

Il Thesaurus del Nuovo soggettario interpreta SKOS

Marta Motta and Dario Rodighiero

Sommario Il consorzio W3C propone SKOS come standard per tutti i vocabolari controllati in modo da creare un linguaggio che permetta di realizzare una rete informativa interoperabile. La Biblioteca nazionale centrale di Firenze ha deciso di sperimentare questo nuovo linguaggio formale. Qui di seguito viene trattato il caso studio della conversione in SKOS del Thesaurus del Nuovo soggettario, illustrandone vantaggi e problematiche.

1 Introduzione

Negli ultimi anni è ormai assodata l'utilità degli strumenti elaborati nell'ambito delle scienze documentali e bibliotecarie per l'organizzazione della conoscenza e l'architettura dell'informazione. L'avvento di Internet ha cambiato i modi con i quali l'informazione prende forma e viene organizzata e questo ha avuto delle conseguenze anche nel campo dell'indicizzazione semantica e del recupero dell'informazione, in cui diventa sempre più importante elaborare strumenti che favoriscano la comunicazione e l'efficacia informativa non solo dei cataloghi delle biblioteche, ma di sistemi di organizzazione della conoscenza più in generale.

Nell'attuale contesto, caratterizzato da un gran numero di risorse informative virtualmente disponibili, gli strumenti per l'organizzazione dell'informazione e della conoscenza devono essere in grado di raggiungere i massimi livelli di efficacia e di efficienza, ossia devono produrre il risultato desiderato dall'utente nella maniera meno dispendiosa in termini di procedure, tempo e costi.

Marta Motta

Biblioteca nazionale centrale di Firenze, piazza Cavalleggeri, 1, 50122 Firenze e-mail: motta.marta@gmail.com

Dario Rodighiero

University of Milano-Bicocca c/o Joint Research Centre - European Commission, via Enrico Fermi, 2749 I-21027 Ispra, Italy e-mail: dario.rodighiero@ext.jrc.ec.europa.eu

Il trasferimento di informazioni sul contenuto concettuale dei documenti mediante strutture formalizzate è un'attività che riguarda molti campi del sapere, come viene confermato anche a livello internazionale, grazie alle esperienze che mirano ad allestire sistemi che possano integrarsi fra loro pur nella diversità delle singole applicazioni. È quindi sempre più urgente l'interconnessione di linguaggi documentari tradizionali con tecniche di indicizzazione e recupero per soggetto (anche di risorse digitali) nell'ambito delle nuove prospettive di interoperabilità offerte dal web. Come sostiene Elaine Svenonius nel suo libro *The intellectual foundation of information organization*¹, un'intelligente e corretta organizzazione dell'informazione è indispensabile non solo per la comunità scientifica, ma anche per i singoli, per l'economia e per la società in generale. Inoltre i principi, gli obiettivi e le tecniche fino ad oggi sviluppate per organizzare l'informazione nel campo degli studi bibliotecomici costituiscono senza dubbio un corpo di conoscenza ampiamente applicabile anche al nuovo contesto dell'era digitale.

Gli attori in gioco sono vari, si va dai thesauri alle tassonomie, dagli schemi di classificazione alle ontologie, ormai definiti come sistemi di organizzazione della conoscenza (KOS). Essi sono gli strumenti a disposizione per ottenere comprensione e comunicazione tra operatori e utenti di diverse aree linguistiche, ma anche per connettere risorse informative di differenti aree tematiche o domini disciplinari.

Gli strumenti di controllo terminologico non solo orientano gli utenti nel recupero delle informazioni (information retrieval), ma soprattutto normalizzano la terminologia attraverso metodi di analisi concettuale coerenti con l'assetto delle specifiche aree disciplinari, creando strutture basilari per l'organizzazione e la gestione delle informazioni.

In questo contesto i thesauri si sono dimostrati come eccellenti strumenti per assegnare metadati semantici², dal momento che la loro struttura consente una gestione efficace di questi metadati. Le relazioni dei thesauri, soprattutto quando questi sono integrati con cataloghi online di biblioteche, portali e siti web, sono utilissime per navigare nelle collezioni di risorse documentali e consentono di esplorare i materiali collegati allargando o restringendo la ricerca, attraverso termini più generici (BT) e termini più specifici (NT), oppure anche proponendo itinerari di avvicinamento attraverso i termini correlati (RT). La stessa presentazione sistematica e a volte classificatoria dei thesauri sembra fatta apposta per favorire la navigazione, particolarmente per le risorse digitali attraverso collegamenti ipertestuali.

Il nostro contributo vuole presentare i risultati di una sperimentazione: la conversione in SKOS (Simple Knowledge Organization System) del Thesaurus del Nuovo soggettario, di cui ci siamo occupati grazie alla collaborazione di Maria Grazia Pepe del settore Servizi informatici della Biblioteca nazionale centrale di Firenze. SKOS è lo standard del W3C per rendere disponibili sul web i sistemi di organizzazione della conoscenza. Il Thesaurus del Nuovo soggettario costituisce la componente terminologica del nuovo sistema di indicizzazione per soggetto impiegabile da biblioteche, archivi, mediateche, ecc.

