2007-05-21</dcterms:modified>
    <USE>20E</USE>
  </NonPreferredTerm>
  <HasHierRelConcept Role="BT">C002</HasHierRelConcept>
  <HasHierRelConcept Role="NT">C005</HasHierRelConcept>
  <HasHierRelConcept Role="NT">C006</HasHierRelConcept>
  <HasHierRelConcept Role="NT">C007</HasHierRelConcept>
</ThesaurusConcept>
<ThesaurusConcept dc:identifier="C005">
 <dcterms:created>2004-06-18</dcterms:created>
 <Notation>FF130</Notation>
 <PreferredTerm xml:lang="en" dc:identifier="37">
   <LexicalValue>fungicides</LexicalValue>
   <dcterms:created>2004-06-18</dcterms:created>
   <dcterms:modified>2007-05-21</dcterms:modified>
 </PreferredTerm>
```

Figure A.7 **Example B: Data encoded in DD 8723-5 format** (*continued*)

```
  <PreferredTerm xml:lang="es" dc:identifier="37E">
    <LexicalValue>fungicidas</LexicalValue>
    <dcterms:created>2004-06-18</dcterms:created>
    <dcterms:modified>2007-05-21</dcterms:modified>
  </PreferredTerm>
  <HasHierRelConcept Role="BT">C004</HasHierRelConcept>
</ThesaurusConcept>
<ThesaurusConcept dc:identifier="C006">
  <dcterms:created>2004-06-18</dcterms:created>
  <Notation>FF140</Notation>
  <PreferredTerm xml:lang="en" dc:identifier="36">
    <LexicalValue>herbicides</LexicalValue>
    <dcterms:created>2004-06-18</dcterms:created>
    <dcterms:modified>2007-05-21</dcterms:modified>
    <UF>53</UF>
  </PreferredTerm>
  <PreferredTerm xml:lang="es" dc:identifier="36E">
    <LexicalValue>herbicidas</LexicalValue>
    <dcterms:created>2004-06-18</dcterms:created>
    <dcterms:modified>2007-05-21</dcterms:modified>
  </PreferredTerm>
  <NonPreferredTerm xml:lang="en" dc:identifier="53">
    <LexicalValue>weedkillers</LexicalValue>
    <dcterms:created>2007-05-20</dcterms:created>
    <dcterms:modified>2007-05-21</dcterms:modified>
    <HistoryNote xml:lang="en">
      <LexicalValue>Synonym added in May 2007</LexicalValue>
    </HistoryNote>
    <USE>36</USE>
  </NonPreferredTerm>
  <HasHierRelConcept Role="BT">C004</HasHierRelConcept>
</ThesaurusConcept>
<ThesaurusConcept dc:identifier="C007">
  <dcterms:created>2004-06-18</dcterms:created>
  <Notation>FF150</Notation>
  <PreferredTerm xml:lang="en" dc:identifier="35">
    <LexicalValue>insecticides</LexicalValue>
    <dcterms:created>2004-06-18</dcterms:created>
    <dcterms:modified>2007-05-21</dcterms:modified>
  </PreferredTerm>
  <PreferredTerm xml:lang="es" dc:identifier="35E">
    <LexicalValue>insecticidas</LexicalValue>
    <dcterms:created>2004-06-18</dcterms:created>
    <dcterms:modified>2007-05-21</dcterms:modified>
  </PreferredTerm>
  <HasHierRelConcept Role="BT">C004</HasHierRelConcept>
</ThesaurusConcept>
<ThesaurusConcept dc:identifier="C008">
  <dcterms:created>2004-06-18</dcterms:created>
  <Notation>FM100</Notation>
  <PreferredTerm xml:lang="en" dc:identifier="22">
    <LexicalValue>feeds</LexicalValue>
    <dcterms:created>2004-06-18</dcterms:created>
    <dcterms:modified>2007-05-21</dcterms:modified>
  </PreferredTerm>
  <PreferredTerm xml:lang="es" dc:identifier="22E">
    <LexicalValue>piensos</LexicalValue>
    <dcterms:created>2004-06-18</dcterms:created>
    <dcterms:modified>2007-05-21</dcterms:modified>
    <UF>56E</UF>
  </PreferredTerm>
```

Figure A.7 **Example B: Data encoded in DD 8723-5 format** (*continued*)

```
    <NonPreferredTerm xml:lang="es" dc:identifier="56E">
      <LexicalValue>alimentos para animales</LexicalValue>
      <dcterms:created>2007-05-20</dcterms:created>
