

**UNIVERSITÀ CA' FOSCARI DI VENEZIA**

**Facoltà di Scienze Matematiche, Fisiche e Naturali**

**Corso di Laurea Triennale in Informatica**

**Tesi di Laurea**

**Laureando: Matteo Ballarin**

**SKOS**

**Un sistema per l'organizzazione della conoscenza**

**Relatore: Chiar.mo Prof. Renzo Orsini**

**ANNO ACCADEMICO 2004-2005**



# Ringraziamenti

*Ringrazio tutti coloro che mi hanno fornito un aiuto alla realizzazione della presente tesi.*

*Ringrazio in particolare il professor Renzo Orsini che mi ha proposto l'argomento, per la gentilezza e la disponibilità dimostrata.*

*Ringrazio il Dr. Douglas Tudhope per aver sempre risposto alle mie email e per avermi fornito informazioni preziose.*

*Ringrazio il Dr. Claudio Gnoli per la disponibilità e l'interesse verso questa tesi.*

*Ringrazio inoltre le persone che mi sono state vicine e mi hanno incoraggiato a non mollare mai.*

# INDICE

<b>Capitolo 1: Verso il Web Semantico.....</b>	<b>1</b>
1.1 Introduzione.....	1
1.2 Il Web Semantico.....	2
1.3 Il reperimento delle informazioni.....	3
1.4 Gli strumenti per il Web Semantico: RDF e RDFS.....	3
1.4.1 RDF Data Model.....	4
1.4.2 RDF Schema o RDFS.....	6
1.5 I THESAURUS: Definizione, standard e diversi tipi di strutture esistenti.....	6
1.5.1 Definizione e standard.....	6
1.5.2 Stato dell'arte: lo standard britannico BS 8723.....	7
1.5.3 Le diverse strutture di thesaurus esistenti.....	8
<b>Capitolo 2: SKOS.....</b>	<b>12</b>
2.1 SKOS (Simple Knowledge Organisation System).....	12
2.2 Il Meta-Modello di SKOS.....	14
2.3 SKOS Core:Primi passi.....	15
2.3.1 Definire i concetti: La classe Concept.....	16
2.3.2 Etichettare i concetti: Le etichette lessicali.....	17
2.3.3 Etichette multilingua.....	18
2.3.4 Aggiungere altre informazioni utili: Proprietà per la documentazione...	19
2.3.4.1 Il Glossario.....	19
2.3.5 Le relazioni semantiche.....	21
2.3.5.1 Tassonomia.....	22
2.4 Verso un esempio completo: La classe Concept Scheme.....	25
2.5 La classe Collection e l'organizzazione a faccette.....	28
2.6 Utilizzo avanzato di SKOS Core:L'integrazione di SKOS con altri vocabolari e il Subject Indexing.....	32
2.6.1 L'integrazione di SKOS con Dublin Core e FOAF.....	32
2.6.2 Subject Indexing.....	34
2.7 Proprietà più avanzate: SKOS Extension Vocabulary Specification.....	36
2.8 Estendere e personalizzare le proprietà.....	38

<b>Capitolo 3: Altri modelli per la rappresentazione della conoscenza.....</b>	41
3.1 Zthes.....	41
3.2 Topic Map.....	50
<b>Capitolo 4: Applicazioni che utilizzano la tecnologia SKOS..</b>	55
4.1 GEMET.....	55
4.2 UKAT: UK Archival Thesaurus.....	59
4.3 Il portale SWED.....	63
4.4 Un'applicazione con Web Service: CSA-NBII Biocomplexity Thesaurus.....	66
<b>Capitolo 5: Conclusioni.....</b>	69
<b>Bibliografia.....</b>	71

# CAPITOLO 1

## Verso il Web Semantico

### 1.1 Introduzione

“...quando Pete rispose al telefono il volume dello stereo si abbassò: era Lucy, la sorella, in linea dallo studio del medico; la madre aveva bisogno di cure specialistiche e Lucy aveva appena impostato il suo agente in modo che le fissasse appuntamento con un fisioterapista vicino a lei e non troppo lontano dal fratello, Pete istruì a sua volta l’agente in modo che collaborasse con quello di Lucy; in pochi minuti, seguendo le esigenze di entrambi, l’agente fissò l’appuntamento.[BHL01]

In questo esempio l’autore desidera esprimere uno dei possibili scenari futuri: agenti software che collaborano tra loro per decidere quale sia la soluzione migliore, tra le varie possibili, rispettando le preferenze indicate dall’utente.

Affinché questo possa essere realizzabile deve nascere il Semantic Web [B98S, WSE]] ovvero un’estensione del web attuale, che dia struttura e contenuto significativo alle pagine, permettendone la comprensione sia alle persone che ai pc.

Grazie al Web semantico le ricerche diventeranno più affidabili, precise e quindi più veloci; i motori di ricerca saranno capaci di estrarre la semantica dai documenti e quindi forniranno risultati che corrispondono esattamente al termine e al significato cercato; attraverso l’uso di opportune ontologie si potranno creare basi di conoscenza comune e condivisa che permetteranno di scambiare informazioni tra soggetti diversi, in lingue diverse, con culture diverse. Agenti diversi riusciranno a comprendersi, dando un senso alle informazioni; questo grazie all’uso di linguaggi d’inferenza che permetteranno il confronto e la “traduzione” tra termini con diversi ma con lo stesso significato.

Il risultato ottenuto da un agente potrà essere verificato dall’utente che potrà chiedere di motivare quel risultato: sarà così possibile verificare il “ragionamento” fatto dall’agente ed eventualmente “addestrarlo” a raffinare o estendere la ricerca sulla base di specifiche esigenze.

Per ciascuno di tali aspetti sono necessarie opportune tecnologie, molte delle quali sono già disponibili e saranno oggetto di questa tesi; per le altre sicuramente ci saranno ulteriori sviluppi nel futuro.

Avere un Web semantico significherà quindi non solo avere strumenti più evoluti per la ricerca e la gestione della conoscenza, ma sarà un ulteriore passo verso una rete veramente universale, superando ogni ostacolo e limite in maniera da portare il Web ad essere una fonte di informazione accessibile a tutti.

## 1.2 Il Web Semantico

Negli ultimi anni abbiamo assistito ad un aumento esponenziale a livello mondiale sia nella crescita, sia nell'utilizzo del Web.

I browser web sono migliorati tantissimo e si arricchiscono sempre più di nuove funzionalità e il W3C (World Wide Web Consortium) [W3C], l'organismo che si occupa di definire gli standard per Web, sta compiendo un assiduo lavoro per permettere e garantire questo sviluppo.

I motori di ricerca migliorano costantemente le proprie funzionalità, offrono nuovi servizi (ad es. il Desktop Search per la ricerca efficiente dei documenti all'interno del proprio pc), i servizi online proliferano: aste, portali per l'intrattenimento, servizi di home banking, e-commerce e molto altro ancora.

Siamo passati da una prima fase in cui le pagine dei siti erano statiche ad una seconda in cui il contenuto dei documenti è diventato dinamico e questo è stato un grosso ed importante cambiamento che ha permesso l'avanzare di nuovi servizi e contenuti. La prossima evoluzione sarà probabilmente verso un Web Semantico, nel quale l'informazione avrà un significato ben preciso e definito con un accesso basato sui significati.

Questo porterà l'utilizzatore, nel tempo, a formulare domande sempre più ampie (mentre oggi occorre essere capaci a formulare domande molto più focalizzate) del tipo: portami a Londra questo fine settimana anziché digitare volo+Londra.

Allo stato attuale il sistema di ricerca basato su parole chiave funziona fino ad un certo punto: molti risultati di una query effettuata sono irrilevanti, di scarso interesse o addirittura assenti perché magari viene impiegato un sinonimo che non compare nel testo oppure perché la parola cercata ha un significato molto vasto (pensiamo ad esempio a parole usate in diversi contesti come chip, package), oppure non sappiamo quali parole digitare o ancora il risultato ottenuto è totalmente inaspettato (vedi caso di Google di qualche anno fa).

HTML non è in grado di risolvere tali problematiche. Esso è orientato piuttosto verso la formattazione e la presentazione del testo all'interno del documento, i suoi tag non portano alcuna semantica e da qui scaturisce la sua inadeguatezza ad essere "machine processable" cioè letto da utenti non in carne ed ossa.

Sono state proposte allora nuove soluzioni come XML e RDF [S02, S03, B01] che permettono di strutturare il documento e di inserire informazione semantica: questi integrati con altri strumenti quali ontologie (collezioni di termini e relazioni tra questi) ed altri come strumenti per l'organizzazione della conoscenza come thesauri potranno estendere il Web a diventare un Semantic Web.

## 1.3 Il reperimento delle informazioni

Come abbiamo detto, sebbene i motori di ricerca siano in costante miglioramento, la ricerca di molti contenuti presenta numerose limitazioni:

- Molte risorse non contengono testo. Pensiamo ad esempio a immagini, audio, video.
- C'è una bassa precisione quando una parola ha significati ambigui (es. *chocolate chips vs silicon chips*).
- La ricerca è incompleta se non tiene conto dei sinonimi e dei termini "related".
- Non esiste nessun tipo di aiuto fornito all'utente per espandere, tradurre, relazionare il termine cercato dall'utente.  
Ad esempio animali domestici → mammiferi → gatti → felini → *Panthera tigris tigris* (tigre del bengala).
- Difficoltà ad esprimere query tipo il documento X con oggetto Y scritto il giorno Z.

Questa difficoltà deriva dal fatto che la maggior parte delle informazioni sono machine-readable ma non machine-understandable. Affinché il significato sia accessibile anche alle macchine è necessario che ai dati sia associata una struttura che permetta di dedurre l'informazione che essi esprimono: i metadati.

I metadati sono generalmente definiti come "dati riguardanti i dati". Essi rappresentano il fondamento per il reperimento delle informazioni: sono delle informazioni comprensibili dalla macchina relative ad una qualsiasi risorsa (non solo Web). Pensiamo ad esempio quando in biblioteca cerchiamo un certo libro: il sistema software permette di ricercare i libri per autore, titolo, soggetto, descrizione, ecc. Tutti questi sono metadati.

La cosa importante è costituita dal fatto che, essendo comprensibili dalla macchina, possono essere utilizzati dai software agent per fare un uso appropriato e più elegante delle risorse.

Per farne un uso efficace è necessario che siano stabilite delle convenzioni per la struttura, la semantica e la sintassi. I metadati consistono in asserzioni sui dati, le quali vengono quindi rappresentate sotto forma di un nome di asserzione e un insieme di parametri. Gruppi di asserzioni relative alla stessa risorsa prendono spesso la forma di una lista di coppie (attributo-valore).

## 1.4 Gli strumenti per il Web Semantico: RDF e RDFS

RDF (Resource Description Framework) [C03, RDF, B01] è lo strumento base per la codifica, lo scambio e il riutilizzo di metadati e consente l'interoperabilità tra applicazioni che si scambiano sul Web informazioni machine-understandable.

Esso è costituito da due componenti:

- RDF Data Model
- RDF Schema

### 1.4.1 RDF Data Model

RDF Data Model fornisce un modello per descrivere le risorse e permette di definire le relazioni tra loro in termini di *nome: valore*.

RDF definisce una risorsa come un qualsiasi oggetto che è identificabile univocamente mediante un Uniform Resource Identifier (URI).

Il modello RDF è molto semplice, ed è basato su tre tipi di oggetti:

- *Risorse*: qualunque cosa descritta da un'espressione RDF ed identificata univocamente da un URI è detta risorsa (*resource*). Una risorsa può essere una pagina Web, una sua parte ma anche un oggetto non direttamente accessibile via Web (per es. un libro, un dipinto, etc.).
- *Proprietà*: una proprietà (*property*) è una caratteristica, un oggetto specifico che definisce una relazione utilizzata per descrivere una risorsa. Essa ha un significato specifico, definisce i valori ammissibili, i tipi di risorse che può descrivere e le sue relazioni con altre proprietà.  
Le proprietà associate alle risorse sono identificate da un nome e assumono dei valori.
- *Affermazioni o statement*: una risorsa, con una proprietà distinta da un nome e un valore della proprietà per la specifica risorsa, costituisce un RDF *statement*. Uno statement è quindi una tripla composta da un *soggetto* (risorsa), un *predicato* (proprietà) e un *oggetto* (valore).

È possibile quindi esprimere delle affermazioni come: R.L.Stevenson è autore de "L'isola del Tesoro".

Utilizzando i formalismi dell'Intelligenza Artificiale, le relazioni tra risorse, proprietà e valori sono rappresentati attraverso dei grafi orientati: le risorse (e quindi gli URI) rappresentano i nodi, le proprietà gli archi orientati e i valori dei rettangoli.

Abbiamo quindi la tripla:

*Resource* **ISBN 881158735-2**

*Property* **Autore**

*Value* **R.L.Stevenson**

In questo caso come URI viene utilizzato ISBN (International Standard Book Number), un sistema di codifica che identifica univocamente un certo testo.

Lo statement verrebbe quindi rappresentato come:

La risorsa ISBN 881158735-2 ha autore R.L.Stevenson.

Graficamente:

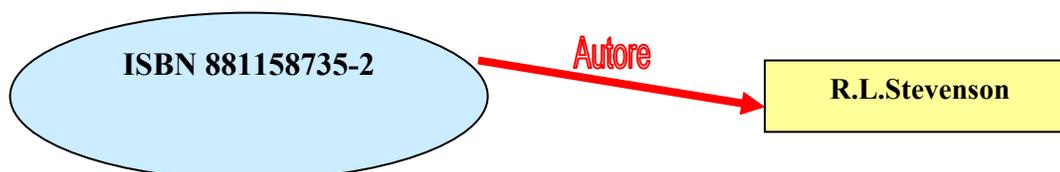


Fig. 1 Rappresentazione grafica dello statement “Stevenson è l’autore dell’Isola del Tesoro”.

RDF consente quindi alle singole comunità di definire la semantica. Tuttavia, non è possibile affidare la semantica semplicemente al nome, che potrebbe avere significati più o meno ampi secondo gli interessi specifici delle singole comunità.

RDF identifica univocamente le proprietà mediante il meccanismo dei namespace. I namespace XML forniscono un metodo per identificare in maniera non ambigua la semantica e le convenzioni che regolano l’utilizzo delle proprietà identificando l’authority che gestisce il vocabolario.

Uno degli esempi più noti è la Dublin Core Initiative [DCHP], un’organizzazione che si occupa di promuovere e diffondere l’utilizzo di metadati standard. Essa definisce un insieme di proprietà che possono essere utilizzate per descrivere determinate risorse. Essa definisce ad esempio rispettivamente le proprietà “*Title, Description e Creator*” nel seguente modo [DDCCE]:

Name: *Title*

Identifier: Title

Description: A name given to the resource.

Comment: Typically, Title will be a name by which the resource is formally known.

Name: *Description*

Identifier: Description

Definition: An account of the content of the resource.

Comment: Examples of Description include, but is not limited to: an abstract, table of contents, reference to a graphical representation of content or a free-text account of the content.

Name: *Creator*

Identifier: Creator

Definition: An entity primarily responsible for making the content of the resource.

Comment: Examples of Creator include a person, an organization, or a service. Typically, the name of a Creator should be used to indicate the entity.

Si può utilizzare quindi un namespace XML per identificare in maniera non ambigua lo schema per il vocabolario Dublin Core, puntando alla risorsa che ne definisce la semantica. Vedremo in seguito un’applicazione di questo fatto.

### 1.4.2 RDF Schema o RDFS

Il data model RDF permette di definire un modello semplice per descrivere le relazioni tra le risorse, in termini di proprietà identificate da un nome e relativi valori.

Tuttavia, RDF data model non fornisce nessun meccanismo per dichiarare queste proprietà, né per definire le relazioni tra queste proprietà ed altre risorse.

RDF Schema (RDFS) permette di definire dei vocabolari, quindi l'insieme delle proprietà semantiche individuata da una particolare comunità.

RDF Schema permette definire significato, caratteristiche e relazioni di un insieme di proprietà, compresi eventuali vincoli sul dominio e sui valori delle singole proprietà.

Inoltre introducendo il concetto di classe e sottoclasse, consente di definire gerarchie di classi, con il conseguente vantaggio che agenti software intelligenti possono utilizzare queste relazioni per svolgere i loro compiti.

RDF definisce quindi la grammatica, le strutture fondamentali per la struttura della frase (soggetto, predicato, oggetto): Matteo Ballarin abita in Veneto.

RDF Schema permette di definire le parole ed i verbi (classi e properties) che desideriamo usare (es. Matteo Ballarin type Person).

RDF e RDFS costituiscono quindi i mattoni su cui fondare il Web Semantico, ma è necessario un ulteriore strato che permetta di descrivere e dichiarare formalmente la semantica delle classi e delle proprietà usate nei documenti.

È necessario che esistano dei "vocabolari" di sapere condivisi in cui siano presenti i termini e i concetti di quel determinato dominio di interesse. Questi vocabolari, detti anche concept schema o vocabolari controllati, non sono altro che delle liste di termini tra i quali sono definite delle relazioni semantiche di:

- Gerarchia
- Preferenza
- Equivalenza

## 1.5 I THESAURUS: Definizione, standard e diversi tipi di strutture esistenti

### 1.5.1 Definizione e standard

Un thesaurus è un insieme di termini, e di relazioni tra essi, che costituiscono il lessico specialistico da usare per descrivere il contenuto dei documenti pubblicati in un certo ambito disciplinare [S05, SE05, PS05].

Il termine thesaurus o tesoro, ha una lunga storia: deriva dal greco e passa attraverso il latino thesaurus: in entrambe le lingue il significato è grosso modo quello di deposito nascosto di beni preziosi, forziere, scrigno. Successivamente il termine è passato nella lingua italiana ad indicare un particolare tipo di dizionario: una classificazione in varie categorie e subcategorie (relazioni astratte, spazio, tempo, proprietà fisiche) di parole ed

espressioni in una determinata lingua, con lo scopo di favorire la ricerca del termine più pertinente ed adatto da esprimere in ogni circostanza.

Attualmente sono tre gli standard per i tesauri che godono di maggior diffusione a livello internazionale [F05]:

- ISO 2788-1986 Guide to establishment and development of monolingual thesauri
- ISO 5964-1985 Guide to establishment and development of multilingual thesauri
- ANSI Z39.19-1974 Guidelines for thesaurus structures and use

I thesauri sono largamente impiegati nel campo dell'information retrieval. Sono usati in diversi modi: come aiuto per la ricerca, fornendo termini per un determinato dominio, come strumenti per la traduzione in più lingue e come strumenti per la classificazione automatica di documenti. Esistono diversi thesauri, molti dei quali sono stati sviluppati nel corso di molti anni da persone con un'approfondita conoscenza della materia. Questi strumenti rappresentano e forniscono una risorsa molto importante per il web semantico.

Un nuovo standard britannico, BS 8723, Structured vocabulary for information retrieval - Guide, rappresenterà la normativa per la costruzione e la gestione di thesauri di nuova generazione.

### **1.5.2 Stato dell'arte: lo standard britannico BS 8723**

Il nuovo standard BS 8723 Structured vocabulary for information retrieval - Guide è in fase di elaborazione. Ecco, in sintesi, la struttura dello standard e il punto sullo stato dell'arte dei lavori:

- *Parte 1: Definitions, symbols and abbreviations*

Fornisce definizioni e concetti chiave comuni a tutti i sistemi di information retrieval che fanno uso di vocabolari strutturati.

- *Parte 2: Thesauri*

Si occupa dei thesauri, riprendendo tutti i concetti trattati negli standard precedenti oltre a fornire linee guida sulle funzioni elettroniche e sui software di gestione.

E' riservata attenzione particolare all'analisi a faccette, appena menzionata negli standard correnti nonostante l'importante ruolo che essa riveste nella costruzione di thesauri.

- *Parte 3: Vocabularies other than thesauri*

Tratta altre tipologie di vocabolari strutturati (schemi di classificazione, thesauri di ricerca, tassonomie e ontologie).

- *Parte 4: Interoperability between vocabularies*

È dedicata all'interoperabilità tra vocabolari; vengono prese in esame possibili situazioni che si possono verificare nello stabilire equivalenze tra diversi vocabolari strutturati.

- *Parte 5: Interoperability between vocabularies and other components of information storage and retrieval systems*

Definisce i protocolli e i formati necessari allo scambio di dati.

Le Parti 1 e 2 sono state completate nel 2004 e in questo momento si stanno raccogliendo le osservazioni della comunità internazionale per apportare eventuali modifiche al testo. Le altre parti sono in lavorazione, ma a breve dovrebbe uscire la versione definitiva.

La sfida per lo standard BS 8723 consiste nel fornire, alla comunità eterogenea di utenti ai quali esso si rivolge, definizioni chiare e comprensibili delle varie tipologie di strumenti di indicizzazione attualmente in uso (tassonomie, ontologie, schemi di classificazione, etc.), evidenziando analogie e differenze e individuando per ciascuno di essi delle ipotesi di linee guida.

### 1.5.3 Le diverse strutture di thesaurus esistenti

La struttura di un thesauro può essere strutturata secondo due diverse alternative:

- Struttura basata sui termini
- Struttura basata sui concetti

Nel primo caso il thesaurus non è altro che una collezione di termini: i termini sono le uniche entità prese in considerazione. Ovviamente essi possono essere in relazione con altri termini, ciò avviene utilizzando relazioni come “broader” (BT, termine più generale), “narrower” (NT, termine più specifico), “related” (RT, in relazione), “use for” (UF, usato per), “use” (US, uso).

Un esempio di thesaurus con struttura basata sui termini è il seguente:

<b>AIDS</b>	
<b>UF</b>	Sindrome da immunodeficienza acquisita
<b>BT</b>	Malattia del sistema immunitario
<b>RT</b>	Persone sieropositive
<b>ACQUA</b>	
<b>UF</b>	H <sub>2</sub> O
<b>RT</b>	Inquinamento delle acque
<b>INQUINAMENTO DELLE ACQUE</b>	
<b>BT</b>	Inquinamento
<b>RT</b>	Acqua
<b>RAFFREDDORE</b>	
<b>US</b>	Malattia delle prime vie respiratorie

Fig 2: Esempio di term-based thesaurus

Una struttura basata sui concetti invece consiste di due diversi tipi di entità: concetti e termini. Un concetto viene definito come un unità base, un qualcosa che è insito nella mente di una persona. Relazioni come “broader”, “related”, “narrower”, avvengono tra concetto e concetto e contengono informazioni riguardanti la struttura che descrivono. In altre parole trasmettono quello che viene identificato come significato. I termini invece non sono altro che “etichette” per identificare un concetto. Le relazioni tra termini contengono solo informazione di tipo lessicale, per esempio un’abbreviazione per un altro termine. Relazioni tra termini e concetti trasmettono un’informazione che riguarda come un termine implica un concetto, cioè il significato vero e proprio.