¹ SVENONIUS [2000]

² BROUGHTON [2008]

La nostra riflessione sulle problematiche e le criticità riscontrate durante il processo di conversione si è basata anche, comparativamente, sulle soluzioni adottate da altri sistemi di indicizzazione come LCSH³, RAMEAU⁴ e thesauri specialistici come LIUC⁵ e MESH⁶, che negli ultimi anni hanno intrapreso la strada della traduzione in SKOS.

2 Il Thesaurus del Nuovo soggettario

Il Nuovo soggettario è un progetto nato all'interno della Biblioteca nazionale centrale di Firenze (BNCF). È un sistema di indicizzazione semantica costituito da regole codificate⁷ (inerenti sintassi e terminologia) e da un Thesaurus generale. La BNCF ha sempre avuto un ruolo rilevante nell'elaborazione e nell'aggiornamento di strumenti di indicizzazione. L'obiettivo era di rinnovare il precedente strumento, il *Soggettario per i cataloghi delle biblioteche italiane*,⁸ e di adeguarlo a principi e standard internazionali alla luce delle esigenze poste dalla nuova generazione di utenti.

Nel 2007 è stato pubblicato il volume contenente le regole di indicizzazione e un prototipo del Thesaurus in continua evoluzione accessibile sul web. Il Thesaurus, ormai lontano dall'essere un prototipo, è accessibile gratuitamente dalle postazioni di ricerca della BNCF e tramite abbonamento dall'esterno, ma da Luglio 2010 sarà consultabile liberamente online⁹.

Il Thesaurus si accresce costantemente grazie al gruppo di lavoro della BNCF e alla collaborazione esterna di alcune istituzioni iniziata nel 2009¹⁰. Dal 2007 la Bibliografia nazionale italiana (BNI), sempre a cura della BNCF, ha iniziato ad im-

³ LCSH (Library of Congress subject headings) è il repertorio delle voci di soggetto utilizzate dalla Library of Congress a partire dal 1898.

⁴ RAMEAU (Répertoire d'autorité-matière encyclopédique et alphabétique unifié) è la lista controllata delle voci di soggetto assegnate dal 1980 ai record bibliografici inseriti nel catalogo in linea della Bibliothèque nationale de France, integrata con voci proposte da biblioteche universitarie, di pubblica lettura e di ricerca. Le voci di RAMEAU derivano in massima parte dal Répertoire de vedettes-matière, elaborato dalla biblioteca dell'Université de Laval (Québec) sulla base delle LCSH

⁵ LIUC è il thesaurus di economia e scienze sociali dell'Università Carlo Cattaneo di Castellanza.

⁶ MESH (Medical Subject Headings) è il thesaurus usato per indicizzare gli articoli delle oltre 5000 riviste mediche presenti nel database bibliografico Medline/PubMed.

⁷ BNCF [2006]

⁸ BNCF [1956]

⁹ Il Thesaurus è raggiungibile dall'indirizzo: http://www.edigeo-online.it/Nuovo_soggettario/#?rigamenu=Nuovosoggettario.Thesaurus

¹⁰ Le istituzioni che finora collaborano al progetto sono: Biblioteca dell'Università Bocconi di Milano, Università degli studi di Milano, l'Università di Pisa. Area bibliotecaria, archivistica e museale, Biblioteca Mario Rostoni dell'Università Carlo Cattaneo di Castellanza (LIUC), l'Istituto di teorie e tecniche dell'informazione giuridica (ITTIG) del Consiglio nazionale delle ricerche (CNR), il Coordinamento biblioteche speciali e specialistiche di Torino (CoBis) e la Società IDEST

piegare il Nuovo soggettario e gradualmente anche altre biblioteche italiane stanno transitando dal vecchio al nuovo linguaggio.

Il Nuovo soggettario è predisposto per un'indicizzazione pre-coordinata (la BNI utilizza i termini del Thesaurus per la costruzione di stringhe di soggetto), ma può essere usato anche in modalità post-coordinata, come dimostrano alcune esperienze già in atto¹¹.

Le quattro componenti che costituiscono la struttura del sistema sono:

1. il Thesaurus,
2. le Norme per il controllo terminologico e per la costruzione delle stringhe di soggetto;
3. il Manuale applicativo (pubblicato online da Febbraio 2010, contenente istruzioni ed esempi applicativi, sia di tipo semantico che sintattico);
4. l'archivio delle stringhe di soggetto derivanti dall'uso del nuovo linguaggio.

L'interfaccia di ricerca del Thesaurus consente di consultare anche la versione digitale del vecchio strumento di indicizzazione (Soggettario del 1956 e suoi aggiornamenti), offrendo quindi un completo accesso a tutto il patrimonio terminologico della tradizione italiana.

Il Thesaurus del Nuovo soggettario è dunque un vocabolario generale che si sta costantemente accrescendo e che attualmente comprende circa 35.000 termini. È generale sia dal punto di vista dei domini coinvolti, sia perché applicabile all'indicizzazione di risorse di varia natura¹².

