

## Soggetti: migrare senza fermarsi al 1956

di Andrea Fabbrizzi<sup>1</sup>

A cinquant'anni dalla pubblicazione del *Soggettario*, l'indicizzazione per soggetto in Italia sta voltando pagina, perché da tempo la Biblioteca nazionale di Firenze ha messo in cantiere il *Nuovo Soggettario*, col quale l'indicizzatore non dovrà più seguire un procedimento di tipo analogico sulla base di un repertorio di voci, nell'assenza di criteri esplicitamente formulati, ma potrà invece contare su una rigorosa distinzione tra sintassi e semantica, e su principi e criteri logici per la formulazione delle stringhe di soggetto e per la costruzione del vocabolario controllato.

Unità di base del nuovo linguaggio è il *termine di indicizzazione*, cioè la rappresentazione lessicale di un concetto unitario, oppure il riferimento a un'entità individuale. Su di esso si fondano sia la costruzione della stringa di soggetto, sia il controllo terminologico, strutturato per ogni termine di indicizzazione in base alle tre relazioni proprie dei thesauri: le relazioni di equivalenza, gerarchiche, associativa.

E' necessario che Aleph sia adeguato a questa nuova modalità di indicizzazione, che con l'adozione da parte della Biblioteca nazionale di Firenze si diffonderà nelle biblioteche italiane. Questo obiettivo può essere realizzato nel rispetto dei formati internazionali Unimarc per lo scambio di dati bibliografici e di autorità.

In *Unimarc/Authorities* e *Unimarc/Bibliographic* per le diverse categorie di entità c'è piena corrispondenza tra i campi coinvolti nell'espressione della relazione di soggetto, e identica distinzione in sottocampi delle intestazioni:

<b>Categorie di entità</b>	<b>Tag U./Authorities</b>	<b>Tag U./Bibliographic</b>
Nome di persona	200	600
Nome di ente collettivo	210	601
Nome geografico	215	607
Nome di marca commerciale	216	616
Nome di famiglia	220	602
Titolo uniforme	230	605
Nome/titolo	240	604
Soggetto generico	250	606
Luogo formalizzato	260	620
Forma	280	608

Nel *Nuovo Soggettario* l'attività di controllo si effettua esclusivamente sui singoli termini di indicizzazione, non sulle stringhe di soggetto. Dunque, nell'archivio di autorità i sottocampi per le suddivisioni (\$x, \$y, \$z, \$j) non devono essere compilati, anche se i campi del blocco 2xx di *Unimarc/Authorities*, come quelli corrispondenti del blocco 6xx di *Unimarc/Bibliographic*, li prevedono.

Nell'archivio bibliografico e nell'archivio di autorità, rispettivamente le stringhe di soggetto e i termini di indicizzazione trovano posto nel campo relativo alla categoria propria dell'entità indicata

---

<sup>1</sup> Università di Firenze, Gruppo di lavoro per il Progetto migrazione ad Aleph.

o rappresentata; per le registrazioni nell'archivio bibliografico il campo compilato è quello corrispondente alla categoria dell'entità che si trova in prima posizione.

Alcuni esempi. Per la stringa di soggetto:

Leopardi, Giacomo – Lettere e carteggi

nella quale il primo termine di indicizzazione è un nome proprio per un'entità individuale che è una persona (Leopardi, Giacomo), e il secondo termine di indicizzazione è un'espressione per un concetto formale, si compilano i seguenti campi rispettivamente nelle registrazioni bibliografiche e di autorità:

**600 #1 \$aLeopardi\$bGiacomo\$jLettere e carteggi**

**200 #1 \$aLeopardi\$bGiacomo**

**280 ## \$aLettere e carteggi**

Per una stringa di soggetto composta da un termine di indicizzazione per un'entità individuale che è un ente collettivo, nell'archivio bibliografico e nell'archivio di autorità si compilano i campi per i nomi di enti collettivi:

**601 12 \$aConcilio di Trento\$c1545-1564**

**210 12 \$aConcilio di Trento\$c1545-1564**

Per la stringa di soggetto:

Enti locali – Personale – Formazione professionale – Italia – Inchieste

che è composta da tre termini di indicizzazione per concetti generici (Enti locali, Personale, Formazione professionale), da un termine per un'entità individuale territoriale (Italia), da un termine per un concetto formale (Inchieste), si compilano i seguenti campi nelle registrazioni dell'archivio bibliografico e dell'archivio di autorità:

**606 ## \$aEnti locali\$xPersonale\$xFormazione professionale\$yItalia\$jInchieste**

**250 ## \$aEnti locali**

**250 ## \$aPersonale**

**250 ## \$aFormazione professionale**

**215 ## \$aItalia**

**280 ## \$aInchieste**

Nell'archivio di autorità il controllo deve avvenire mediante lo sviluppo di un thesaurus sulla base delle *relazioni di equivalenza* (USE, UF), *gerarchiche* (BT, NT), *associativa* (RT).