## Information

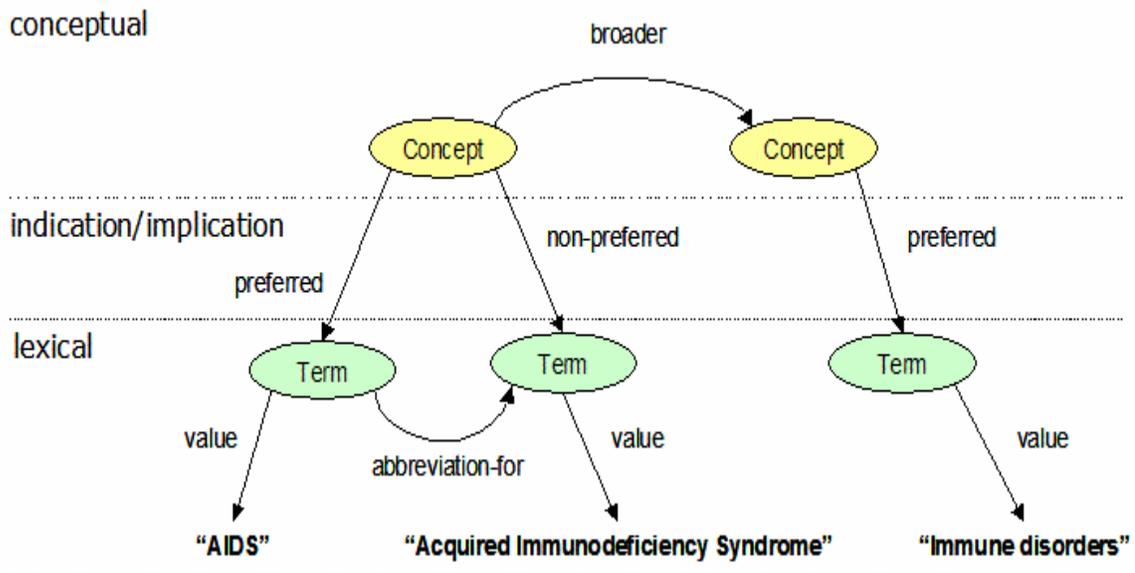


Fig 3 Esempio di struttura concept-based

Quest’ultimo tipo di modellazione offre una più precisa descrizione dell’informazione contenuta all’interno del thesauro. Infatti, sono distinti esplicitamente due tipi di informazione: concettuale e lessicale; vi è così una maggiore chiarezza e minor rischio di creare confusione.

Il primo modello, molto più tradizionale e a noi familiare invece offre una maggiore compattezza dello schema e del formato dei dati.

In alcuni thesauri i termini ed i concetti sono organizzati in categorie. Ad esempio:

(dispositivi)	
A001	. dispositivi
A002	. . . dispositivi fissi
A003	. . . dispositivi portatili
A004	. . . dispositivi elettrici
A005	. . . dispositivi manuali
(persone)	
A100	. persone
A101	. . . neonati
A102	. . . bamini
A103	. . . adulti
(proprietà)	
A200	. proprietà
A201	. . proprietà chimiche
A202	. . proprietà fisiche
A203	. . . proprietà ottiche
A204	. . . . colore
A205	. . . . contrasto
A206	. . . . luminanza
A207	. . . . luminosità
A208	. . . pressione
A209	. . . temperatura

Fig. 4 Esempio di termini raggruppati per categoria

(dispositivi)	
A001	. dispositivi
	. . <dispositivi per portabilità>
A002	. . . dispositivi fissi
A003	. . . dispositivi portatili
	. . <dispositivi per alimentazione>
A004	. . . dispositivi alimentati elettricamente
A005	. . . dispositivi manuali
	. . <dispositivi per uso>
	. . . dispositivi per l'immersione
	. . . dispositivi fotografici
	. . . dispositivi per misure fisiche
(persone)	
A100	. persone
	. . <persone per età>
A101	. . . neonati
A102	. . . bambini
A103	. . . adulti
	. . <persone per professione>
	. . . sommozzatori
	. . . modelle
	. . . fotografi
	. . . fisici

Fig. 5 Altro esempio con un diverso tipo di raggruppamento

Questo tipo di schema di classificazione introduce maggiore informazione riguardo la struttura del dominio che vogliamo definire, facilitandoci il compito per cui è stata pensata: la successiva consultazione ed il reperimento delle informazioni.

Questi tipi di strumenti sono quindi fondamentali per il Web Semantico e in generale anche per il Web di oggi. A questo proposito possiamo citare la digital library community, un gruppo di esperti che, interagendo insieme nell'ambito di specifiche comunità di utenti, sviluppano thesauri per specifici argomenti.

Come esempi possiamo citare AAT (Art & Architecture Thesaurus) o DMOZ (Directory MOZilla): quest'ultimo è un archivio di siti selezionato e mantenuto da migliaia di esperti che verificano e censiscono risorse sul Web in base al settore di esperienza. DMOZ è organizzata gerarchicamente in categorie tematiche, partendo da argomenti generali e proseguendo via via verso temi sempre più specifici e dettagliati. Essa stessa è una comunità e offre a tutti l'opportunità di contribuire al progetto.

È infatti possibile inserire gratuitamente il proprio sito in una data categoria dell'archivio seguendo le istruzioni e i suggerimenti forniti online.

Inoltre chiunque sia interessato a collaborare, partecipando attivamente alla costruzione della directory, può scegliere un argomento nel quale si sente ferrato e fare richiesta dalla pagina della categoria corrispondente.

Bisogna inoltre considerare che gli standard elencati prima, risalenti a venti anni fa, sono stati sviluppati per un mondo non informatizzato quindi non tengono conto dell'impatto dei mezzi digitali.

A questo proposito il British Standard Institute sta rivedendo gli standard e alcuni membri dell'organizzazione contribuiscono al progetto SKOS per assicurare che le due iniziative possano convergere.

C'è quindi bisogno di uno strumento per esprimere insiemi di concetti e vocabolari che sia sufficientemente potente da supportare la ricerca semantica attraverso metadati, ma allo stesso tempo facile da usare e a basso costo. Questo strumento basato su RDF e RDFS è SKOS.

Nel capitolo successivo verrà introdotto il framework SKOS [M05, MMW04] (Simple Knowledge Organisation System) e verranno descritte le proprietà ed i costrutti più significativi modellandoli su degli esempi concreti di concept-schema.

Nel terzo capitolo sono descritti due sistemi "alternativi" per l'organizzazione della conoscenza: Zhtes [ZTHES] e Topic Map [TAO, TMF].

Nel quarto capitolo vengono mostrate alcune applicazioni Web che utilizzano questa tecnologia: si tratta di portali di informazione scientifica, nei quali vengono realizzati ed utilizzati uno o più thesaurus per organizzare dati di diversa natura. Questo strumento in molti casi viene reso disponibile per il download in modo che chiunque possa utilizzarlo all'interno di una propria applicazione web oppure vi è la possibilità di accedervi come web service.

Nel quinto capitolo infine si traggono le conclusioni di questo lavoro, evidenziando come questa tecnologia possa essere d'aiuto e come possa essere utilizzata in pratica.

# CAPITOLO 2

## SKOS

*However, the construction of useful thesauruses is an art rather than a science and requires extensive knowledge of the particular subject area under consideration and of the record collections to be processed* (G. Salton)



### 2.1 SKOS (Simple Knowledge Organisation System)

SKOS (Simple Knowledge Organisation System) è un'area di lavoro che sviluppa specifiche e standard per supportare l'uso dei sistemi d'organizzazione della conoscenza (KOS) nel quadro del web semantico [SKOS 05].

Sviluppato inizialmente dal luglio del 2003 (come progetto open-source attraverso pubbliche mailing-list e attraverso wiki) all'interno del progetto SWAD-E (Semantic Web Advanced Development for Europe) intendeva definire un modello per thesaurus compatibile con gli standard ISO più importanti (ISO 2788, ISO 5964). Dal Settembre 2004 la responsabilità per i successivi sviluppi è passata al W3C e il 10 Maggio 2005 è stato pubblicato il primo Working Draft. Attualmente grazie al contributo di numerosi sviluppatori e gruppi di discussione il 2 Novembre è stata pubblicata la seconda revisione del documento [SKOSCG05].

SKOS Core Vocabulary [SKOSCVS05] è stato progettato con l'idea di essere facilmente estensibile: esso vuole fare da connettore tra i più tradizionali sistemi di organizzazione della conoscenza usati nelle biblioteche, musei, archivi come i thesauri e i sistemi di classificazione e le nuove strutture pensate per il Web come le open-directory, i blog topic exchange.

SKOS Core Vocabulary consiste in una serie di classi RDFS e proprietà di RDF che sono usate per rappresentare il contenuto e la struttura di base dei cosiddetti concept-schema cioè thesaurus, schemi di classificazione, elenchi di titoli d'argomento, tassonomie, terminologie, glossari e altri tipi di vocabolario. Poiché sono in formato RDF, questi possono essere manipolati, trasformati, archiviati, e vi si possono fare ricerche usando gli strumenti RDF. Pertanto SKOS può essere usato come formato d'interscambio attraverso le librerie digitali e per il recupero dati. Inoltre è possibile l'interazione con altri strumenti e rappresentazioni usati nel web semantico.

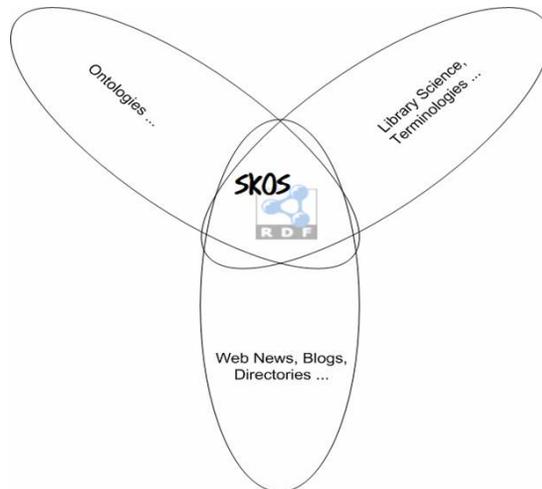


Fig 6: La collocazione di SKOS ed il suo ruolo nell'ambito del semantic web.

La possibilità di essere esteso con altri vocabolari RDF come DCMI Metadata Terms e FOAF [FOAF] (Friend of a Friend) ne aumenta enormemente le potenzialità in quanto, in quei casi in cui gli strumenti di SKOS Core non dovesse soddisfare i requisiti o essere sufficienti, è possibile definire nuove classi. Per questo motivo SKOS Core offre una naturale flessibilità che fornisce una base per l'interoperabilità anche nelle situazioni in cui i concept-schema sono stati sviluppati utilizzando diverse specifiche.

Un'altra caratteristica che SKOS Core eredita da RDF è il meccanismo fornito dalle sub-classes e dalle sub-property in RDFS: per supportare questo tipo di estensione le proprietà di SKOS Core Vocabulary sono raggruppate in famiglie:

- Properties per le etichette lessicali
- Properties per le etichette per la documentazione
- Properties per le relazioni semantiche
- Properties per le etichette simboliche

Le properties all'interno di queste famiglie sono organizzate in maniera gerarchica in modo che sia possibile estendere quella più appropriata per le proprie necessità. Ecco un esempio che riguarda le properties per le relazioni semantiche: `semanticRelation` è la radice e i livelli successivi costituiscono delle "specializzazioni" del nodo padre.

```
skos:semanticRelation
|
+-- skos:broader
|
+-- skos:narrower
|
+-- skos:related
```

SKOS Core può essere usato quindi per:

- Trasferire i sistemi KOS già esistenti nel mondo del Web
- Costruire da zero semplici concept-schemes per il Web

Esso fornisce un framework per creare dei “link” tra i concetti e le parole e le frasi che usiamo nel comune linguaggio. Una volta che questo tipo di informazione è catturata può essere usata in moltissime applicazioni come la classificazione automatica dei documenti su Web o traduzioni multilingua di interi glossari.

## 2.2 Il Meta-Modello di SKOS

SKOS Core permette di definire concetti e concept-schemes. Un concetto è un oggetto atomico, un unità di pensiero che può in qualche modo essere descritto o definito, mentre un concept-schema è una collezione, un insieme di concetti. Come esempio di concetto possiamo prendere qualsiasi “parola del vocabolario”, mentre come esempio di concept-schema possiamo pensare al vocabolario stesso inteso come lista di termini.

Ogni concetto è identificato da un termine o se vogliamo usare la terminologia introdotta da SKOS, da una label, un’etichetta che ci permette così di “accedere” al concetto.

Oltre alla label principale o meglio “preferred label”, lo stesso concetto può avere altre numerose label le cosiddette “alternative labels” con cui ci può riferire: stiamo parlando dei cosiddetti sinonimi.

Le relazioni tra concetti appartenenti allo stesso concept-schema sono definite come relazioni semantiche, mentre relazioni tra concetti appartenenti a concept-schema diversi sono definite come “semantic mapping”[SKOSTUT].

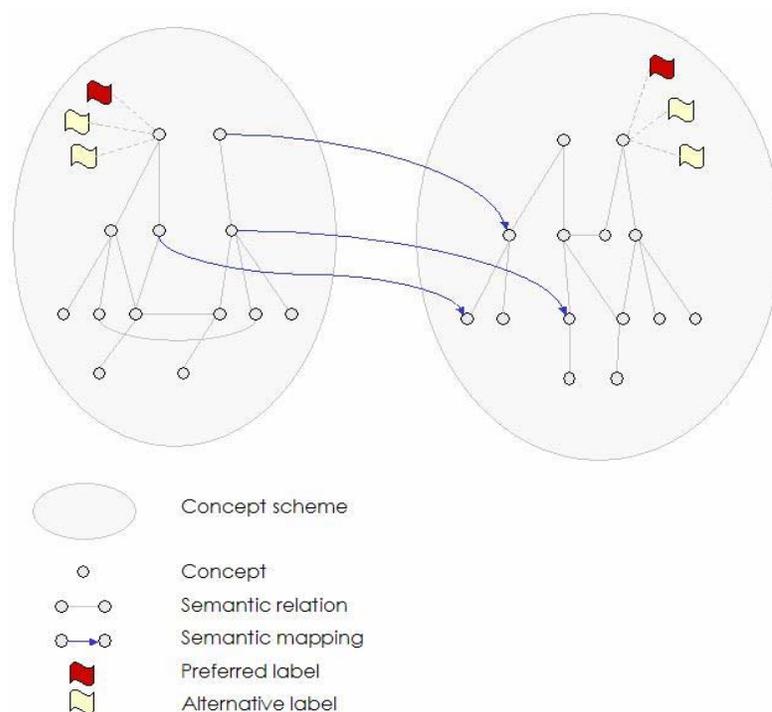


Fig.7 Il meta modello di SKOS

## 2.3 SKOS Core: Primi passi

Abbiamo visto prima come esprimere delle affermazioni usando i grafi con RDF. Vediamo un altro esempio [SKOSTUT]:

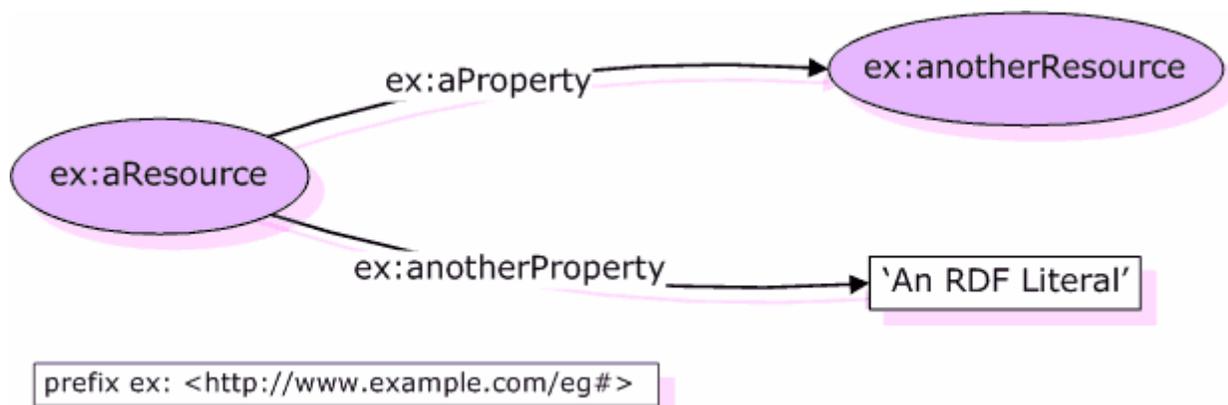


Fig. 8 Esempio di grafo RDF

Questa riportata qui sotto è la serializzazione del grafo:

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:ex="http://www.example.com/eg#">

  <rdf:Description rdf:about="http://www.example.com/eg#aResource">
    <ex:aProperty
rdf:resource="http://www.example.com/eg#anotherResource"/>
    <ex:anotherProperty>An RDF Literal</ex:anotherProperty>
  </rdf:Description>

</rdf:RDF>
```

Notiamo come in questo esempio sono presenti due namespace (URI) referenziati dai prefissi 'rdf' e 'ex' che forniscono un metodo per identificare in maniera non ambigua la semantica e le convenzioni che regolano l'utilizzo delle proprietà identificando l'authority che gestisce il vocabolario.

### 2.3.1 Definire i concetti: La classe Concept

SKOS Vocabulary considera “Concept” l’unità fondamentale di un thesaurus o di qualunque altro concept-schema. La classe Concept permette di modellare una determinata risorsa in maniera da esprimere questa come un concetto. Un concetto può essere immaginato come un qualcosa che può essere definito o descritto.

Ogni concetto può avere una sola descrizione o meglio un solo “preferred term” mentre può avere illimitate descrizioni alternative o “alternative tags”.

Il modo più semplice per definire un concetto in RDF è assegnargli un URI e utilizzare la forma `skos:Concept`.

Eccone un esempio [SKOSTUT]: qui la tripla `love rdf:type Concept` indica che `love` è un’istanza di `Concept` cioè la risorsa stessa; in questo caso `love` è un concetto.

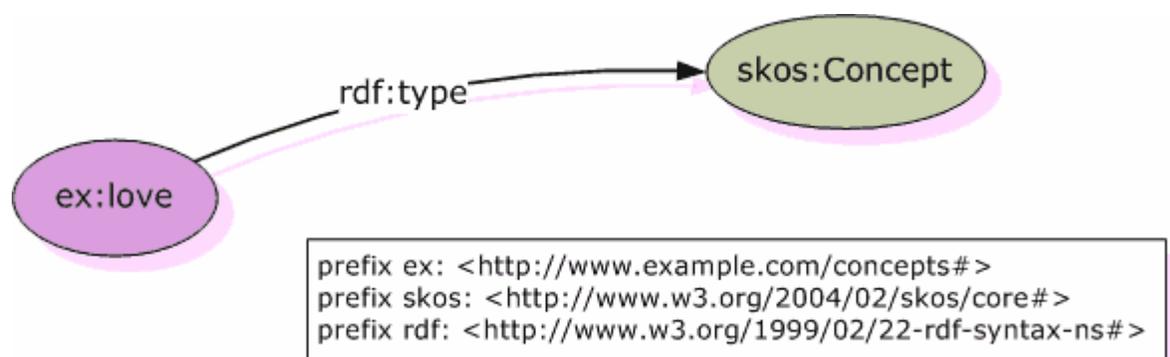


Fig. 9 La classe `love` come istanza di `Concept`

Ecco qui di seguito, la sintassi RDF. Notiamo l’URI che identifica il namespace di SKOS e l’URI che identifica il concetto di “love”. Questo ci da una totale indipendenza da ogni thesauro. L’URI di SKOS Core Vocabulary è `http://www.w3.org/2004/02/skos/core` e i rispettivi URI per le classi e le proprietà aggiungendo un identificatore al termine della stringa ad esempio

`http://www.w3.org/2004/02/skos/core#ConceptName`

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#">

  <skos:Concept rdf:about="http://www.example.com/concepts#love"/>

</rdf:RDF>
```

### 2.3.2 Etichettare i concetti: *Le etichette lessicali*

Le etichette servono a denotare le risorse utilizzando il linguaggio comune. Attraverso le etichette `skos:prefLabel`, `skos:altLabel` vengono indicati rispettivamente il termine preferito o più comunemente usato per un certo concetto e altri sinonimi.

Ecco un esempio [SKOSTUT]:

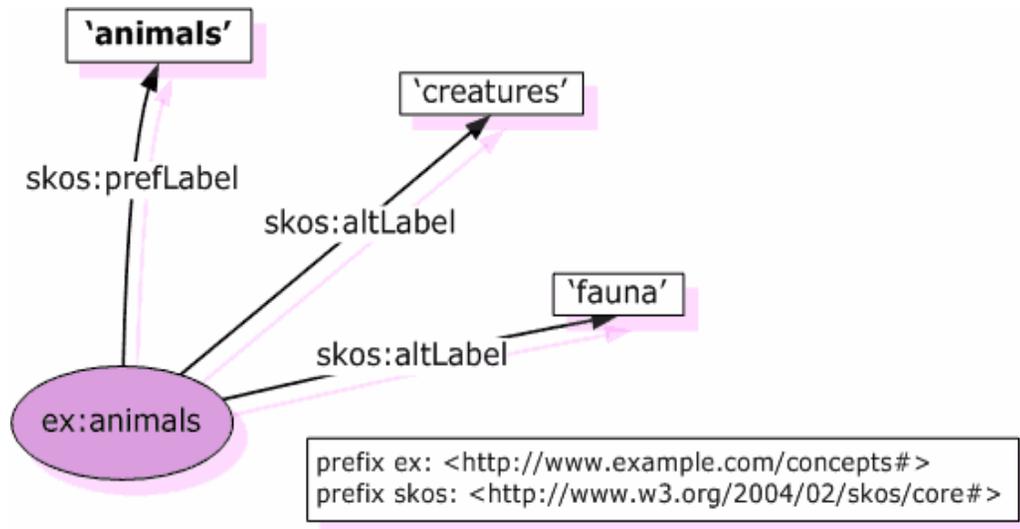


Fig. 10 Le etichette lessicali

#### La serializzazione in RDF

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#">

  <skos:Concept rdf:about="http://www.example.com/concepts#animals">
    <skos:prefLabel>animals</skos:prefLabel>
    <skos:altLabel>creatures</skos:altLabel>
    <skos:altLabel>fauna</skos:altLabel>
  </skos:Concept>

</rdf:RDF>
```

Esistono poi altre tre etichette:

- `skos:hiddenLabel`: un'etichetta "nascosta" che normalmente non è visibile ma che permette l'accesso al contenuto ad applicazioni che fanno ricerche basate sul testo. Normalmente questa etichetta è utilizzata per includere varianti sintatticamente incorrette dello stesso termine: ad es. elettricità, elettricità ecc.
- `skos:prefSymbol`: un'etichetta che permette di includere un simbolo o un'immagine di una certa risorsa
- `skos:altSymbol`: un simbolo o un'immagine alternativa per una risorsa

### 2.3.3 Etichette multilingua

SKOS Core fornisce il modo di rappresentare i concetti in altre lingue attraverso delle etichette.

Questa è una possibilità molto importante, in quanto permette di definire lo stesso termine in più lingue e se, immaginiamo un futuro in cui il web semantico è una realtà, la possibilità di ricercare lo stesso termine in più lingue aumenta moltissimo le potenzialità [SKOSTUT].

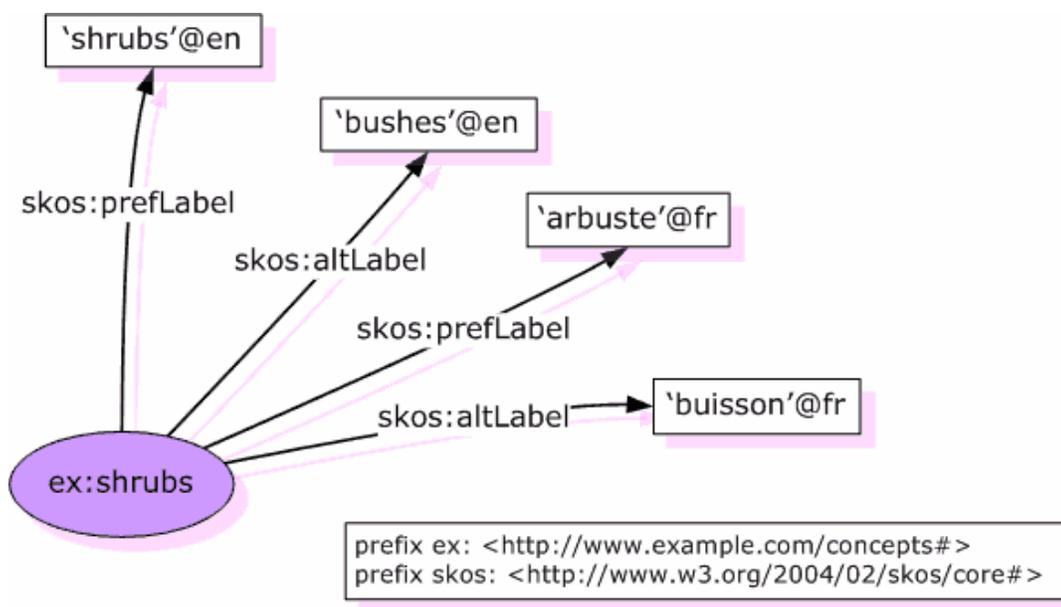


Fig. 11 Etichette multilingua

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"

  <skos:Concept rdf:about="http://www.example.com/concepts#shrubs">
    <skos:prefLabel xml:lang="en">shrubs</skos:prefLabel>
    <skos:altLabel xml:lang="en">bushes</skos:altLabel>
    <skos:prefLabel xml:lang="fr">arbuste</skos:prefLabel>
    <skos:altLabel xml:lang="fr">buisson</skos:altLabel>
  </skos:Concept>

</rdf:RDF
```

### 2.3.4 Aggiungere altre informazioni utili: *Proprietà per la documentazione*

Ci sono 7 proprietà che si possono usare per aggiungere dell'informazione per la descrizione di un concetto. Queste proprietà hanno come superclasse **skos:note**: essa può essere usata per fornire una qualunque tipo di descrizione utile a qualsiasi scopo.

