BS 8723 advances to encompass interoperability

Stella G Dextre Clarke

Convenor, IDT/2/2 Working Group of BSI

Thesaurus standards back in 2001

- ISO 2788-1986 Guidelines for the establishment and development of monolingual thesauri

  = BS 5723:1987

- ISO 5964-1985 Guidelines for the establishment and development of multilingual thesauri

  = BS 6723:1985

Procedural aspects

- Constituting a working group for an international standard takes a while
- It seemed easier to work first on the British Standards; then offer these up to the international community
- The BSI committee must be UK-based, but invites international input wherever possible without offending BSI/ISO

Procedure in detail

- Small UK Working Group of unpaid volunteers submits drafts to BSI committee IDT/2/2

- Comments are invited informally from everyone interested

- Progress is very slow… but worthwhile, we hope

Outline of new standard

BS 8723: Structured vocabularies for information retrieval – Guide

- 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

Progress to date

- Parts 1 and 2 published November 2005
- Drafts of Parts 3 and 4 have had first round of circulation and feedback; revised drafts will go out for public consultation Autumn 2006
- Draft of Part 5 currently in development

New features in BS 8723-2: [monolingual] Thesauri

- Clearer guidance on applying facet analysis to thesauri
- Some changes to the 'rules' for compound terms
- More guidance on managing thesaurus development and maintenance
- Functional specification for software to manage thesauri
- General overhaul in all areas

BS 8723-4: Interoperability between vocabularies

- Huge demand for accessing information that has been indexed with another language and/or vocabulary. The buzzword is 'Mapping'. The Semantic Web is just one application.
- Part 4 to include multilingual thesauri as a special case of mapping between vocabularies
- Other challenges:
  - Accommodating vocabularies other than thesauri, e.g. classification schemes, subject headings, ontologies
  - mapping between precoordinate and postcoordinate schemes
  - mapping when notations are not all enumerated, but built up by rules for synthesis?

Chapters of BS 8723-3: Vocabularies other than thesauri

- Classification schemes
- Business classification schemes for records management
- Taxonomies
- Subject heading schemes
- Ontologies
- Authority lists (?)

Challenges for BS 8723-3

- Accommodating mixed objectives:
  - Providing *normative* advice where needed
  - *Describing* existing schemes for which it may be too late to apply a new standard
- Taxonomies chapter to be mostly normative
- Ontologies and Authority lists purely descriptive
- The others have a mix

BS 8723-5: Exchange formats and protocols for interoperability

- Vocabularies must work with
    - Search engines
    - Content Management Systems
    - Web publishing software, etc.
- Build on existing formats and protocols for data exchange, e.g. Z39.50 and Zthes, XML schema, MARC, SKOS Core, Topic Maps, ADL thesaurus protocol and others?
- Workshop in London, October 2006

International adoption

- To turn BS 8723 into an ISO standard, we need an international committee formed of delegates from the corresponding committees of national standards bodies (ISO TC46/SC9)

- Interest already expressed by AENOR, AFNOR and others. More needed!

- It is the opportunity to make a lot of improvements

- The campaign starts formally in December 2006. Just let me know if you want to join in.

Want to get involved?

- If you would like to contribute to the next stages, please get in touch!
- Contact SDClarke@LukeHouse.demon.co.uk