È conforme all'ISO 2788-1986¹³ e, in attesa della pubblicazione definitiva del nuovo ISO 25964 che lo sostituirà, sono state considerate anche le indicazioni di ANSI/NISO Z39.19 del 2005 e dello standard britannico BS 8723:2005-2008¹⁴.

Per quanto riguarda le caratteristiche strutturali del vocabolario, si tratta di un Thesaurus prevalentemente monogerarchico (la maggior parte dei termini ha un solo BT). La struttura gerarchica viene costruita applicando l'analisi a faccette e criteri di precedenza definiti. La poligerarchia è adottata solo raramente (secondo criteri espressamente indicati) e comunque sempre nel rispetto dei principi che regolano le relazioni gerarchiche. L'analisi a faccette, assume il ruolo fondamentale di organizzare i termini in una struttura classificatoria basata su quattro categorie principali (Agenti, Azioni, Cose, Tempo) e su ulteriori caratteristiche di divisione (13 faccette: Persone e gruppi, Attività, Processi, Strumenti, ecc.)¹⁵. Come sostiene Vanda Broughton "l'analisi a faccette rende molto più facile la gestione dell'intera gamma di concetti, compresi quelli relativi a idee astratte. Ad oggi, questo tipo di analisi,

¹¹ La società IDEST che cura la serie speciale della Bibliografia nazionale dei libri per ragazzi, utilizza i termini del Thesaurus per accessi da semplici parole chiave.

¹² Presso la Biblioteca nazionale di Firenze, il Nuovo soggettario è stato applicato con successo per indicizzare fotografie di un prestigioso fondo del Novecento (Fondo Pannunzio). Cfr. LUCARELLI [2008]

¹³ ISO 2788/1986

¹⁴ ANSI/NISO Z.39.19; BS 8723:2005-2008. Cfr. MOTTA - TIBERI [2009]


¹⁵ Cfr. CHETI - PARADISI [2008]

intesa come base per la creazione di un insieme integrato di strumenti lessicali, non ha rivali”¹⁶.

Il sistema del Nuovo soggettario si basa su un modello analitico-sintetico, quindi non enumera a priori tutti i possibili termini e tutte le loro possibili combinazioni all’interno di stringhe di soggetto, ma consente di inserirne di nuovi al bisogno. L’indicizzatore può proporre nuovi termini in fase di catalogazione; ne deriva un sistema aperto, facilmente scalabile e flessibile in fase di indicizzazione.

Si tratta dunque di un Thesaurus che è anche in grado di gestire il passaggio fra vecchie e nuove forme terminologiche ed ha anche la caratteristica di presentare dati che esplicitano all’utente il complesso lavoro di analisi semantica svolto da chi lo allestisce. Uno dei suoi punti di forza è senza dubbio l’ampio apparato di informazioni associato ad ogni termine e la presenza di collegamenti a risorse elettroniche esterne.

Come si può vedere nell’esempio del termine “Acculturazione” (Fig. 1), i record dei termini prevedono numerose note: note di definizione; note d’ambito, note di orientamento all’interno delle note d’ambito; note storiche (con informazioni su precedenti forme, significati, usi del termine); note sintattiche (note che danno istruzioni all’indicizzatore sull’uso del termine nella costruzione delle stringhe di soggetto).

Acculturazione 

Categoria/Faccetta: Azioni:Processi
Nota d'ambito: Mutamento profondo e integrazione fra culture diverse di gruppi o popoli entrati in contatto tra loro in vari modi (conquista o dominio politico, assimilazione, immigrazione, ecc.)

UF Meticcio culturale
TT Processi
BT [Processi di ambito culturale e socioculturale]
RT Cultura
 Divulgazione
 Evoluzione culturale
 Meticcio
 Popoli

Nota storica: Precedentemente anche con il significato di acquisizione, da parte di gruppi sociali o di singoli, di un grado più avanzato di cultura
Fonte: BNI 1956-1985; VLI; →VTonline; DeM; PT; DDC21
DDC (Ed. 22): 303.482

Agenzia catalografica / proponente: BNI
Status del record: Termine strutturato
Identificativo: 14320

Figura 1 Il termine Acculturazione secondo il Nuovo soggettario

Nell’interfaccia di ricerca gli utenti possono facilmente individuare, nel campo Fonte, tutte le risorse utilizzate per controllare significati, morfologia e garanzia bibliografica dei termini. Queste fonti repertoriali possono essere liste di intestazioni di soggetto (ad es. LCSH, RAMEAU), thesauri generali e specializzati (ad es. AAT;

¹⁶ BROUGHTON [2008], p. 257

EUROVOC; MESH), dizionari (ad es. Dizionario Treccani online), enciclopedie (Enciclopedia britannica, Enciclopedia italiana Treccani, ecc.). Se le fonti citate sono disponibili liberamente sul web, è stato inserito un link diretto dal campo Fonte del termine.

