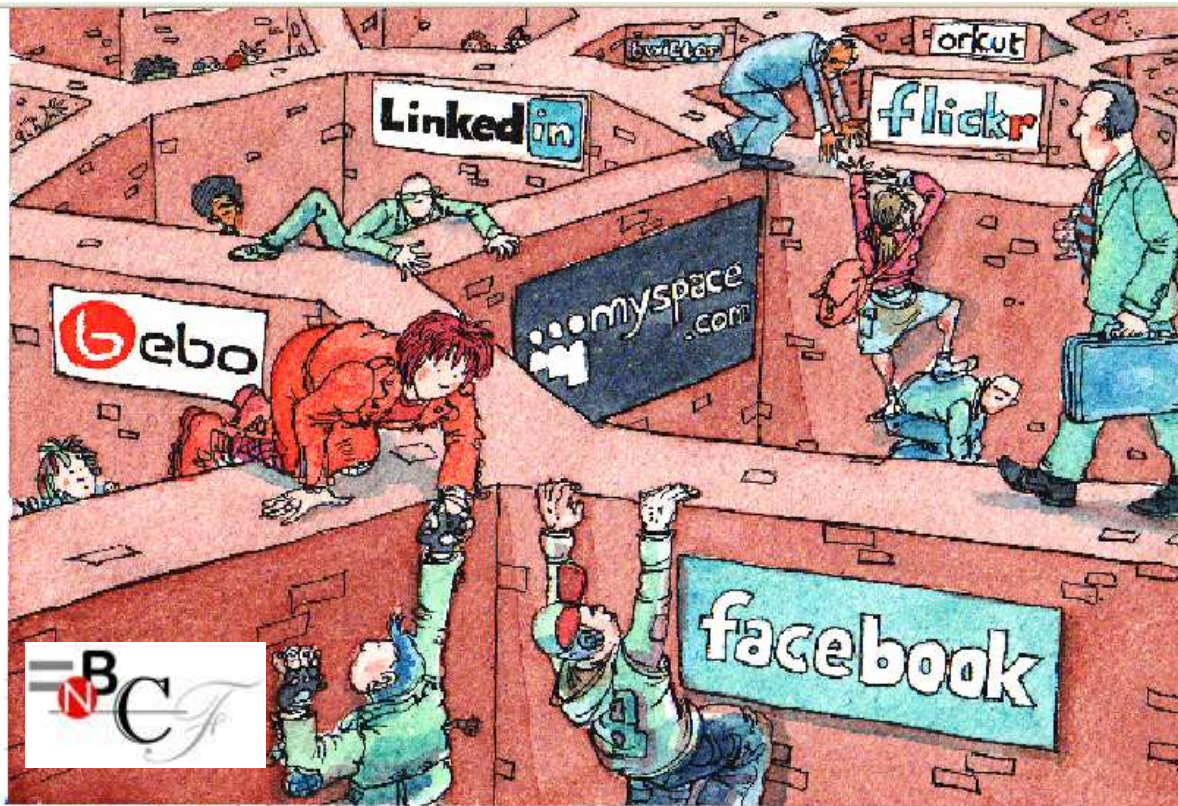


Esperienze di linked data e web semantico in BNCF



Anna Lucarelli

anna.lucarelli@beniculturali.it

6' Incontro ISKO Italia

Firenze 20 maggio 2013

**Biblioteca nazionale centrale
Firenze**

Da Tim Berners Lee (2009):

[http://www.w3.org/2009/Talks/0204-ted-tbl/#\(22](http://www.w3.org/2009/Talks/0204-ted-tbl/#(22)



Biblioteche nazionali europee e “servizi” di linked data



<http://www.bl.uk/bibliographic/datafree.html>

data.bnf.fr

<http://data.bnf.fr/semanticweb>

Web sémantique et modèle de données



Linked Data Service of the German National Library

http://www.dnb.de/DE/Service/DigitaleDienste/LinkedData/linkedata_node.html;jsessionid=3AE715724BAB65B1D23D73615263CE8F.prod-worker2



