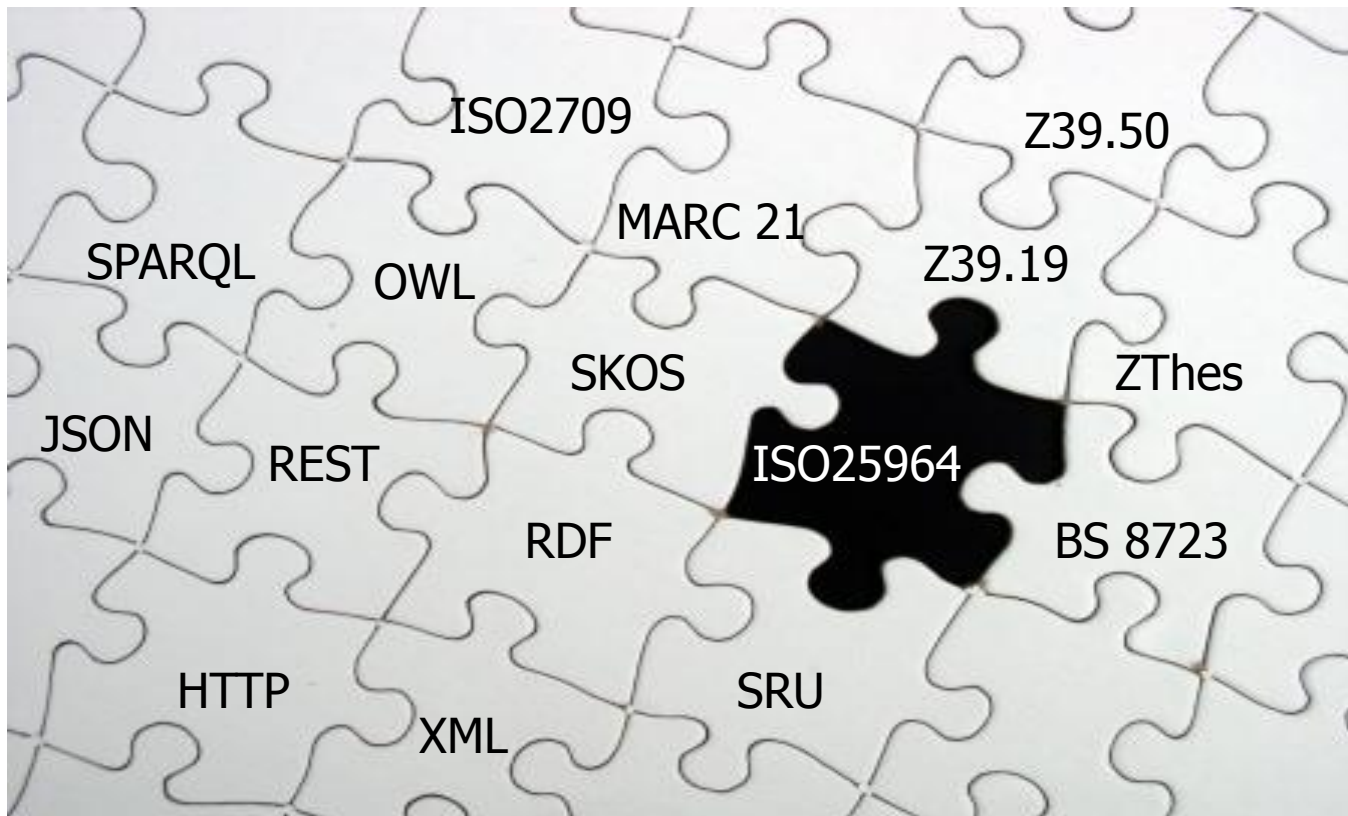




ISO25964 as a prelude to DDC mapping projects

Stella Dextre Clarke
Convenor, ISO TC46/SC9 WG8

In a networked world, standards underpin everything





Introducing ISO 25964

ISO 25964: Thesauri and interoperability with other vocabularies

- Part 1: Thesauri for information retrieval *thesaurus in isolation*
- Part 2: Interoperability with other vocabularies *in a networked world*
- It updates ISO 2788 and ISO 5964
- Shares common ground with ANSI/NISO Z39.19
- Part 1, **published in 2011**, covers monolingual and multilingual thesauri
- Part 2, **published in 2013**, covers mapping between thesauri and other types of vocabulary
- information retrieval seen as main application; mapping applies to index terms or to search terms
- See more at <<http://www.niso.org/schemas/iso25964/>>