Il Thesaurus può essere consultato grazie ad un'applicazione che consente la ricerca dei termini, la visualizzazione delle schede, la presentazione dell'intera struttura gerarchica e la navigazione delle relazioni attraverso hyperlink. Questa funzionalità può essere utile nel momento in cui il vocabolario viene utilizzato come strumento per organizzare risorse digitali e la visualizzazione gerarchica può essere navigata per raggiungere direttamente queste risorse.

Il Thesaurus del Nuovo soggettoario ha un'architettura flessibile sotto diversi aspetti che permette:

1. una costante crescita quantitativa;
2. di accogliere anche pacchetti di terminologia specialistica e settoriale pur in un contesto che rimane comunque generale;
3. di dialogare e a interagire con thesauri specializzati, con altri strumenti di indicizzazione, con strumenti lessicografici ed enciclopedici, attraverso modelli strutturali di vario tipo, come i recenti standard bibliografici hanno delineato¹⁷.

3 Conversione in SKOS del Thesaurus del Nuovo soggettoario

Il Nuovo soggettoario si propone dunque come un sistema per l'organizzazione della conoscenza, in cui ogni funzionalità è finalizzata al recupero dell'informazione. L'utilità di questi sistemi per lo sviluppo del web è ormai indiscussa, nonostante in passato linguaggi di questo tipo siano stati considerati molto complessi da costruire e usare. Essendo nati come strumenti biblioteconomici tradizionali, inizialmente sembravano inappropriati per un ambiente digitale e non necessari rispetto a quanto consentito dalla potenza nel processare informazioni dei moderni motori di ricerca. L'esperienza del Nuovo soggettoario ha dimostrato che i due mondi possono reciprocamente potenziarsi. Infatti se da un lato il Thesaurus viene reso disponibile come applicazione web accessibile all'esterno, a sua volta il sistema ha la possibilità di accedere a risorse esterne attraverso il protocollo Zthes¹⁸.

In quest'ottica SKOS nasce come linguaggio comune per vocabolari controllati, come thesauri, schemi di classificazione, tassonomie e soggettoari. Non va dimenticato però che SKOS si basa su degli standard già esistenti: RDF (Resource Description Framework) e RDFS (RDF Schema), che si propongono come linguaggi di interoperabilità più generici. Quindi SKOS si colloca in un contesto di interoperabilità

¹⁷ Cfr. i modelli *Backbone structure* e *Satellite vocabularies* descritti rispettivamente in BS 8723 e ANSI/NISO Z39.19. "Il Thesaurus del Nuovo soggettoario può proporsi come una sorta di meta-thesaurus da cui possono derivare tanti altri thesauri specialistici, oppure definirsi come punto di scambio e raccordo fra molteplici thesauri specializzati in varie discipline" MOTTA - TIBERI [2009]

¹⁸ Zthes, <http://zthes.z3950.org/>

tra vocabolari controllati che si pone in un contesto di interoperabilità ancora più ampio.

Quindi la decisione di sperimentare la conversione del Thesaurus del Nuovo soggettario in SKOS può essere vista come la volontà di compiere un ulteriore passo verso una maggiore interoperabilità con altri sistemi. Questa è la strada intrapresa anche da altri importanti sistemi di indicizzazione: le americane LCSH, il sistema francese RAMEAU, le regole tedesche RSWK, thesauri specialistici come AGRO-VOC, MESH, e in ambito italiano il thesaurus LIUC dell'Università di Castellanza, ma anche schemi di classificazione come la Classificazione decimale Dewey (DDC), la Classificazione decimale universale (UDC), e la Classificazione Bliss (BC).

RDF, RDFS e SKOS sono linguaggi basati su triple di tipo Soggetto – Predicato – Oggetto, che si basano sui concetti. Tuttavia tra i vocabolari controllati ci sono sia strumenti che sposano l'adozione di termini – come il Nuovo soggettario –, che quella di concetti. Questa caratteristica che SKOS eredita dai suoi padri (RDF e RDFS), oltre ad essere discutibile a livello teorico, può provocare nel processo di conversione una perdita di informazioni¹⁹, soprattutto nel caso di thesauri multilingui²⁰.

Per esempio nell'ambito della traduzione del thesaurus LIUC, il primo thesaurus italiano disponibile sul web a sperimentare la conversione, la scelta di SKOS che un concetto può essere rappresentato da termini di lingue diverse crea problemi di conversione²¹.

Come detto in precedenza, il Thesaurus del Nuovo soggettario è basato sullo standard per thesauri monolingui ISO 2788 compatibile con SKOS. Il codice identificativo di ogni termine del Thesaurus è stato usato per generare il codice di identificazione RDF. Questa soluzione è stata scelta anche nella conversione delle LCSH, in cui per identificare i concetti di SKOS è stato usato il numero LCCN (Library of Congress Control Number) che si trova nel campo 001 del linguaggio MARC. Il numero è stato incorporato in un indirizzo URL del tipo <http://lcsch.info/{lccn}#concept>.