Datos enlazados en la BNE

<http://www.bne.es/es/Inicio/Perfiles/Bibliotecarios/DatosEnlazados/index.html>



Biblioteca Nazionale svizzera > Servizi > **Strategia di open data: metadati di Helveticat**

<http://www.nb.admin.ch/aktuelles/03147/03950/04156/index.html?lang=it>



<http://www.theeuropeanlibrary.org/tel4/lod>

“The European Library aims to promote the wider use and exposure of the collections it aggregates from its member libraries. With this purpose in mind, one of the API services provided by The European Library is to make available Linked Open Data”

Model and Vocabularies

For the bibliographic resources, the dataset follows a data model based on the [Europeana Data Model](#) [EDM]

The data is represented using classes and properties from the following vocabularies

- Multilingual Dictionary of Cataloguing Terms and Concepts -
- RDA Group 2 Elements
- Europeana Data Model
- OWL Web Ontology Language
- Dublin Core
- Dublin Core Terms
- SKOS Simple Knowledge Organization System
- WGS84 Geo Positioning
- XML
- OAI ORE Terms
- FOAF
- ISO 639-2 Languages



License

The linked open data set is available under the Creative Commons [CC0 1.0 Universal](#) license

Future Work

- Alignment of the data model with the ongoing activities of the library domain and linked open data
- Expressing library data using [RDA element sets and vocabularies](#)
- Expressing links between bibliographic resources
- Expressing library data by linking to existing vocabularies
- Expressing links between bibliographic resources aggregated by The European Library
- Expressing links to external open datasets
- This work will be carried out jointly with the partner libraries through The European Library's Technical and Interoperability Working Group


La strada dei linked data in BNCF avviata dal Nuovo soggettario

Due settori della BNCF al lavoro:

1. “Ricerche e strumenti di indicizzazione semantica” (Lucarelli, Franci, Francioni, Marzocca, Viti)
2. “Servizi informatici” (Bergamin, Pepe)

- si è evoluto nell'ottica del Web semantico sviluppando un'interoperabilità tecnica
Per ora unico strumento della BNCF in rdf; primo nel mondo a attivare link con Wikipedia
- è disponibile in SKOS/RDF, formati e linguaggi dei linked data
- colloquia con altri strumenti non solo di ambito bibliotecario (Menduni, [Come abbiamo ordinato gli open data di Firenze con l'aiuto di una biblioteca](#))
- accoglie terminologia proposta da biblioteche, istituzioni, archivi
- è “open” su aspetti qualitativi, quantitativi, metadati

Il primo Thesaurus italiano generale (“universale”)

- disponibile sul Web dal 2007: <http://thes.bncf.firenze.sbn.it/ricerca.php>
da aprile 2013 con interfaccia anche in inglese: 

- strumento “in progress”, oggi oltre **47.800 termini**

- costruito secondo gli standard ISO

- integrato con l’OPAC della BNCF

- interoperabile con altri thesauri, enciclopedie, ecc.

- prevede accessi multilingue (equiv. LCSH)



Date	Termini
Gennaio 2007	13.000 c.
Settembre 2007	16.000 c.
Febbraio 2008	19.200 c.
Settembre 2008	25.000 c.
Giugno 2009	31.000 c.
Febbraio 2010	34.300 c.
Giugno 2010	35.700 c.
Febbraio 2011	38.000 c.
Settembre 2011	43.700 c.
Febbraio 2012	45.500 c.
Settembre 2012	46.300 c.
Marzo 2013	47.800 c.

Le più recenti linee di sviluppo

In linea con quanto si sta realizzando in altri Paesi:

1. Interoperabilità

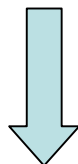
- metadati del Thesaurus disponibili in:
 - Zthes dal 2007
 - **SKOS/RDF dal 2010** <http://thes.bncf.firenze.sbn.it/thes-dati.htm>
- Implementazione di collegamenti con altri strumenti online
<http://digitale.bncf.firenze.sbn.it/NS-SPARQL/> (sperimentale)
- uso in ambiti non bibliotecari nello spirito del Web semantico (es. Open data del Comune di Firenze; Soprintendenza Archivistica; Cesvot)

