



C C L R C



SKOS

Simple Knowledge Organisation with
the Semantic Web

- The Semantic Web
 - Vision
 - Implementation
- How do I put my KOS on the semantic web?
 - SWAD-E Thesaurus Activity
 - SKOS is the answer!
- What do I get out of doing it?
 - Business case for using semantic web technologies
- The future ...
 - Community driven development
 - Participation



- The semantic web is a **vision**

- *“What if the world wide web were one giant database, linking both human-readable documents and machine-readable data in a way useful to both mankind and machine?”*

Scenario A: I want to go see Al Miles' presentation at the DELOS workshop in Lund ...

Event data (time, date, location) ...

Calendar (resolve conflict) ...

Dentists availability ...

New appointment ...

Flight search ...

Hotel search ...

etc. ...

Scenario B: Find me scientific papers in Swedish about the compound C₂H₄N ...

Locate journals services ...

Locate thesaurus service ...

Translate query terms ...

The semantic web can help wherever data is useful ***across applications***



- The semantic web ***will not*** replace the current web, but will add a new layer to it
 - Current web = web of information for human consumption
 - Semantic web = web of data for machine consumption
- Semantic web search engines will not replace Google ...
 - But they may make it smarter!
- The **semantic** web
 - Allows you to express yourself in a way that computers can understand
- The semantic **web**
 - Allows you to distributed data across the internet
- Publishing
 - Current web – **publish** information
 - Semantic web – **publish** data

- Continuing the analogy ...
 - Language is the framework for transferring meaning between people ...
 - ⇒ The semantic web depends on a framework for transferring meaning between computer programs ...
- RDF (Resource Description Framework) ...
 - Provides the grammar, the basic sentence structure (subject, predicate, object) (MyWebPage title 'Lund Presentation') (AlistairMiles livesIn Oxfordshire)
- RDF Schema ...
 - Allows you to define the nouns and verbs (i.e. 'Classes' and 'Properties') you want to use e.g. (AlistairMiles type Person)
- OWL (Web Ontology Language) ...
 - Builds on RDF and RDFS
 - Express even more ... e.g. constraints (a Person can only live in one place) (a Person cannot also be a Car)

- Use RDF Schema (RDFS) and OWL to define '**RDF Vocabularies**'
- (N.B. an 'RDF Schema' is an RDF vocabulary)
- (N.B. an 'OWL Ontology' is an RDF vocabulary)

- Current RDF vocabularies supporting RDF description of commonly used data ...
 - VCard (People, contact details)
 - FOAF (People, social networks)
 - Dublin Core (Metadata)
 - RDF-Cal (Calendar)
 - OWL (Ontologies)
- Open community developments
 - Semantic web is comfortable with plurality but ...
 - Huge benefits when schemas are widely adopted ...
 - Community development promotes consensus.

- Thesauri extremely useful and widely used resource
 -
 - 1998 DESIRE (Koch, Brickley)
 - 2000 LIMBER (Miller, Matthews, Wilson)
 - 2004 SWAD-E ...
- Also other efforts
 - CERES, CALL, ETBT ...
 - See SWAD-Europe Thesaurus Activity Deliverable 8.2: Review of RDF Thesaurus Work <http://www.w3.org/2001/sw/Europe/reports/thes/8.2/>
- Semantic Web foundations are now laid ...
 - RDF is a W3C rec, OWL is a W3C rec.
 - Tools are mature e.g. Jena, Sesame, Redland.