Overview of this Presentation

- Thesaurus fundamentals
- Interoperability objectives/assumptions
- Models for mapping
- Mapping types
- Symbols, tags, conventions
- Handling the data
- Vocabulary types in ISO 25964-2
- How to handle pre-coordination



Simple fundamentals of the thesaurus (Part 1 Clause 4)

- Concepts are represented by terms
- aim is to guide indexer and searcher....to choose the same term for the same concept [This is why we need to establish relationships between terms and concepts]
- As one consequence, users need to accept a degree of artificiality in the vocabulary



Relationship tags

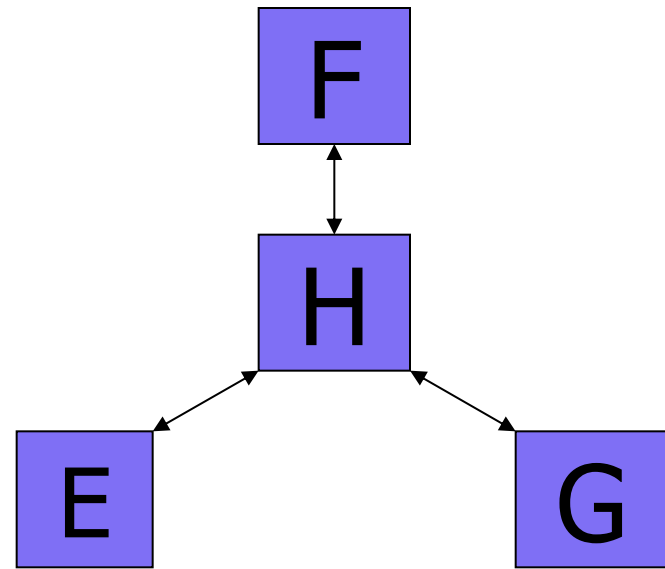
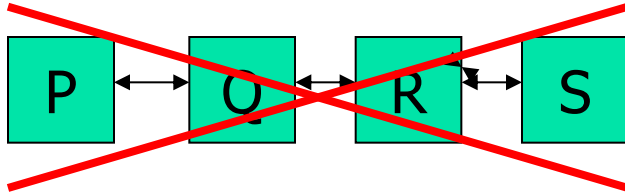
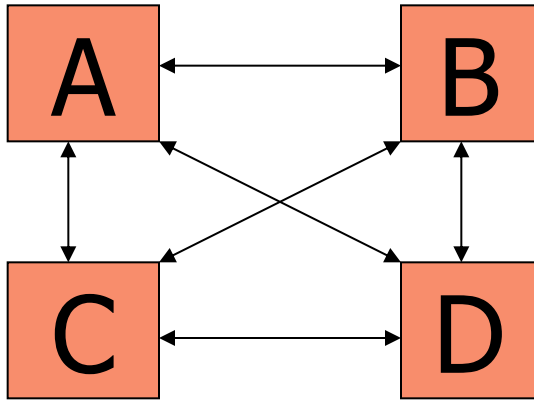
Tag	Meaning
USE	Use [before the corresponding preferred term]
UF	Used for [before the corresponding non-preferred term]
USE [A] + [B]	Use [before combinations of terms]
UF+	Used for [reciprocal of term combinations]
BT	Broader term
NT	Narrower term
RT	Related term

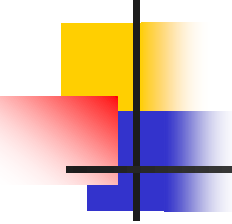


Interoperability objectives/assumptions

- ISO 25964 deals with interoperability (not just mapping)
- Definitions of mapping (for gerund/noun respectively):
 - process of establishing relationships between the **concepts** of one vocabulary and those of another
 - relationship between a **concept** in one vocabulary and one or more **concepts** in another
- Context is information retrieval (conversion of either search terms or metadata)
- Boolean methods are assumed (probably too much)
- Semantic Web expectations motivate interoperability
- What is your own context? assumptions? objectives?