2. Indicizzazione automatica

Impiego automatico del Thesaurus nell'indicizzazione di risorse digitali

L'impiego di SKOS per i Thesauri

- Simple Knowledge Organisation System
sviluppato dal W3C Semantic Web Deployment Working Group (SWDWG)
- Si autodefinisce “*data model*”
I suoi oggetti non sono le cose (*entità*) ma i concetti e le gerarchie concettuali presi in considerazione senza un diretto riferimento alle cose (*entità*); i concetti sono considerati indipendentemente dalla loro espressione linguistica/letterale
- Utilizza il linguaggio RDF (*Resource Description Framework*)
- Usato per migrare Thesauri (conformi a ISO) nel mondo dei *Linked data*



*L'assegnazione di un valore univoco ai concetti (URI)
facilita l'interoperabilità tra KOS differenti, cioè la possibilità di mappare
entità semantiche di schemi concettuali diversi*

Etichette SKOS, metadati,

Lo standard definisce alcune “etichette” impiegate per codificare le relazioni tra concetti:

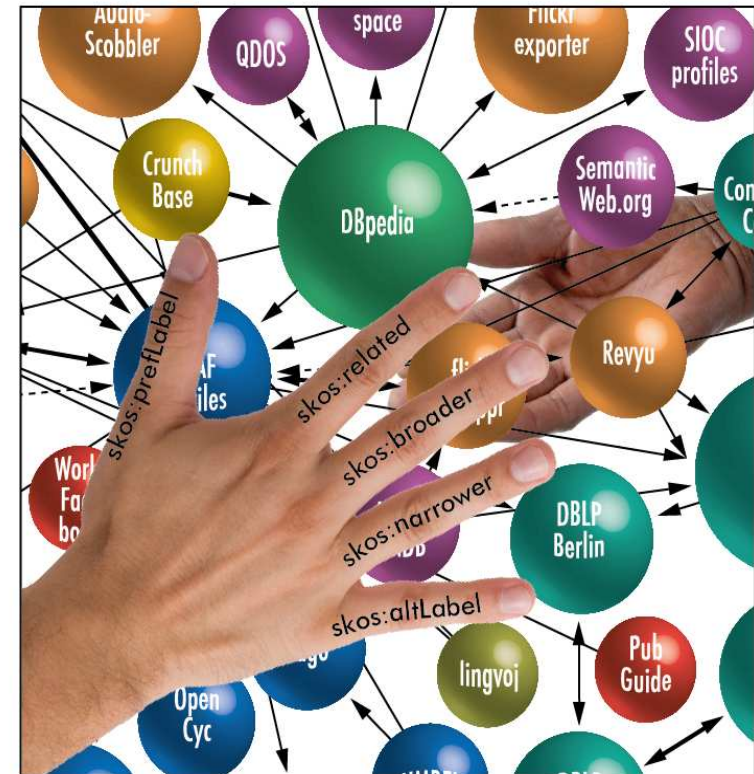
skos:closeMatch / skos:exactMatch

(diversi livelli di equivalenza)

skos:broaderMatch

skos:narrowerMatch

skos:relatedMatch



Giving and taking of linked data with SKOS

Esempi di termini del Nuovo soggettoario che attivano interoperabilità

[Web semantico](#)

[Motori di ricerca](#)

[Mediatori](#)

[Cartolarizzazione](#)

Gli esempi evidenziano collegamenti con altri strumenti come Thesauri, altri sistemi di indicizzazione (es. LCSH), Enciclopedie on-line (es. Treccani, Wikipedia), ecc.

In particolare:

- [Web semantico](#) DDC23 ; Wikipedia; Rameau; LCSH
- [Motori di ricerca](#) proposto da Bocconi; relazioni recipr. LIUC /NS; da rdf: Treccani non espone i dati in SKOS
- [Mediatori](#) relazioni reciproche LIUC / NS
- [Cartolarizzazione](#) relazioni reciproche Acc. Crusca / NS

Il Thesaurus del *Nuovo soggettario* nel formato SKOS

- Prima mappatura nel 2010
- Il lavoro ha seguito varie fasi ed è in continua evoluzione

Nessun problema	Problemi incontrati
Note <ul style="list-style-type: none">• Definizione• Scope note• History note• Fonti• DDC, ecc.	– Collezioni / etichette di nodo > Risolto – Relazione di variante storica > Risolto (cambio di forma rispetto al passato) – Relazione di scomposizione
Hanno tutte trovato un'adeguata esplicitazione in SKOS	La relazione di Scomposizione non ha trovato adeguata esplicitazione in SKOS

Implementazione dei *linked data*

Come abbiamo potuto vedere dagli esempi mostrati, nel NS le corrispondenze possono essere attivate mediante:

- “Citazione” nel campo “Fonte” dello strumento che si desidera collegare (es. Agrovoc, Treccani ecc).

