Presentation: ISO 25964-1: a new standard for development of thesauri and exchange of thesaurus data

Presenters: Stella Dextre Clarke, Johan De Smedt.

Date: September 28, 2011

Session 1:

Abstract:

ISO 25964-1 is the newly published ISO standard "Information and documentation — Thesauri and interoperability with other vocabularies — Part 1: Thesauri for information retrieval", covering mono- and multilingual thesauri. After a general introduction to the standard, this presentation will focus on the XML schema that was developed for exchange of thesaurus data. The schema is underpinned by a UML data model that necessarily provides for all the features a thesaurus may include, e.g. terms, relations among terms, concepts, relations among concepts, arrays, concept groups, etc. The presentation discusses how the XML schema deals with the several thesaurus aspects, including some custom extensions on the basic formal framework. Some constraints have been incorporated in the schema to ensure the quality of the represented thesauri. A comparison is made with the SKOS schema (developed in the context of the semantic web), and some conclusions are drawn regarding the differences and the application domains. This helps to identify SKOS extensions that may capture thesaurus specifics that are not handled currently.