

# L'applicazione del Nuovo Soggettario in SBN

## Introduzione

L'obiettivo del Gruppo è stato quello di individuare le modalità per rendere possibile la più ampia **diffusione ed applicazione nella rete SBN** del Nuovo Soggettario (2007) che, non soltanto dovrà **sostituire** il Soggettario (1956), ma per la sua stessa impostazione, si configura come un Thesaurus generale, uno strumento di riferimento a livello nazionale per qualsiasi tipo di materiale che tende a svilupparsi acquisendo anche l'esperienza e la collaborazione di biblioteche, Università e di altri centri di eccellenza.

La consapevolezza della complessità della rete SBN, sia per la quantità e la diversa tipologia delle biblioteche aderenti, sia per una tradizione di "autonomia" dei poli e delle biblioteche nella catalogazione semantica, ha prospettato come unica strada percorribile quella del passaggio graduale all'utilizzo del nuovo strumento, garantito dalla massima **compatibilità** e possibilità di coesistenza con quanto finora prodotto (o che potrà ancora essere prodotto) mediante l'utilizzo del Soggettario (1956).

Il Nuovo Soggettario in effetti non si configura come uno strumento completamente diverso dal precedente, ma come il prodotto di una sua evoluzione: molte stringhe di soggetto create con il vecchio strumento saranno valide anche per il nuovo e gran parte della terminologia del vecchio strumento è confluita nel nuovo.

Parallelamente si è effettuata una riflessione sulle nuove possibilità offerte dal **protocollo SBNMARC**, sviluppate nella prospettiva di rendere sempre più l'Indice SBN un'**infrastruttura** al servizio delle biblioteche partecipanti.

Nell'Indice SBN, prima della sua evoluzione, i soggetti e le classi erano gestiti come fossero attributi del documento: non era infatti prevista né la funzione di ricerca del soggetto (o della classe) né la possibilità di un trasferimento da Indice a Polo di queste informazioni, se non a partire dal record del documento cui erano legate.

Con il progetto di Evoluzione dell'Indice SBN si sono implementate nuove funzionalità - come la possibilità di alimentare i diversi archivi di Authority (autori, marche, soggetti, etc.) indipendentemente da eventuali legami a documenti - al fine di offrire alle biblioteche un maggior numero di **strumenti di ausilio per la catalogazione**: tale obiettivo ha orientato tutte le scelte effettuate per lo sviluppo del nuovo protocollo di colloquio SBNMARC, come,

ad esempio, quella di registrare e trasmettere le immagini legate alle marche, o i dati specifici dei record di Autorità precedentemente gestiti su una base dati autonoma e non trasmessi dal protocollo SBN.

Così anche nel settore della catalogazione semantica, ed in particolare per i soggetti, sono state previste le seguenti, nuove funzionalità:

- ricerca del soggetto e possibilità di navigazione sui titoli collegati;
- ricerca del descrittore e possibilità di visualizzare i soggetti collegati;
- eventuale importazione su Polo di un soggetto/i indipendentemente dal legame al titolo per poterlo riutilizzare nella soggettazione di altri documenti;
- creazione e modifica di stringhe di soggetto condivise con l'Indice<sup>1</sup>.

### **1. Gestione del Soggettario (1956) e del Nuovo Soggettario nell'Indice SBN**

Partendo dalle considerazioni sopra esposte - ed in particolare prevedendo la coesistenza in Indice di stringhe formulate in base ai due diversi strumenti - si è ritenuto opportuno, dal punto di vista applicativo e funzionale, far confluire in un unico archivio tutte le stringhe di soggetto, indicando con il codice Tipo Soggettario lo strumento utilizzato per la loro formulazione: FI (Soggettario 1956) e FN (Nuovo Soggettario 2007); si è ritenuto inoltre di non duplicare stringhe con uguale contenuto e diverso codice soggettario. Infatti, l'eventuale duplicazione delle stringhe, oltre a comportare dispendio di energie e di tempo per i catalogatori, impedirebbe loro la consultazione, attraverso le funzioni di navigazione, di tutti i titoli con lo stesso contenuto semantico, costringendo ad una doppia interrogazione. Le stringhe che rimarranno invariate nel passaggio dal vecchio al nuovo strumento verranno identificate da un altro codice FE (Firenze Entrambi).