Le sottoclassi invece sono:

- **skos:definition**: utile per dare una spiegazione completa e precisa della risorsa. Ad esempio “frutto della pianta della famiglia delle Bromeliaceae”.
- **skos:scopeNote**: utilizzato restringere od aumentare la descrizione de concetto  
Es: “Frequenze delle microne: da 1GHz a 300 GHz.
- **skos:example**: utilizzato per fornire degli esempi dell'uso del termine o del concetto
- **skos:historyNote**: utilizzato per indicare significativi cambiamenti di significato.  
Es: “Pere era stato inserito come concetto più specifico di verdura anziché frutta”.
- **skos:editorialNote**: utilizzato per fornire informazioni di tipo amministrativo  
Es: “Rivolgersi al Dr.Miles per una definizione più completa del concetto”.
- **skos:changeNote**: utilizzato per tener traccia delle varie modifiche occorse.  
Es: “Spostata la classe Pera da verdura a frutta. Effettuato il 15/12/2004 da Dr.Miles”.

Già con questo insieme ridotto di proprietà e di classi è possibile modellare un semplice esempio di concept schema un glossario.

#### 2.3.4.1 Il Glossario

La parola glossario deriva dal greco “glossa” e significava “lingua”, intesa come organo muscolare.

Il significato odierno lo definisce come in una lista di termini con le rispettive definizioni. Normalmente compare alla fine di un libro o manuale con lo scopo di introdurre il significato dei termini nuovi o di cui il lettore può non esserne a conoscenza. I termini possono riguardare un certo soggetto (glossario dei termini usati in un determinato documento) o un certo lavoro (glossario per la protezione e tutela dell'ambiente, ad esempio). In generale il glossario è specifico di un dato dominio e difficilmente include sinonimi: generalmente consiste in una lista di termini e definizioni.

Vogliamo quindi rappresentare i seguenti termini estratti dal glossario presente sul sito di Ateneo dell'Università di Ca'Foscari nella sezione orientamento:

**Piano di studio**: ‘insieme di insegnamenti che lo studente deve sostenere per ottenere il titolo di studio’

**Semestre**: ‘arco di tempo in cui vengono tenuti gli insegnamenti’

**Tutor:** ‘studente senior o docente a cui lo studente può rivolgersi avere per delle indicazioni’

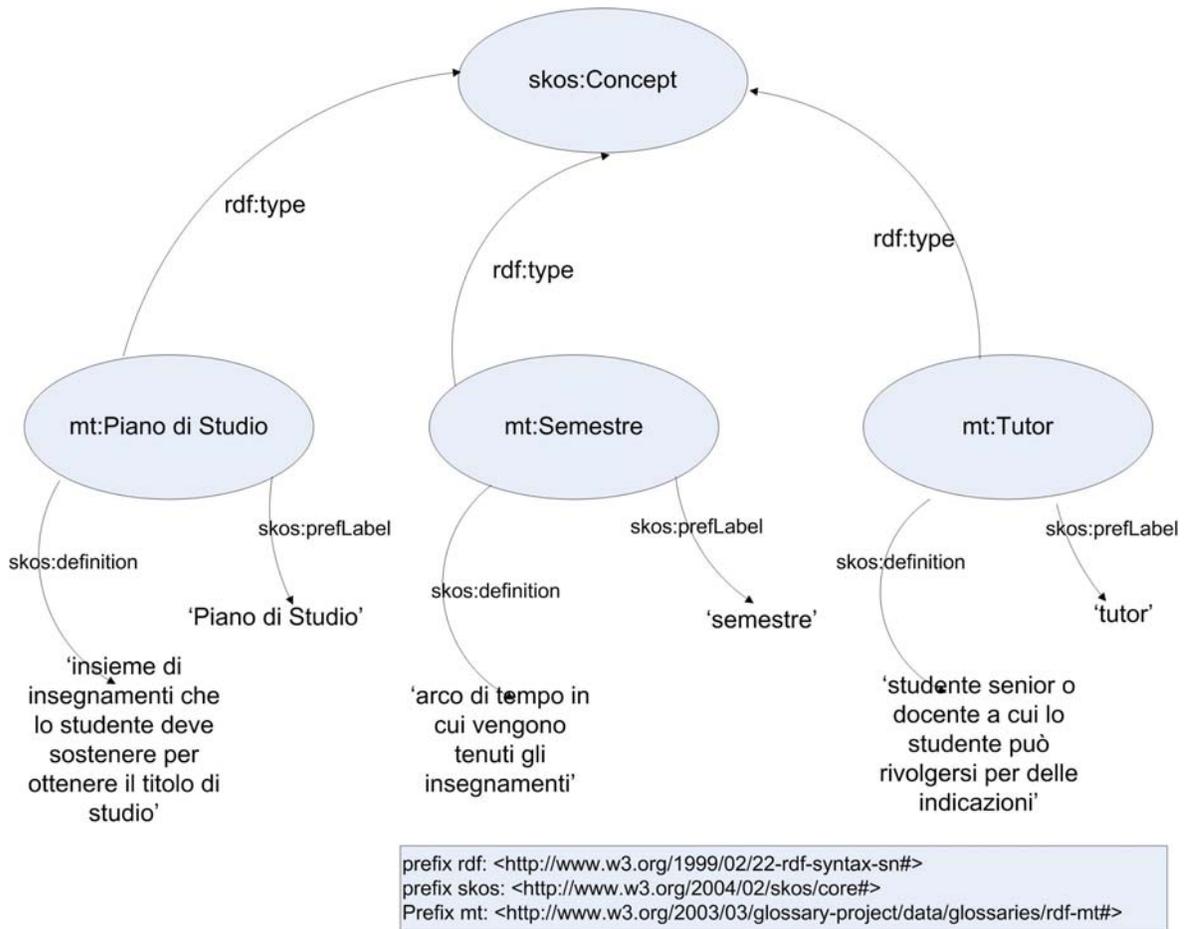


Fig. 12 Rappresentazione grafica del glossario estratto dalla guida dello studente

La codifica del grafo in RDF/XML è presentata qui di seguito

```
<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  xmlns:mt="http://www.w3.org/2003/03/glossary-
project/data/glossaires/rdf-mt#"

  <skos:Concept
rdf:about="http://www.w3.org/2003/03/glossaryproject/data/glossaries/rdf-
mt#Piano di Studio">
  <skos:prefLabel>Piano di studio</skos:prefLabel>
  <skos:definition>Insieme di insegnamenti che lo studente deve
sostenere per ottenere il titolo di studio
</skos:definition>
</skos:Concept>

  <skos:Concept rdf:about="http://www.w3.org/2003/03/glossary-
```

```

project/data/glossaries/rdf-mt#Semestre">
  <skos:prefLabel>Semestre</skos:prefLabel>
  <skos:definition>Arco di tempo in cui vengono tenuti gli insegnamenti
</skos:definition>
</skos:Concept>

  <skos:Concept rdf:about="http://www.w3.org/2003/03/glossary-
project/data/glossaries/rdf-mt#Tutor">
  <skos:prefLabel>Tutor</skos:prefLabel>
  <skos:definition>studente senior o docente a cui lo studente può
rivolgersi per delle indicazioni
  </skos:definition>
  </skos:Concept>
</rdf:RDF>

```

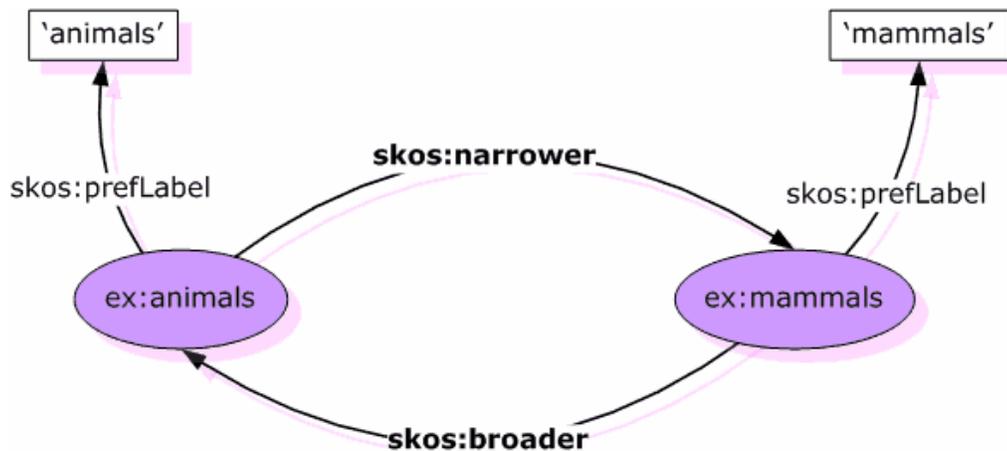
### 2.3.5 Le relazioni semantiche

Come detto in precedenza SKOS Core Vocabulary è un applicazione RDF disegnata per creare “concept schema”, dove con questo termine ci riferiamo a tutte quelle strutture come thesauri, vocabolari, tassonomie che integrano all’interno della propria struttura un determinato schema con delle relazioni semantiche per classificare, ordinare, regolare diversi concetti. Una delle cose fondamentali che questo framework offre è la possibilità di definire relazioni semantiche tra i concetti, un fatto molto familiare nella realtà e nella mente umana.

SKOS Core Vocabulary include delle proprietà per definire delle relazioni semantiche tra concetti diversi. Queste relazioni sono: `skos:broader`, `skos:narrower` e `skos:related`.

Per dire che un concetto ha un significato più generale di un altro si usa `skos:broader`, mentre per dire l’esatto inverso ovvero che il significato è più specifico si usa `skos:narrower`. Le due proprietà sono una la duale dell’altra.

Ecco un esempio [SKOSTUT]:



```

prefix ex: <http://www.example.com/concepts#>
prefix skos: <http://www.w3.org/2004/02/skos/core#>

```

Fig. 13 Proprietà per le relazioni semantiche: narrower e broader

Ed ecco la serializzazione in RDF:

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#">

  <skos:Concept rdf:about="http://www.example.com/concepts#mammals">
    <skos:prefLabel>mammals</skos:prefLabel>
    <skos:broader
rdf:resource="http://www.example.com/concepts#animals"/>
  </skos:Concept>

  <skos:Concept rdf:about="http://www.example.com/concepts#animals">
    <skos:prefLabel>animals</skos:prefLabel>
    <skos:narrower
rdf:resource="http://www.example.com/concepts#mammals"/>
  </skos:Concept>

</rdf:RDF>
```

**Skos:related** è più generico: si usa per dire che un determinato concetto ha una qualche relazione con un altro. A questo punto è opportuno introdurre una nuova classe che permette di relazionare diversi termini tra loro.

Vediamo ora un esempio che fa uso di queste nuove relazioni introdotte:

### 2.3.5.1 Tassonomia

Il termine tassonomia deriva dal greco “tassein” cioè classificare e nomos cioè legge, scienza. Inizialmente la tassonomia era solamente la scienza che si occupava di classificare gli organismi, ma successivamente la parola ha cominciato ad essere usata per classificazioni di altro tipo e oggi giorno essa può essere utilizzata per qualsiasi oggetto.

La sua caratteristica fondamentale è la struttura gerarchica:

Può essere organizzata organizzando gli oggetti in gruppi, ma più spesso è organizzata secondo una struttura ad albero: come esempio consideriamo lo schema di classificazione scientifico (quello inventato da Linneo) il nodo padre è Organismi e poi in ordine gerarchico Regno, Phylum, Classe, Ordine, Famiglia, Genere, Specie.

L'esempio di tassonomia presentato è estratto da un manuale di zoologia e illustra la classificazione scientifica del leopardo.

- Regno: Animali
  - Philum:Cordati
    - Classe: Mammiferi
      - Ordine: Carnivori
        - Famiglia: Felidi
          - Specie: Pantera Pardus

Per brevità nel grafo sottostante non sono stati riportati tutti i vari ordini.

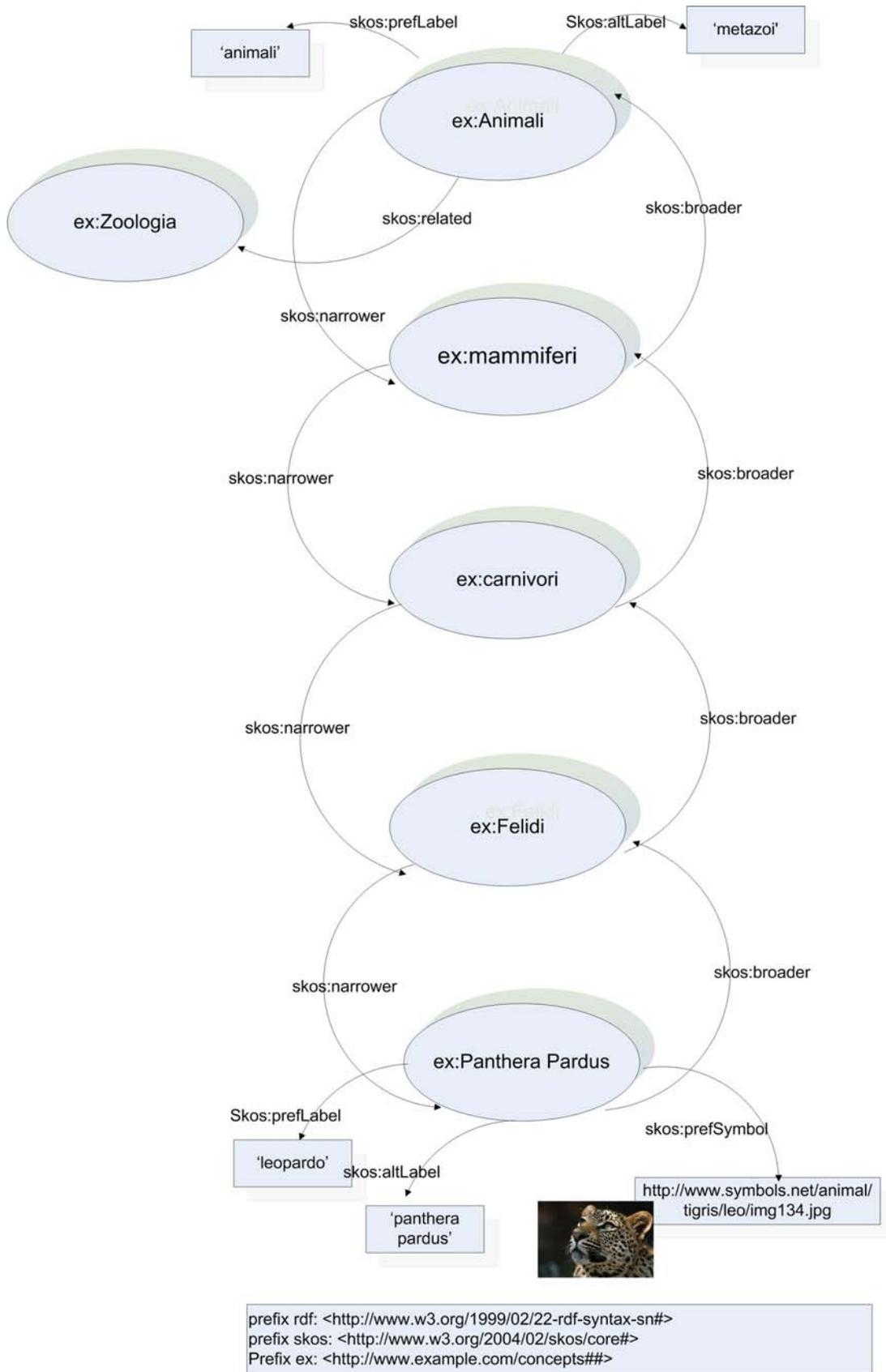


Fig. 14 Esempio di estratto di tassonomia del leopardo

## La rispettiva codifica in XML/RDF

```
<<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  xmlns:ex="http://www.example.com/eg#">

  <skos:Concept rdf:about="http://www.example.com/concepts#animali">
    <skos:prefLabel>Animali</skos:prefLabel>
    <skos:altLabel>Metazoi</skos:altLabel>
  </skos:Concept>

  <skos:Concept rdf:about="http://www.example.com/concepts#mammiferi">
    <skos:prefLabel>mammiferi</skos:prefLabel>
    <skos:narrower
rdf:resource="http://www.example.com/concept#carnivori">
    <skos:broader
rdf:resource="http://www.example.com/concepts#animali"/>
  </skos:Concept>

  <skos:Concept rdf:about="http://www.example.com/concepts#carnivori">
    <skos:prefLabel>carnivori</skos:prefLabel>
    <skos:narrower rdf:resource="http://www.example.com/concept#felidi">
    <skos:broader
rdf:resource="http://www.example.com/concepts#mammiferi"/>
  </skos:Concept>

  <skos:Concept rdf:about="http://www.example.com/concepts#felidi">
    <skos:prefLabel>Felidi</skos:prefLabel>
    <skos:narrower
rdf:resource=http://www.example.com/concept#Leopardo"/>
    <skos:broader
rdf:resource="http://www.example.com/concepts#carnivori"/>
  </skos:Concept>

  <skos:Concept rdf:about="http://www.example.com/concepts#leopardo">
<skos:prefLabel>Leopardo</skos:prefLabel>
<skos:prefSymbol>rdf:resource=http://www.symbols.net/animal/tigris/img1
34.jpg"/>
    <skos:broader
rdf:resource="http://www.example.com/concepts#felidi"/>
  </skos:Concept>

</rdf:RDF>
```

Notiamo la definizione dei vari concetti con la rispettiva definizione di URI. La forma gerarchica attribuita alla tassonomia è stata effettuata utilizzando l'uso di narrower e broader. Attraverso l'attributo rdf:resource viene indicata la risorsa con senso più ampio o più specifico.

## 2.4 Verso un esempio completo: La classe Concept Scheme

Abbiamo appena visto che spesso i concetti sono definiti in relazione con altri concetti. I casi presentati precedentemente erano degli estratti presi da strutture per l'organizzazione della conoscenza.

Quindi le classi e le proprietà viste finora permettevano solamente di estendere schemi di classificazione già esistenti.

Una cosa che molto spesso vorremmo fare e che, il framework di SKOS mette a disposizione, è costruire nuovi schemi partendo da zero.

Oltre alle relazioni semantiche offerte delle proprietà `skos:narrower` e `skos:broader` esiste la classe `skos:Concept Scheme`.

Questa classe permette di definire un concept-schema personalizzato, basato sulle nostre effettive necessità e che, grazie al meccanismo degli URI, potremmo usare insieme con altre strutture già esistenti per descrivere un certo dominio essendo sicuri di non portare nessuna ambiguità.

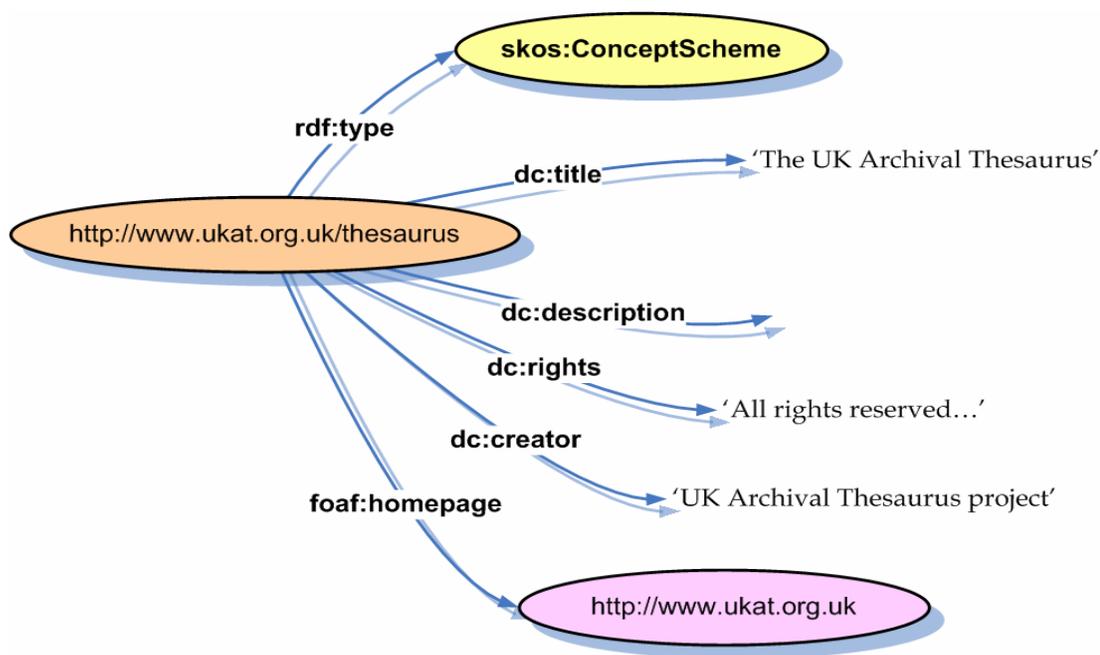


Fig.15 Rappresentazione di concept schema

Ecco presentato qui sopra un esempio di thesaurus che utilizza delle meta properties di DUBLIN CORE e FOAF per aggiungere dei metadati riguardanti lo schema: il titolo, l'autore e la descrizione.

Se ad esempio volessimo definire un semplice thesaurus in RDF/XML avremmo una codifica del genere:

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">

  <skos:ConceptScheme rdf:about="http://example.com/thesaurus">
    <dc:title>Un semplice esempio di thesauro</dc:title>
    <dc:description>Un esempio di thesauro per mostrare l'uso
di SKOS-Core schema</dc:description>
    <dc:creator>Matteo Ballarin</dc:creator>
  </skos:ConceptScheme>

</rdf:RDF>
```

xmlns:dc="http://purl.org/dc/elements/1.1/" è l'URI di Dublin Core che identifica le proprietà title, description e creator.