Le relazioni gerarchiche (BT e NT), quelle associative (RT) e di equivalenza (USE/UF) tra i termini del Thesaurus del Nuovo soggettario sono state convertite negli elementi corrispondenti del *data model* di SKOS: rispettivamente *skos:broader*, *skos:narrower*, *skos:relatedTerm* e *skos:altLabel*. Sono stati riscontrati dei problemi invece per esprimere due particolari relazioni del Thesaurus, che hanno lo scopo di connettere le precedenti forme terminologiche con le nuove:

1. la relazione di variante storica, una relazione non standard del Thesaurus del Nuovo Soggettario che collega termini preferiti del vocabolario a termini del vecchio strumento non più accettati. Il legame espresso con la sigla “Historical see for” (HSF) è di estrema utilità perché consente di caricare in modo automatico nei cataloghi in linea delle biblioteche i legami fra vecchi e nuovi termini, senza correggere ogni singola intestazione di soggetto mutata;

¹⁹ VAN ASSEM et alii [2006]

²⁰ BALLESTRA [2008]

²¹ CAVALERI [2009]

2. la relazione di scomposizione espressa dal simbolo USE+ e dal suo inverso UF+. Questa relazione collega termini composti a termini preferiti derivati dalla loro scomposizione e viceversa.

Per risolvere il primo problema (la variante storica) abbiamo pensato di usare una *subpropriety* dell'etichetta di SKOS *hiddenLabel* che viene trasformata in una sottoetichetta del tipo *nuovosoggettario:obsoleteTerm = X*²². In questo caso, quindi, bisognerebbe avanzare al W3C una modifica dello standard, altrimenti questa soluzione si tradurrebbe in una mera personalizzazione del linguaggio che gli altri sistemi non riuscirebbero a interpretare a scapito dell'interoperabilità stessa, il principale obiettivo di SKOS.

Nel secondo caso, invece, abbiamo ipotizzato di usare come riferimento quanto è stato fatto per il Thesaurus MESH. Durante il processo di conversione del MESH per un problema simile è stata creata ad hoc l'etichetta *mesh:preferredCombination*²³, stiamo valutando se usare questa stessa soluzione anche per il Thesaurus del Nuovo soggettario. Per ora la relazione di scomposizione compare come una semplice *AltLabel* del termine preferito del Thesaurus, con una evidente perdita di informazioni.

Un caso analogo si è verificato anche durante il passaggio in SKOS di LCSH e di RAMEAU. Le americane LCSH così come il sistema francese RAMEAU, possono presentare intestazioni di soggetto formate da due concetti composti (per esempio "Teatro – 17. sec.") che sono stati resi in SKOS uniti con il semplice trattino: ma in questo modo i due termini non vengono interpretati, come separati e combinati insieme al momento della costruzione della stringa di soggetto, ma come un unico concetto composto. Anche in questo esempio, quindi, la struttura estremamente semplificata di SKOS provoca una perdita di informazioni utili.

Come abbiamo visto in precedenza, oltre alle relazioni semantiche, ogni termine del Thesaurus presenta un ricco apparato di informazioni: vari tipi di note, la maggior parte delle quali trovano una adeguata etichetta in SKOS, per esempio *skos:definition*, *skos:scopeNote*, *skos:historyNote*, *skos:editorialNote*. Quest'ultima è stata usata, nella nostra sperimentazione, per indicare, anche più volte nello stesso record del termine: il campo dei repertori consultati per il controllo terminologico, il campo per l'agenzia bibliografica che propone il termine, il campo per lo status di lavorazione del termine.

Abbiamo usato l'etichetta *skos:example* per introdurre la nota sintattica che dà indicazioni per la costruzione delle stringhe di soggetto secondo le regole del sistema di indicizzazione.

Inoltre, molti termini del Thesaurus del Nuovo soggettario sono stati mappati con le corrispettive notazioni della Classificazione decimale Dewey (DDC)²⁴. Questo consente di costruire un ponte verso altri sistemi che, pur basandosi su diversi

²² VAN ASSEM et alii [2006], p. 7

²³ VAN ASSEM et alii [2006], p. 13

²⁴ È allo studio la fattibilità di una versione italiana di WebDewey <http://www.oclc.org/dewey/versions/webdewey>, anche al fine di creare una interoperabilità effettiva e non solo virtuale tra il Thesaurus del Nuovo soggettario e la CDD

metodi di indicizzazione per soggetto o su diverse lingue, si affiancano all'uso della DDC e ne prevedono il collegamento. Questo *mapping* è stato mantenuto anche nel passaggio del Thesaurus in SKOS usando l'etichetta *skos:notation*.

Caratteristica interessante di SKOS è la sua interoperabilità con altri schemi di metadati come Dublin Core e FOAF (Friend of A Friend). Nel Thesaurus del Nuovo soggetto abbiamo usato per alcune informazioni di servizio, come la data di creazione del termine, lo schema Dublin Core.

4 Conclusioni

Che possibilità apre per il Thesaurus del Nuovo soggetto la conversione in SKOS?

I benefici più immediati sono senza dubbio:

1. la maggiore interoperabilità con altri sistemi che sono stati convertiti in SKOS;
2. il vantaggio di avere ogni concetto/termine del Thesaurus identificato unicamente da un URI.