Nel formato *Unimarc/Authorities* sono previste le entità e le relazioni di un thesaurus strutturato secondo la norma *ISO 2788*. Pertanto, se Aleph è in grado di gestire correttamente gli archivi di autorità secondo *Unimarc/Authorities*, è in grado anche di gestire thesauri a tre relazioni.

Un archivio di autorità può dirsi tale se è “costituito da registrazioni di autorità contenenti, oltre la forma scelta per l’instestazione, le forme non preferite e associate, e note che permettono di identificare la forma preferita, di spiegare le relazioni tra questa e le forme non preferite e associate, di indicare le fonti consultate”.<sup>2</sup> Si dovrebbero pertanto superare eventuali incompletezze di Aleph rispetto ad *Unimarc/Authorities*, in particolare riguardo alle note per il pubblico (campi 3xx), ai rinvii vedi anche (campi 5xx), alle indispensabili note del catalogatore (campi 810-836), alle registrazioni generali esplicative (label, posizione 6=z).

### Compilazione della stringa di soggetto e controllo dei termini di indicizzazione in Aleph

Il controllo di un termine di indicizzazione deve essere efficace per ogni stringa di soggetto di cui esso è parte: quindi, ogni modifica di un termine di indicizzazione nell’archivio di autorità deve riflettersi in ogni stringa di soggetto che contiene quel termine nell’archivio bibliografico.

Per realizzare mediante l’archivio di autorità il controllo dei termini di indicizzazione contenuti nelle stringhe di soggetto registrate nell’archivio bibliografico, si propone che le stringhe di soggetto nel record bibliografico non siano compilate direttamente, ma siano *generate* a partire da tanti *campi intermedi* quanti sono i termini di indicizzazione che dovranno comporre la stringa di soggetto.

Per esempio, per generare la stringa di soggetto:

Enti locali – Personale – Formazione professionale – Italia – Inchieste

nel record bibliografico occorre compilare i seguenti campi intermedi, nello stesso ordine che si vuole ottenere nella stringa di soggetto:

**606 ##\$aEnti locali**

**606 ##\$aPersonale**

**606 ##\$aFormazione professionale**

**607 ##\$aItalia**

**608 ##\$aInchieste**

A questi campi corrispondono cinque registrazioni di autorità nell’archivio di controllo, tre delle quali recano il campo 250 (corrispondente al 606 nell’archivio bibliografico) per concetti generici, una il campo 215 (corrispondente al 607) per un’entità individuale territoriale, e una il campo 280 (corrispondente al 608) per un concetto formale.<sup>3</sup> La corrispondenza tra i campi intermedi per i singoli termini di indicizzazione nel record bibliografico e i campi nelle registrazioni di autorità che contengono gli stessi termini di indicizzazione, rende possibile il controllo dei termini di indicizzazione nel record bibliografico da parte dell’archivio di autorità.

La compilazione dei campi intermedi per i singoli termini di indicizzazione nel record bibliografico può avvenire in tre modi:

(1) cattura del termine di indicizzazione già presente nell’archivio di autorità;

<sup>2</sup> Marcelle Beaudiquez; Françoise Bourdon. *Management and use of name authority files: personal names, corporate bodies and uniform titles: evaluation and prospects*. München: Saur, 1991, p. 28.

<sup>3</sup> Cf. la tabella a p. 1.

- (2) creazione del termine di indicizzazione nell'archivio di autorità e sua cattura nel record bibliografico;
- (3) creazione del termine di indicizzazione nel record bibliografico e sua riscrittura in un nuovo record dell'archivio di autorità.

Una volta ultimata la compilazione dei campi intermedi per i singoli termini di indicizzazione, con apposita istruzione il catalogatore determina la *generazione automatica* della stringa di soggetto completa, in un campo 6xx del record bibliografico composto con i dati contenuti nei campi intermedi. Il tag del campo per la stringa di soggetto corrisponde alla categoria dell'entità registrata nel primo campo intermedio.