In merito all'estrazione dei descrittori verrà effettuata una modifica nella procedura di Indice per riconoscere le parentesi quadre, aperta e chiusa, come segno di separazione tra un descrittore e l'altro; i connettivi racchiusi tra parentesi quadre non saranno registrati come descrittori.

Per favorire la transizione dal vecchio al Nuovo Soggettario la scelta tecnico/politica è stata quella di non precludere, attraverso controlli vincolanti da parte dell'applicativo di Indice, l'utilizzo, anche contemporaneo, dei due soggettari da parte dei Poli. Infatti alcuni Poli potrebbero trovarsi nella situazione di avere contemporaneamente biblioteche in

---

<sup>1</sup> Il trattamento dei soggetti sarebbe in tutto simile a quello degli autori se non fosse per il fatto che i soggetti, a differenza degli autori, non sono localizzati per gestione in Indice e non sono quindi oggetto di allineamento, mentre sono oggetto di allineamento i legami tra documenti e soggetti, nel caso in cui la configurazione del Polo preveda l'invio da parte dell'Indice dei legami stessi

grado di applicare il Nuovo Soggettario e biblioteche il cui personale, non ancora adeguatamente formato, continuerebbe ad utilizzare il Soggettario del '56. Pertanto l'indicazione, da parte del Polo, sul Tipo soggettario utilizzato sarà utile come informazione sulla stringa di soggetto e per altre funzioni, come ad esempio, la realizzazione di statistiche, etc.; ma non comporterà automaticamente controlli da parte dell'Indice. Sarà lasciata agli applicativi di Polo la scelta di utilizzare o meno filtri e controlli in relazione al codice Tipo di soggettario.

Le modalità e gli strumenti per la realizzazione di questi obiettivi sono descritte nel paragrafo 1 dell'Allegato A.

## **2. Importazione in Indice di alcune delle relazioni fra termini presenti nel Thesaurus del Nuovo Soggettario**

Nel corso dell'attività del gruppo la Biblioteca nazionale centrale di Firenze ha ribadito la sua disponibilità a fornire all'Indice le relazioni di rinvio, di scomposizione e di variante storica create nel Thesaurus. Grazie a tale disponibilità l'Indice SBN potrà offrire un importante **sussidio alle biblioteche** per la catalogazione semantica, promuovendo **l'utilizzo diffuso** del nuovo strumento.

Poiché le relazioni tra descrittori non sono gestite nel colloquio polo-Indice, è stata sviluppata una procedura che, a partire da un file testuale inviato dalla BNCF, ha già caricato sperimentalmente in Indice, in modalità batch, le relazioni sopra citate.

La procedura è stata lanciata in ambiente di test ed ha inserito in base dati **7.849** nuovi descrittori e **11.586** nuovi legami descrittore-descrittore

Le modalità e gli strumenti per la realizzazione di questo obiettivo sono descritte nel paragrafo 2 dell'Allegato A

## **3. La cooperazione in Indice nell'ambito della catalogazione semantica**

La diffusione del Nuovo Soggettario e le ulteriori funzionalità presenti nel colloquio Polo / Indice con il Protocollo SBNMARC, descritte nella parte introduttiva, concorrono a favorire un processo di maggiore "standardizzazione" e condivisione dei record nella catalogazione semantica in SBN. Pur rispettoso dell'autonomia e delle specificità delle biblioteche in questo ambito, e consapevole delle difficoltà che esse potrebbero avere nell'abbandonare eventuali altri strumenti di indicizzazione, il gruppo auspica che questa tendenza vada affermandosi sempre di più nella rete.

