

Crosswalking from Terminology to Terminology

**Leveraging Semantic Information
across
Communities of Practice**

Sue Ellen Wright & David Sommers

**LPCC 2008
mappakech**

Terminology ≠ Terminology

- Controlled vocabularies are made up of *terms* and are widely called *terminologies*: a list of terms that have been enumerated explicitly (ANSI Z39-19)
- Concept oriented lexical databases used for processing of spoken and written discourse using special languages are also called *terminologies*. (ISO 704, 1087)

Original Premise

- Desirability of exchanging or leveraging information on concept relations and rigorous definitions embedded in terminological databases and expressed in the TermBase eXchange format (TBX) and controlled vocabularies (SKOS)
- Hypothesis: critical data categories used in the two systems should be isomorphic enough to enable interactivity

LEEC 2008
marrakech

Terminologies

- *SKOS: Subject-purposed vocabularies = Subject Language Terminologies*
 - *Svenonius: SLTs*
 - *E.g., thesauri: controlled vocabularies*
 - *Information storage & retrieval*
- *TBX: Language-purposed vocabularies*
 - *Tudhope*
 - *Terminological Databases (TDBs)*
 - *Uncontrolled vocabularies; discourse oriented*

Hypothesis: TDB Term = SKOS Term (Label)?

SKOS classes and properties	12620 data category → SKOS	12620 data category ≠ Not in SKOS
Labels and Terms		
label	N/A	term (A.1)

- Svenonius:
 - SLTs designed for the special purpose of retrieving information; extension = all *documents about a subject* (e.g., all documents about butterflies)
 - TDBs designed to document terms used in discourse; extension = the class consisting of *all objects covered by the concept* (e.g., all butterflies)

LEEC 2008
marrakech

Further *Label-Term* Issues

Labels and Terms		
label	N/A	term (A.1)
prefLabel	term (A.1), status = preferred term	
altLabel	variant, abbreviation, full form, etc.	term, status = admitted term (A.2.9.1)
hiddenLabel	term status = deprecated term (A.2.9.1)	term, status = deprecated term (not recommended) (A.2.9.1)
	Misspelling	(not included in TBX)

- *Term* is always present in TDBs and *optionally* designated with a status marker
- *Label* does not occur without a status marker
- Solution: Treat *term* as *preferred* unless otherwise indicated?

LPCC 2008
marrakech

Problems with Proposed Solution

- Unlike in SLTs, TDBs can feature several preferred or admitted terms in one entry, with other factors (such as customer subsets, project identifiers, etc.) used to discriminate the contextual relativity of usage in human discourse.
- Thus the solution is problematic and not straight-forward.

TermType Issues

Labels and Terms		
symbol	term termType = symbol (A.2.1)	
altSymbol	Ditto above, status = admitted term (A.2.9.1)	
prefSymbol	Ditto above, status = preferred term (A.2.9.1)	

- SKOS makes allowances for the use of an alternative symbol or preferred symbol.
- TBX allows a wide variety (approx. 20) different possible term types (abbreviations, etc.)
- These may alternate contextually with regard to their “preferred” status

LEEC 2008
marrakech

Hypothesis: Definition = Definition?

Definitions		
definition	definition (A.5.1)	(See discussion)
scopeNote	explanation (A.5.2)	(See discussion)

- SKOS definition: *A statement or formal explanation of the meaning of a concept.*
- SKOS scopeNote: *A note that helps to clarify the meaning of a concept.*
- These declarations coincide ostensibly with TBX */definition/* and */explanation/*, respectively.

LPCC 2008
marrakech

Problems with Proposed Solution

- Skos examples:

```
<skos:Concept rdf:about="http://my.example.org/GCL/702#scopeNote">  
  <skos:prefLabel xml:lang="en"> Competitiveness</skos:prefLabel>  
  <skos:scopeNote xml:lang="en">The ability of businesses to compete in local,  
national or international markets.</skos:scopeNote>  
</skos:Concept>
```

```
<skos:Concept rdf:about="http://www.example.com/concepts#banana">  
  <skos:prefLabel xml:lang="en"> banana</skos:prefLabel>  
  <skos:definition xml:lang="en">A long curved fruit with a yellow skin and soft, sweet  
white flesh inside.</skos:definition>  
</skos:Concept>
```

- Problem: Both of these items are formal definitions in the sense of TBX (ISO 704)

LREC 2008
marrakech

Further SKOS Examples

```
<skos:definition>A feature type category for places such as the Erie Canal
```

```
</skos:definition>
```

```
<skos:scopeNote>Manmade waterway used by watercraft or for drainage, irrigation,  
mining, or water power</skos:scopeNote> (Mikhalenko 2005)
```

- This approach, if followed, would suggest flipping the crosswalk assignment, with TBX */definition/* mapping to SKOS *scopeNote*, and TBX */explanation/* mapping to SKOS *definition*!
- Consultation with SKOS experts is desirable to pin down actual practice in light of these two approaches, but the second style conforms to Sevonius's suggested role of definitions for SLTs in that it defines a feature type for purposes of retrieval rather than a concept associated with a discourse-oriented term.