      <dcterms:modified>2007-05-21</dcterms:modified>
      <HistoryNote xml:lang="es">
        <LexicalValue>Sinónimo agregado en mayo de 2007</LexicalValue>
      </HistoryNote>
      <USE>22E</USE>
    </NonPreferredTerm>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C009">
    <dcterms:created>2004-06-18</dcterms:created>
    <Notation>FP100</Notation>
    <PreferredTerm xml:lang="en" dc:identifier="23">
      <LexicalValue>fuel</LexicalValue>
      <dcterms:created>2004-06-18</dcterms:created>
      <dcterms:modified>2007-05-21</dcterms:modified>
    </PreferredTerm>
    <PreferredTerm xml:lang="es" dc:identifier="23E">
      <LexicalValue>combustibles</LexicalValue>
      <dcterms:created>2004-06-18</dcterms:created>
      <dcterms:modified>2007-05-21</dcterms:modified>
   </PreferredTerm>
  </ThesaurusConcept>
</Thesaurus>
```

A.4 Example C: Compound terms

This example addresses compound terms quite specifically, so it has very few features or relationships other than USE+ and UF+. All the preferred terms have concept identifiers (tagged CID), but since the compound terms represent concepts that are not really “in” the thesaurus, these do not have concept identifiers. All the terms do have term identifiers (TID).

Figure A.8 shows the sample data as an alphabetical display. A classified display is irrelevant for this example, and a hierarchical display seems trivial, but for completeness the latter has been included as Figure A.9. Figure A.10 shows the data encoded in DD 8723-5 format.

Figure A.8 **Example C: Alphabetical display**

austenitic chromium manganese steel
TID 3
USE austenitic steel
\+ chromium steel
\+ manganese steel

austenitic steel
CID C05
TID 5
UF+ austenitic chromium manganese steel
BT steel

chromium steel
CID C06
TID 6
UF+ austenitic chromium manganese steel
BT steel

coal
CID C01
TID 1
UF+ coal mining
UF+ opencast coal mining

coal mining
TID 4
USE coal
\+ mining

manganese steel
CID C07
TID 7
UF+ austenitic chromium manganese steel
BT steel

mining
CID C02
TID 2
UF+ coal mining
NT opencast mining

opencast coal mining
TID 8
USE coal
\+ opencast mining

opencast mining
CID C09
TID 9
UF+ opencast coal mining
BT mining

steel
CID C10
TID 10
NT austenitic steel
chromium steel
manganese steel

Figure A.9 **Example C: Hierarchical display**

coal

mining
. opencast mining

steel
. austenitic steel
. chromium steel
. manganese steel

Figure A.10 **Example C: Data encoded in DD 8723-5 format**

```
<Thesaurus xsi:schemaLocation="http://www.bsigroup.com/resources/standards/bs8723/
../XmlSchema/DD8723-5.xsd">
  <dc:identifier>http://schemas.bs8723.org/Examples/ExampleCompounds/</dc:identifier>
  <dc:language>en</dc:language>
  <ThesaurusConcept dc:identifier="C01">
    <dcterms:created>2007-05-24</dcterms:created>
    <PreferredTerm dc:identifier="1">
      <LexicalValue>coal</LexicalValue>
      <dcterms:created>2007-05-24</dcterms:created>
      <UF Role="UF +">4</UF>
      <UF Role="UF +">8</UF>
    </PreferredTerm>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C02">
    <dcterms:created>2007-05-24</dcterms:created>
    <PreferredTerm dc:identifier="2">
      <LexicalValue>mining</LexicalValue>
      <dcterms:created>2007-05-24</dcterms:created>
      <UF Role="UF +">4</UF>
    </PreferredTerm>
    <HasHierRelConcept Role="NT">C09</HasHierRelConcept>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C05">
    <dcterms:created>2007-05-24</dcterms:created>
    <PreferredTerm dc:identifier="5">