Perché ciò possa accadere è però necessario affrontare alcuni aspetti quali:

- la formazione dei bibliotecari sulle regole del nuovo strumento;

- la necessità di dare ai bibliotecari che soggettano in Indice delle indicazioni di carattere tecnico-biblioteconomico su alcune attività che possono svolgere nel catalogo centrale, come ad esempio sulle correzioni di stringhe di soggetto; alcune di queste indicazioni sono descritte nell'Allegato 2; le indicazioni relative alle attività di allineamento dei Poli a fronte di correzioni su soggetti presenti in Indice sono descritte nell'allegato 1;
- la proposta di istituzione di una "Struttura" che, similmente a quella a suo tempo creata per la pulizia dell'archivio titoli e dell'archivio autori dell'Indice SBN, si faccia carico di una serie di problemi connessi alla catalogazione semantica in Indice;
- la necessità di rivedere ed ampliare la gerarchia dei Poli che possono inviare soggetti in Indice e che dovrebbero indicare i componenti della nuova struttura.

Per quanto riguarda questo ultimo aspetto, pur restando valida la scala di priorità dei Poli decisa nel 1997 in occasione della nascita dell'OPAC SBN, si ritiene tuttavia necessario riesaminare il problema alla luce dei cambiamenti avvenuti nell'arco di 12 anni. Da un'indagine effettuata ai primi di Aprile 2009, risulta che l'Indice contiene 9.887.346 documenti, divisi in 8.848.192 monografie di base, 712.520 volumi senza titolo significativo e 326.634 periodici. I soggetti sono in tutto 1.244.190 e i legami titolo/soggetto 2.014.709. Anche ipotizzando (ma sappiamo che non è così) che ad ogni notizia corrisponda un unico soggetto, il rapporto tra la quantità di documenti soggettati rispetto alla quantità complessiva dei documenti è poco più del 20%. E' stato verificato inoltre che, da gennaio 2008 ad aprile 2009 sono state effettuate in tutto 73.479 attività di creazione soggetto; indubbiamente il dato non può indicare in assoluto una diminuzione delle attività di soggettazione in Indice, poiché parte dei documenti potrebbe essere stata soggettata con stringhe già esistenti; ma l'esiguità delle creazioni effettuate è comunque un elemento significativo. In base a tali indagini il Gruppo ritiene necessario che si riveda e si ampli la lista dei Poli soggettatori/classificatori; l'introduzione del Nuovo Soggettario può essere l'occasione per rilanciare l'importanza della Catalogazione semantica in SBN e creare una prima rosa di nuovi Poli intenzionati ad alimentare il catalogo collettivo, dopo accurati corsi di formazione. Per quanto riguarda il livello di partecipazione, il protocollo SBNMARC, tramite il controllo delle abilitazioni, consente già di distinguere tra le attività di Creazione di un Authority (Soggetti) e la creazione di un legame titolo/soggetto con un soggetto già esistente e, tramite i livelli di autorità, consente di diversificare i Poli che possono o meno intervenire su soggetti già presenti.

Anche se tali aspetti non facevano parte del mandato attribuito al Gruppo di lavoro, si è ritenuto comunque utile farli presenti e proporli al Comitato di gestione SBN come oggetto di discussione di un'eventuale altra riunione. In questo documento il Gruppo ha appena abbozzato alcune proposte, ma ritiene che gli argomenti accennati debbano essere ulteriormente analizzati ed approfonditi in sedi specifiche.

## **Allegato 1**

### **Indicazioni tecniche per gli sviluppatori degli applicativi di Polo**

Alla luce delle considerazioni esposte nel documento diretto al Comitato di gestione SBN, il Gruppo di lavoro ha elaborato le seguenti indicazioni relative alla gestione del Soggettario (1956) e del Nuovo Soggettario nell'Indice SBN.

A tali indicazioni dovranno adeguarsi gli applicativi di Polo e il client di Interfaccia Diretta nell'ambito del colloquio con l'Indice SBN.

