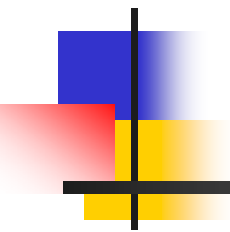


The BS8723 thesaurus data model and exchange format, and its relationship to SKOS



Stella G Dextre Clarke (Convenor, IDT/2/2
Working Group of BSI and Project Leader for ISO NP 25964)

Leonard D Will

Nicolas Cochard



Summary

- Background on BS8723 and its relationships with other key standards
- Closer look at relationship with SKOS
- Leonard Will to discuss the data modelling challenges
- Nic Cochard will compare DD8723-5 format with SKOS
- A look forward to the opportunities offered by ISO 25964 development



Evolution background: 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](#)
- None of them had any vision of the web world we now inhabit



Strategy of BSI working group

- Revision and extension of ISO 2788 and ISO 5964 was the long-term goal
- Revision of BS 5723 and BS 6723 a shorter-term tactic. Bring them together in a new BS 8723
- BSI drafting committee IDT/2/2/1 comprises Leonard Will, Alan Gilchrist, Ron Davies and convenor Stella Dextre Clarke. Nic Cochard was co-opted later
- Comments were invited and received internationally wherever possible
- On completion, BS 8723 was offered for adoption as an international standard [to become ISO 25964]



Overview of BS 8723

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

Motivation throughout is “interoperability”



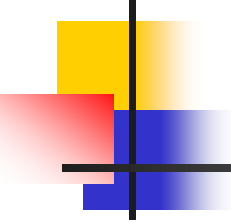
BS8723 Part 5: Exchange formats and protocols for interoperability

- In Oct 2006 a workshop was held, to seek advice from experts
- Key conclusions:
 - We need a data model
 - We should recommend one standard format, not several, and it should be based on XML
 - The format should provide only for thesauri
 - We should not develop a new protocol
- Model and schema at <http://schemas.bs8723.org/>
- DD 8723-5 is to be published July 2008 (DD stands for Draft for Development)



BS8723-5 in relation to SKOS

- Early work on SKOS took account of BS8723-2, leading to much commonality
- But there is a difference of objective
 - SKOS geared towards Semantic Web function
 - BS8723-5 to ease the exchange of thesauri conforming to BS8723-2 and BS8723-4



ISO NP 25964 (adoption of BS 8723 as an ISO standard)

- The proposal to revise ISO 2788 and ISO 5964, basing the work on BS 8723, was submitted to ISO TC 46/SC 9 members in April 2007
- Project now approved
- At least 11 countries participating: Canada, Denmark, France, Germany, Finland, New Zealand, South Africa, Sweden, UK, Ukraine, USA . Maybe China too.



Next steps towards ISO 25964

- The ISO 25964 WG has only recently got going.
- Significant evolution from BS 8723 is expected.
- It will rearrange the 5 Parts inherited from BS 8723 into 2 Parts as follows:
- Thesauri and interoperability with other vocabularies
 - Part 1: Thesauri for information retrieval
 - Part 2: Interoperability with other vocabularies
- Leonard Will is leading the work on data modelling and formats. This gives us the opportunity to make adjustments where useful.



So how should the relationship with SKOS evolve?

- Adjust ISO 25964 to match SKOS?
- Adjust SKOS to match ISO 25964?
- Keep them distinct, but provide easy means of converting between them?
- And what about other players, e.g. MARC, Zthes, etc?
- Over to you!