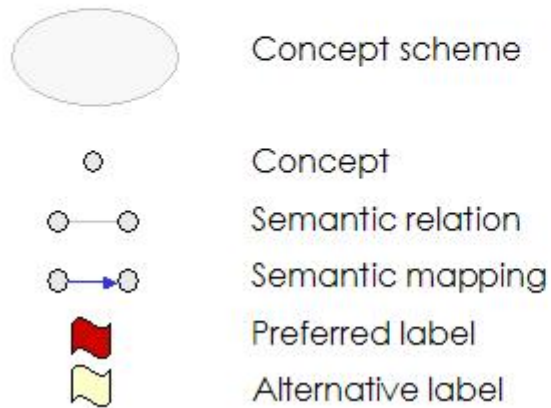
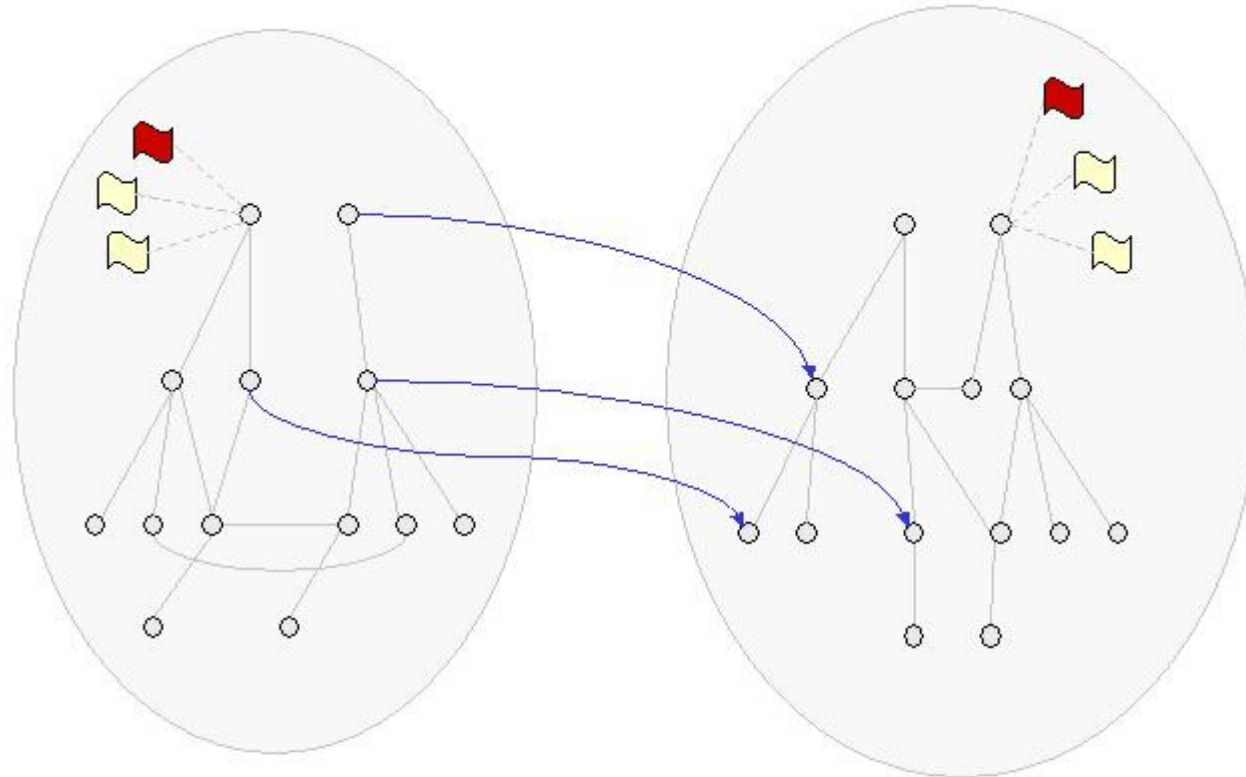


- Semantic Web Advanced Development for Europe
 - EU project
 - Supporting Semantic Web Activity of W3C
 - Applications, reference implementations, guidelines, best practises ...
 - Partners: CCLRC, HP Labs, ILRT, ERCIM, Stilo
- SWAD-E Thesaurus Activity exploring ...
 - RDF description of ...
 - Standard monolingual thesauri
 - Multilingual thesauri
 - Inter-thesaurus mapping
 - Classification schemes
 - Migration guidelines
 - Technology demo ...
 - SOAP web service API and implementation for thesaurus access



- How do I put my KOS on the semantic web?
- SKOS-Core ...
 - Simple Knowledge Organisation Systems
 - Focus on thesauri
 - Awareness of similar KOS types (at lower end of complexity axis)
 - Complement to OWL (perfect for high-end of complexity axis)
 - N.B. As Doug said – more complex **does not** mean better
 - ⇒ All types of KOS are useful, depends on use scenario
- SKOS-Core 1.0 public release ...
 - See SKOS-Core 1.0 Guide <http://www.w3.org/2001/sw/Europe/reports/thes/1.0/guide/>
 - See SKOS-Core 1.0 Guidelines for Migration <http://www.w3.org/2001/sw/Europe/reports/thes/1.0/migrate/>
 - Covers standard and non-standard thesauri