La conversione in SKOS dei vocabolari controllati usati per l'indicizzazione per soggetto in ambito bibliotecario può essere usata come base per la mappatura semantica di sistemi di organizzazione della conoscenza che strutturano domini disciplinari differenti o che usano lingue diverse. Quindi l'interoperabilità tecnica diviene la base per l'interoperabilità semantica fra i vari strumenti.

In quest'ottica il Nuovo soggetto ha la possibilità di allinearsi con progetti in campo europeo di condivisione di accessi semantici come il progetto MACS²⁵, che ha lo scopo di sviluppare un sistema per l'accesso multilingue per soggetto tramite la mappatura semantica dei sistemi d'indicizzazione RAMEAU (per la lingua francese), LCSH (per la lingua inglese) e SWD (per la lingua tedesca), che hanno tutti una versione in SKOS. Entrare a far parte del progetto MACS vuol dire ricoprire un ruolo importante anche nell'arricchimento del recupero semantico di Europea (European digital library, museum and archive) attraverso il progetto TELPlus²⁶, che ha ripreso i risultati di MACS in questo ambito.

A livello italiano, il Thesaurus del Nuovo soggetto potrebbe proporsi anche come vocabolario controllato di riferimento per il controllo di uniformità e l'organizzazione dei contenuti nei siti web della Pubblica amministrazione. La recente pubblicazione, in versione *draft*, delle nuove *Linee guida per i siti web della PA*²⁷, prevede infatti che "i sistemi di classificazione utilizzati per le risorse dei siti web della Pubblica amministrazione debbano consentire l'interoperabilità semantica, ovvero la possibilità di individuare in modo omogeneo gli attributi che caratterizzano una

²⁵ MACS (Multilingual Access to Subjects) <https://macs.vub.ac.be/pub>

²⁶ TELPlus <http://www.theeuropeanlibrary.org/portal/organisation/cooperation/telplus/index.php>

²⁷ *Linee guida per i siti web della PA* http://www.innovazionepa.gov.it/media/367125/linee_guida_siti_web_pa.pdf

risorsa (metadati) e i valori che gli attributi possono assumere (vocabolari) quando si descrivono i contenuti. Sistemi tecnologicamente interoperabili in assenza di interoperabilità semantica non possono scambiare e condividere dati, documentazione e servizi”²⁸.

Il Thesaurus del Nuovo soggetto potrebbe diventare un “servizio terminologico” da affiancare e integrare con quelli già in uso nei siti degli enti della Pubblica amministrazione, anche grazie alla partecipazione attiva degli stessi enti interessati che potrebbero proporre nuovi termini. In questo modo tutti i siti pubblici offrirebbero una modalità di accesso semantico comune, attraverso lo stesso linguaggio condiviso, con un evidente vantaggio per l’usabilità stessa dei siti e, eventualmente, il controllo su ulteriori tag proposti dagli stessi utenti, validati da chi allestisce il Thesaurus.

Esempio del termine “Biblioteche” del Thesaurus del Nuovo soggetto convertito in SKOS

```
<rdf:RDF>
<skos:Concept rdf:about="termine?id=780">
<skos:prefLabel>Biblioteche</skos:prefLabel>
<skos:broader rdf:resource="termine?id=1944">Organizzazioni culturali</skos:broader>
<skos:narrower rdf:resource="termine?id=779">Biblioteche speciali</skos:narrower>
<skos:narrower rdf:resource="termine?id=9221">Emeroteche</skos:narrower>
<skos:narrower rdf:resource="termine?id=9588">Biblioteche pubbliche</skos:narrower>
<skos:narrower>
<skos:collection>
<rdfs:label tipo="20662">[Biblioteche secondo l'appartenenza]</rdfs:label>
<skos:member rdf:resource="termine?id=7449">Biblioteche private</skos:member>
<skos:member rdf:resource="termine?id=9582">Biblioteche ecclesiastiche</skos:member>
<skos:member rdf:resource="termine?id=9586">Biblioteche comunali</skos:member>
<skos:member rdf:resource="termine?id=9597">Biblioteche statali</skos:member>
<skos:member rdf:resource="termine?id=9607">Biblioteche provinciali</skos:member>
<skos:member rdf:resource="termine?id=9609">Biblioteche regionali</skos:member>
</skos:collection>
</skos:narrower>
<skos:narrower>
<skos:collection>
<rdfs:label tipo="20664"> [Biblioteche secondo le finalità, gli obiettivi e le funzioni]</rdfs:label>
<skos:member rdf:resource="termine?id=7394">Biblioteche circolanti</skos:member>
<skos:member rdf:resource="termine?id=9378">Gabinetti di lettura</skos:member>
<skos:member rdf:resource="termine?id=9606">Biblioteche popolari</skos:member>
<skos:member rdf:resource="termine?id=9612">Biblioteche specializzate</skos:member>
<skos:member rdf:resource="termine?id=22007">Posti di prestito</skos:member>
</skos:collection>
</skos:narrower>
<skos:narrower>
<skos:collection>
<rdfs:label tipo="20665">[Biblioteche secondo l'utenza]</rdfs:label>
<skos:member rdf:resource="termine?id=9584">Biblioteche carcerarie</skos:member>
<skos:member rdf:resource="termine?id=9598">Biblioteche per ciechi</skos:member>
<skos:member rdf:resource="termine?id=9603">Biblioteche per ragazzi</skos:member>
<skos:member rdf:resource="termine?id=9610">Biblioteche scolastiche</skos:member>
<skos:member rdf:resource="termine?id=9616">Biblioteche ospedaliere</skos:member>
<skos:member rdf:resource="termine?id=9617">Biblioteche parlamentari</skos:member>
<skos:member rdf:resource="termine?id=9621">Biblioteche universitarie</skos:member>
</skos:collection>
</skos:narrower>
<skos:altLabel xml:lang="en">Libraries</skos:altLabel>
<skos:altLabel>Impiego nelle biblioteche</skos:altLabel>
<skos:altLabel>Biblioteche nella letteratura narrativa</skos:altLabel>
<skos:altLabel>Biblioteche aziendali</skos:altLabel>
<skos:altLabel>Biblioteche degli enti locali</skos:altLabel>
<skos:altLabel>Biblioteche ministeriali</skos:altLabel>
<skos:altLabel>Biblioteche conventuali</skos:altLabel>
<skos:altLabel>Animatori di biblioteche</skos:altLabel>
<skos:related rdf:resource="termine?id=303">Libri</skos:related>
<skos:related rdf:resource="termine?id=584">Documenti</skos:related>
<skos:related rdf:resource="termine?id=851">Bibliotecari</skos:related>
<skos:related rdf:resource="termine?id=1818">Biblioteconomia</skos:related>
<skos:related rdf:resource="termine?id=3112">Mediateche</skos:related>
<skos:related rdf:resource="termine?id=9108">Cataloghi bibliografici</skos:related>
<skos:related rdf:resource="termine?id=9860">Inventari</skos:related>