In base ai tag dei campi intermedi per i termini di indicizzazione, il programma ricostruisce correttamente i sottocampi del campo generato automaticamente, nella sequenza determinata dall'ordine di compilazione dei campi intermedi. Riprendendo l'esempio precedente, nel campo per la stringa di soggetto (tag 606 corrispondente al primo campo intermedio), si generano i seguenti sottocampi:

Enti locali: tag 606 (concetto) -> sottocampo a (primo elemento)  
 Personale: tag 606 (concetto) -> sottocampo x (suddivisione generica)  
 Formazione professionale: tag 606 (concetto) -> sottocampo x (suddivisione generica)  
 Italia: tag 607 (nome geografico) -> sottocampo y (suddivisione geografica)  
 Inchieste: tag 608 (forma) -> sottocampo j (suddivisione formale):

**606 ##\$aEnti locali\$xPersonale\$xFormazione professionale\$yItalia\$jInchieste**

Qualsiasi modifica nel record di autorità per un termine di indicizzazione registrato in uno o più record bibliografici, deve automaticamente produrre per quei record bibliografici un'istruzione simile a quella con la quale il catalogatore genera il campo per la stringa di soggetto a partire dai campi intermedi per i singoli termini di indicizzazione, determinando l'aggiornamento sia dei corrispondenti campi intermedi che dei campi per le stringhe di soggetto.

### Connettivi

Nelle stringhe di soggetto compilate secondo il *Nuovo Soggettario* sono previste espressioni linguistiche, denominate *connettivi*, che hanno la funzione di favorire la corretta interpretazione della stringa di soggetto; queste espressioni non sono termini di indicizzazione, e non sono controllate nell'archivio di autorità.

Per esempio, nelle seguenti stringhe di soggetto i connettivi sono indicati in grassetto:

Farmaci – Consumo **da parte degli** anziani  
 Diagnostica – Impiego **della** risonanza magnetica nucleare  
 Architettura – **In** relazione **alla** musica  
 Acciai speciali – Elasticità **e** flessibilità

Pertanto, nei campi dei record bibliografici per le stringhe di soggetto, generati automaticamente, dovrebbe essere possibile l'aggiunta manuale di espressioni linguistiche non comprese nei sottocampi corrispondenti ai campi intermedi. Tali espressioni in seguito non dovrebbero essere

soggette a modifiche automatiche nel caso di cambiamenti riguardanti i sottocampi per i termini di indicizzazione.

### Migrazione dei descrittori e dei soggetti

I descrittori SBN (UFID) corrispondono solo in parte ai termini di indicizzazione. Inoltre, tra i descrittori è attualmente impossibile distinguere le diverse categorie di entità (persone, enti ecc.). Quindi in fase di migrazione non è opportuno utilizzare i campi Unimarc standard (200-280; 600-608), che dovrebbero essere compilati a regime.

La migrazione dovrebbe dar luogo a tanti record di autorità quanti sono i descrittori, utilizzando per essi un nuovo campo non standard, es. il *tag 290*, con struttura identica al tag 250 (ad eccezione dei sottocampi \$x, \$y, \$z, \$j, anche in questo caso inutili perché non ci possono essere suddivisioni). Nell'archivio bibliografico i descrittori dovrebbero essere registrati in un unico campo intermedio ripetibile, il tag 610, che è appropriato perché riservato a espressioni singole non controllate. Contemporaneamente, la stringa di soggetto completa (UFIC) dovrebbe essere registrata in un campo 606.

La corrispondenza fra i campi 610 nei record bibliografici e campo 290 nel record di autorità dovrebbe rendere immediatamente possibili le operazioni di controllo sui descrittori e il successivo aggiornamento delle relative stringhe.

La migrazione dei descrittori potrebbe anche avvenire in un archivio di autorità provvisorio esclusivamente dedicato ai descrittori, che in seguito dovrebbe essere progressivamente svuotato, sostituendo i campi 290 per descrittori con gli appropriati campi 200-280 per termini di indicizzazione.

### Inizio del lavoro corrente

Nell'archivio di autorità si avvia la compilazione dei campi standard (200-280); contestualmente, nell'archivio bibliografico si utilizzano campi intermedi con tag standard (600-608) per generare le stringhe di soggetto nei campi 6xx.