Se il “citato” è disponibile in SKOS/RDF, NS si arricchisce della relazione *skos:closeMatch*

- Indicazione esplicita di equivalenza nel campo apposito *Equiv. LCSH* (*Library of Congress Subject Headings*).

Impiegata la relazione *closeMatch*, concettualmente più ampia di *exactMatch* inizialmente impiegata

- Creato sempre un *deep link* allo strumento “citato”

- Le diverse scelte e “politiche” di LoC; BNF; DNB: es. [Basque language](#)

Equivalenze del Thesaurus attive ad oggi

Link tramite il campo Fonte (in odine decrescente)

Strumenti collegati	Numero collegamenti	Strumenti di indicizzazione: Thesauri, Classificaz.,ecc.	Enciclopedie	Altro
Wikipedia (DBPEDIA)	11760		X	
Treccani.it	8210		X	
LCSH	2950	X		
AGROVOC	1600	X		
LIUC /Thes	670	X		
RAMEAU	660	X		
EUROVOC	135	X		
DOGI (Ittig)	120	X		
Acc. Crusca	3			X

Equivalenze del Thesaurus attive tramite *skos:closeMatch*

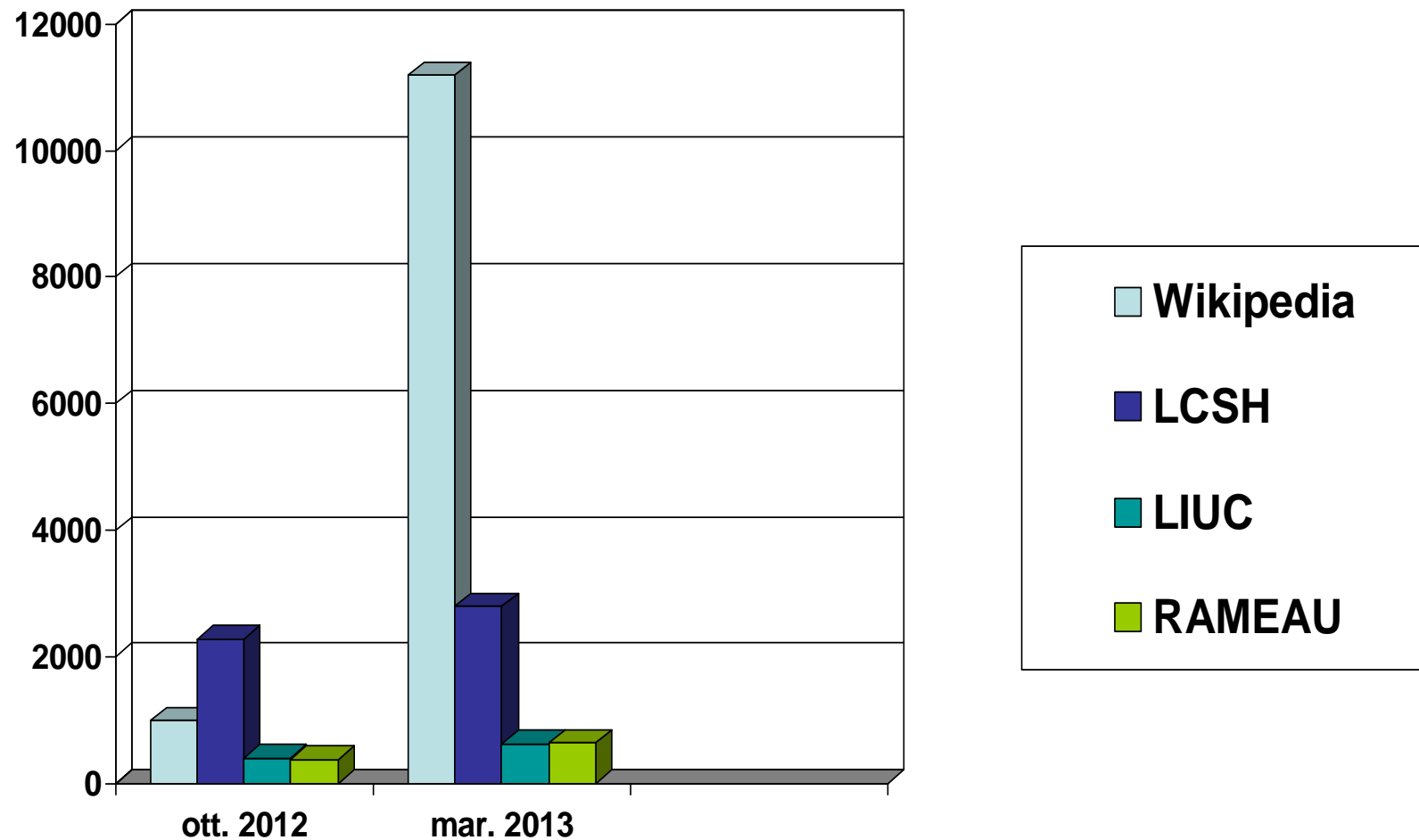
DBPedia (Wikipedia)	11185	ns-dbpedia	Scaricati i dati da DBPedia Italia , riversati in un DB. Sperimentazione automatismo per creare i link. In futuro useremo endpoint SparQL
LCSH	2790	ns-lcsh	E' già predisposto un DB con i dati in formato SKOS disponibile sul sito della LoC (agg. ogni 6 mesi)
AGROVOC	1080	ns-agrovoc	Unico caso di link via software attraverso interrogaz. diretta, attraverso loro endpoint SparQL
RAMEAU	630	ns-rameau	Trovati i dati in formato SKOS, e procedura come per LCSH. Problemi con i caratteri speciali....
LIUC/Thes	620	ns-liuc	<u>Interoperabilità reciproca NS/LIUC</u> Scarichiamo il loro SKOS come nel caso della LoC . Sperimentato, attraverso lo SKOS, prototipo di link di vario livello