```

²⁸ *Linee guida per i siti web della PA* http://www.innovazionepa.gov.it/media/367125/linee_guida_siti_web_pa.pdf

```

<skos:related rdf:resource="termine?id=10323">Soprintendenze bibliografiche</skos:related>
<skos:related rdf:resource="termine?id=21523">Fondi speciali</skos:related>
<skos:related rdf:resource="termine?id=27835">Edifici per la cultura</skos:related>
<skos:related rdf:resource="termine?id=30683">Biblioteche digitali</skos:related>
<skos:historyNote>Precedentemente sia al singolare che plurale: Biblioteca (Soggettario); Biblioteche (Soggettario);
precedentemente anche con il significato di Biblioteche private di persone, famiglie, ecc. </skos:historyNote>
<skos:example>PARTE/PROPR. Segue i termini che indicano un ente (in questi casi nella forma variante Biblioteca),
p.e."Accademia dei Geogofili - Biblioteca" (CG)</skos:example>
<skos:editorialNote> FONTE: EB (Soggettario); VLI; PT; DeM; GBV; GBD; DDC22</skos:editorialNote>
</skos:Concept>
</rdf:RDF>

```

Riferimenti bibliografici

1. American National Standard Institute - National Information Standards Organization. AN-SI/NISO Z39.19 [2005], *Guidelines for the construction, format and management of monolingual controlled vocabularies*, Bethesda (USA), NISO Press, 2005.
2. BALLESTRA, Laura [2008], *Multiculturalità e thesauri multilingue: problemi e prospettive alla luce delle Guidelines for multilingual thesauri di IFLA*. Intervento all'11. Workshop Teca del Mediterraneo, 2008, <http://www.bcr.puglia.it/tdm/documenti/workshop/2008/Ballestra.pdf>
3. Biblioteca Nazionale Centrale di Firenze (BNCF) [2006], *Nuovo soggetto. Guida al sistema italiano di indicizzazione per soggetto. Prototipo del Thesaurus*, Milano, Editrice bibliografica, 2006.
4. Biblioteca nazionale centrale di Firenze (BNCF) [1956], *Soggettario per i cataloghi delle biblioteche italiane*, Firenze, Stamperia Il cenacolo, 1956.
5. BERGAMIN, Giovanni [2007], *Speciale Nuovo soggetto. Applicazioni web*, in *Biblioteche oggi*, 25 (2007), n. 6, p. 94-95.
6. British Standards Institution (BSI), BS 8723:2005-2008, *Structured vocabularies for information retrieval*, London: British Standards Institution, 2005-2008.
7. BROUGHTON, Vanda [2008], *Costruire thesauri. Strumenti per indicizzazione e metadati semantici*, a cura di Piero Cavaleri, Laura Ballestra, Luisa Venuti, Milano, Editrice bibliografica, 2008.
8. CAVALERI, Piero [2009], *Il thesauri di economia e scienze sociali della Biblioteca Rostoni e SKOS*. Intervento al convegno "I thesauri tra cataloghi e Web", Firenze, Istituto degli Innocenti, 6 Febbraio 2009, <http://www.iskoi.org/doc/thesauri4.htm>
9. CHETI, Alberto – PARADISI, Federica [2008], *Facet analysis in the development of a general controlled vocabulary*, in *Axiomathes* 18 (2008), n. 2 giugno 2008, p. 223–241.
10. FRANCI, Luciana – LUCARELLI, Anna – MOTTA, Marta – ROLLE, Massimo [2009], *The Nuovo soggetto Thesaurus: Structural features and web application projects*. Intervento al Florence, Satellite Meeting "Looking at the Past and Preparing for the Future", IFLA 2009.
11. GNOLI, Claudio, MARINO, Vittorio, ROSATI, Luca [2006] *Organizzare la conoscenza : dalle biblioteche all'architettura dell'informazione per il web*, Milano, Hops Tecniche nuove, 2006
12. GNOLI, Claudio – SCOGNAMIGLIO, Carlo [2008], *Ontologia e organizzazione della conoscenza : introduzione ai fondamenti teorici dell'indicizzazione semantica*, Lecce, Pensa multimedia, 2008
13. HARPER, Corey A. [2006], *Encoding Library of Congress Subject Headings in SKOS: Authority Control for the Semantic Web*, 2006, <https://scholarsbank.uoregon.edu/xmlui/bitstream/handle/1794/3268/dc2006.pdf>; [jsessionId=DB77B536B4ED99629484CD4905B4BA28?sequence=1](https://scholarsbank.uoregon.edu/xmlui/bitstream/handle/1794/3268/dc2006.pdf)
14. International organization for standardization. ISO 2788/1986, *Documentation. Guidelines for the establishment and development of monolingual thesauri*, [Geneva]: ISO, 1986.
15. *Library of Congress Subject Headings(LCSH)* <http://id.loc.gov/authorities/>