#### **1. Gestione del Soggettario (1956) e del Nuovo Soggettario nell'Indice SBN**

1. nel catalogo a soggetto dell'Indice SBN coesisteranno stringhe di soggetto formulate sia secondo le regole del Soggettario (1956) che del Nuovo Soggettario; i record relativi a soggetti presenteranno in base dati Indice, come già previsto, il codice Tipo Soggettario che potrà assumere i seguenti valori: "FI"=Soggettario 1956, "FN"=Nuovo Soggettario o "FE"= Firenze-Entrambi, nel caso in cui la formulazione delle stringhe risulti uguale con entrambi gli strumenti;
2. differentemente da quanto avviene oggi, il codice Tipo soggettario, come attributo del soggetto, dovrà essere modificabile per variare il valore del codice (da FI a FE) via via che stringhe già esistenti verranno "recuperate" ed utilizzate per la catalogazione semantica con il nuovo strumento; anche le nuove stringhe che saranno inserite con il codice FN (Nuovo Soggettario) potranno essere modificate con il codice FE, ma si auspica che siano già create con tale valore, se valide per tutti e due gli strumenti;
3. per ottenere tale scopo si prospettano le seguenti ipotesi:
  - a. in fase di inserimento da parte di un Polo di una stringa di soggetto con codice Tipo soggettario FI o FN, l'Indice, se trova nel proprio archivio la stessa stringa con diverso codice Tipo soggettario, rifiuta l'inserimento, modificando in FE il codice tipo soggettario del record già presente e rinviandolo al Polo.
  - b. la modifica del codice Tipo soggettario non sarà oggetto di una nuova specifica richiesta del protocollo SBNMARC, ma utilizzerà la richiesta già esistente di Modifica soggetto. In base all'abilitazione del Polo, l'Indice aggiornerà sia la stringa che il codice tipo soggettario, se il Polo è abilitato alla modifica del soggetto, fatti

salvi i controlli sul livello di autorità; oppure aggiornerà soltanto il codice tipo soggettario, se il Polo è abilitato alla creazione, ma non alla modifica del soggetto.

- c. Per quanto riguarda i descrittori, che non sono inviati dai Poli all'Indice, ma calcolati autonomamente dal sistema centrale, si ritiene opportuno che, per evitare duplicazioni, sia l'Indice ad attribuire ad essi il codice Tipo soggettario (FI, FN o FE) in base al codice delle stringhe dalle quali sono estratti, o modificandone il valore in FE, nel caso in cui il descrittore estratto sia presente sia in stringhe formulate sulla base del Soggettario 1956 che in stringhe formulate con il Nuovo Soggettario.
4. In merito all'estrazione dei descrittori verrà effettuata una modifica nella procedura di Indice per riconoscere le parentesi quadre, aperta e chiusa, come segno di separazione tra un descrittore e l'altro; i connettivi racchiusi tra parentesi quadre non saranno registrati come descrittori.
5. Sia il protocollo SBN che il protocollo SBNMARC saranno messi in grado di ricevere soggetti con la codifica FI, FN e FE; ne consegue che, mentre gli applicativi con Protocollo SBNMARC, che già gestiscono tale informazione, dovranno soltanto aggiungere i nuovi valori ammessi, gli applicativi con protocollo SBN dovranno necessariamente predisporre all'invio ed alla ricezione del codice Tipo soggettario nel colloquio con l'Indice; non sarà invece obbligatorio per tali applicativi sviluppare nuovi algoritmi per l'estrazione dei descrittori originati da stringhe di soggetto basate sul Nuovo Soggettario;
6. L'indicazione, da parte degli applicativi di Polo, sul Tipo di soggettario utilizzato non comporterà automaticamente controlli da parte dell'Indice: così, in fase di ricerca / cattura soggetti in Indice, sarà il Polo a decidere se applicare o meno il filtro e il controllo sul codice soggettario utilizzato; in fase di invio della notizia analitica dall'Indice ai Poli non vi sarà alcun "ritaglio" o mascheratura di soggetti, ma verranno resi visibili tutti i soggetti legati alla notizia: si ritiene infatti opportuno che i Poli abbiano in visione anche soggetti di tipo diverso da quelli corrispondenti al proprio profilo di gestione, per avere completezza informativa; anche in questo caso, in fase di Interrogazione notizia sarà a cura del Polo applicare il controllo per catturare solo i soggetti collegati di tipo corrispondente al proprio profilo.
7. Appena le nuove funzionalità saranno disponibili in Indice e sui Poli, con un processo batch verrà attribuito il codice FN ai soggetti inseriti dalla BNCF a partire dal luglio 2007 con stringhe formulate sulla base del Nuovo Soggettario e ne saranno estratti i