      <LexicalValue>austenitic steel</LexicalValue>
      <dcterms:created>2007-05-24</dcterms:created>
      <UF Role="UF +">3</UF>
    </PreferredTerm>
    <HasHierRelConcept Role="BT">C10</HasHierRelConcept>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C06">
    <dcterms:created>2007-05-24</dcterms:created>
    <PreferredTerm dc:identifier="6">
      <LexicalValue>chromium steel</LexicalValue>
      <dcterms:created>2007-05-24</dcterms:created>
      <UF Role="UF +">3</UF>
    </PreferredTerm>
    <HasHierRelConcept Role="BT">C10</HasHierRelConcept>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C07">
    <dcterms:created>2007-05-24</dcterms:created>
    <PreferredTerm dc:identifier="7">
      <LexicalValue>manganese steel</LexicalValue>
      <dcterms:created>2007-05-24</dcterms:created>
      <UF Role="UF +">3</UF>
    </PreferredTerm>
    <HasHierRelConcept Role="BT">C10</HasHierRelConcept>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C09">
    <dcterms:created>2007-05-24</dcterms:created>
    <PreferredTerm dc:identifier="9">
      <LexicalValue>opencast mining</LexicalValue>
      <dcterms:created>2007-05-24</dcterms:created>
      <UF Role="UF +">8</UF>
    </PreferredTerm>
    <HasHierRelConcept Role="BT">C02</HasHierRelConcept>
  </ThesaurusConcept>
```

Figure A.10 **Example C: Data encoded in DD 8723-5 format** (*continued*)

```
  <ThesaurusConcept dc:identifier="C10">
    <dcterms:created>2007-05-24</dcterms:created>
    <PreferredTerm dc:identifier="10">
      <LexicalValue>steel</LexicalValue>
      <dcterms:created>2007-05-24</dcterms:created>
    </PreferredTerm>
    <HasHierRelConcept Role="NT">C05</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C06</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C07</HasHierRelConcept>
  </ThesaurusConcept>
  <CompoundNonPreferredTerm dc:identifier="3">
    <LexicalValue>austenitic chromium manganese steel</LexicalValue>
    <dcterms:created>2007-05-24</dcterms:created>
    <USE Role="USE +">5</USE>
    <USE Role="USE +">6</USE>
    <USE Role="USE +">7</USE>
  </CompoundNonPreferredTerm>
  <CompoundNonPreferredTerm dc:identifier="4">
    <LexicalValue>coal mining</LexicalValue>
    <dcterms:created>2007-05-24</dcterms:created>
    <USE Role="USE +">1</USE>
    <USE Role="USE +">2</USE>
  </CompoundNonPreferredTerm>
  <CompoundNonPreferredTerm dc:identifier="8">
    <LexicalValue>opencast coal mining</LexicalValue>
    <dcterms:created>2007-05-24</dcterms:created>
    <USE Role="USE +">1</USE>
    <USE Role="USE +">9</USE>
  </CompoundNonPreferredTerm>
</Thesaurus>
```

A.5 Example D: Thesaurus extract, showing top level terms/concepts

This example, slightly improbable in practice, is intended to show how top terms/concepts are handled. Just three of its top terms are included in the extract, namely fruits, wines and hardware. Just one of the hierarchies makes use of node labels. For simplicity, other features such as notation have been omitted from the displays.

Figure A.11 shows an alphabetical display for Example D; Figure A.12 a classified display and Figure A.13 a hierarchical display. Example D is comparable with Example B in that the data in the alphabetical display need to be complemented by the extra information carried in the classified display, in order to capture the whole content and structure. When encoded as in Figure A.14, however, one XML file is capable of conveying all the data needed to build any of the displays.