- SKOS-Core is based on strictly concept-oriented (C-O) model of thesaurus
- C-O Model ...
 - A thesaurus (concept scheme) consists of a set of concepts. Each concept may have a single preferred label, any number of alternative labels. Concepts within a thesaurus may be related via semantic relations (e.g. broader, narrower, related). Concepts from different thesauri may be related via semantic mappings.
- Cf. Term-Oriented (T-O) Model ...
 - A thesaurus consists of a set of terms. There are preferred terms and non-preferred terms. PTs may be related via BT NT RT, PT/NPT related via USE/UF.
- NB.
 - A thesaurus is a thesaurus ... C-O and T-O models are different ways of describing (viewing) the same thing ... there is no conflict.



- Ease of maintenance ...
 - On the semantic web concepts are identified by URIs ...
 - Can switch which labels are pref and alt without having to alter URI of concept
- Compatibility with other KOS types
 - E.g. classification schemes
- Reduce ambiguity and confusion to absolute minimum ...
 - E.g. When you say ‘term’ are you referring to a piece of meaning (a concept) or a string of characters? (This confusion has big impact on thesaurus mapping, among other things)





- Equally possible to label a concept with a character string or an image ...
- Application in symbolic concept schemes for web accessibility (people who can't read or write) ...
- [Example data]

- SKOS-Core 1.0 public release (first major release)
 - Feb 2004
 - Current deployments including UKAT, GEMET, GCL (UK), APAIS
 - 4 months of testing ...
- Community effort ...
 - All work done via public mailing list <public-esw-thes@w3.org> & esw wiki
 - Core team from SWAD-E, plus currently around 30 contributing
- Next 12 months ...
 - Expect minor improvements and some new features, but no major structural change (except for modelling top concepts)

- What do I get out of publishing my thesaurus on the semantic web?
- Semantic web is about distributed data ...
 - RDF => linked thesauri
- SKOS-Core ... extensible standard for publishing thesaurus data ...
 - Problem with standards is that they are inflexible
 - RDF Schema => SKOS extensions
 - Capture all unique features AND get interoperability
 - (Have your cake and eat it!)



- Extending classes via `rdfs:subClassOf` ...
 - E.g. `gemet:Group subClassOf skos:Concept`
 - ... means everything that is a `gemet:Group` is also a `skos:Concept`
 - ... standard RDFS inference engine will infer this automatically
 - Extending properties via `rdfs:subPropertyOf` ...
 - E.g. `gemet:broaderGroup subPropertyOf skos:broader`
 - ... means every semantic relation involving the `gemet:broaderGroup` property is also a semantic relation involving the `skos:broader` property
 - ... standard RDFS inference engine will infer this automatically
- ⇒ Can publish thesaurus preserving all unique features and still plug it into standard SKOS software tools

Some RDF ... ('Turtle' format)

skos:narrower

rdfs:subPropertyOf

aos:containsSubstance

```
skos:narrower
  a  rdf:Property.

aos:containsSubstance
  a  rdf:Property;
  rdfs:subPropertyOf  skos:narrower.

ex:conceptX
  aos:containsSubstance  ex:conceptY.
```