I concetti, come abbiamo visto prima, non devono per forza appartenere a un determinato schema, ma possono essere dichiarati come entità a se stante. Se però vogliamo dire che un certo termine appartiene ad un particolare schema, usiamo la proprietà skos:inScheme. Ad esempio:

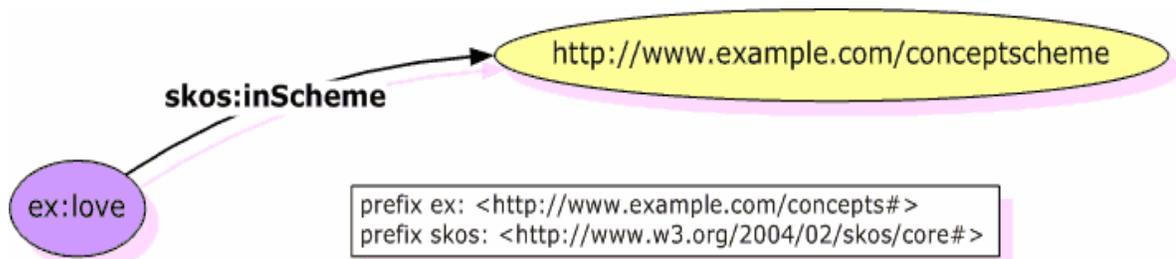


Fig. 16 La proprietà inScheme

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"

  <skos:Concept rdf:about="http://www.example.com/concepts#love">
    <skos:inScheme rdf:resource="http://www.example.com/conceptscheme" />
  </skos:Concept>

</rdf:RDF>
```

Quasi sempre i thesauri sono organizzati raggruppando i concetti in super categorie; delle categorie fondamentali se vogliamo. Per rappresentare un certo concetto come nodo di livello più alto esiste la proprietà `skos:hasTopConcept`. Questa permette di creare una struttura gerarchica come questa qui di seguito [SKOSTUT]:

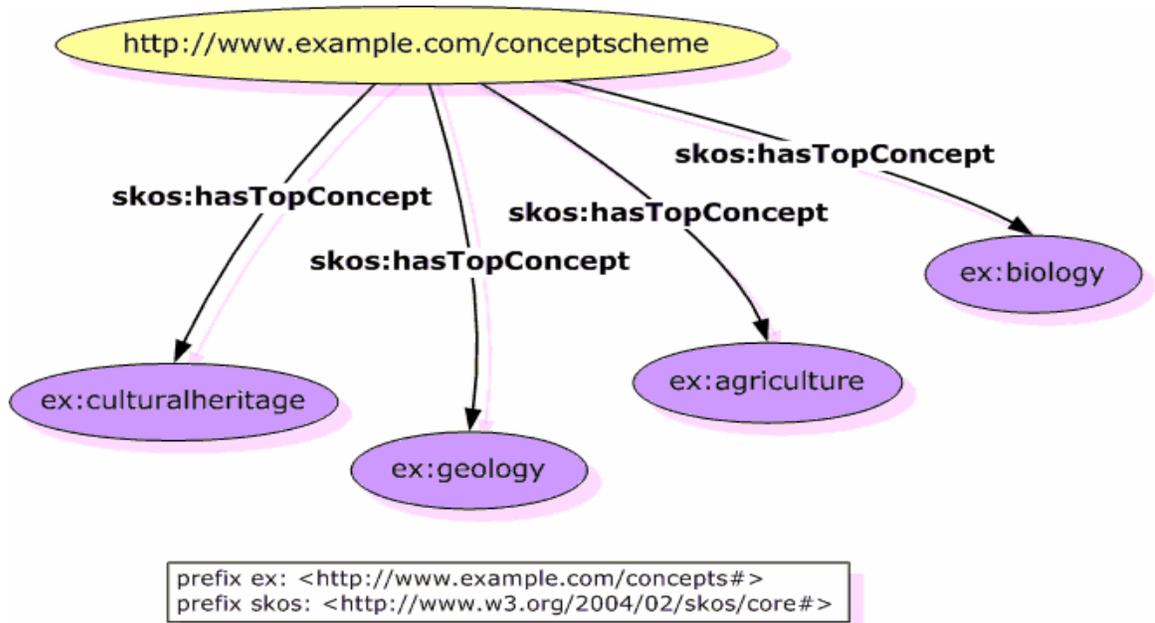


Fig. 17 La proprietà hasTopConcept

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"

  <skos:ConceptScheme rdf:about="http://www.example.com/conceptscheme">
    <skos:hasTopConcept
  rdf:resource="http://www.example.com/concepts#biology"/>
    <skos:hasTopConcept
  rdf:resource="http://www.example.com/concepts#geology"/>
    <skos:hasTopConcept
  rdf:resource="http://www.example.com/concepts#agriculture"/>
    <skos:hasTopConcept
  rdf:resource="http://www.example.com/concepts#culturalheritage"/>
  </skos:ConceptScheme>
</rdf:RDF>
  
```

La proprietà `hasTopConcept` è molto importante in quanto, creando un collegamento tra concept schema e i vari concetti, permette di identificare in maniera immediata i termini padre di un dato termine e questo può essere molto utile in fase di consultazione o di ricerca nel thesauro.

## 2.5 La classe *Collection* e l'organizzazione a faccette

Molti thesauri organizzano i concetti secondo faccette (dall'inglese "facet"): dei raggruppamenti ad alto livello di concetti inerenti alla stessa categoria, disciplina, materiale, luogo, ecc. Ad esempio:

- Animali, mici, batteri, betulle potrebbero appartenere alla faccetta degli *esseri viventi*.
- Scrivere, cucinare, scavare potrebbero appartenere alla faccetta delle *attività*.
- Compleanni, guerre, partite di calcio potrebbero apparire come faccetta di *eventi*.

Questo tipo di collezione di termini in biblioteconomia è chiamato '*schiera*' e l'etichetta che è loro assegnata identifica il '*criterio di raggruppamento*'. Ulteriori raggruppamenti all'interno di una faccetta, secondo ulteriori caratteristiche di divisione nel linguaggio tecnico dei tesauri, sono chiamati *sottofaccette* o *sottoschiere*.

Ecco un esempio in cui viene modellato la categoria del latte. Abbiamo preso in considerazione come elemento di raggruppamento l'origine: in questo caso è il latte prodotto dagli animali, e come elementi della schiera al suo interno troviamo il latte di mucca, il latte di bufala, di pecora e così via.

<milk by source  
animal>

← Etichetta del nodo indicante le caratteristiche del raggruppamento

- Buffalo milk
  - cow milk
  - goat milk
  - sheep milk
- } Schiera

La classe che ci permette di modellare queste particolari tipi di categorie è la classe `skos:Collection`.

Attraverso la proprietà `skos:member` è possibile aggiungere i vari componenti della schiera al nodo principale che descrive la categoria.

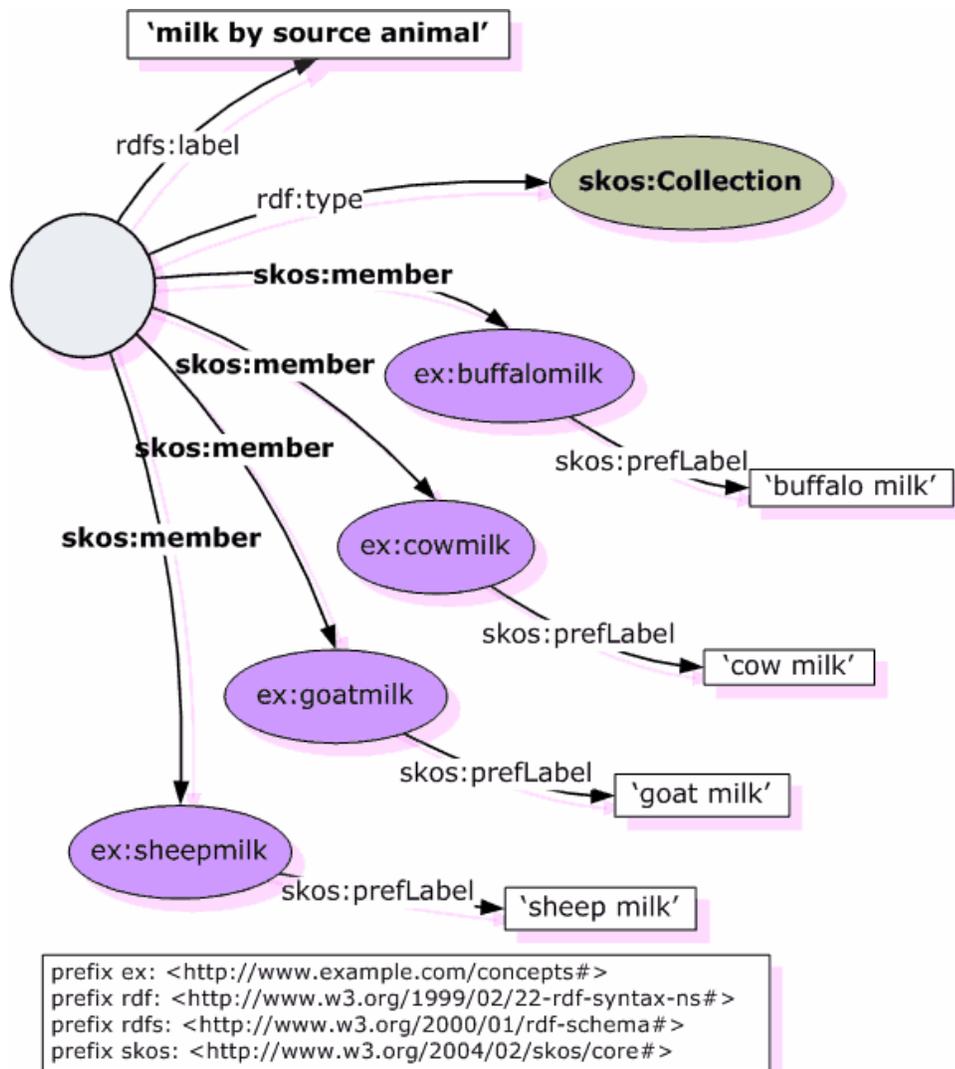


Fig. 18 La proprietà member per modellare schiere di conoscenza

La rispettiva codifica in RDF/XML

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">

  <skos:Collection>
    <rdfs:label>milk by source animal</rdfs:label>
    <skos:member
rdf:resource="http://www.example.com/concepts#buffalomilk"/>
    <skos:member rdf:resource="http://www.example.com/concepts#cowmilk"/>
    <skos:member
rdf:resource="http://www.example.com/concepts#goatmilk"/>
    <skos:member
rdf:resource="http://www.example.com/concepts#sheepmilk"/>
  </skos:Collection>

  <skos:Concept rdf:about="http://www.example.com/concepts#buffalomilk">

```

```

    <skos:prefLabel>buffalo milk</skos:prefLabel>
  </skos:Concept>

  <skos:Concept rdf:about="http://www.example.com/concepts#cowmilk">
    <skos:prefLabel>cow milk</skos:prefLabel>
  </skos:Concept>

  <skos:Concept rdf:about="http://www.example.com/concepts#goatmilk">
    <skos:prefLabel>goat milk</skos:prefLabel>
  </skos:Concept>

  <skos:Concept rdf:about="http://www.example.com/concepts#sheepmilk">
    <skos:prefLabel>sheep milk</skos:prefLabel>
  </skos:Concept>

</rdf:RDF>

```

In alcune schiere l'ordine con cui sono rappresentati i concetti è significativo, quando ad esempio vogliamo rispettare l'ordine alfabetico. La sottoclasse `skos:CollectionOrdered` attraverso la proprietà `skos:memberList` permette di definire collezioni di concetti che rispettano un certo ordine. Ecco un esempio:

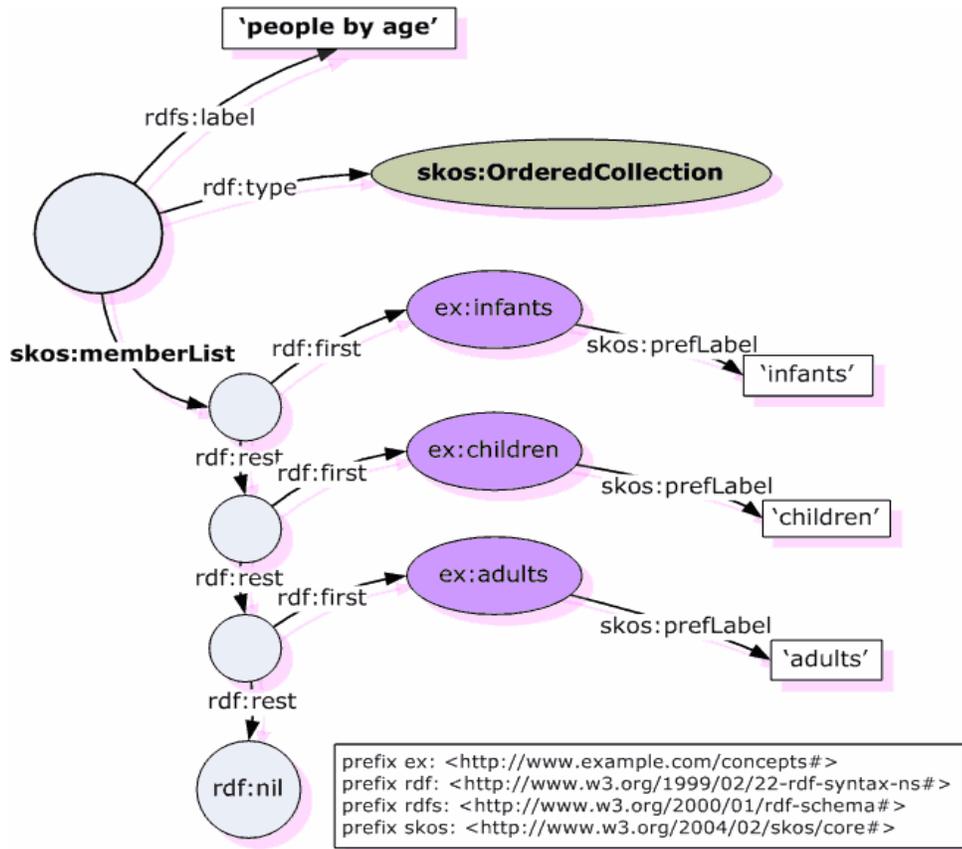


Fig. 19 Rappresentazione di collezione ordinata

## La serializzazione in XML/RDF

```
rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"

  <skos:OrderedCollection>
    <rdfs:label>people by age</rdfs:label>
    <skos:memberList rdf:parseType="Collection">
      <skos:Concept
rdf:about="http://www.example.com/concepts#infants"/>
      <skos:Concept
rdf:about="http://www.example.com/concepts#children"/>
      <skos:Concept rdf:about="http://www.example.com/concepts#adults"/>
    </skos:memberList>
  </skos:OrderedCollection>

  <skos:Concept rdf:about="http://www.example.com/concepts#infants">
    <skos:prefLabel>infants</skos:prefLabel>
  </skos:Concept>

  <skos:Concept rdf:about="http://www.example.com/concepts#children">
    <skos:prefLabel>children</skos:prefLabel>
  </skos:Concept>

  <skos:Concept rdf:about="http://www.example.com/concepts#adults">
    <skos:prefLabel>adults</skos:prefLabel>
  </skos:Concept>

</rdf:RDF>
```

Ovviamente è possibile definire nested collection con cui rappresentare strutture come

```
people
.<people by age>
..<infants by age>
...babies
...toddlers
..children
..adults
```

## La codifica in RDF/XML

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"

  <skos:Concept rdf:about="http://www.example.com/concepts#people">
    <skos:prefLabel>people</skos:prefLabel>
    <skos:narrower>
      <skos:OrderedCollection>
        <rdfs:label>people by age</rdfs:label>
        <skos:memberList rdf:parseType="Collection">
```

```

    <skos:OrderedCollection>
      <rdfs:label>infants by age</rdfs:label>
      <skos:memberList rdf:parseType="Collection">
        <skos:Concept
rdf:about="http://www.example.com/concepts#babies"/>
        <skos:Concept
rdf:about="http://www.example.com/concepts#toddlers"/>
      </skos:memberList>
    </skos:OrderedCollection>
    <skos:Concept
rdf:about="http://www.example.com/concepts#children"/>
    <skos:Concept
rdf:about="http://www.example.com/concepts#adults"/>
    </skos:memberList>
  </skos:OrderedCollection>
</skos:narrower>
</skos:Concept>

<skos:Concept rdf:about="http://www.example.com/concepts#babies">
  <skos:prefLabel>babies</skos:prefLabel>
</skos:Concept>

<skos:Concept rdf:about="http://www.example.com/concepts#toddlers">
  <skos:prefLabel>toddlers</skos:prefLabel>
</skos:Concept>

</rdf:RDF>

```

## ***2.6 Utilizzo avanzato di SKOS Core: L'integrazione di SKOS con altri vocabolari e il Subject Indexing***

### ***2.6.1 L'integrazione di SKOS con Dublin Core e FOAF***

Come già anticipato, una delle caratteristiche fondamentali e più interessanti del framework di SKOS è la naturale estensibilità ed interoperabilità di SKOS Core Vocabulary con altri vocabolari per il web semantico.

**Dublin Core**, come accennato, si propone come uno standard di descrizione delle risorse in formato elettronico ed è costituito da diversi elementi descrittivi come: Title (titolo), Creator (Creatore), Subject (Soggetto) ecc.

Data la sua semplicità è ampiamente utilizzato e praticamente tutti gli standard e i progetti fanno riferimento ad esso.

L'esempio qui di seguito illustra come SKOS Core possa essere utilizzato in combinazione con altri vocabolari per il web semantico per creare descrizioni più ricche ed esaustive di un certo dominio.

Supponiamo quindi di voler modellare il concetto di laptop. Abbiamo supposto che da una certa data in poi, il termine preferito o più comunemente usato per riferirsi ad un pc portatile, sia passato da laptop a notebook. Abbiamo quindi utilizzato la proprietà “*change*

*note*” introdotta all’inizio e abbiamo ritenuto carino e opportuno indicare quando è avvenuto questo cambiamento e da chi è stato fatto. Un altro vocabolario utilizzato in questo esempio oltre a Dublin Core è **FOAF** (Friend Of A Friend).

**FOAF** (Friend-of-a-Friend) è il nome-acronimo del personaggio misterioso a cui capitano tutte le vicende raccontate nelle leggende metropolitane di mezzo mondo; ma è anche il nome di un progetto (FOAF-project.org) di semantic web che punta a creare una rete mondiale di amici-di-amici utilizzando un vocabolario comune, un vocabolario RDF chiamato FOAF: a essere codificati in formato standard sono dati personali e relazioni tra persone (gruppi di persone, comunità, aziende). In un file FOAF vengono inseriti la descrizione di un singolo e una lista di contatti (*friends*): amici, collaboratori, colleghi, ecc. L'idea di fondo è quella che per cercare una persona si possa fare come nella vita di tutti i giorni, muovendosi tra le conoscenze: questa persona ha collaborato con un amico con cui ho lavorato, per cui potrebbe interessarmi contattarlo... Il tutto si basa su un file di metadati RDF legato ad una pagina personale, che oltre a specificare dati e caratteristiche della persona, permette di ricostruire un elenco di altre pagine e dei relativi metadati (la rete di amici).

Attraverso quest'altra applicazione di RDF abbiamo aggiunto altri metadati importanti che rendono più complessa e strutturata la risorsa che stiamo descrivendo.

Conoscere il "percorso" che una determinato termine ha seguito da una comunità all'altra può aiutarci inoltre a comprendere la natura del termine.

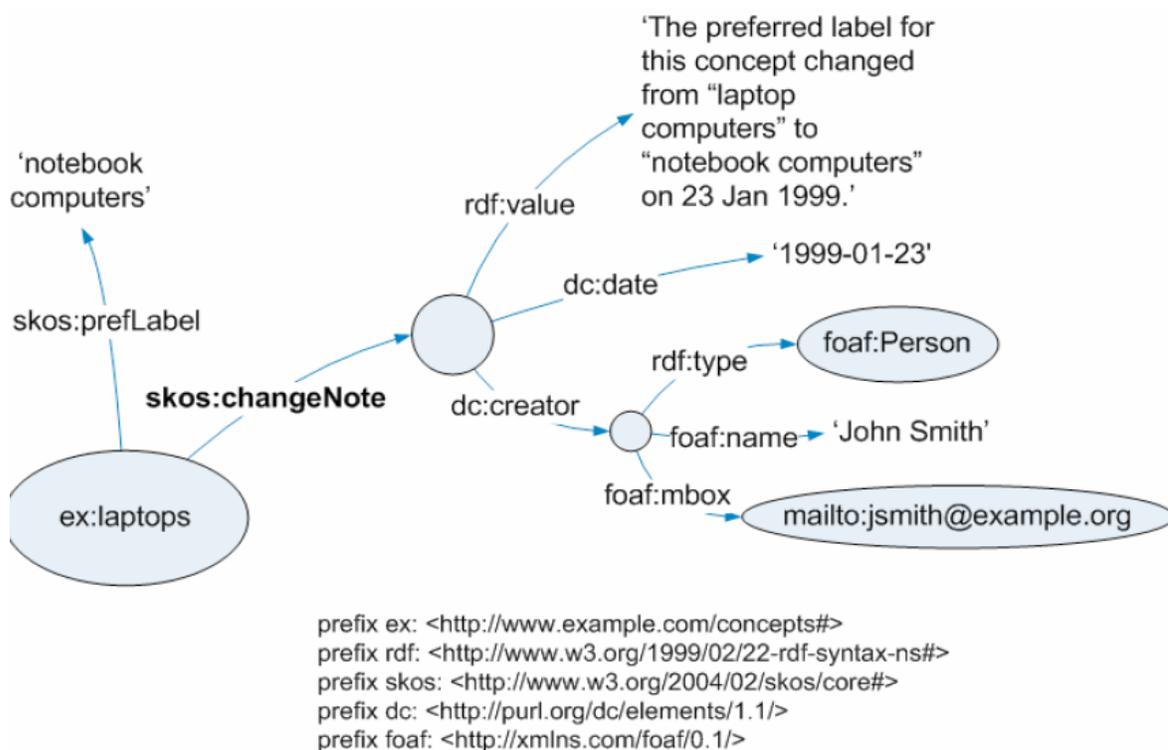


Fig. 20 Un esempio di integrazione di diversi vocabolari  
 Il prefisso dc identifica una properties di Dublin Core, mentre foaf indica una property di FOAF

Ecco la serializzazione in XML/RDF del grafo

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:foaf="http://xmlns.com/foaf/0.1/">

  <skos:Concept rdf:about="http://www.example.org/concepts#laptops">
    <skos:prefLabel>notebook computers</skos:prefLabel>
    <skos:changeNote rdf:parseType="Resource">
      <rdf:value>The preferred label for this concept changed from
'laptop computers' to
'notebook computers' on 23 Jan 1999.</rdf:value>
      <dc:creator>
        <foaf:Person>
          <foaf:name>John Smith</foaf:name>
          <foaf:mbox rdf:resource="mailto:jsmith@example.org"/>
        </foaf:Person>
      </dc:creator>
      <dc:date>1999-01-23</dc:date>
    </skos:changeNote>
  </skos:Concept>