Figure A.11 **Example D: Alphabetical display**

bananas
CID: C001
TID: 9
TT: fruits
BT: tropical fruits

bar utensils
CID: C002
TID: 16
TT: hardware
BT: hardware
NT: corkscrews

calamondins
CID: C003
TID: 8
TT: fruits
BT: citrus fruits
tropical fruits

Chardonnay wines
CID: C004
TID: 20
TT: wines
BT: wines

citrus fruits
CID: C005
TID: 4
TT: fruits
BT: fruits
NT: calamondins
lemons
oranges

corkscrews
CID: C006
TID: 18
TT: hardware
BT: bar utensils

French wines
CID: C007
TID: 11
TT: wines
BT: wines

fruits
CID: C008
TID: 1
NT: citrus fruits
tropical fruits

German wines
CID: C009
TID: 12
TT: wines
BT: wines
NT: Liebfraumilch

glassware
CID: C010
TID: 15
TT: hardware
BT: hardware
NT: wineglasses

hardware
CID: C011
TID: 3
NT: bar utensils
glassware

lemons
CID: C012
TID: 7
TT: fruits
BT: citrus fruits

Liebfraumilch
CID: C013
TID: 19
TT: wines
BT: German wines
Riesling wines
white wines

mangoes
CID: C014
TID: 10
TT: fruits
BT: tropical fruits

oranges
CID: C015
TID: 6
TT: fruits
BT: citrus fruits

red wines
CID: C016
TID: 13
TT: wines
BT: wines

Riesling wines
CID: C017
TID: 21
TT: wines
BT: wines
NT: Liebfraumilch

tropical fruits
CID: C018
TID: 5
TT: fruits
BT: fruits
NT: bananas
calamondins
mangoes

white wines
CID: C019
TID: 14
TT: wines
BT: wines
NT: Liebfraumilch

wineglasses
CID: C020
TID: 17
TT: hardware
BT: glassware

wines
CID: C021
TID: 2
NT: Chardonnay wines
French wines
German wines
red wines
Riesling wines
white wines

Figure A.12 **Example D: Classified display**

fruits
. citrus fruits
. . calamondins
. . lemons
. . oranges
. tropical fruits
. . bananas
. . calamondins
. . mangoes

hardware
. bar utensils
. . corkscrews
. glassware
. . wineglasses

wines
. (wines by colour)
. . red wines
. . white wines
. . . Liebfraumilch
. (wines by country of origin)
. . French wines
. . German wines
. . . Liebfraumilch
. (wines by grape)
. . Chardonnay wines
. . Riesling wines
. . . Liebfraumilch
. (accessories)
. . corkscrews
. . wineglasses

Figure A.13 **Example D: Hierarchical display**

fruits
. citrus fruits
. . calamondins
. . lemons
. . oranges
. tropical fruits
. . bananas
. . calamondins
. . mangoes

hardware
. bar utensils
. . corkscrews
. glassware
. . wineglasses

wines
. Chardonnay wines
. French wines
. German wines
. . Liebfraumilch
. red wines
. Riesling wines
. . Liebfraumilch
. white wines
. . Liebfraumilch

Figure A.14 **Example D: Data encoded in DD 8723-5 format**

```
<Thesaurus xsi:schemaLocation="http://www.bsigroup.com/resources/standards/bs8723/
../XmlSchema/DD8723-5.xsd">
  <dc:identifier>http://schemas.bs8723.org/Examples/ExampleTopConcept/</dc:identifier>
  <dc:language>en</dc:language>
  <ThesaurusArray dc:identifier="A008">
    <IsLabelledBy>
      <LexicalValue>wines by colour</LexicalValue>
    </IsLabelledBy>
    <HasSuperOrdinate>C021</HasSuperOrdinate>
    <HasMember>C016</HasMember>
    <HasMember>C019</HasMember>
  </ThesaurusArray>
  <ThesaurusArray dc:identifier="A010">
    <IsLabelledBy>
      <LexicalValue>wines by country of origin</LexicalValue>
    </IsLabelledBy>
    <HasSuperOrdinate>C021</HasSuperOrdinate>
    <HasMember>C007</HasMember>
    <HasMember>C009</HasMember>
  </ThesaurusArray>
  <ThesaurusArray dc:identifier="A012">
    <IsLabelledBy>
      <LexicalValue>wines by grape</LexicalValue>
    </IsLabelledBy>
    <HasSuperOrdinate>C021</HasSuperOrdinate>
    <HasMember>C004</HasMember>
    <HasMember>C017</HasMember>
  </ThesaurusArray>
  <ThesaurusArray dc:identifier="A014">
    <IsLabelledBy>
      <LexicalValue>accessories</LexicalValue>
    </IsLabelledBy>
    <HasSuperOrdinate>C021</HasSuperOrdinate>
    <HasMember>C006</HasMember>
    <HasMember>C020</HasMember>
  </ThesaurusArray>
  <ThesaurusConcept dc:identifier="C001">
    <dcterms:created>2007-06-05</dcterms:created>
    <PreferredTerm dc:identifier="9">
      <LexicalValue>bananas</LexicalValue>
      <dcterms:created>2007-06-05</dcterms:created>
    </PreferredTerm>
    <HasHierRelConcept Role="BT">C018</HasHierRelConcept>
    <HasTopConcept>C008</HasTopConcept>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C002">
    <dcterms:created>2007-06-05</dcterms:created>
    <PreferredTerm dc:identifier="16">
      <LexicalValue>bar utensils</LexicalValue>