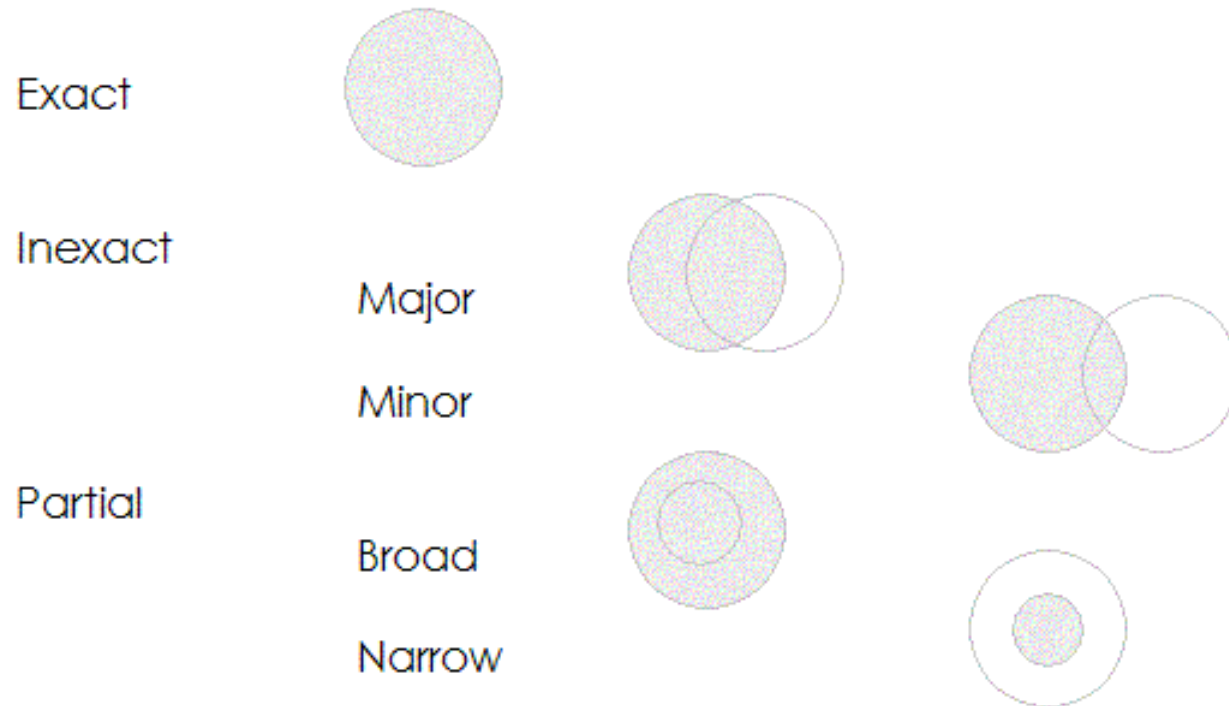
Standard RDFS inference

```
ex:conceptX
  skos:narrower  ex:conceptY.
```

(Although TK mentioned inconsistencies)

- SKOS-Mapping: an RDF vocabulary for describing inter-thesaurus mappings
 - Semantic mapping = mapping between concepts
 - See SWAD-Europe Thesaurus Activity Deliverable 8.4: Inter-Thesaurus Mapping <http://www.w3.org/2001/sw/Europe/reports/thes/8.4/>

Imagine two concepts, A and B, from different thesauri. Concept A retrieves document set $d(a)$ from a collection. Concept B retrieves document set $d(b)$ from a collection.

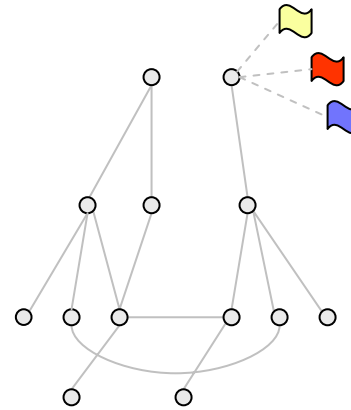


- SKOS-Mapping designed to support notion of **complete mapping** ...
 - Mapping to guarantee complete recall
 - Mapping to improve relevance ranking (precision)
- Other features ...
 - Ordered mappings for best match and ranking
 - Concept combinations (AND, OR, NOT) as mapping targets
- Status ...
 - Pre-release of SKOS-Mapping available for testing