</rdf:RDF>
```

Da notare l'uso dell'attributo `rdf:parseType="Resource"`, un'abbreviazione possibile nella sintassi RDF/XML per rappresentare i nodi vuoti.

## 2.6.2 Subject Indexing

Scopo principale della creazione di un concept schema è fornire un aiuto nell'organizzazione di un insieme di risorse, ad esempio per un insieme di documenti Web. Spesso di un documento è interessante conoscere il soggetto e i concetti che vengono rappresentati all'interno di esso. Il processo di Subject Indexing o indicizzazione per soggetto, consiste in un'analisi ragionata del contenuto del documento al fine di identificare i concetti in esso presenti, ed assegnare loro una qualche etichetta testuale, per permettere successivamente un rapido ritrovamento dell'informazione per un uso futuro. Subject Indexing properties è in insieme di proprietà messe a disposizione da SKOS Core che permette l'estensione della proprietà `dc:subject` di Dublin Core per definire relazioni tra risorse e concetti.

Le proprietà messe a disposizione da SKOS per il subject indexing sono:

- `skos:subject`
- `skos:isSubjectOf`
- `skos:primarySubject`
- `skos:isPrimarySubjectOf`.

Skos:subject è una sub-property di dc:subject. Il significato è lo stesso ma mentre dc:subject si utilizza con qualsiasi tipo di risorsa, il campo di azione di skos:subject è ristretto alla classe Concept. Essa è quindi utilizzata per dire che un certo concetto è il soggetto di una determinata risorsa. Ad esempio

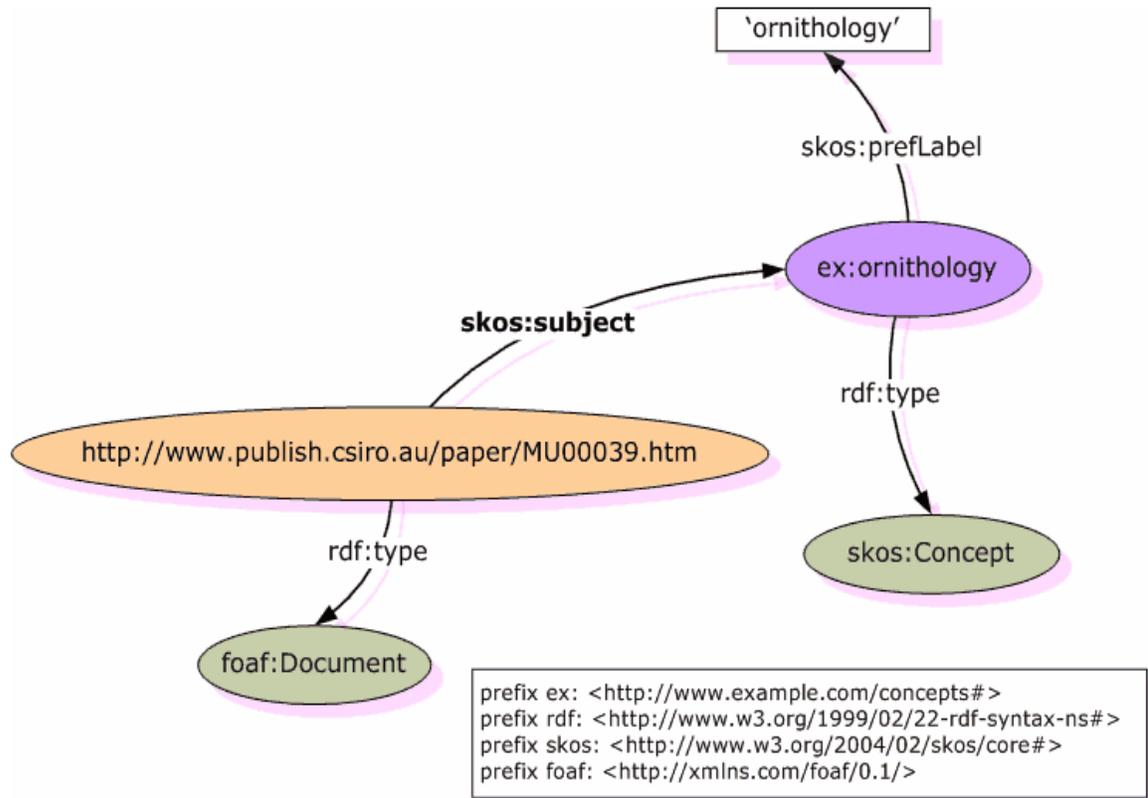


Fig 21 La proprietà member subject

Questa proprietà supporta la regola di inferenza  $[(?d \text{ skos:subject } ?x) (?x \text{ skos:broader } ?y) \text{ implica } (?d \text{ skos:subject } ?y)]$  cioè che se x è il soggetto di una certo predicato d e y ha significato più ampio (broader) allora anche y è soggetto di d.

Skos:isSubjectOf è invece la proprietà inversa di skos:subject.

La proprietà skos:PrimarySubject (e la proprietà inversa skos:isPrimarySubjectOf) è una sub-property di skos:subject. Essa è utile per affermare, in quei documenti o risorse in cui compaiono diversi soggetti, qual'è il soggetto principale.

Queste viste finora alcune delle principali classi ed etichette di SKOS Core Vocabulary.

Molte di esse sono ancora in fase di testing da parte della comunità del W3C, altre necessitano di una revisione.

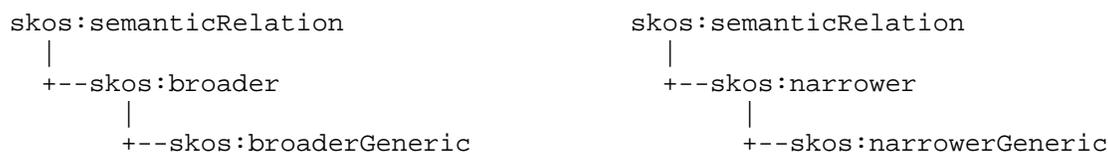
## 2.7 Proprietà più avanzate: SKOS Extension Vocabulary Specification

Oltre a queste viste finora, esistono delle altre proprietà pensate per estendere le funzionalità di SKOS Core Vocabulary e per supportare alcune caratteristiche dei sistemi per l'organizzazione della conoscenza, in particolare dei thesauri: SKOS Extensions [SKOSE04]. Di queste saranno dati alcuni accenni senza scendere nel dettaglio in quanto sono "unstable" e necessitano di revisione da parte del gruppo di lavoro.

La proprietà `skos:semanticRelation` corrisponde al nodo radice di tutte le proprietà utilizzate per "linkare" semanticamente diversi concetti all'interno dello stesso concept-schema. Al primo livello di estensione, abbiamo visto prima, troviamo le proprietà `skos:broader`, `skos:narrower` e `skos:related`. A loro volta queste prevedono un ulteriore livello di estensione.

Vediamo la struttura:

### *BroaderGeneric / NarrowerGeneric*



Queste due proprietà permettono di "incasellare" in una più specifica o più generale categoria un concetto (sono l'equivalente semantico di `rdfs:subClassOf`). Vediamo un esempio:

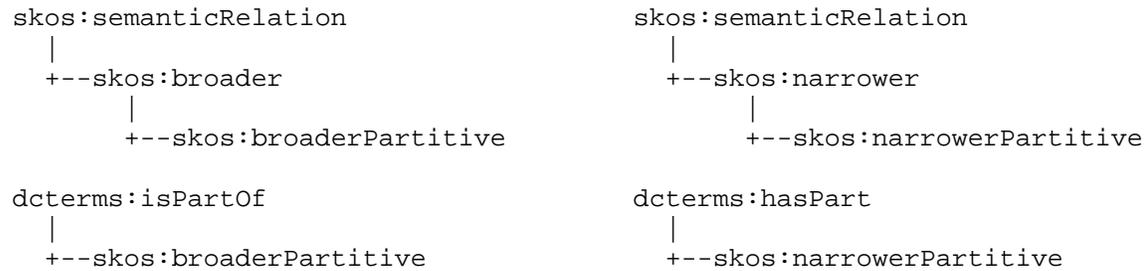
```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  xmlns:sex="http://www.w3.org/2004/02/skos/extensions#">

  <skos:Concept rdf:about="http://example.com/Concept/0004">
    <skos:prefLabel>English pubs, clubs and
bars</skos:prefLabel>
    <skos:altLabel>English drinking
establishments</skos:altLabel>
    <skos:inScheme
rdf:resource="http://example.com/thesaurus"/>
    <sex:narrowerGeneric
rdf:resource="http://example.com/Concept/0005"/>
  </skos:Concept>

  <skos:Concept rdf:about="http://example.com/Concept/0005">
    <skos:prefLabel>English pubs</skos:prefLabel>
    <skos:altLabel>English public houses</skos:altLabel>
    <skos:inScheme
rdf:resource="http://example.com/thesaurus"/>
    <sex:broaderGeneric
rdf:resource="http://example.com/Concept/0004"/>
  </skos:Concept>

</rdf:RDF>
```

## ***BroaderPartitive / NarrowerPartitive***



Queste properties si utilizzano per esprimere relazioni partitive (semanticamente equivalente a `rdf:Type`). Ad esempio l'Oxfordshire è parte dell'Inghilterra: ecco la rappresentazione con SKOS:

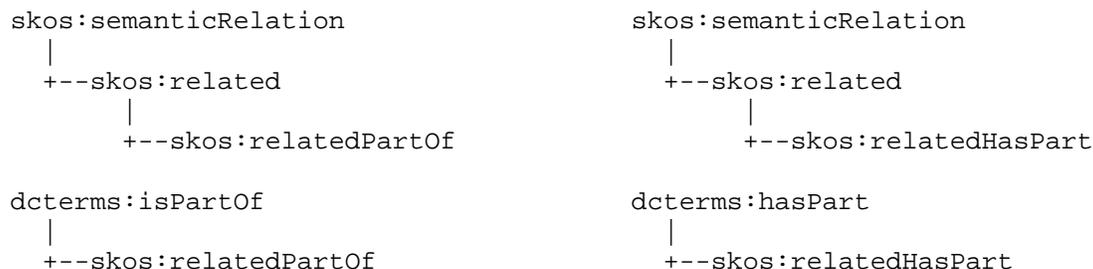
```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#">

  <skos:Concept rdf:about="http://example.com/concept/0007">
    <skos:prefLabel>Oxfordshire county</skos:prefLabel>
    <skos:altLabel>Oxon</skos:altLabel>
    <skos:inScheme
rdf:resource="http://example.com/thesaurus"/>
    <skos:broaderPartitive
rdf:resource="http://example.com/concept/0008"/>
  </Concept>

  <skos:Concept rdf:about="http://example.com/concept/0008">
    <skos:prefLabel>England</skos:prefLabel>
    <skos:inScheme
rdf:resource="http://example.com/thesaurus"/>
    <skos:narrowerPartitive
rdf:resource="http://example.com/concept/0007"/>
  </Concept>

</rdf:RDF>
```

## ***RelatedPartOf / RelatedHasPart***



È simile alle properties precedenti serve per esprimere relazioni partitive. È stato introdotto nel framework di SKOS per permettere una completa interoperabilità tra thesauri che trattano in maniera diversa le relazioni semantiche di tipo partitivo.

## **2.8 Estendere e personalizzare le proprietà**

Tutte le proprietà che collegano i concetti tra di loro fanno parte di una gerarchia derivata dalla proprietà fondamentale `skos:semanticRelation`. È possibile allora personalizzare ed estendere SKOS Vocabulary definendo delle altre relazioni semantiche secondo le proprie necessità.

È possibile inoltre “specializzare” le proprietà per le etichette in maniera da avere preferred e lexical label diverse a seconda del tipo di utenti. Ad esempio per medici e pazienti, per bambini ed adulti, per scienziati e non.

Le nuove proprietà potranno avere un significato più specifico in modo da soddisfare un requisito particolare e locale, senza peraltro sacrificare il beneficio dell’interoperabilità offerta dall’utilizzo di una rappresentazione standard per i thesaurus.

Vediamo subito un esempio di questa importante ed interessante caratteristica.

Supponiamo di voler definire l’etichetta *prefScientificLabel*: ecco come è possibile farlo:

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"

  <rdf:Property rdf:about="http://www.example.com/skos-core-
extension#prefScientificLabel">
    rdfs:label 'preferred scientific label';
    skos:definition 'The preferred lexical label for scientists.';
    rdfs:subPropertyOf skos:prefLabel;
  </rdf:Property>

  <rdf:Property rdf:about="http://www.example.com/skos-core-
extension#altScientificLabel">
    rdfs:label 'alternative scientific label';
    skos:definition 'An alternative lexical label for scientists.';
    rdfs:subPropertyOf skos:altLabel;
  </rdf:Property>

</rdf:RDF>
```

Ecco un possibile utilizzo:

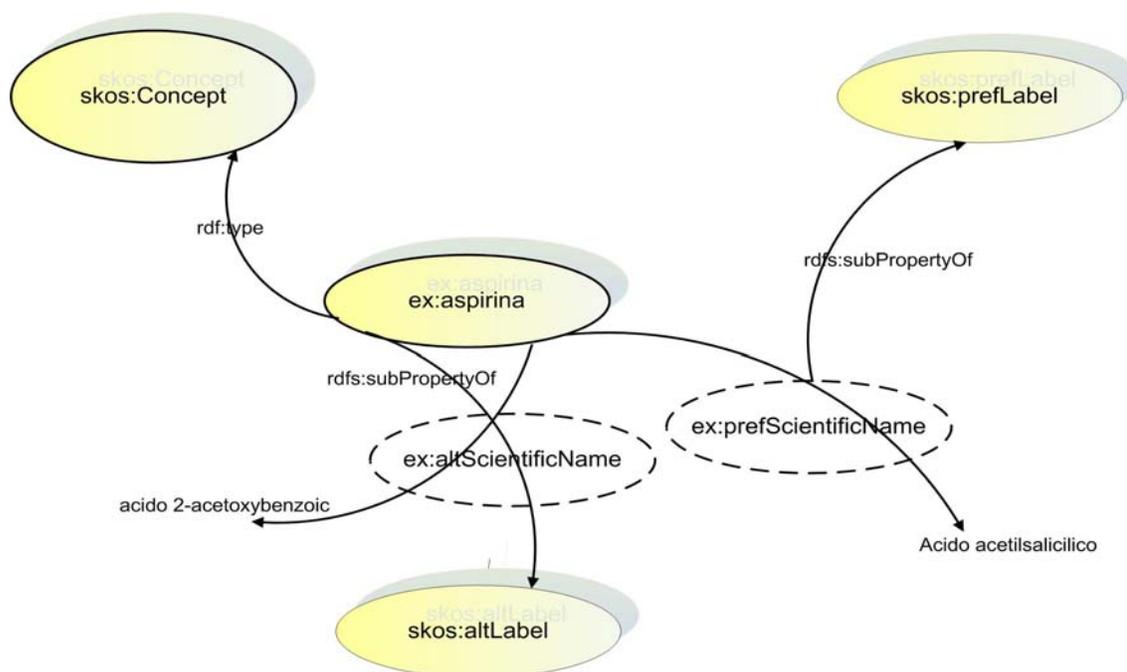


Fig. 22 Estensione e raffinamento di etichette in SKOS

Notiamo come la proprietà “*prefScientificLabel*” è stata dichiarata come sub-property di `skos:prefLabel` mentre le altre come sub-properties di `skos:altLabel`. Questo per evitare di violare il vincolo secondo cui per ogni singolo concetto può esistere una sola preferred label.

Vediamo adesso un altro esempio riguardante l’estensione delle relazioni semantiche. Come già visto è necessario per creare un’estensione di una proprietà è necessario dichiararla.

Vogliamo quindi modellare la relazione “*isPhysicalPartOf*” e la sua duale “*hasPhysicalPart*”.

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"

  <rdf:Property rdf:about="http://www.example.com/skos-core-
extensions#isPhysicalPartOf">
    <rdfs:comment>Expresses a physical has-part relationship between two
concepts.</rdfs:comment>
    <rdfs:subPropertyOf
rdf:resource="http://www.w3.org/2004/02/skos/core#broader"/>
    <owl:inverseOf rdf:resource="http://www.example.com/skos-core-
extensions#hasPhysicalPart"/>
  </rdf:Property>

```

```

<rdf:Property rdf:about="http://www.example.com/skos-core-
extensions#hasPhysicalPart">
  <rdfs:comment>Expresses a physical is-part-of relationship between
two concepts.</rdfs:comment>
  <rdfs:subPropertyOf
rdf:resource="http://www.w3.org/2004/02/skos/core#narrower"/>
  <owl:inverseOf rdf:resource="http://www.example.com/skos-core-
extensions#isPhysicalPartOf"/>
</rdf:Property>

```

In questo esempio è interessante notare l'utilizzo di alcuni costrutti OWL [OWLHP] (Ontology Web Language).

OWL fornisce un linguaggio per definire ontologie strutturate basate sul Web che permettono un'integrazione ed un'interoperabilità maggiore di dati tra comunità che descrivono il loro dominio di conoscenza.

OWL si basa su "RDF Model and Schema" e aggiunge un vocabolario più ampio per descrivere proprietà e classi: tra le altre, relazioni tra classi (ad esempio disgiunzione), cardinalità (ad esempio "esattamente uno"), uguaglianza, tipizzazione più ricca di proprietà, caratteristiche di proprietà (ad esempio simmetria) e classi enumerate.

In questo esempio la proprietà `owl:inverse` viene usata per dichiarare la relazione inversa tra `isPhysicalPart` e `hasPhysicalPart`.

Vediamo ora come queste nuove proprietà possono essere utilizzate per modellare il Regno Unito e l'Irlanda del Nord come stato appartenente all'Europa e la proprietà duale.

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  xmlns:ext="http://www.example.com/skos-core-extensions#">

  <skos:Concept rdf:about="http://www.example.com/geography#UK">
    <skos:prefLabel>United Kingdom of Great Britain and Northern
Ireland</skos:prefLabel>
    <ext:isPhysicalPartOf
rdf:resource="http://www.example.com/geography#Europe"/>
  </skos:Concept>

  <skos:Concept rdf:about="http://www.example.com/geography#Europe">
    <skos:prefLabel>Europe</skos:prefLabel>
    <ext:hasPhysicalPart
rdf:resource="http://www.example.com/geography#UK"/>
  </skos:Concept>

```

In questa maniera è possibile estendere o specializzare qualsiasi altra proprietà di SKOS e attraverso l'utilizzo di OWL, dichiarare significati particolari in maniera formale definendo delle restrizioni o delle proprietà particolari.

## CAPITOLO 3

### Altri modelli per la rappresentazione della conoscenza

Sebbene il soggetto principale di questa tesi sia SKOS, è opportuno citare per completezza alcuni tra i vari sistemi esistenti per l'organizzazione della conoscenza.

#### 3.1 Zthes

Zthes [ZTHES] descrive un modello astratto per la rappresentazione e la ricerca nei thesauri prendendo come riferimento quanto previsto dallo standard ISO 2788. L'idea fondamentale di questo progetto è di proporre un modello che permetta l'implementazione di thesauri a cui si possa accedere attraverso il protocollo Z39.50 e SRW (Search/Retrieve Web Service), quantunque il modello stesso sia sufficientemente generale da permettere implementazioni con altri protocolli e formato dei dati.

Lo standard Z39.50 è un protocollo inizialmente sviluppato all'interno del WAIS (Wide Area Information Service), software nato contemporaneamente al WWW e al Gopher come risposta al bisogno di organizzare l'informazione presente in rete e di renderla più facilmente reperibile, che in seguito ha continuato ad evolversi singolarmente specializzandosi nello sviluppo di applicazioni di information retrieval distribuite. E' uno standard sviluppato dall'organizzazione statunitense ANSI e diffusamente utilizzato dalla Library of Congress.

Scopo del protocollo Z39.50 è fornire un servizio per l'interrogazione di cataloghi o basi di dati fra loro eterogenee e consentire l'elaborazione di interfacce di interrogazione rivolte all'utente finale.

Lo standard di basa su un'architettura client-server. Nella maggior parte delle applicazioni, normalmente il server Z39.50 non contiene la base di dati, ma è in grado di accedervi a richiesta del client. In altre parole risiede ad un livello esterno e ha la funzione di filtrare le richieste che arrivano da client remoti e di indirizzarle al DBMS che esegue la ricerca ed invia la risposta al server Z39.50 che la reinoltrerà al client richiedente.

Alla risposta finale si può arrivare attraverso successive approssimazioni, in quanto successive richieste del client possono basarsi sulle precedenti risposte ottenute dal server. Per ciascuna richiesta sono creati dei gruppi ("set") di registrazioni che soddisfano alla richiesta e questi gruppi costituiscono la base per ulteriori elaborazioni.

La sua peculiarità sta nel fatto che permette di interrogare più archivi, uno per volta o contemporaneamente, attraverso una singola interfaccia e un singolo linguaggio di interrogazione, ottenendo risultati omogenei.

Sono ormai numerose le applicazioni dello standard in campo bibliografico. Quasi tutti i principali produttori di sistemi di automazione per biblioteca e distributori di informazione in linea offrono server Z39. 50.

L'adozione di questo protocollo risolverebbe, quindi, il problema della moltiplicazione delle interfacce, evitando, da un lato, all'utente il senso di smarrimento al momento della ricerca delle informazioni desiderate e, dall'altro, la crescita di mole di lavoro ai bibliotecari, interpellati sempre più frequentemente per illustrare il funzionamento dei programmi più svariati.

I servizi offerti da Zthes sono:

- ricerca di una rappresentazione all'interno del thesaurus di un certo termine o concetto
- ricerca di tutti i termini con significato più ampio (broader term) di un dato concetto
- ricerca di tutti i termini con significato più ristretto (narrower term) di un dato concetto
- ricerca del preferred term di un dato concetto
- ricerca dei non preferred terms di un dato concetto
- ricerca di termini equivalenti in altre lingue
- ricerca dei termini related di un dato concetto

Ecco un esempio di struttura di thesaurus estratto <http://zthes.z3950.org/xml/thatt.txt.gz>

```
<Zthes>
  <term>
    <termId>N0865</termId>
    <termName>THIRD WORLD</termName>
    <termType>ND</termType>
    <relation>
      <relationType>USE</relationType>
      <termId>R0352</termId>
      <termName>DEVELOPING COUNTRIES</termName>
    </relation>
  </term>

  <term>
    <termId>R0352</termId>
    <termName>DEVELOPING COUNTRIES</termName>
    <termType>PT</termType>
    <relation>
      <relationType>UF</relationType>
      <termId>N0886</termId>
      <termName>UNDERDEVELOPED COUNTRIES</termName>
    </relation>
    <relation>
      <relationType>RT</relationType>
      <termId>R0658</termId>
      <termName>INTERNATIONAL AID</termName>
    </relation>
  </term>

  <relation>
    <relationType>RT</relationType>
    <termId>387</termId>
    <termName>ECONOMIC DEVELOPMENT</termName>
  </relation>

  <term>
    <termId>N0886</termId>
    <termName>UNDERDEVELOPED COUNTRIES</termName>
    <termType>ND</termType>
    <relation>
      <relationType>USE</relationType>
      <termId>0352</termId>
```

```
<termName>UNDERDEVELOPED COUNTRIES</termName>
</relation>
</term>

<term>
<termId>387</termId>
<termName>ECONOMIC DEVELOPMENT</termName>
<termType>PT</termType>
<relation>
<relationType>NT</relationType>
<termId>R1855</termId>
<termName>SUSTINABLE DEVELOPMENT</termName>
</relation>
</term>

<term>
<termId>R1855</termId>
<termName> SUSTINABLE DEVELOPMENT </termName>
<termType>PT</termType>
</term>

<term>
<termId>R0658</termId>
<termName>INTERNATIONAL AID</termName>
<termType>PT</termType>
</term>
```

Ed ecco una rappresentazione grafica attraverso un grafo orientato. I concetti sono rappresentati attraverso delle ellissi gialle e le relazioni semantiche attraverso delle frecce in colore rosso. Come per SKOS, ad ogni concetto è assegnata un'etichetta utilizzata come termine da ricercare all'interno del database: il *termName* ed un *termId*, una stringa di caratteri che lo identifica univocamente all'interno del thesaurus.