      <dcterms:created>2007-06-05</dcterms:created>
    </PreferredTerm>
    <HasHierRelConcept Role="BT">C011</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C006</HasHierRelConcept>
    <HasTopConcept>C011</HasTopConcept>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C003">
    <dcterms:created>2007-06-05</dcterms:created>
    <PreferredTerm dc:identifier="8">
      <LexicalValue>calamondins</LexicalValue>
      <dcterms:created>2007-06-05</dcterms:created>
      </PreferredTerm>
    <HasHierRelConcept Role="BT">C005</HasHierRelConcept>
    <HasHierRelConcept Role="BT">C018</HasHierRelConcept>
    <HasTopConcept>C008</HasTopConcept>
  </ThesaurusConcept>
```

Figure A.14 **Example D: Data encoded in DD 8723-5 format** (*continued*)

```
<ThesaurusConcept dc:identifier="C004">
 <dcterms:created>2007-06-05</dcterms:created>
  <PreferredTerm dc:identifier="20">
    <LexicalValue>Chardonnay wines</LexicalValue>
    <dcterms:created>2007-06-05</dcterms:created>
  </PreferredTerm>
  <HasHierRelConcept Role="BT">C021</HasHierRelConcept>
  <HasTopConcept>C021</HasTopConcept>
</ThesaurusConcept>
<ThesaurusConcept dc:identifier="C005">
  <dcterms:created>2007-06-05</dcterms:created>
  <PreferredTerm dc:identifier="4">
    <LexicalValue>citrus fruits</LexicalValue>
    <dcterms:created>2007-06-05</dcterms:created>
  </PreferredTerm>
  <HasHierRelConcept Role="BT">C008</HasHierRelConcept>
  <HasHierRelConcept Role="NT">C003</HasHierRelConcept>
  <HasHierRelConcept Role="NT">C012</HasHierRelConcept>
  <HasHierRelConcept Role="NT">C015</HasHierRelConcept>
  <HasTopConcept>C008</HasTopConcept>
</ThesaurusConcept>
<ThesaurusConcept dc:identifier="C006">
  <dcterms:created>2007-06-05</dcterms:created>
  <PreferredTerm dc:identifier="18">
    <LexicalValue>corkscrews</LexicalValue>
    <dcterms:created>2007-06-05</dcterms:created>
  </PreferredTerm>
  <HasHierRelConcept Role="BT">C002</HasHierRelConcept>
  <HasTopConcept>C011</HasTopConcept>
</ThesaurusConcept>
<ThesaurusConcept dc:identifier="C007">
  <dcterms:created>2007-06-05</dcterms:created>
  <PreferredTerm dc:identifier="11">
    <LexicalValue>French wines</LexicalValue>
    <dcterms:created>2007-06-05</dcterms:created>
  </PreferredTerm>
  <HasHierRelConcept Role="BT">C021</HasHierRelConcept>
  <HasTopConcept>C021</HasTopConcept>
</ThesaurusConcept>
<ThesaurusConcept dc:identifier="C008">
  <dcterms:created>2007-06-05</dcterms:created>
  <TopConcept>true</TopConcept>
  <PreferredTerm dc:identifier="1">
    <LexicalValue>fruits</LexicalValue>
    <dcterms:created>2007-06-05</dcterms:created>
  </PreferredTerm>
  <HasHierRelConcept Role="NT">C005</HasHierRelConcept>
  <HasHierRelConcept Role="NT">C018</HasHierRelConcept>
</ThesaurusConcept>
<ThesaurusConcept dc:identifier="C009">
  <dcterms:created>2007-06-05</dcterms:created>
  <PreferredTerm dc:identifier="12">
    <LexicalValue>German wines</LexicalValue>
    <dcterms:created>2007-06-05</dcterms:created>
  </PreferredTerm>
  <HasHierRelConcept Role="BT">C021</HasHierRelConcept>
  <HasHierRelConcept Role="NT">C013</HasHierRelConcept>
  <HasTopConcept>C021</HasTopConcept>
</ThesaurusConcept>
```

Figure A.14 **Example D: Data encoded in DD 8723-5 format** (*continued*)

```
<ThesaurusConcept dc:identifier="C010">
  <dcterms:created>2007-06-05</dcterms:created>
  <PreferredTerm dc:identifier="15">
    <LexicalValue>glassware</LexicalValue>
    <dcterms:created>2007-06-05</dcterms:created>
  </PreferredTerm>
  <HasHierRelConcept Role="BT">C011</HasHierRelConcept>
  <HasHierRelConcept Role="NT">C020</HasHierRelConcept>
  <HasTopConcept>C011</HasTopConcept>
</ThesaurusConcept>
<ThesaurusConcept dc:identifier="C011">
  <dcterms:created>2007-06-05</dcterms:created>
  <TopConcept>true</TopConcept>
  <PreferredTerm dc:identifier="3">
    <LexicalValue>hardware</LexicalValue>
    <dcterms:created>2007-06-05</dcterms:created>
  </PreferredTerm>
  <HasHierRelConcept Role="NT">C002</HasHierRelConcept>
  <HasHierRelConcept Role="NT">C010</HasHierRelConcept>
</ThesaurusConcept>
<ThesaurusConcept dc:identifier="C012">
  <dcterms:created>2007-06-05</dcterms:created>
  <PreferredTerm dc:identifier="7">
    <LexicalValue>lemons</LexicalValue>
    <dcterms:created>2007-06-05</dcterms:created>