Una routine riconosce nell'archivio di autorità (oppure nei due archivi di autorità, uno dei quali provvisorio) due forme identiche in campi diversi e, (1) elimina il campo 290, (2) nei record bibliografici collegati replica il termine di indicizzazione nel campo corrispondente di Unimarc/Bibliographic al posto del descrittore precedentemente in 610, e dà l'istruzione di aggiornare o sostituire con altri campi 6xx i campi 606 in cui sono registrate le stringhe collegate.

### Tag dei campi intermedi

Se necessario, i campi intermedi che servono al controllo dei termini di indicizzazione e alla generazione della stringa di soggetto, ma che non sono oggetto di esportazione ISO 2709, possono essere campi non standard. I nuovi campi non standard devono conservare lo stesso significato dei corrispondenti campi standard, con indicatori e sottocampi conformi (ad eccezione degli inutili sottocampi per le suddivisioni \$x, \$y, \$z, \$j).

Un'ipotesi di tag non standard per i campi intermedi nel record bibliografico:

600 Nome di persona -> 690  
 601 Nome di ente collettivo -> 691  
 602 Nome di famiglia -> 692  
 604 Nome/titolo -> 694  
 605 Titolo -> 695  
 606 Soggetto generico -> 696  
 607 Nome geografico -> 697  
 608 Forma -> 698

In questo caso, la corrispondenza tra i campi dell'archivio di autorità e i campi dell'archivio bibliografico sarebbe rappresentata nella seguente tabella:

<b>Categorie di entità</b>	<b>U./A.</b>	<b>U./B. campi intermedi</b>	<b>U./B. stringhe di soggetto</b>
Nome di persona	200	690	600
Nome di ente collettivo	210	691	601
Nome geografico	215	697	607
Nome di famiglia	220	692	602
Titolo uniforme	230	695	605
Nome/titolo	240	694	604
Soggetto generico	250	696	606
Forma	280	698	608

#### Posizione del termine di indicizzazione

Nel record bibliografico, per garantire la corretta attribuzione ad una particolare stringa di soggetto di un termine di indicizzazione registrato in uno dei campi intermedi, e stabilirne l'esatta posizione all'interno di ogni stringa di soggetto, ad ognuno dei campi non standard possono essere aggiunti due indicatori, che il catalogatore dovrà compilare:

(1) *Numero della stringa di soggetto*

nel quale è indicato, da 1 a 9, il numero attribuito alla stringa di soggetto;

(2) *Posizione del termine di indicizzazione*

nel quale è indicata, da 1 a 9, la posizione del termine di indicizzazione all'interno della stringa di soggetto definita dal primo indicatore.

Per la stringa presa ad esempio, si avrebbero i seguenti campi non standard con indicatori aggiunti:

**696 11##\$aEnti locali**

**696 12##\$aPersonale**

**696 13##\$aFormazione professionale**

**697 14##\$aItalia**

**698 15##\$aInchieste**

i campi intermedi generano il campo 606 per la stringa di soggetto, corrispondente al tag 696 del primo campo non standard:

**606 ##\$aEnti locali\$xPersonale\$xFormazione professionale\$yItalia\$yInchieste**

Se nello stesso record bibliografico si dovesse registrare un secondo soggetto, ad esempio:

Aziende sanitarie locali – Infermieri – Formazione professionale

esso sarebbe generato dai seguenti campi intermedi non standard:

**696 21##\$aAziende sanitarie locali**

**696 22##\$aInfermieri**

**696 23##\$aFormazione professionale**

**606 ##\$aAziende sanitarie locali\$xInfermieri\$xFormazione professionale**

In alternativa, ad ognuno dei campi non standard si può aggiungere un sottocampo, es. \$i, da compilare con due caratteri:

(1) in posizione 0, il numero della stringa di soggetto cui appartiene il termine di indicizzazione, da 1 a 9;

(2) in posizione 1, la posizione del termine di indicizzazione all'interno della stringa di soggetto individuata col carattere in posizione 0, da 1 a 9.

Ad esempio, i campi intermedi per i termini di indicizzazione di due stringhe di soggetto compresenti nella stessa registrazione bibliografica sarebbero così strutturati:

**696 ##\$aEnti locali\$i1**

**696 ##\$aPersonale\$i2**

**696 ##\$aFormazione professionale\$i3**

**697 ##\$aItalia\$i4**

**698 ##\$aInchieste\$i5**

**696 ##\$aAziende sanitarie locali\$i21**

**696 ##\$aInfermieri\$i22**

**696 ##\$aFormazione professionale\$i23**

Firenze, 9 giugno 2006