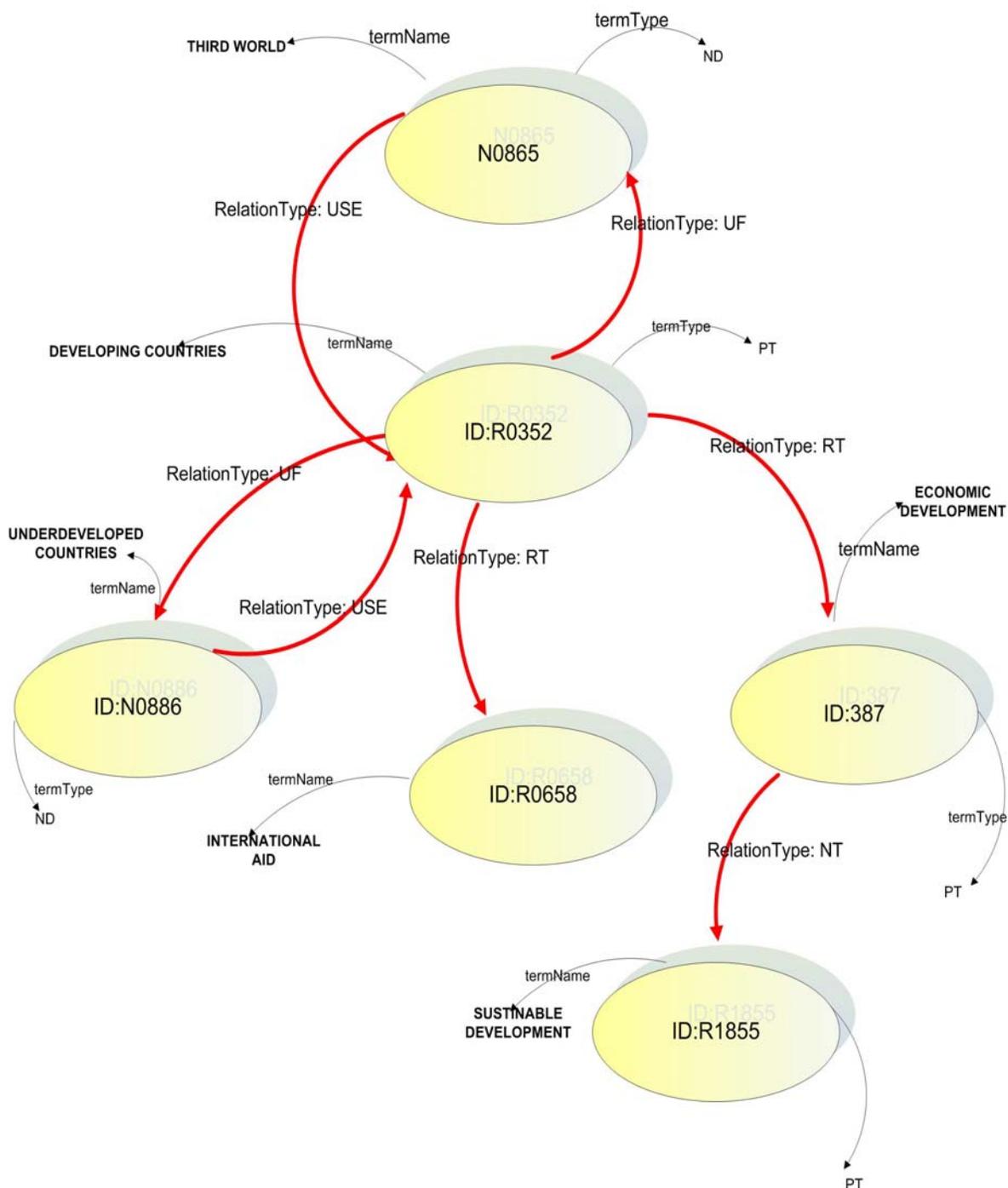


Fig. 23 Il grafo che modella l'esempio sopra descritto utilizzando Zthes

Per alcuni termini esiste inoltre un termType che ne indica il tipo. I tipi possono essere:

- “PT”: Preferred term
- “ND”: Non descriptor, cioè un non preferred term
- “NL”: Node Label cioè un termine “fantoccio” (dummy) utilizzato all’interno della gerarchia per un’organizzazione a faccette.

I tipi di relazione semantica sono sostanzialmente uguali a quelli di SKOS Core e possono assumere i seguenti valori:

- “NT”: Narrower term
- “BT”: Broader term
- “USE”: Use instead: il termine relativo dovrebbe essere usato in preferenza al corrente
- “UF”: Use for
- “RT”: Related term
- “LE”: Linguistic equivalent: utilizzato per esprimere il concetto in altri idiomi

Esistono inoltre degli altri elementi come in SKOS per la documentazione.

Vediamo un esempio di applicazione basata su Zthes: il thesaurus APAIS [APAIS] (Australian Public Affairs Information Service) raggiungibile all’indirizzo <http://www.nla.gov.au/apais/thesaurus/about.html>.

Questo thesaurus accessibile tramite un comune browser Web, contiene i termini utilizzati per indicizzare gli articoli di APAIS, una guida alla letteratura scientifica e umanistica. Il thesaurus è stato introdotto come database online nel 1978 e fu reso pubblico dalla National Library of Australia nel 1980.

Ecco la pagina principale di ricerca:

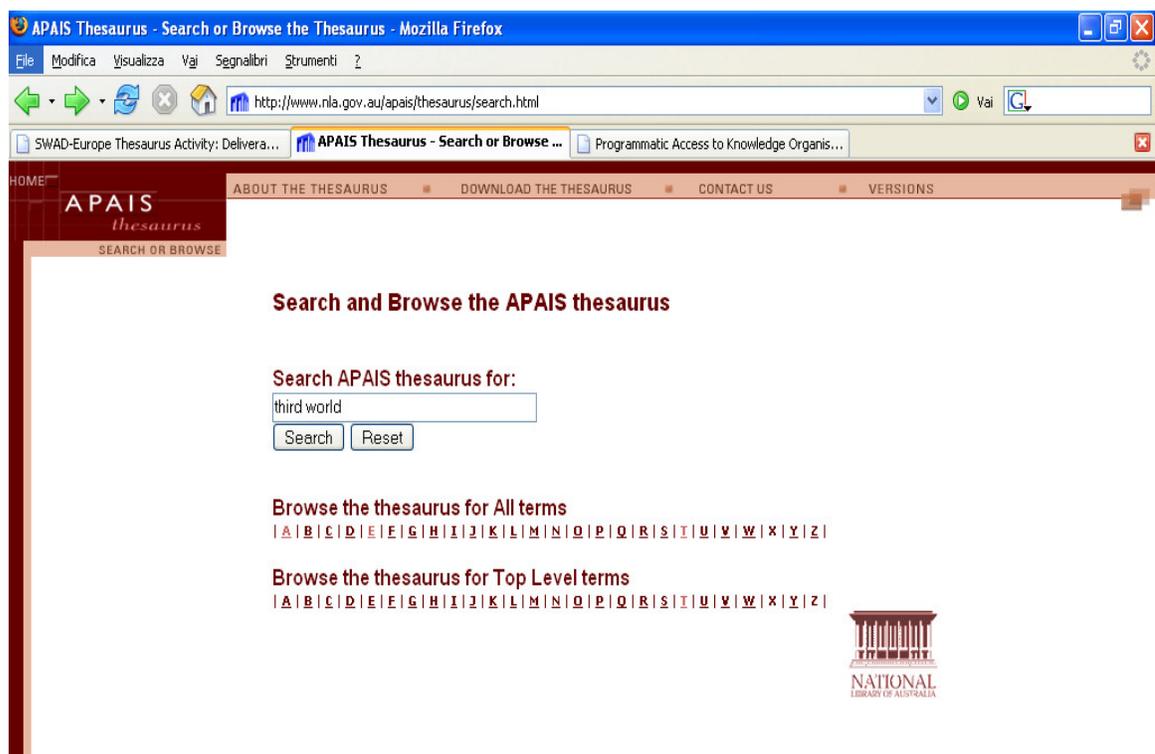


Fig 24 La pagina principale di ricerca del portale APAIS sviluppato con tecnologia Zthes

I risultati della ricerca sono visualizzati in ordine di rilevanza.

Results for 'third world' - Mozilla Firefox

File Modifica Visualizza Vai Segnalibri Strumenti ?

http://search.nla.gov.au/query.html?rq=1&qm=1&col=apais&style=apais&q=third+world&style=apais&submit=Search

SWAD-Europe Thesaurus Activity: Delivera... Results for 'third world' Programmatic Access to Knowledge Organism...

## APAIS Thesaurus Search

[Skip to content](#)

Results for: third world Document count: third (4) world (15) third world (2)

17 results found, sorted by relevance [score using date](#) [hide summaries](#) [group by location](#) 1-10 ▶

<b><u>Third world</u></b> ... Third world ... Third world ... <a href="http://www.nla.gov.au/apais/thesaurus/html/N0865.html">http://www.nla.gov.au/apais/thesaurus/html/N0865.html</a> - 2.7KB	89% ██████████ 13 Apr 05 <a href="#">Find Similar</a> <a href="#">Highlight</a>
<b><u>Developing countries</u></b> ... Third world ... <a href="http://www.nla.gov.au/apais/thesaurus/html/R0352.html">http://www.nla.gov.au/apais/thesaurus/html/R0352.html</a> - 3.0KB	69% ██████████ 13 Apr 05 <a href="#">Find Similar</a> <a href="#">Highlight</a>
<b><u>World War 2</u></b> ... World War 2 ... World War 2 ... <a href="http://www.nla.gov.au/apais/thesaurus/html/R1696.html">http://www.nla.gov.au/apais/thesaurus/html/R1696.html</a> - 2.9KB	45% ██████████ 13 Apr 05 <a href="#">Find Similar</a> <a href="#">Highlight</a>
<b><u>World War 1</u></b> ... World War 1 ... World War 1 ... <a href="http://www.nla.gov.au/apais/thesaurus/html/R1695.html">http://www.nla.gov.au/apais/thesaurus/html/R1695.html</a> - 2.9KB	45% ██████████ 13 Apr 05 <a href="#">Find Similar</a> <a href="#">Highlight</a>

Fig 25 Il risultato della ricerca per “third world”

Ecco il risultato della ricerca dell'esempio presentato: per brevità sono presentati solo i primi 2 nodi del grafo.

**Third world**  
Term: **Third world**  
Term ID: **N0865**  
Term Type: **ND (Non-preferred Term)**  
Term Notes: -  
Date Created: -  
Date Last Modified: **31/07/2000**  
**Other Terms**  
*Use instead:* [Developing countries](#)



**Developing countries**  
Term: **Developing countries**  
Term ID: **R0352**  
Term Type: **PT (Preferred Term)**  
Term Notes: -  
Date Created: **1979**  
Date Last Modified: **9/04/2002**  
**Other Terms**  
*Use For:* [Third world](#)  
[Underdeveloped countries](#)  
*Related Terms:* [Economic development](#)  
[International aid](#)



Fig 26 Visualizzazione dettagliata del termine "third world"

Fig 27 Il termine alternativo "developing countries"

Di seguito invece compare la finestra riguardante il termine Economic policy. È stato ritenuto opportuno presentare lo screen-shot di questa pagina web per evidenziare le varie relazioni presenti in questo termine.

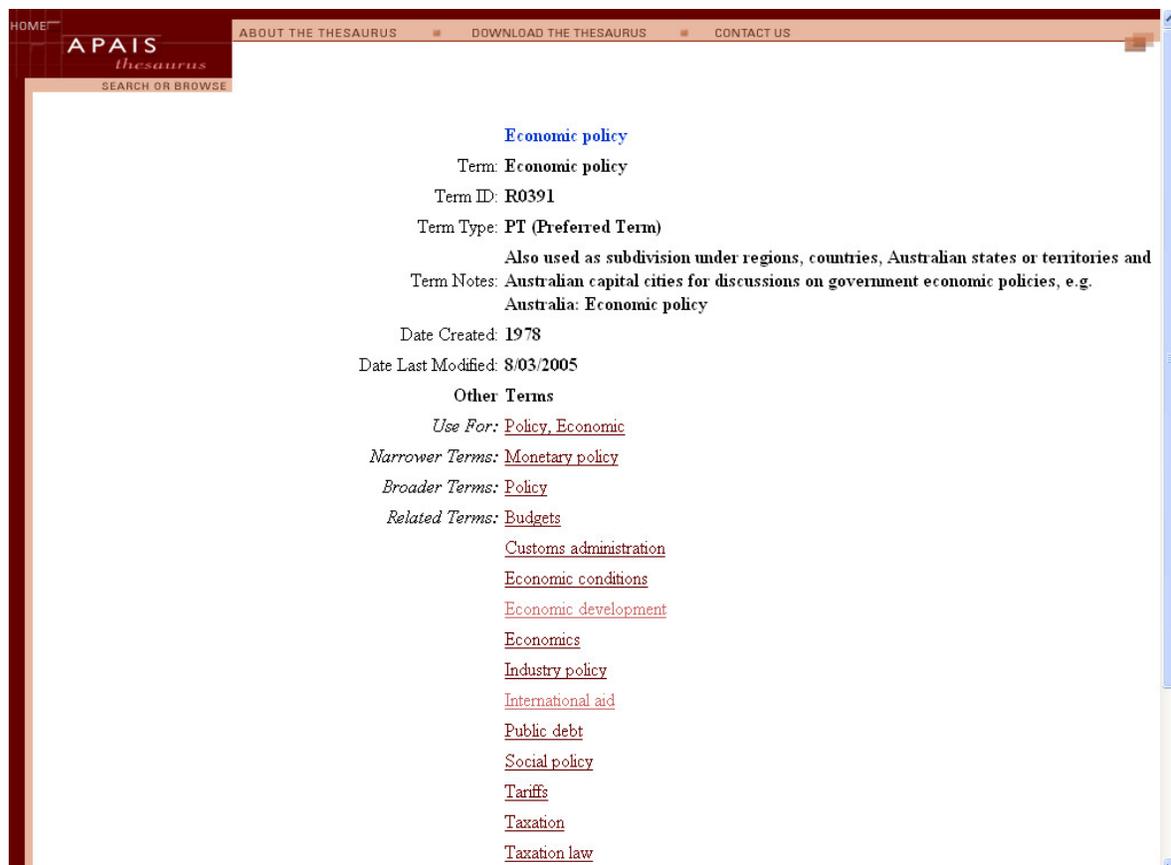


Fig 28 Visualizzazione del termine "Economy policy". Da notare la presenza delle etichette Narrower e Broader Terms

## 3.2 Topic Map

Le mappe tematiche [TAO, UTM, TMF] (formalmente chiamate topic navigation map) rappresentano un altro meccanismo per la rappresentazione della conoscenza divenute standard ISO 13250 nel 2000.

Una mappa tematica consiste in una collezione di temi (topic). Il concetto di topic è molto generale e simile a quello di **Concept** in SKOS. Viene definito come la rappresentazione all'interno di una mappa tematica di una qualsiasi cosa: una persona, un'entità, un concetto, su cui possano essere fatte delle asserzioni; sia esso un oggetto del mondo reale o una risorsa elettronica accessibile dal calcolatore. L'unica limitazione è che un topic può rappresentare uno ed un solo argomento. E' quindi possibile costruire una mappa a partire da qualsiasi tipo di informazione.

Consideriamo ad esempio la seguente frase: <<Il dipartimento di Informatica è situato in via Torino a Mestre>> e analizziamone le caratteristiche.

Gli argomenti e quindi i topic che si possono individuare sono tre: "Il dipartimento di Informatica", "via Torino" e "Mestre": questi sono i "*names*" o "*nomi*", le etichette da utilizzare come riferimento.

I nomi sono quindi delle etichette che permettono di identificare i topic all'interno della mappa. Queste etichette vengono usate dalle applicazioni per ordinare e presentare i topic. Un'etichetta può essere una qualsiasi stringa ma anche un'immagine e, per ogni topic, oltre al nome principale possono essere presenti una o più varianti.

I topic spesso vengono "tipizzati", ovvero classificati in tipi o categorie generali: i topic types. "Mestre" potrebbe appartenere al topic type di *città* mentre "via Torino" al topic type *rete viaria*.

Un topic può essere collegato con altre risorse informative di qualsiasi genere, in qualche modo rilevante o collegata a quell'argomento. Questo tipo di risorse vengono chiamate "*occurrences*" o "*occorrenze*". Un'occorrenza potrebbe essere il sito internet del dipartimento di informatica, una foto dell'edificio o un video.

Definiti gli argomenti si passa alle relazioni (association), che mettono in collegamento i vari topic tra loro.

Nell'esempio precedente, "è situato" rappresenta un'association.

Ogni topic che partecipa ad una association riveste un determinato "ruolo" o "role". Nel caso della relazione "il dipartimento di Informatica è situato a Mestre" espressa dall'associazione tra il dipartimento di Informatica e Mestre, i ruoli potrebbero essere "edificio" e "luogo". Quindi come per i topic anche gli association roles possono essere tipizzati e il tipo stesso rappresenta un topic.

Per verificare che a un *topic* corrisponda uno e un solo argomento è previsto l'utilizzo della *subject identity*. Se due *topic* hanno la stessa identità semantica allora fanno riferimento allo stesso soggetto, devono essere uniti e le rispettive caratteristiche, nomi, associazioni e occorrenze, vengono sommate tra loro. Ciò è molto utile quando si effettua il *merge*, la fusione tra due o più *topic map*, in modo da poter coprire un più vasto settore di informazioni, per non avere ridondanza nella definizione dei *topic*. Per questo motivo la *subject identity* deve essere una risorsa informativa identificabile in maniera univoca, solitamente un URI.

Vediamo una rappresentazione grafica dei concetti espressi finora:

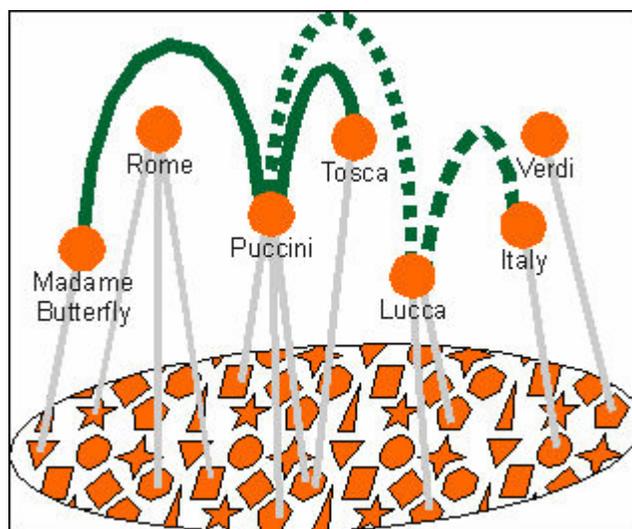


Fig. 29: Diagramma della struttura di una mappa tematica

Nella figura i *topic* sono i cerchi in alto contraddistinti da dei nomi e sovrapposti a un gruppo di risorse informative, disegnate in basso come cerchi, quadrati, stelle e triangoli racchiusi in un'ellisse. Le linee rette che collegano le figure geometriche in basso con i cerchi in alto di conseguenza rappresentano le occorrenze di un argomento, mentre le parabole tra i vari *topic* sono le relazioni esistenti tra loro.

Anche per le mappe tematiche, la necessità di una lingua franca per l'interscambio dei documenti tra i vari programmi, ha reso necessaria la creazione di uno standard conforme alle specifiche di XML.

Il progetto XTM (XLM Topic Maps) è iniziato nel 2000 da parte dell'organizzazione TopicMaps.org e nel 2001 è stata resa disponibile la prima versione, direttamente usabile in Internet e compatibile con lo standard ISO 13250 in grado di supportare ed essere supportato facilmente da un vasto numero di applicazioni.

L'esempio seguente mostrerà la rappresentazione di una semplice mappa tematica secondo lo standard XTM.

```
<?xml version="1.0"?>
<topicMap xmlns="http://www.topicmaps.org/xtm/1.0/"
  xmlns="http://www.w3.org/1999/xlink">

  <topic id="zzyz">
    <baseName>
      <baseNameString>Sony Computer</baseNameString>
    </baseName>
    <occurrence>
```

```
<resourceRef xlink:href="http://www.sony-computer.com/" />
</occurrence>
<occurrence>
  <resourceData>1977</resourceData>
</occurrence>
</topic>
</topicMap>
```

Nell'esempio notiamo che:

- La radice di una XML Topic Map è l'elemento <topicMap>
- L'elemento <topic> ha un attributo id. Attraverso questo attributo è possibile riferirsi in seguito al topic descritto.
- Il nome principale di un topic viene memorizzato in <baseName> che a sua volta si serve di <baseNameString>. Per gli eventuali altri nomi sono presenti <variant>, <variantName> e <parameters>.
- Il parametro <resourceRef> è un indicatore di argomento utilizzato per riferirsi ad altre risorse che si serve del parametro <occurrence>
- Con <resourceData> è possibile inserire descrizioni e informazioni testuali.

Così come per gli altri sistemi per l'organizzazione della conoscenza, lo scopo delle mappe tematiche è di rendere il Web più ordinato e di aiutare i programmi a ricercare informazioni più intelligenti.

In rete è possibile trovare diversi esempi che danno l'idea delle possibilità di utilizzo.

Un knowledge browser è Omnigator, realizzato da Ontopia con tecnologia J2EE.

Omnigator permette di navigare un qualsiasi file contenete una mappa tematica: uno degli esempi più completi proposti riguarda l'opera italiana.

Navigando *The Italian opera* con Omnigator nella versione XTM è possibile scegliere tra diverse modalità di visualizzazione. La prima, *ontology*, crea degli elenchi separati dividendo per *topic types*, *association types*, *association role types* e *occurrence types*. La successiva, *master index*, mostra tutti i *topic* in ordine alfabetico. Le due ultime opzioni, *index of individuals* e *index of themes*, mostrano rispettivamente un'organizzazione per *topic* "reali", ossia quelli che non rappresentano una categoria astratta, ma un argomento concreto, qualsiasi esso sia, mentre con l'ultima opzione, vengono visualizzati i motivi musicali, le arie. Indifferentemente dalla modalità scelta, per ogni *topic* è possibile vedere i nomi, le occorrenze interne ed esterne alla mappa e le associazioni esistenti.

Seguendo il percorso composer ->Puccini ->(Association) La rondine arriviamo alla seguente pagina:

The screenshot shows a web browser window with the Omnigator interface. The main content area features a thematic map of Italy with nodes for 'Contaxx' and 'Stanzichin'. Surrounding the map are several data panels:

- Names (2)**
  - La rondine
    - Rondine - Scope: Sort
  - The Swallow - Scope: English
    - Swallow - Scope: English; Sort
- Internal Occurrences (2)**
  - Première date
    - 1917-03-27
  - Audio recording
    - 5 56338 2
- External Occurrences (2)**
  - Illustration
    - file:///usr/local/oks-online-demo/jakarta-tomcat/webapps/omnigator/WEB-INF/topic
  - Web page
    - http://opera.stanford.edu/opera/Puccini/LaRondine/main.html - Scope: Online; Ope
- Associations (20)**
  - Composed by
    - Puccini, Giacomo
  - Contains aria
    - Chi il bel sogno di Doretta

Fig 30 Visualizzazione della pagina Web che illustra un esempio di mappa tematica

Il software nella sua ultima versione include Vizigator che permette una rappresentazione grafica della mappa tematica. È possibile quindi navigare all'interno della mappa cliccando sui vari nodi.