  </PreferredTerm>
  <HasHierRelConcept Role="BT">C005</HasHierRelConcept>
  <HasTopConcept>C008</HasTopConcept>
</ThesaurusConcept>
<ThesaurusConcept dc:identifier="C013">
  <dcterms:created>2007-06-05</dcterms:created>
  <PreferredTerm dc:identifier="19">
    <LexicalValue>Liebfraumilch</LexicalValue>
    <dcterms:created>2007-06-05</dcterms:created>
  </PreferredTerm>
  <HasHierRelConcept Role="BT">C009</HasHierRelConcept>
  <HasHierRelConcept Role="BT">C017</HasHierRelConcept>
  <HasHierRelConcept Role="BT">C019</HasHierRelConcept>
  <HasTopConcept>C021</HasTopConcept>
</ThesaurusConcept>
<ThesaurusConcept dc:identifier="C014">
  <dcterms:created>2007-06-05</dcterms:created>
  <PreferredTerm dc:identifier="10">
    <LexicalValue>mangoes</LexicalValue>
    <dcterms:created>2007-06-05</dcterms:created>
  </PreferredTerm>
  <HasHierRelConcept Role="BT">C018</HasHierRelConcept>
  <HasTopConcept>C008</HasTopConcept>
</ThesaurusConcept>
<ThesaurusConcept dc:identifier="C015">
  <dcterms:created>2007-06-05</dcterms:created>
  <PreferredTerm dc:identifier="6">
    <LexicalValue>oranges</LexicalValue>
    <dcterms:created>2007-06-05</dcterms:created>
  </PreferredTerm>
  <HasHierRelConcept Role="BT">C005</HasHierRelConcept>
  <HasTopConcept>C008</HasTopConcept>
</ThesaurusConcept>
<ThesaurusConcept dc:identifier="C016">
  <dcterms:created>2007-06-05</dcterms:created>
  <PreferredTerm dc:identifier="13">
    <LexicalValue>red wines</LexicalValue>
    <dcterms:created>2007-06-05</dcterms:created>
  </PreferredTerm>
  <HasHierRelConcept Role="BT">C021</HasHierRelConcept>
  <HasTopConcept>C021</HasTopConcept>
</ThesaurusConcept>
```

Figure A.14 **Example D: Data encoded in DD 8723-5 format** (*continued*)

```
  <ThesaurusConcept dc:identifier="C017">
    <dcterms:created>2007-06-05</dcterms:created>
    <PreferredTerm dc:identifier="21">
      <LexicalValue>Riesling wines</LexicalValue>
      <dcterms:created>2007-06-05</dcterms:created>
    </PreferredTerm>
    <HasHierRelConcept Role="BT">C021</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C013</HasHierRelConcept>
    <HasTopConcept>C021</HasTopConcept>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C018">
    <dcterms:created>2007-06-05</dcterms:created>
    <PreferredTerm dc:identifier="5">
      <LexicalValue>tropical fruits</LexicalValue>
      <dcterms:created>2007-06-05</dcterms:created>
    </PreferredTerm>
    <HasHierRelConcept Role="BT">C008</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C001</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C003</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C014</HasHierRelConcept>
    <HasTopConcept>C008</HasTopConcept>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C019">
    <dcterms:created>2007-06-05</dcterms:created>
    <PreferredTerm dc:identifier="14">
      <LexicalValue>white wines</LexicalValue>
      <dcterms:created>2007-06-05</dcterms:created>
    </PreferredTerm>
    <HasHierRelConcept Role="BT">C021</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C013</HasHierRelConcept>
    <HasTopConcept>C021</HasTopConcept>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C020">
    <dcterms:created>2007-06-05</dcterms:created>
    <PreferredTerm dc:identifier="17">
      <LexicalValue>wineglasses</LexicalValue>
      <dcterms:created>2007-06-05</dcterms:created>
    </PreferredTerm>
    <HasHierRelConcept Role="BT">C010</HasHierRelConcept>
    <HasTopConcept>C011</HasTopConcept>
  </ThesaurusConcept>
  <ThesaurusConcept dc:identifier="C021">
    <dcterms:created>2007-06-05</dcterms:created>
    <TopConcept>true</TopConcept>
    <PreferredTerm dc:identifier="2">
      <LexicalValue>wines</LexicalValue>
      <dcterms:created>2007-06-05</dcterms:created>
    </PreferredTerm>
    <HasHierRelConcept Role="NT">C004</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C007</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C009</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C016</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C017</HasHierRelConcept>
    <HasHierRelConcept Role="NT">C019</HasHierRelConcept>
  </ThesaurusConcept>
</Thesaurus>
```

Bibliography

NOTE All websites accessed 23 May 2008.