DoGi (Ittig)	105	ns-dogi	<u>Interoperabilità reciproca NS/DoGi</u> L' Ittig sta lavorando all'arricchimento del loro SKOS già a nostra disposizione
EUROVOC	100	ns-eurovoc	Propongono una soluzione per la gestione della scomposizione
DDC	7870	ns-DDC	Caso a parte: manualmente si inseriscono i corrispondenti e si ci si collega al sito DDC che espone gratuitamente le vedette

Incremento delle equivalenze negli ultimi 6 mesi (ottobre 2012- marzo 2013)

	10/2012	03/2013
DBPedia (Wikipedia)	990	11185
LCSH	2280	2790
LIUC/Thes	400	620
RAMEAU	370	630

Incremento delle equivalenze negli ultimi 6 mesi (ottobre 2012- marzo 2013)

Implementazione di procedure automatiche per link con Wikipedia



Dai metadati del Nuovo soggettario ad una futura disponibilità come linked data dei record bibliografici

- Trasparenza informativa e pubblicazione dei dati open source
<http://thes.bncf.firenze.sbn.it/thes-dati.htm>
- Licenza Creative Commons
(dati usabili liberamente purché ne venga citata la paternità BNCF)
<http://creativecommons.org/licenses/by/2.5/it/deed.it>
- Registrazione non obbligatoria a chi scarica il formato SKOS
- Visibilità di:
 - [statistiche sugli accessi al server](#) (pagina Intranet)
 - [quali e quanti utenti lo scaricano](#), anche mese per mese, anno per anno

Se estenderemo ai record bibliografici nel loro insieme quanto avviato nell'ambito del *Nuovo soggettario*, la BNCF si potrà allineare ai servizi di altre Biblioteche nazionali per i *linked data*.