Recommended "Models for mapping"





Full range of ISO 25964-2 mapping types

- Basic mapping types:

- Equivalence

- Simple

- Compound**

- Intersecting compound equivalence**

- Cumulative compound equivalence**

- Hierarchical

- Broader

- Narrower

- Associative

- **Simple equivalence can be marked as "Exact" or "Inexact"**

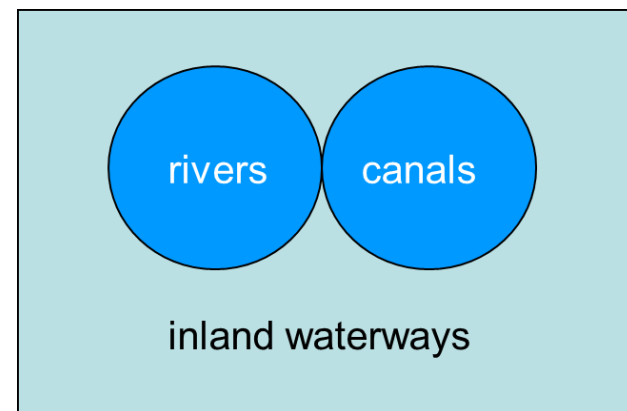
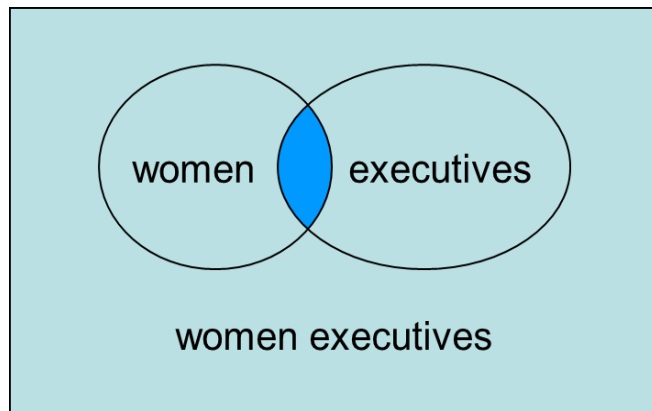


Full range of ISO 25964-2 mapping types with examples

- Basic mapping types:
 - Equivalence
 - Simple: Laptop computers EQ Notebook computers
 - Compound
 - Intersecting compound equivalence: Women executives EQ Women + Executives
 - Cumulative compound equivalence: Inland waterways EQ Rivers | Canals
 - Hierarchical
 - Broader: Streets BM Roads
 - Narrower: Roads NM Streets
 - Associative: e-Learning RM Distance education
- Exact equivalence: Aubergines =EQ Egg-plants
- Inexact equivalence: Horticulture ~EQ Gardening

Intersecting versus cumulative equivalence

Women executives EQ Women + Executives
Inland waterways EQ rivers | canals





Symbols, tags, conventions

Relationship tag	Corresponding tag for mappings	Optional subdivision
USE	EQ	=EQ ~EQ
UF	[not applicable]	
[not applicable] USE [A] + [B]	EQ [A] [B] EQ [A] + [B]	
UF+	[not applicable]	
BT	BM	
NT	NM	
RT	RM	



Another complication: mappings to more than one vocabulary

dairy products

UF milk products

BT animal products

NT butter
milk

RT dairies

VOC1 =EQ milk products

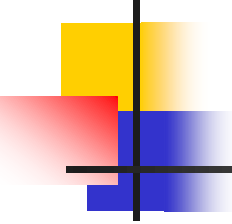
VOC2 BM animal products

VOC2 NM milk

Subjectivity pervades mapping projects!