Ecco uno screen-shot dell'applet java:

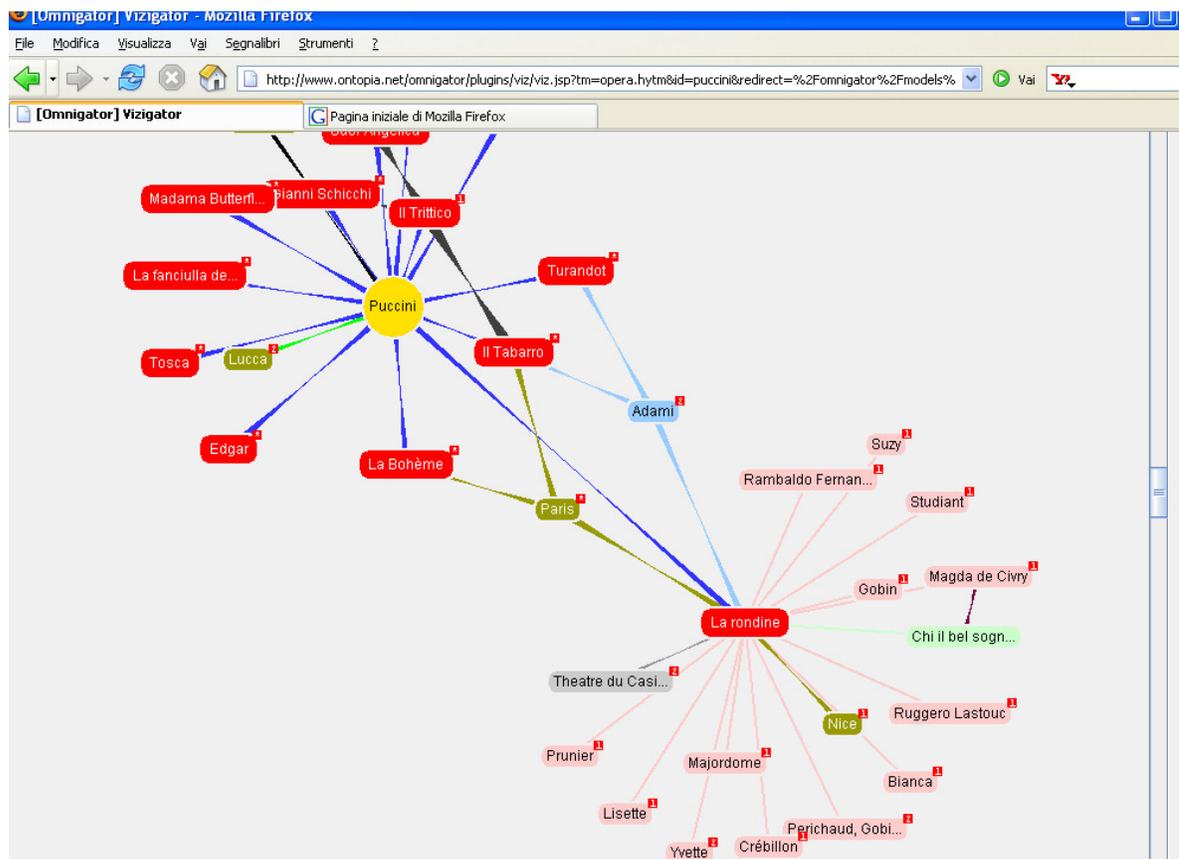


Fig 31 Il software Vizigator che permette una rappresentazione grafica di qualsiasi file contenente una mappa tematica

# CAPITOLO 4

## Applicazioni che utilizzano la tecnologia SKOS

### 4.1 GEMET

GEMET [GEMET] (General Multilingual Environmental Thesaurus) è un tesoro multilingua prodotto nell'ambito di un progetto finanziato dall'Agenzia Europea per l'ambiente (EIONET). Quest'organizzazione fornisce informazioni che sono di supporto alle decisioni che riguardano il miglioramento ambientale europeo.

Il thesaurus è disponibile in 16 diverse lingue tra cui italiano, tedesco, danese, inglese e americano: i file XML condividono la stessa struttura di markup e gli stessi nomi per gli elementi, cambia solo il contenuto secondo la lingua.

È possibile accedere al contenuto utilizzando un comune browser web e navigando attraverso delle comuni pagine html o altrimenti è disponibile per utilizzo da applicazioni remote come web service.

La classificazione ad alto livello del thesaurus si compone di 3 supergruppi, 35 Gruppi e 106 Top Terms, ovvero termini descrittivi di livello gerarchico più alto. L'insieme di tutti i termini è costituito da 6592 termini.

Ecco la pagina principale di navigazione del thesaurus accessibile all'indirizzo [http://www.eionet.eu.int/gemet/index\\_html?langcode=it](http://www.eionet.eu.int/gemet/index_html?langcode=it)

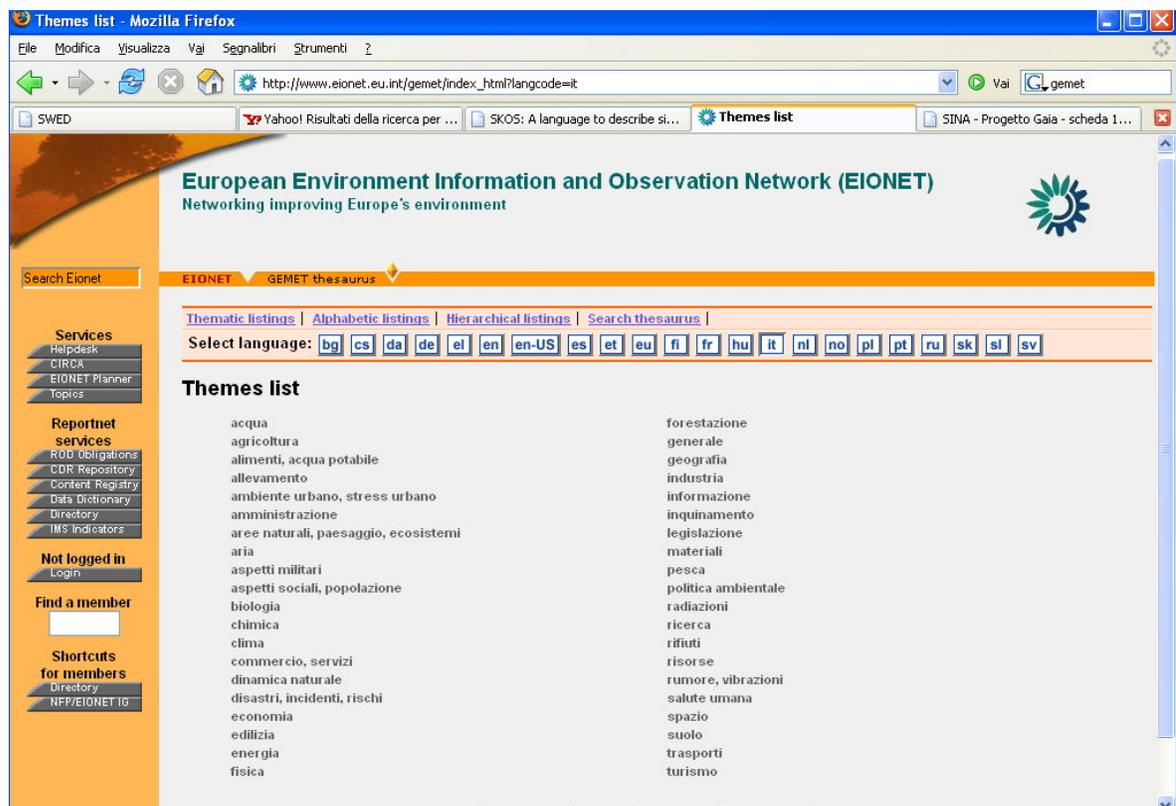


Fig.32 Pagina principale del portale GEMET

Il thesaurus è organizzato per:

- *Categorie tematiche*: agricoltura, allevamento, industria, pesca, ecc
- *Ordine gerarchico*: 4 grandi gruppi: aspetti sociali, ambiente naturale, attività e termini miscellanei
- *Ordine alfabetico*
- *Ricerca diretta*: digitando un preciso termine

Qui di seguito è presentata la pagina visualizzata cercando come topic il concetto di “abbattimento dell’inquinamento”.

Nella parte destra compaiono le varie traduzioni disponibili nelle altre lingue del termine, mentre a sinistra sono evidenziate le relazioni semantiche.

Oltre alle tradizionali relazioni presenti nei thesauri (relazioni di tipo broader o narrower), GEMET presenta altre caratteristiche addizionali, non standard: un termine può essere in relazione con un tema, un gruppo ed un supergruppo fornendo così “un'altra dimensione” da cui si può accedere al termine.

Nell'esempio che stiamo considerando i temi sono: chimica, inquinamento, politica ambientale, mentre i gruppi sono: rifiuti, inquinanti, inquinamento.

Sia i gruppi, i supergruppi che i temi hanno delle etichette che li rappresentano.

Vediamo ora come sono definite queste nuove classi:

```
@prefix gemet: <http://www.eionet.eu.int/gemet/schema#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .

gemet:Theme a rdfs:Class;
  rdfs:label 'Theme';
  rdfs:subClassOf skos:Concept.

gemet:Group a rdfs:Class;
  rdfs:label 'Group';
  rdfs:subClassOf skos:Concept.

gemet:SuperGroup a rdfs:Class;
  rdfs:label 'Super Group';
  rdfs:subClassOf skos:TopConcept.
```

La classe `gemet:SuperGroup` è definita come sottoclasse di `skos:TopConcept` in quanto, come abbiamo visto precedentemente, questa rappresenta il top level della gerarchia, mentre la classi `gemet:Theme` e `gemet:SuperGroup` vengono definite come sottoclasse di `skos:Concept`.

The screenshot shows the GEMET multilingual thesaurus interface. At the top, there is a language selection bar with buttons for various languages: bg, cs, da, de, el, en, en-US, es, et, eu, fi, fr, hu, it, nl, no, pl, pt, ru, sk, sl, sv. The main heading is "abbattimento dell'inquinamento". Below this, the "Concept definition:" is provided in English: "The reduction in degree or intensity of pollution in soil, rivers, lakes, seas, atmosphere, etc. (Source: PHC / LANDY)". A note indicates that the definition is not available for the current language. The interface is divided into three main sections: "broader terms", "narrower terms", and "related terms".

- broader terms:**
  - inquinamento
- narrower terms:**
  - abbattimento dell'inquinamento da petrolio
  - bacino di inquinamento
  - riduzione del rumore
  - riduzione delle emissioni
- related terms:**
  - apparecchiatura per il disinquinamento
  - messa al bando dei clorofluorocarburi e degli halons

On the right side, there is a list of translations for the concept in various languages:

- Български:** Намаляване на замърсяването
- Čeština:** omezování znečištění
- Dansk:** forureningsbekæmpelse
- Deutsch:** Bekämpfung der Verschmutzung
- Ελληνικά:** αντιρρυπαντική προστασία
- English:** pollution abatement
- English (US):** pollution abatement
- Español:** mitigación de la contaminación
- Eesti keel:** reostustõrje, saastetõrje
- Euskara:** poluzioaren arintze; kutsaduraren arintze
- Suomi:** saasteen vähentäminen
- Français:** réduction de la pollution
- Magyar:** szennyezés elhárítás
- Nederlands:** bestrijding van (milieu)verontreiniging
- Norsk:** forurensningsbekjemping
- Polski:** redukcja zanieczyszczeń
- Português:** diminuição progressiva da poluição
- Русский:** снижение уровня загрязнения
- Slovenčina:** redukcija znečistenia
- Slovenščina:** zmanjševanje onesnaženja
- Svenska:** föroreningsbekämpning; bekämpande av föroreningar

At the bottom left, there are navigation options: "Scope note: scope note is not available", "Themes: chimica, inquinamento, politica ambientale", and "Groups: RIFIUTI, INQUINANTI, INQUINAMENTO".

Fig. 33 Pagina con i risultati della ricerca per “abbattimento dell’inquinamento”

Siccome GEMET è un tesaurus multilingua, la codifica in RDF viene fatta in due parti: vi è una struttura portante, in RDF, che contiene degli identificatori numerici per i concetti a cui sono associati delle etichette per i gruppi e i concetti per ogni diversa lingua. In questo modo, sfruttando il vantaggio della struttura comune, non vi è la necessità che essa sia ripetuta per ogni diverso linguaggio ma semplicemente ricostruita.

Ecco qui di seguito un estratto della struttura fondamentale di GEMET espressa in RDF:

```

<rdf:RDF
  xmlns:gemet="http://www.eionet.eu.int/GEMET/skos-ext#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xml:base="http://www.eionet.eu.int/GEMET/" >
  <rdf:Description rdf:about="c_204">
    <skos:inScheme rdf:resource="../GEMET"/>
    <rdf:type
rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
    <skos:narrower rdf:resource="c_10217"/>
    <gemet:broaderTheme rdf:resource="t_23"/>
    <skos:broader rdf:resource="c_4648"/>
    <gemet:broaderTheme rdf:resource="t_2"/>
  </rdf:Description>
  <rdf:Description rdf:about="c_11786">
    <skos:broader rdf:resource="c_11124"/>
    <skos:inScheme rdf:resource="../GEMET"/>
    <gemet:broaderTheme rdf:resource="t_4"/>
    <rdf:type
rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>

```

```

</rdf:Description>
<rdf:Description rdf:about="c_7962">
  <skos:related rdf:resource="c_7969"/>
  <skos:inScheme rdf:resource="../GEMET"/>
  <gemet:broaderTheme rdf:resource="t_36"/>
  <rdf:type
rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
  <skos:narrower rdf:resource="c_7452"/>
  <gemet:broaderGroup rdf:resource="g_7956"/>
  <skos:related rdf:resource="c_7970"/>
</rdf:Description>
<rdf:Description rdf:about="g_14979">
  <gemet:subGroupOf rdf:resource="sg_5499"/>
  <rdf:type rdf:resource="skos-ext#Group"/>
  <skos:inScheme rdf:resource="../GEMET"/>
</rdf:Description>
<rdf:Description rdf:about="t_34">
  <rdf:type rdf:resource="skos-ext#Theme"/>
  <skos:inScheme rdf:resource="../GEMET"/>
</rdf:Description>
</rdf:RDF>

```

### Un esempio estratto dal thesaurus GEMET in portoghese

```

<rdf:RDF
  xmlns:gemet="http://www.eionet.eu.int/GEMET/skos-ext#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  xml:base="http://www.eionet.eu.int/GEMET/" >
  <rdf:Description rdf:about="c_204">
    <skos:prefLabel xml:lang="pt">paisagens agrícolas</skos:prefLabel>
  </rdf:Description>
  <rdf:Description rdf:about="c_11786">
    <skos:prefLabel xml:lang="pt">índices bióticos</skos:prefLabel>
  </rdf:Description>
  <rdf:Description rdf:about="c_4657">
    <skos:prefLabel xml:lang="pt">ecologia paisagística</skos:prefLabel>
  </rdf:Description>
  <rdf:Description rdf:about="g_8575">
    <skos:prefLabel xml:lang="pt">COMÉRCIO, SERVIÇOS</skos:prefLabel>
  </rdf:Description>
  <rdf:Description rdf:about="t_29">
    <gemet:acronymLabel>REC</gemet:acronymLabel>
    <skos:prefLabel xml:lang="pt">turismo</skos:prefLabel>
  </rdf:Description>
</rdf:RDF>

```

Si noti come nella codifica RDF compare solo l'identificativo del termine e l'etichetta nella lingua portoghese.

## 4.2 UKAT: UK Archival Thesaurus

Un altro rappresentativo esempio di thesaurus, costruito utilizzando il framework di SKOS, è UKAT [UKAT] raggiungibile all'indirizzo web <http://www.ukat.org.uk>.

Lo scopo di questo progetto era quello di creare un vocabolario controllato a cui potesse contribuire potenzialmente chiunque avesse determinate conoscenze di un certo argomento o disciplina.

Come punto di partenza per la strutturazione del thesaurus è stato preso UNESCO Thesaurus a cui sono stati aggiunti altri termini provenienti da altri vocabolari.

UKAT, è suddiviso in 7 aree di conoscenza:

1. Education
2. Science
3. Culture
4. Social and human science
5. Information and communication
6. Politics, law and economics
7. Events

Ognuna di queste aree è suddivisa ulteriormente in microthesauri. Questi ultimi contengono un insieme di termini in relazione tra loro da un punto di vista concettuale.

Ad esempio nell'area "Science" compaiono i seguenti microthesauri:

### Main thesaurus: browse by hierarchy

#### 2. Science

- [2.05 Scientific approach](#)
- [2.10 Science and research management](#)
- [2.15 Mathematics and statistics](#)
- [2.20 Physical sciences](#)
- [2.25 Chemical sciences](#)
- [2.30 Space sciences](#)
- [2.35 Earth sciences](#)
- [2.40 Geography and oceanography](#)
- [2.45 Meteorology](#)
- [2.50 Hydrology](#)
- [2.55 Environmental sciences and engineering](#)
- [2.60 Pollution, disasters and safety](#)
- [2.65 Natural resources](#)
- [2.70 Biology](#)
- [2.75 Natural sciences](#)
- [2.80 Medical sciences](#)
- [2.85 Pathology](#)

[Main thesaurus: A-Z listing](#) | [Search](#) | [Need some help?](#)

Fig. 34 Microthesauri della categoria Science

Ogni microthesaurus è preceduto da un tre cifre. Il numero prima del punto indica l'area principale di conoscenza. Scegliendo ad esempio il terzo elemento otteniamo la seguente visualizzazione:

## Mathematics

**Source:** UNESCO

**Status:** Approved

### Used For

[Mathematical sciences](#)

### Microthesaurus:

[2.15 Mathematics and statistics](#)

#### Narrower terms

[Algebra](#)

[Arithmetic](#)

[Combinatorial mathematics](#)

[Equations](#)

[Geometry](#)

[Logarithms](#)

[Mathematical analysis](#)

[Mathematical logic](#)

[Number theory](#)

[Set theory](#)

[Topology](#)

[Trigonometry](#)

#### Related terms

[Mathematical instruments](#)

[Mathematical physics](#)

[Mathematicians](#)

[Statistics](#)

Fig. 35 Termini relativi alla voce "mathematics"

Sono evidenziati i termini con significato più ampio o termini che hanno una certa attinenza (related).

UKAT mette a disposizione inoltre un form per effettuare delle ricerche avanzate all'interno del thesaurus:

## Advanced search

Search for term

Matching

*To perform advanced searches select the filters and your search options below.*

	Filter by	Value	Matching
<input type="button" value="and"/>	Microthesaurus	select field of knowledge <input type="text"/> select microthesarus <input type="text"/>	
<input type="button" value="and"/>	Source	select source <input type="text"/> Other <input type="text"/>	
<input type="button" value="and"/>	Scope Notes	<input type="text"/>	Exact match <input type="text"/>
<input type="button" value="and"/>	Status	select status <input type="text"/>	
<b>Number of results</b>		10 <input type="text"/> per page	<input type="button" value="Search"/>

[Reset values](#)

Fig. 36 Form di ricerca avanzato di UKAT

Attraverso questo form è possibile effettuare delle ricerche mirate in maniera da visualizzare solo determinati risultati, specificando ad esempio quale sorgente di conoscenza utilizzare.

Qui di seguito è visualizzato un estratto del thesaurus UKAT riguardante il concetto “Mathematics”. Notiamo come compaiono i riferimenti ai concetti identificati numericamente riguardanti le relazioni semantiche.

```
<skos:Concept rdf:about="http://www.ukat.org.uk/thesaurus/concept/2079">
  <skos:prefLabel>Mathematics</skos:prefLabel>
  <skos:altLabel>Mathematical sciences</skos:altLabel>
  <skos:inScheme rdf:resource="http://www.ukat.org.uk/thesaurus"/>
  <skos:inScheme
rdf:resource="http://www.ukat.org.uk/thesaurus/micro/215"/>
  <skos:narrower
rdf:resource="http://www.ukat.org.uk/thesaurus/concept/197"/>
  <skos:narrower
```

```
rdf:resource="http://www.ukat.org.uk/thesaurus/concept/342"/>
  <skos:narrower
rdf:resource="http://www.ukat.org.uk/thesaurus/concept/986"/>
  <skos:narrower
rdf:resource="http://www.ukat.org.uk/thesaurus/concept/2116"/>
  <skos:narrower
rdf:resource="http://www.ukat.org.uk/thesaurus/concept/2518"/>
  <skos:narrower
rdf:resource="http://www.ukat.org.uk/thesaurus/concept/8532"/>
  <skos:narrower
rdf:resource="http://www.ukat.org.uk/thesaurus/concept/6168"/>
  <skos:narrower
rdf:resource="http://www.ukat.org.uk/thesaurus/concept/523"/>
  <skos:narrower
rdf:resource="http://www.ukat.org.uk/thesaurus/concept/4962"/>
  <skos:narrower
rdf:resource="http://www.ukat.org.uk/thesaurus/concept/2077"/>
  <skos:narrower
rdf:resource="http://www.ukat.org.uk/thesaurus/concept/4127"/>
  <skos:narrower
rdf:resource="http://www.ukat.org.uk/thesaurus/concept/9307"/>
  <skos:related
rdf:resource="http://www.ukat.org.uk/thesaurus/concept/19021"/>
  <skos:related
rdf:resource="http://www.ukat.org.uk/thesaurus/concept/13478"/>
  <skos:related
rdf:resource="http://www.ukat.org.uk/thesaurus/concept/3698"/>
  <skos:related
rdf:resource="http://www.ukat.org.uk/thesaurus/concept/1357"/>
  <rdfs:seeAlso
rdf:resource="http://www.ukat.org.uk/thesaurus/term.php?i=2079"/>
  </skos:Concept>
```

Il thesaurus è disponibile liberamente per il download e quindi può essere utilizzato per creare portali di informazione come quello che verrà illustrato nel prossimo paragrafo.

## 4.3 Il portale SWED

SKOS è utilizzato anche nell'ambito di SWED [SWED] (Semantic Web Environmental Directory). Lo scopo è quello di creare una directory di organizzazioni che si occupano di ambiente all'interno del Regno Unito.

Essenzialmente SWED fornisce una vista globale sui dati, forniti dalle singole organizzazioni effettuando un' "operazione di raccolta" tra i diversi siti presenti sul Web. I vari membri delle organizzazioni quindi, possiedono, controllano e pubblicano le loro informazioni. Queste informazioni sono raccolte, confrontate e pubblicate sul portale.

In questo modo, con una così grande ricchezza di informazioni ben indicizzate è molto più semplice, rispetto alle tradizionali directory, cercare e trovare informazioni rilevanti al nostro scopo.

Oltre al prototipo di portale raggiungibile all'indirizzo

<http://www.swed.org.uk/swed/index.html>, il progetto comprende anche del software scaricabile gratuitamente che permette di creare il proprio portale personalizzato.

Qui di seguito viene mostrata la pagina principale di SWED.