ЛРСС 2008
marrakech

Hypothesis: Note Maps to Note?

Notes		
note	note (A.8)	
changeNote	note inside transacGrp, type="modification"	
editorialNote	note (A.8) within terminology management transactions (A.10.1)	
example	example (A.5.4)	
historyNote	history note	(not included in TBX)
scopeNote	explanation (A.5.2)	(See discussion above)

- ScopeNote and Definition (also listed as a note in SKOS) accounted for
- TBX */example/* can map to SKOS *example*
- But unqualified SKOS *note* is only a superordinate placeholder

LEEC 2008
marrakech

Hypothesis: Note Maps to Note?

Notes		
note	note (A.8)	
changeNote	note inside transacGrp, type="modification"	
editorialNote	note (A.8) within terminology management transactions (A.10.1)	
historyNote	history note	(not included in TBX)

- Other notes:
 - TBX notes are differentiated syntactically through embedding in different structures.
 - E.g, *changeNote* would be embedded in a TBX `<transacGrp>` of type *modification*
 - *historyNote* is interesting and might in some cases be a useful addition to TBX

Leec 2008
marrakech

Concept-Related Mappings

Concept System/Scheme		
ConceptScheme	concept system (A.7.1)	
Concept	entry identifier (A.10.15)	
inScheme	[pointer to A.7.1]	(suggested for TBX)

- *SKOS conceptSchemes* are only *optionally* based on statements of concept relations.
- TBX termbases also are optional with regards to statements of relations.
- But TBX *concept systems* by their very definition state relations (superordinate, subordinate, coordinate, etc.)

LPCC 2008
mappakech

Proposal for Addition to TBX

- *inScheme* points from a concept entry to the identifier for a concept scheme/concept system
- This would be extremely useful for use with termbases because within any given TDB, there may be numerous facets represented by different concept systems, so any given relational statement is only relevant within the framework of one specific ordering system.

LPCC 2008
māppakēch

Subject = Subject Field?

Subject references		subject field (A.4.5)
subject		(not broken down in TBX; see discussion above)
primarySubject		But multiples allowed
subjectIndicator		
isSubjectOf		
isPrimarySubjectOf		

- SKOS orientation relates to subjects treated in knowledge objects.
- TBX orientation relates to subject fields in which a concept occurs.
- Both allow for multiple references, but the nature of the reference is different, although some analogies may be useful.

ЛРСС 2008
marrakech

Relations

Relations		
semanticRelation	N/A	Concept relations
hasTopConcept	Potentially: broader concept generic (A.7.2.1)	
broader	superordinate concept generic (A.7.2.2)	
narrower	subordinate concept generic (A.7.2.3)	

- The notion of *semantic relation* is subsumed in the TBX notion of a *concept relation*.
- *Top concept* is only potentially related to *broader concept generic* because the broader concept in TBX is several abstractions higher, but not necessarily the top concept.
- Broader and narrower fit superordinate and subordinate.

Collections = Coordinate Concepts?

Relations		
CollectableProperty	N/A	[Embedded as genus element in rigorous definitions]
Collection		[Any superordinate concept that could become the subject of a collection]
member	coordinate concept (A.7.2.4)	[Any subordinate concept that could become a member of a collection]
OrderedCollection	Ordered thesaurus (A.9.6) ??	
memberList	ordered coordinate concept (A.7.2.4.1)	

- Collectable property not explicit as a separate element in TBX.
- Collection: can be generated as all children of a single superordinate concept (coordinate; concepts)
- Ordered collection: not accounted for; items listed are just suggestion.

LPCC 2008
māppakēch

Other Relations

Relations		
related	related concept (A.7.2.5)	
		broader concept partitive
		superordinate concept partitive
		subordinate concept partitive
		coordinate concept partitive
		temporally related concept
		spatially related concept
		associated concept
		(Nuopponen categories)
		antonym (is disjoint with) (10.18.6.1)

- TBX has options that go beyond generic relations.
- For this purpose we propose that SKOS is inadequate for our purposes.
- Proposal to move to OWL-Lite.

LPCC 2008
marrakech

Leveraging Concept-Oriented Resources

- Rigorous definitions reside in many different kinds of resources
- Variety of xml-based formats and markup languages
- Ability to reference authoritative terms and concepts for reuse in other environments
- Ability to build external relational registers that use authoritative documentation to support ontological claims

LEEC 2008
marrakech

Probable Venues

- NLP research
- Industry players seeking to link terminology, taxonomy, and ontology resources
- Government agencies and bureaus creating knowledge management resources

Lee 2008
marrakech

Contacts

- Sue Ellen Wright
Kent State University
sellenwright@gmail.com
- Dave Summers
Kent State University
dgsummers@gmail.com

LEEC 2008
māppakech