[1] WORLD WIDE WEB CONSORTIUM. *HTML 4.01 specification*. W3C specification, 24 December 1999. [Latest version available at http://www.w3.org/TR/html401.]

[2] WORLD WIDE WEB CONSORTIUM. *XML Schema – Part 0: Primer Second Edition*. W3C Recommendation, 28 October 2004. [Available at http://www.w3.org/TR/xmlschema-0/.]

[3] WORLD WIDE WEB CONSORTIUM. *XML Schema – Part 1: Structures Second Edition*. W3C Recommendation, 28 October 2004. [Available at http://www.w3.org/TR/xmlschema-1/.]

[4] WORLD WIDE WEB CONSORTIUM. *XML Schema – Part 2: Datatypes Second Edition*. W3C Recommendation, 28 October 2004. [Available at http://www.w3.org/TR/xmlschema-2/.]

[5] POSTEL, J., and REYNOLDS, R. *File Transfer Protocol*. RFC 0959. Internet Engineering Task Force, October 1985. [Online version available at http://tools.ietf.org/html/0959.]

[6] OBJECT MANAGEMENT GROUP. *Unified Modeling Language™ : UML® resource page*. Needham, Ma.: Object Management Group, 2008. [Available at http://www.uml.org/.]

[7] OBJECT MANAGEMENT GROUP. *Unified Modeling Language Specification*. Version 1.4.2, January 2005. [Available at http://www.omg.org/docs/formal/05-04-01.pdf. Also available from ISO as ISO/IEC 19501.]

[8] OBJECT MANAGEMENT GROUP. *Unified Modeling Language (UML), version 2.1.2*. Needham, Ma.: Object Management Group, 2007. [Available at http://www.omg.org/technology/documents/formal/uml.htm.]

[9] DUBLIN CORE METADATA INITIATIVE. *Dublin core metadata element set, version 1.1*. DCMI recommendation, 14 January 2008. [Latest version available at http://uk.dublincore.org/documents/dces/. Full set of DCMI standards available at http://dublincore.org/.]

[10] UK CABINET OFFICE. *e-Government Schema Guidelines for XML*. Version 3.1. London: Cabinet Office, 6 February 2004. [Available at http://www.govtalk.gov.uk/schemasstandards/developerguide_document.asp?docnum=946.]

[11] WORLD WIDE WEB CONSORTIUM. *Extensible Markup Language (XML) 1.0 (Fourth Edition)*. W3C Recommendation, 16 August 2006. [Latest version available at http://www.w3.org/TR/REC-xml.]

[12] WORLD WIDE WEB CONSORTIUM. *Extensible Markup Language (XML) 1.1 (Second Edition)*. W3C recommendation, 16 August 2006.
[Latest version available at http://www.w3.org/TR/xml11.]

[13] UNICODE CONSORTIUM, ed. JOAN ALIPRAND et al. *The Unicode standard, version 4.0*. Boston, Ma.: Addison-Wesley, 2003. ISBN 0-321-18578-1.
[Available at http://www.unicode.org/versions/Unicode4.0.0/.]

[14] US LIBRARY OF CONGRESS NETWORK DEVELOPMENT AND MARC STANDARDS OFFICE. *MARC 21 format for authority data*. Washington, D.C.: Library of Congress.
[Concise version available online at http://www.loc.gov/marc/authority/.]