descrittori con l'algoritmo appositamente sviluppato che tiene conto della nuova punteggiatura;

## **2. Relazioni tra descrittori: Importazione in Indice di alcune delle relazioni fra termini presenti nel Nuovo Soggettario**

La procedura di caricamento batch dei legami fra descrittori effettua le seguenti attività:

- normalizza le descrizioni calcolando le chiavi dei descrittori
- associa alle chiavi normalizzate gli identificativi di descrittori già presenti in Indice con stessa chiave
- assegna nuovi identificativi per chiavi descrittori non presenti
- inserisce i nuovi descrittori nella relativa tabella
- inserisce i nuovi legami nella tabella dei legami tra descrittori

La procedura è stata lanciata in ambiente di test ed ha inserito in base dati **7.849** nuovi descrittori (ovviamente non collegati a soggetti) e **11.586** nuovi legami descrittore-descrittore

Le relazioni inviate dalla BNCF riguardano i legami di rinvio (USE/UF), di variante storica (HSF/HSF+) e di scomposizione (USE+/UF+); il protocollo SBNMARC attualmente invia le relazioni di un descrittore soltanto se si interroga un soggetto che lo include mentre non le fornisce in risposta ad una ricerca del descrittore; si richiederà un intervento evolutivo sul protocollo SBNMARC per ottenere la risposta più esaustiva nella fase di consultazione dei descrittori.

Poiché, come già esposto, le relazioni tra descrittori sono precaricate sulla base dati e riportate solo nei messaggi di risposta da parte dell'Indice, gli applicativi di Polo sono liberi di implementare i nuovi valori (HSF/HSF+ e USE+/UF+) rendendo possibile la consultazione di tali relazioni, oppure di limitarsi alla visualizzazione della stringa di soggetto, senza prospettare i descrittori e le loro relazioni.

## **3. Modifica della stringa di soggetto**

La funzione di modifica sarà oggetto di un'apposita abilitazione che verrà attribuita secondo criteri da stabilire: quali poli, in base a quale livello di autorità del soggetto, confrontato con il livello di autorità del Polo sui soggetti, etc; sarà possibile correggere

soggetti inseriti dal polo stesso o utilizzati come legami a notizie possedute. Negli altri casi, se necessaria, la modifica di un soggetto sarà operata tramite le funzionalità di Interfaccia Diretta.

Si fa presente che nei test per la certificazione di conformità al protocollo SBNMARC la modifica del soggetto non è mai stata oggetto di verifica. Pertanto gli applicativi che vorranno dotarsi di tale funzionalità, potranno contattare l'ICCU, che incaricherà il Gruppo per il mantenimento del protocollo SBNMARC di predisporre dei casi di test appropriati.

#### **4. Allineamenti a fronte di una correzione della stringa di soggetto**

Dal punto di vista dell'applicazione e del colloquio Polo/Indice, la modifica di una stringa di soggetto in Indice non genera allineamenti in quanto i soggetti non sono localizzati per gestione; tuttavia l'informazione verrà comunque comunicata indirettamente: infatti, poiché il soggetto fa parte del reticolo del titolo, l'Indice accenderà i flag di allineamento per variazione legami per tutti i titoli legati e i poli che li possiedono. I Poli che nel loro profilo hanno scelto la possibilità di ricevere i soggetti nel reticolo del documento, riceveranno il nuovo reticolo con il soggetto variato. Il polo potrà scegliere se ignorare o meno i legami semantici inviati dagli allineamenti.