- What's the difference between an associative mapping and an inexact equivalence?
- Examples:
 - e-learning RM distance education
 - horticulture ~EQ gardening



Techniques for identifying candidate mappings

- General procedure – the basic intellectual process
- Computer-assisted direct matching
- Co-occurrence mapping
- Other methods – look for them!



Managing the data

- What to record – at three levels (individual mappings; sets of mappings between 2 vocabs; mapping clusters)
- Storing the data – where and how
- Maintenance and noting the nature of changes



Other vocabulary types in ISO 25964-2

- Mapping (linking) to/from various vocabulary types:
 - other thesauri
 - classification schemes
 - file plans (Classification schemes used for records management)
 - taxonomies
 - subject heading schemes
 - ontologies
 - terminologies
 - name authority lists
 - synonym rings

General prospects for mapping to/from a thesaurus

↔ thesaurus	mapping relatively straightforward
↔ name authority list	mapping usually straightforward but common concepts few
↔ classification scheme ↔ file plan ↔ taxonomy ↔ subject heading scheme	concept mapping useful in IR, pre-coordination common
- synonym rings - terminology - ontology	concept mapping rarely useful; complementary uses are a more likely prospect

General prospects for mapping to/from a thesaurus

↔ thesaurus	mapping relatively straightforward	Hierarchies should be logical
↔ name authority list	mapping usually straightforward but common concepts few	
↔ classification scheme ↔ file plan ↔ taxonomy ↔ subject heading scheme	concept mapping useful in IR, pre-coordination common	Hierarchies follow user convenience
- synonym rings - terminology - ontology	concept mapping rarely useful; complementary uses are a more likely prospect	



SKOS – a special ontology case

- SKOS complements ISO 25964 (was developed in parallel)
- ISO 25964 tells how to build a thesaurus; **SKOS tells how to publish it on the World Wide Web**
- ISO 25964 includes some thesaurus features not provided for in SKOS; **SKOS applies to other vocabulary types as well as thesauri**



Mapping properties in SKOS (Simple Knowledge Organization System)

Basic “properties” (skos:mappingRelation):

- skos:closeMatch (symmetric)
 - skos:exactMatch (symmetric, transitive)
- skos:relatedMatch (symmetric)
- skos:broadMatch (inverse of narrowmatch)
- skos:narrowMatch (inverse of broadmatch)
- No provision for compound mappings
- But much attention to reciprocity
- SKOS was developed for the Semantic Web



ISO25964 → SKOS conversion

ISO 25964 mapping tags	SKOS mapping property
EQ	skos:closeMatch
=EQ	skos:exactMatch
~EQ	skos:closeMatch
EQ [A] [B] EQ [A] + [B]	? 2 x skos:narrowMatch ? ? 2 x skos:broadMatch ?
BM	skos:broadMatch
NM	skos:narrowMatch
RM	skos:relatedMatch



Linking ISO 25964 data model with SKOS

- The ISO 25964 data model is broadly compatible with SKOS, especially when the SKOS-XL extension is used. See correspondence table at http://www.niso.org/apps/group_public/download.php/12351/Correspondence%20ISO25964-SKOSXL-MADS-2013-12-11.pdf
- An RDF schema encapsulating the mappings between models is at <http://purl.org/iso25964/skos-thes>



Need a copy of ISO 25964 ?

- Download it from ISO at
http://www.iso.org/iso/home/store/catalogue_detail.htm?csnumber=53657
http://www.iso.org/iso/home/store/catalogue_detail.htm?csnumber=53658
- Order it from your national standards body (e.g. BSI, DIN, ANSI, AFNOR)
- Some public/academic reference libraries may stock it
- It is not cheap to purchase ☹
- However, the **XML schema** for exchange of thesaurus data is in an Annex which is available online **without charge or password control**. Go to <http://www.niso.org/schemas/iso25964/>