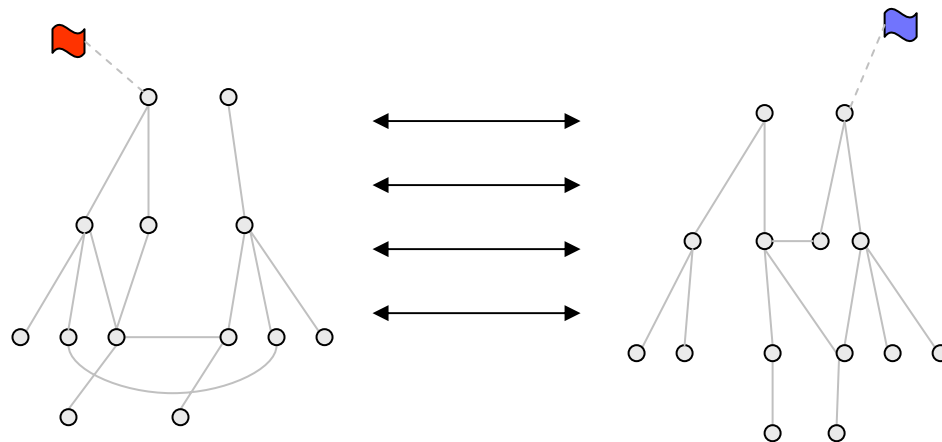


Analyse each language component

Multilingual Labelling



Interlingual Mapping



- Example of multilingual label data
 - See SWAD-E Thesaurus Activity Deliverable 8.3: RDF Encoding of Multilingual Thesauri <http://www.w3.org/2001/sw/Europe/reports/thes/8.3/> for full explanation of multilingual labelling vs. inter-lingual mapping (although beware code examples use a deprecated version of SKOS-Core!)

- Portuguese translation of the SKOS-Core 1.0 Guide ...

“People become librarians because they know too much. Their knowledge extends beyond mere categories. They cannot be confined to disciplines. Librarians are all-knowing and all-seeing. They bring order to chaos. They bring wisdom and culture to the masses. They preserve every aspect of human knowledge. Librarians rule. And they will kick the crap out of anyone who says otherwise.”

<http://www.librarianavengers.org/worship.html>



- SKOS API
 - Define functionality of a web service providing access to a thesaurus
 - Abstract, independent of implementation
- SOAP implementation
 - Javadoc ...
- Motivation ...
 - Standard interface to thesaurus service means maximal re-use of software components and maximal interoperability of data
- See
 - <http://www.w3.org/2001/sw/Europe/reports/thes/skosapi.html>
 - Links to Javadoc and WSDL

Web Client

SOAP

Apache AXIS

Sesame RDF Server

Relational DB

Alistair Miles, CCLRC

- Semantic Web is a web of data for machines
- SKOS-Core
 - Extensible standard for publishing thesauri on the semantic web
 - Standard/Non-standard, monolingual/multilingual
- SKOS-Mapping
 - Extensible standard for publishing inter-thesaurus mappings on the semantic web
 - Semantic mappings supporting information retrieval across collections
- SKOS API
 - Standard web service interface to thesaurus service (access only)
 - Standard => 'Plug and play' for distributed IR systems



- Community driven development ...
 - So far informal collaboration via public mailing list
 - Expand community of contributors to SKOS schemas and APIs
 - Improve support for wider collaboration (wikis, change log)
 - Improve support for users tracking updates
- N.B. these are YOUR standards!
 - All contributors are welcome!
- Semantic Web Best Practises and Deployment Working Group (SWBP-WG)
 - RDF Thesaurus Task Force
 - Take SKOS-Core publish as W3C note

- Semantic web is a web of data ...
 - Ontologies next to thesauri next to classification schemes next to topic maps next to web directories next to bookmark schemes next to ...
 - Exploring the common ground ...
 - Web directories (e.g. DMOZ: Open Directory, Yahoo)
 - Bookmark (favourites) schemes (e.g. Amaya shared bookmarks)
 - Topic hierarchies for weblogging (Semantic blogging, Semantic community portals)
 - N.B. these are not ‘thesauri’ as such but do have thesaurus like features ...
 - RDF vocabularies can be used in part, in combination ...
 - Maximise re-use and interoperability
- ⇒ The semantic web is an interesting place to be!



- Explore application of SKOS for classification schemes ...
 - (Often semantic ambiguity in concept hierarchy => CANNOT go straight to OWL class hierarchy)
 - Meaning of resource classifications ...
- Semantic disambiguation
- As-is publication



- SWAD-E Thesaurus Activity
 - <http://www.w3.org/2001/sw/Europe/reports/thes/>
- SKOS-Core 1.0 Guide
 - <http://www.w3.org/2001/sw/Europe/reports/thes/1.0/guide/>
- SKOS-Core 1.0 Guidelines for Migration
 - <http://www.w3.org/2001/sw/Europe/reports/thes/1.0/migrate/>
- Public Developers Mailing List
 - <mailto:public-esw-thes@w3.org>
 - <http://lists.w3.org/Archives/Public/public-esw-thes/>

- W3C Semantic Web Activity
 - <http://www.w3.org/2001/sw/>
- RDF Primer
 - <http://www.w3.org/TR/rdf-primer/>
- OWL Guide
 - <http://www.w3.org/TR/owl-guide/>