[15] US LIBRARY OF CONGRESS NETWORK DEVELOPMENT AND MARC STANDARDS OFFICE. *MARC 21 format for classification data*. Washington, D.C.: Library of Congress.
[Concise version available online at http://www.loc.gov/marc/classification/.]

[16] US LIBRARY OF CONGRESS NETWORK DEVELOPMENT AND MARC STANDARDS OFFICE. *MARC 21 XML schema*. Washington, D.C.: Library of Congress.
[Available at http://www.loc.gov/standards/marcxml/.]

[17] WORLD WIDE WEB CONSORTIUM. *SKOS Simple Knowledge Organization System Primer*. W3C Working Draft 21 February 2008.
[Latest version available at http://www.w3.org/TR/swbp-skos-core-guide/.]

[18] WORLD WIDE WEB CONSORTIUM. *SKOS Simple Knowledge Organization System Reference*. W3C Working Draft 25 January 2008.
[Latest version available at http://www.w3.org/TR/swbp-skos-core-spec/.]

[19] TAYLOR, HAMMER, LEVAN, LYNCH, PLACE, STEVENS and TICE. *The Zthes specifications for thesaurus representation, access and navigation.* 17 February 2006.
[Available at http://zthes.z3950.org/.]

[20] JANÉE, G, IKEDA, S. and HILL, LL. *The ADL Thesaurus Protocol – Version 1.0*.
[Latest version available at http://www.alexandria.ucsb.edu/thesaurus/specification.html.]

[21] WORLD WIDE WEB CONSORTIUM, INFORMATION SOCIETY TECHNOLOGIES and SWAD EUROPE. *SKOS API (Simple Knowledge Organization Systems Application Programming Interface)*.
[Latest version available at http://www.w3.org/2001/sw/Europe/reports/thes/skosapi.html.]

[22] US LIBRARY OF CONGRESS. *SRU: Search/Retrieve via URL*. Washington, D.C.: Library of Congress. [Online version available at http://www.loc.gov/standards/sru/.]

[23] US LIBRARY OF CONGRESS. *Z39.50*. Maintenance Agency page. Washington, D.C.: Library of Congress. [Online version available at http://www.loc.gov/z3950/agency/.]

Index (DD 8723-5)

Key to references in index
F Reference to figure
T Reference to table

This page deliberately left blank

BSI – British Standards Institution

BSI is the independent national body responsible for preparing British Standards. It presents the UK view on standards in Europe and at the international level. It is incorporated by Royal Charter.

Revisions

British Standards are updated by amendment or revision. Users of British Standards should make sure that they possess the latest amendments or editions.

It is the constant aim of BSI to improve the quality of our products and services. We would be grateful if anyone finding an inaccuracy or ambiguity while using this British Standard would inform the Secretary of the technical committee responsible, the identity of which can be found on the inside front cover.
Tel: +44 (0)20 8996 9000 Fax: +44 (0)20 8996 7400

BSI offers members an individual updating service called PLUS which ensures that subscribers automatically receive the latest editions of standards.

Buying standards

Orders for all BSI, international and foreign standards publications should be addressed to Customer Services. Tel: +44 (0)20 8996 9001
Fax: +44 (0)20 8996 7001 Email: orders@bsigroup.com

You may also buy directly using a debit/credit card from the BSI Shop on the Website http://www.bsigroup.com/shop.

In response to orders for international standards, it is BSI policy to supply the BSI implementation of those that have been published as British Standards, unless otherwise requested.

Information on standards

BSI provides a wide range of information on national, European and international standards through its Library and its Technical Help to Exporters Service. Various BSI electronic information services are also available which give details on all its products and services. Contact the Information Centre. Tel: +44 (0)20 8996 7111
Fax: +44 (0)20 8996 7048 Email: info@bsigroup.com

Subscribing members of BSI are kept up to date with standards developments and receive substantial discounts on the purchase price of standards. For details of these and other benefits contact Membership Administration. Tel: +44 (0)20 8996 7002
Fax: +44 (0)20 8996 7001 Email: membership@bsigroup.com

Information regarding online access to British Standards via British Standards Online can be found at http://www.bsigroup.com/BSOL.

Further information about BSI is available on the BSI website at http://www.bsigroup.com.

BSI Group Headquarters
389 Chiswick High Road,
London W4 4AL, UK
Tel +44 (0)20 8996 9001
Fax +44 (0)20 8996 7001
www.bsigroup.com/standards

Copyright