16. *Linee guida per i siti web della PA*, http://www.innovazionepa.gov.it/media/367125/linee_guida_siti_web_pa.pdf
17. Linked Data, [urlhttp://linkeddata.org/home](http://linkeddata.org/home)
18. *LIUC [Thesaurus]*, <http://www.biblio.liuc.it/zwebsvr/zetesis.asp?WCI=Browse&WCE=BROWSE&WCU=THESAURUS>
19. LUCARELLI, Anna [2008], *Quando una collezione speciale chiede ospitalità ad una grande Biblioteca*, in *Piccoli scritti di Biblioteconomia per Luigi Crocetti. 10 marzo 2007-10 marzo 2008*, cura di C. Cavallaro, P. Innocenti, Manziana (Roma), Vecchiarelli, 2008, p. 183-201.
20. LUCARELLI, Anna [2009], *Nuovo soggettoario: una sintesi sullo stato dell'arte*, in *Rapporto sulle biblioteche italiane 2007-2008*, a cura di Vittorio Ponzani, Roma: AIB, 2009, p. 108-114.
21. *Medical subject headings (MESH)*, <http://www.ncbi.nlm.nih.gov/mesh>
22. MILES, ALISTAIR [2008], *The Simple Knowledge Organization System (SKOS) in the Context of Semantic Web Deployment*, 2008, http://www.iskouk.org/presentations/miles_web_and_skos_200807.pdf
23. MOTTA, Marta - TIBERI, Melissa [2009], *Riflessi dello standard britannico BS 8723 nel Thesaurus del Nuovo soggettoario* in "Bollettino AIB" 49 (2009), n. 3, p. 325-340.
24. *Multilingual Access to Subjects (MACS)*, <https://macs.vub.ac.be/pub>
25. NKOS (Networked KOS), <http://nkos.slis.kent.edu/>
26. *Répertoire d'autorité-matière encyclopédique et alphabétique unifié (RAMEAU)* <http://rameau.bnf.fr/>
27. RIDI, Riccardo [2010] *Il mondo dei documenti*, Roma-Bari, Laterza, 2010
28. SIGNORE, Oreste [2003], *Strutturare la conoscenza: XML, RDF, Semantic Web*, 2003 <http://dit.unitn.it/~bouquet/kr-2005/materiale/ck2003.pdf>
29. *Simple Knowledge Organization System (SKOS)*, <http://www.w3.org/2004/02/skos/>
30. SUMMERS, Ed, ISAAC, Antoine, REDDING, Clay, KRECH, Dan [2008], *LCSH, SKOS and Linked Data*, 2008, <http://arxiv.org/abs/0805.2855>
31. SVENONIUS, Elaine [2000], *The Intellectual foundation of information organization*, Cambridge, Mass, The MIT Pres, 2000.
32. *TELPlus*, <http://www.theeuropeanlibrary.org/portal/organisation/cooperation/telplus/index.php>
33. VAN ASSEM, Mark, MALAISE, Veronique, MILES, Alistair, SCHREIBER, Guus [2006], *A Method to Convert Thesauri to SKOS*, 2006, <http://www.cs.vu.nl/~mark/papers/Assem06b.pdf>
34. W3C, *SKOS use cases and requirements*, 19 Agosto 2009, <http://www.w3.org/skos-ucr>
35. W3C, *SKOS Simple Knowledge organization system. Reference*, 18 Agosto 2009, <http://www.w3.org/skos-reference>
36. *Zthes*, <http://zthes.z3950.org/>