The screenshot shows the main page of the SWED portal. At the top, there is a green banner with the 'swed semantic web environmental directory' logo. Below this is a search bar with a 'go' button and radio buttons for 'all items' and 'within current results'. A 'help?' link is also visible. On the left side, there are navigation links for 'Home', 'Administration', and 'Recently visited:'. The main content area is titled 'Browse' and contains a descriptive paragraph: 'Organisations and projects are categorised using six characteristics or facets which are shown below. The numbers in the brackets show how many results are in that facet. To start browsing pick an option from one of the facets. Click + to see full lists of options for each facet.' There are six facet boxes, each with a '+' icon and a 'What is this facet?' link. The facets are: 1. Topic of interest (11), 2. Organisation type (3), 3. Activity (6), 4. Project type (1), 5. Operational area (75), and 6. Name (5). Each facet lists several sub-categories with their respective counts in parentheses and a 'More options' link. The footer contains the SWAD Europe logo and a 'Terms of use' link.

Fig.37 La pagina principale del portale

Si vede subito da questa pagina come i dati siano organizzati attraverso sei faccette. I numeri all'interno delle parentesi indicano la cardinalità di quella determinata faccetta.

Il vantaggio derivante da questa organizzazione, è l'utilizzo di un insieme di metadati mutuamente esclusivi rappresentanti ciascuno un aspetto o proprietà capaci di descrivere l'oggetto stesso. In questo modo ogni oggetto è classificato secondo una pluralità di attributi e gode di un meccanismo di pluridimensionalità[MV04]. Non è necessario

conoscere il nome della categoria in cui l'oggetto è collocato per poterlo ritrovare e il criterio di ordinamento non è necessariamente gerarchico.

I vocabolari utilizzati per organizzare i dati secondo diverse gerarchie sono pubblicati utilizzando OWL per le ontologie e SKOS per i thesauri. Nella pagina principale troviamo gli argomenti suddivisi in diverse categorie che facilitano la ricerca fungendo da filtro per le informazioni che non interessano.

Selezionando ad esempio dalla categoria “Topic Of Interest”, il link “Animal Welfare” otteniamo il seguente risultato:

The screenshot shows the SWED search results page. The header includes the SWED logo and the text 'semantic web environmental directory'. Below the header is a search bar with a 'go' button and a filter section with 'all items' selected. The main content area is titled 'Current Search Results' and shows 'Topic of interest > Animal Welfare [X]'. It lists 11 results, with the first three visible: 'Advocates for Animals (AFA)', 'Barn Owl Trust', and 'Humane Slaughter Association (HSA)'. Each result includes a brief description and its 'Type of Organisation'.

Fig. 38 Risultato della query per “animal welfare”

Otteniamo cioè un elenco delle organizzazioni che come scopo hanno la tutela e la protezione di determinate specie animali.

SKOS in questo ambito, viene impiegato per costruire una tassonomia utilizzata rispettivamente per classificare i tipi di organizzazione esistenti ed un thesaurus utilizzato per classificare i “topic of interest” delle diverse organizzazioni.

Qui di seguito è visualizzata un'istantanea della parte di thesaurus che definisce il topic “animal welfare” e il topic of interest “pollution\_control\_remediation”.

```
@prefix rdfs:      <http://www.w3.org/2000/01/rdf-schema#> .
@prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix :         <http://www.swed.org.uk/2004/06/swed_toi#> .
@prefix owl:    <http://www.w3.org/2002/07/owl#> .
@prefix skos:     <http://www.w3c.rl.ac.uk/2003/11/21-skos-core#> .

:animal_welfare
  a      skos:Concept ;
```

```

    rdfs:comment "" ;
    rdfs:label "Animal Welfare"@en ;
    skos:broader :topics_of_interest ;
    skos:externalID "0AW0" ;
    skos:prefLabel "Animal Welfare"@en .

:pollution_control_remediation
    a          skos:Concept ;
    rdfs:comment "" ;
    rdfs:label "Pollution (inc. Control and Remediation)"@en ;
    skos:broader :topics_of_interest ;
    skos:externalID "0P0" ;
    skos:prefLabel "Pollution (inc. Control and Remediation)"@en .

```

Qui invece viene mostrato un estratto della tassonomia utilizzata per classificare i diversi tipi di organizzazione.

```

@prefix rdfs:    <http://www.w3.org/2000/01/rdf-schema#> .
@prefix rdf:    <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix :      <http://www.swed.org.uk/2004/02/swed/org_type#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix skos:  <http://www.w3c.rl.ac.uk/2003/11/21-skos-core#> .

:funding_organisation
    a          skos:Concept ;
    rdfs:comment "Organisations that provide funding for other
organisation for projects" ;
    rdfs:label "Funding"@en ;
    skos:broader :organisation ;
    skos:prefLabel "Funding"@en .

:uk_national_government_organisation
    a          skos:Concept ;
    rdfs:comment "Temporary documentation in the place of a full scope
note for this term" ;
    rdfs:label "UK National Government"@en ;
    skos:broader :public_sector_organisation ;
    skos:prefLabel "UK National Government"@en .

```

Questa applicazione quindi, illustra e dimostra, come i thesauri possono essere utilizzati nell'ambito del Semantic Web, in maniera da poter essere condivisi e consultati per catalogare risorse e per essere usati in combinazioni con i motori di ricerca.

## 4.4 Un'applicazione con Web Service: CSA-NBII Biocomplexity Thesaurus

Uno degli sforzi del gruppo di ricerca che sta sviluppando SKOS è rivolto all'utilizzo dei thesauri nell'ambito del Web Semantico per poter essere utilizzati per la catalogazione, il reperimento delle informazioni e l'utilizzo in congiunzione ai motori di ricerca.

A questo scopo sono state sviluppate delle API in Java per permettere la pubblicazione e la ricerca dei concetti all'interno dei thesauri.

Queste API sono state implementate attraverso un'interfaccia web service che fornisce un insieme di metodi per accedere e interrogare thesaurus basati su SKOS Core Schema utilizzando i meccanismi e formati standard SOAP e WSDL.

Applicazioni software, scritte in diversi linguaggi di programmazione e implementate su diverse piattaforme hardware, possono beneficiare dell'interoperabilità fra diversi software e accedere in questo modo a "fonti di sapere condiviso" permettendo un accesso alle informazioni più razionale ed efficiente nell'ottica di un Web Semantico.

Ecco alcuni esempi di metodi del package **org.w3.y2001.sw.Europe.skos**

getConcept (uri)

getConceptsMatchingKeyword/Regex (string)

getAllConceptRelatives (concept)

getSupportedSemanticRelations

getAllConceptRelatives (concept, relation)

getAllConceptsByPath (concept, relation, distance)

Questi sono alcuni dei metodi dell'interfaccia SKOSThesaurus che provvede a generare la descrizione del Web Service in WSDL.

Altre classi importanti presenti nel package sono:

- La classe Concept: permette di istanziare i concetti definendo una preferred label, degli esempi, un URI ecc.
- La classe Relation: permette di definire le relazioni semantiche
- La classe ConceptRelatives: permette di relazionare i concetti tra loro definendo il tipo di relazione tra esistente tra loro e la distanza.

Un esempio di questa applicazione è raggiungibile all'indirizzo

<http://thesaurus.nbio.gov/SearchNBIIThesaurus/> in cui è possibile ricercare termini presenti nel thesaurus del portale NBII (National Biological Information Infrastructure).

Questa qui sotto è la pagina principale di ricerca: cliccando su check thesaurus si dà via alla ricerca e successivamente vengono presentati i termini trovati con indicati a fianco le relazioni semantiche.

NBII Home

- About NBII
- Current Biological Issues
- Biological Disciplines
- Geographic Perspectives
- Teacher Resources
- Data & Information Resources
- Search
- Contact Us

## Browse the CSA-NBII Biocomplexity Thesaurus

The Biocomplexity Thesaurus was developed through a partnership between the NBII and CSA, a leading bibliographic database provider.

Please enter a term to search in the CSA-NBII Thesaurus

[Check Thesaurus](#)

Results:

- [Food organisms](#) BT [Organisms](#)
- NT [Food plants](#)
- RT [Bioactive compounds](#)
- RT [Diets](#)
- RT [Food](#)
- RT [Food availability](#)
- RT [Food chains](#)
- RT [Food fishes](#)
- RT [Host plants](#)
- RT [Phytoplankton](#)
- RT [Prey](#)
- RT [Zooplankton](#)
- SC ASF
- SC LSC

Fig. 39 La pagina principale di ricerca del portale NII

Nello stesso portale è disponibile inoltre una dimostrazione del thesaurus come web service. Attraverso questa applicazione è possibile non solo ricercare un termine, ma specificare il tipo di relazione semantica e cercare ad esempio solo i termini con significato più ampio.

Ecco come si presenta il prototipo della pagina di ricerca:



[Main >>](#)

## NBII Thesaurus Web Service Prototype - Using SKOS

Enter a Term:

Choose a Term Group:

- All
- BT
- NT
- RT

Fig.40 Esempio di applicazione web service

Ed ecco i risultati della ricerca del termine “plant”

Results for term(s) containing Plants	
Term Group	Terms
<b>Plants</b>	
Broader Terms:	Organisms
Narrower Terms:	Alfalfa , Algae , Crops , Legumes , Aquatic plants , Cryptogams , Chlorophyta , Desmids , Diatoms , Endolithic algae , Heterocysts , Phycobionts , Phytoplankton , Seaweeds , Zooxanthellae , Angiosperms , Arecaceae , Grasses , Fruit bodies , Aquatic hyphomycetes , Periphyton , Fungi , Freshwater plants , Macrophytes , Seagrasses , Beans , Bryophyta , Cash crops , Chamaephytes , Conifers , Coprophilous fungi , Gymnosperms , Trees , Grains , Industrial crops , Tree crops , Cultivars , Cultivated plants , Cycads , Decay fungi , Deciduous trees , Ferns , Lichens , Desert plants , Endangered plants , Endemic plants , Endophytes , Entomopathogenic fungi , Epiphytes , Extinct plants , Foliage , Food plants , Forbs , Fruit trees , Fungus balls , Fungus gardens , Herbs , Mesophilic fungi , Molds , Mycelia , Mycobionts , Nematophagous fungi , Phytopathogenic fungi , Xerophilic fungi , Yeasts , Gametophytes , Ginkgoes , Gnetophyta , Halophytes , Hardwoods , Horsetails , Host plants , Indigenous plants , Introduced plants , Invasive plants , Weeds , Lianas , Lycopodiophyta , Mangroves , Tropical plants , Medicinal plants , Ornamental plants , Plantlets , Protected plants , Sclerophylls , Seedlings , Sporophytes , Tetrasporophytes , Transgenic plants , Vines , Wildflowers , Woody plants , Xerophytes , Shrubs , Softwoods
Related Terms:	Planting , Horticulture , Apical dominance , Arboreta , Fungi , Plant structures , Biogeography , Botany , Plant biomass , Botanical gardens , Plant breeding , Cardiac glycosides , Plant cells , Cotton , Leaves , Plant debris , Plant diseases , Plant communities , Endosperm , Vegetation , Epinasty , Phytoestrogens , Plant extracts , Sap , Plant production , Hypocotyls , Kelp beds , Plant preparations , Plant resources , Parenchyma (plants) , Phloem , Vascular system (plant anatomy) , Pith , Plant diversity , Plant physiology , Plant populations , Plant protection , Plant species reintroduction , Plant viruses , Rare species , Veins (plant anatomy) , Xylem
Subject Category:	LSC , POL , ASF
Scope Note:	Use of a more specific term is recommended.
Use For:	Vascular plants
Use:	No use terms

Fig. 41 Visualizzazione dei termini trovati per “plants”

## CAPITOLO 5

### CONCLUSIONI

Lo sviluppo del Web Semantico coinvolge non solo diverse aree del World Wide Web come web service, motori di ricerca, ontologie, ma anche ambiti più diversi come biblioteconomia e altre discipline che si occupano da centinaia di anni dell'organizzazione della conoscenza.

Le raccolte di conoscenza costituiscono un componente fondamentale nell'infrastruttura dell'informazione nella vita moderna. Essa fornisce una risorsa essenziale per le persone e per gli specialisti, utilizzata per riferimenti e per la ricerca.

Ad ogni modo, l'ambiente dell'organizzazione della conoscenza sta cambiando. I costi per l'acquisto e l'archiviazione di libri e giornali aumentano e lo spazio richiesto per contenere fisicamente collezioni cartacee di documenti sta diventando così oneroso che solo poche istituzioni pubbliche possono permetterselo.

Allo stesso tempo, l'uso di mezzi digitali ha favorito la produzione di materiale in forma elettronica economica e con minori vincoli di spazio richiesto. Inoltre il WWW e la banda larga hanno rivoluzionato la distribuzione e l'accesso alle risorse informative da ogni parte del mondo.

Tutti questi aspetti hanno favorito la nascita di un'organizzazione digitale della conoscenza, della nascita di biblioteche digitali.

Il ruolo rimane lo stesso, ma il modo di accedervi sarà diverso. Come detto all'inizio il maggior problema con cui ci si scontra è trovare il materiale cercato tra una grande mole di informazioni disponibili, evidenziare all'utente le informazioni più rilevanti e fornirgli degli strumenti per affinare la sua ricerca.

SKOS ha dato prova di essere un semplice ma potente strumento per presentare e condividere raccolte di termini.

C'è certamente ancora diversa strada da percorrere affinché esso diventi uno standard robusto e utilizzabile. Attualmente ha si è appena conclusa la seconda revisione del Working Draft da parte del W3C. Una terza revisione è prevista per fine Febbraio inizio Marzo 2006. E' necessario che siano costruiti tools per lo sviluppo, la ricerca, il mantenimento e la condivisione dei thesauri. Alcuni di questi sono già disponibili come le API Java, SKOS Validator (un'applicazione Java che permette di controllare la sintassi RDF e la costruzione dello schema), portalCore una libreria contenente sorgenti per costruire e configurare un piccolo portale (disponibile sul sito [www.swed.org.uk](http://www.swed.org.uk)).

SKOS intende essere un complemento rispetto a strumenti e tecnologie per esprimere ontologie come OWL (Web Ontology Language). È più flessibile, più semplice da utilizzare e più adatto ad esprimere sistemi per l'organizzazione della conoscenza. OWL è sicuramente un linguaggio più potente, ma questo non significa necessariamente migliore, i due oggetti si complementano a vicenda.

SKOS è basato fortemente su un modello C-O (Concept Oriented). Secondo questo modello un thesaurus consiste di un insieme di concetti. Ogni concetto ha una sola preferred label e un numero qualsiasi di etichette alternative.

I diversi concetti all'interno del thesaurus possono essere relazionati tra loro attraverso le relazioni semantiche (broader, narrower, related). Questo tipo di approccio Concept Oriented permette una maggiore facilità nel mantenimento, riduce l'ambiguità e la confusione.

Un'area di applicazione particolarmente promettente riguarda il supporto alle cosiddette "folksonomies"; una sorta di sforzo collettivo da parte di comunità di utenti volto a condividere e a classificare risorse di qualsiasi tipo mediante l'utilizzo di metadati.

SKOS si presenta quindi come una tecnologia emergente in grado di combinare la potenza delle strutture per la rappresentazione della conoscenza con la flessibilità e l'estensibilità del Web Semantico.

A parere dello scrivente, SKOS è un framework molto utile ed interessante per collegare strumenti di indicizzazione semantica tradizionali, che sono un patrimonio di tecniche molto sofisticato e valido sviluppato nel corso di oltre un secolo di ricerche, con le recenti tecnologie telematiche ed in particolare il Web.

Spesso purtroppo i due mondi si ignorano a vicenda, con grande spreco di risorse e reinvenzioni di cose già esistenti, solo con nomi diversi magari perché più commerciali.

La domanda che sorge spontanea è: se la biblioteconomia ha studiato per secoli i sistemi per l'organizzazione della conoscenza e ha sviluppato sistemi di classificazione impiegati in tutto il mondo perché non utilizzarli anche per il Web invece di ripartire da zero?

Al di là dell'aspetto che queste tecnologie potranno avere nel Web Semantico, i vantaggi derivanti da un utilizzo dei sistemi per l'organizzazione della conoscenza anche nel web attuale sono indubbi.

L'uso di uno schema di classificazione aiuterebbe sicuramente gli utenti ad orientarsi in quella che spesso è la natura caotica del Web, fornendo un modello facilmente comprensibile dall'utente, senza sacrificare l'integrità e a beneficio dell'usabilità e del reperimento delle informazioni.

Classificare non è un compito semplice. È necessario quindi che il progettista e il content manager attingano al patrimonio della biblioteconomia, che ha già centinaia di anni alle spalle di esperienza, per cercare di impiegare sistemi di navigazione migliori.

SKOS in questo punto, si pone come tecnologia da affiancare ai tradizionali sistemi per l'organizzazione della conoscenza, in grado di migliorare l'organizzazione e il successivo reperimento delle informazioni.

# BIBLIOGRAFIA

## Articoli

[MMW04] Alistair Miles, Brian Matthews, Michael Wilson, Dan Brickley, Dave Beckett, Nikki Rogers – SKOS Core: A language to describe simple knowledge structures for the web, 2004

<http://idealliance.org/proceedings/xtech05/papers/03-04-01/>

[BHL01] Tim Berners-Lee, James Hendler and Ora Lassila - The Semantic Web, 2001

[http://www.sciam.com/print\\_version.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21](http://www.sciam.com/print_version.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21)

[BS98S] Tim Berners-Lee –Semantic Web roadmap, 1998

<http://www.w3.org/DesignIssues/Semantic.html>

[B01] Tim Bray –What is RDF, 2001

<http://xml.com/lpt/a/2001/01/24/rdf.htm>

[C03] Paolo Ceravolo, - I linguaggi del Web Semantico, 2003

[http://pro.html.it/articoli/id\\_334/idcat\\_46/pag\\_1/pag.html](http://pro.html.it/articoli/id_334/idcat_46/pag_1/pag.html)

[F05] Elena Franchini, Nuove prospettive nell'evoluzione dei thesauri: interoperabilità, integrazione, standard, 2005

[VM04] Vittorio Marino, Classificazioni per il Web – I vantaggi dell'adozione di schemi a faccette –

<http://www.aib.it/aib/contr/marino1.htm>

[M05] Peter Mikhaleiko, - Introducing SKOS, 2005

<http://www.xml.com/pub/a/2005/06/22/skos.html>

[SE05] Francesca Severino -What thesaurus to define EU/ACP relations? ISKO Italy- UniMIB meeting : Milan : June 24, 2005

[S02] Oreste Signore – RDF per la rappresentazione della conoscenza  
[www.w3c.it/papers/RDF.pdf](http://www.w3c.it/papers/RDF.pdf)

[S03] Oreste Signore - Strutturare la conoscenza: XML, RDF, Semantic Web, 2003

<http://dit.unitn.it/~bouquet/kr-2005/materiale/ck2003.pdf>

[S05] S. Spinelli, - Introduzione ai thesauri, 2005

<http://mail.biocfarm.unibo.it/~spinelli/indicizzazione/thesauri.htm>

[T05] Joseph T.Tennis – SKOS and the Ontogenesis of Vocabularies, 2005 DC-2005:  
"Vocabularies in Practice".  
[www.slais.ubc.ca/PEOPLE/faculty/tennis-p/Tennis\\_SKOSOnto\\_2005.pdf](http://www.slais.ubc.ca/PEOPLE/faculty/tennis-p/Tennis_SKOSOnto_2005.pdf)

[TB04] Douglas Tudhope, Ceri Binding – Towards Terminology Services:experiences with  
a pilot web service thesaurus browser, 2004 [www.comp.glam.ac.uk/~FACET/TUDHOPE-DC05.DOC](http://www.comp.glam.ac.uk/~FACET/TUDHOPE-DC05.DOC)

[TB05] Douglas Tudhope, Ceri Binding - KOS at your Service: Programmatic Access to  
Knowledge Organisation Systems, 2005  
<http://jodi.ecs.soton.ac.uk/Articles/v04/i04/Binding/#HODGE2000>

[PS05] Josè Ramon Perez Aguera, Rodrigo Sanchez Jimenez – Thesauri managing and  
Software Agent: a proposed Architecture, 2005  
<http://dc2005.uc3m.es/program/presentations/>

## Risorse Web

[APAIS] APAIS. Australian Public Affairs Information Service Thesaurus  
<http://www.nla.gov.au/apais/thesaurus/>

[CSA/NBII] CSA/NBII Biocomplexity Thesaurus Web Services  
<http://nbii-thesaurus.ornl.gov/thesaurus/>

[DDCCE] Description of the Dublin Core Elements  
<http://dublincore.org/documents/dc>

[FOAF] RDFWeb and Friend of a Friend (FOAF), 24-07-2003  
<http://rdfweb.org/mt/foaflog/archives/2003/07/28/12.46.56/index.html>

[GEMET] GEMET GEneral Multilingual Environmental Thesaurus  
<http://www.eionet.eu.int/gemet>

[QGPT] Quick Guide to Publishing a Thesaurus on the Semantic Web, 17-05-2005  
<http://www.w3.org/TR/2005/WD-swbp-thesaurus-pubguide-20050517/>

[RDF] Resource Description Framework (RDF)  
<http://www.w3.org/RDF/>

[SKOS05] Simple Knowledge Organisation System (SKOS), 2005  
<http://www.w3.org/2004/02/skos/>

[SKOSE04] Simple Knowledge Organisation System (SKOS) SKOS Extensions, 2004  
<http://www.w3.org/2004/02/skos/extensions/>

[SKOSCG05] SKOS Core Guide, 02-11-2005  
<http://www.w3.org/TR/2005/WD-swbp-skos-core-guide-20051102/>

[SKOSCVS05] SKOS Core Vocabulary Specification, 02-11-2005  
<http://www.w3.org/TR/2005/WD-swbp-skos-core-spec-20051102/>

[SKOSTUT] SKOS Core Tutorial for DCMI 2005, 2005  
[dc2005.uc3m.es/program/tutorials/tutorial4\\_eng.ppt](http://dc2005.uc3m.es/program/tutorials/tutorial4_eng.ppt)

[SWED] SWED Semantic Web Enviromental Directory  
<http://www.swed.org.uk/swed/index.html>

[OF05] Open Forum 2005 on Metadata Registries (Open Forum 2005), Berlin, Germany,  
11-14 Apr 2005  
<http://epubs.cclrc.ac.uk/bitstream/674/berlin2005a.pdf>

[OWL] OWL Web Ontology Language Overview  
<http://geocities.com/ioannespetrus/traduzioni/w3c/owloverview.html?20052>

[OWLHP] OWL Web Ontology Language  
<http://www.w3.org/TR/owl-features/>

[W3C] World-Wide Web Consortium  
[www.w3.org](http://www.w3.org)

[TMF] Topic Maps F.A.Q.  
<http://www.infoloom.com/tmfaq.htm>

[DCHP] The Dublin Core Home Page  
<http://dublincore.org>

[TAO] The TAO of Topic Maps  
<http://www.ontopia.net/topicmaps/materials/tao.html>

[UKAT] UKAT UK Archival Thesaurus  
<http://www.ukat.org.uk/>

[UTM] Understanding Topic Maps “White Paper”  
<http://www.infoloom.com/whitepaper.htm>

[WSE] Web Semantico: L’estensione del Web Semantico  
<http://www.websemantico.org/articoli/semanticweb.php>

[ZTHES] Zthes: A Z39.50 Profile for Thesaurus Navigation  
<http://zthes.z3950.org/profile/current.html>