Un'altra possibilità data al Polo è quella di acquisire le variazioni effettuate sui soggetti di Indice richiedendo da Interfaccia Diretta l'export SBNMARC dei soggetti collegati a titoli gestiti dal proprio polo e modificati in un intervallo temporale di interesse.

Infine l'interrogazione dei soggetti, prima in Polo e poi in Indice, come accade per gli autori, può evidenziare modifiche in Indice sulle quali il Polo può via via allinearsi, "catturando" la versione più aggiornata.

Tenendo conto delle raccomandazioni volte ai bibliotecari e presentate nell'Allegato 2, si auspica che le attività di correzione di stringhe di soggetto in Indice siano ridotte a quelle effettivamente necessarie.

## Allegato 2

### Indicazioni ai bibliotecari che soggettano in Indice

#### Modifica della stringa di soggetto

Per quanto riguarda la funzione di modifica di una stringa di soggetto in Indice, il gruppo concorda sulle seguenti indicazioni:

- 1) la funzione di modifica sarà oggetto di un'apposita abilitazione che verrà attribuita secondo criteri da stabilire: quali poli, in base a quale livello di autorità del soggetto, confrontato con il livello di autorità del Polo sui soggetti, etc;
- 2) sarà possibile correggere soltanto soggetti inseriti dal polo stesso o utilizzati come legami a notizie possedute. Negli altri casi, se necessaria, la modifica di un soggetto sarà operata tramite le funzionalità di Interfaccia Diretta.
- 3) per i poli dotati dell'abilitazione a correggere, si ribadisce l'indicazione di evitare, se possibile, di effettuare eccessive modifiche e di attenersi alle seguenti regole di comportamento:
  - correzioni di tipo ortografico (errori di battitura, normalizzazione dei caratteri)

Es:

*roma – Guide* → *Roma – Guide*

*vicennza -periodici* → *Vicenza - Periodici*

la correzione è sempre possibile;

- inesattezze nella costruzione formale del soggetto

Es. :

*Vicenza – Chiesa di Santa Corona* → *Vicenza – Santa Corona*

*Medici, Caterina : de' <regina di Francia>* → *Caterina <regina di Francia>*

la correzione è sempre possibile;

- se l'intervento comporta una modifica nell'ordine dei termini presenti nella stringa (ordine di citazione) bisogna verificare
  1. che la stringa sia o meno formalmente corretta
  2. che non ci siano altri documenti collegati.

Se la stringa è formalmente corretta e vi sono legami con altri documenti si cancella il legame errato e si crea un nuovo soggetto. Se vi è un unico legame con il documento che si sta trattando si può invece modificare la stringa

Se la stringa è formulata in modo errato si torna al punto precedente (inesattezze nella costruzione formale del soggetto)

- se l'intervento comporta un cambiamento nel contenuto semantico occorre verificare in Indice i titoli collegati al soggetto in questione:
  - a. se il soggetto è legato al solo titolo sul quale si sta lavorando o se, pur essendo legato ad altri titoli, questi sono posseduti dalla biblioteca e la sua correzione sarebbe giusta e valida anche per questi, la correzione si può effettuare;
  - b. se invece il soggetto è concettualmente errato per il titolo sul quale si sta lavorando, ma corretto per altri titoli ai quali è collegato, si deve cancellare il legame esistente e legare il titolo ad un nuovo soggetto; in conclusione, si vuole evitare che la modifica di una stringa di soggetto, finalizzata alla corretta catalogazione semantica di un titolo, si rifletta negativamente sugli altri titoli ai quali è collegato, stravolgendone il contenuto concettuale