

Una DTD XML per FRBR

di Giovanna Granata

MARC, ISO2709 e XML

La recente introduzione da parte di W3C dell'eXtensible Markup Language (XML) [1] e la fortuna che fin dalla sua comparsa esso ha incontrato hanno dato luogo a una molteplicità di progetti in campi tra loro assai diversi. Oltre alla strutturazione dei documenti, settore per cui è stato originariamente pensato, il linguaggio infatti, in virtù delle sue potenzialità sintattiche, è ormai ampiamente diffuso per numerosi altri impieghi, dalla progettazione di data base fondata sull'uso di DTD (Document Type Definition) e XML Schema per l'impostazione della loro architettura interna, alla verifica della completezza e correttezza dei file dati mediante funzionalità di controllo e di validazione della loro struttura, fino alla definizione di formati logici di trasmissione dati che sfruttano XML come semplice supporto strutturale: la possibilità di articolare l'informazione secondo schemi complessi, rigorosamente annidati grazie all'uso di marcatori di inizio e fine campo, infatti, consente di rappresentare ampie classi di dati in maniera relativamente semplice.

È proprio con riferimento a quest'ultimo aspetto, quello della trasmissione dati, che anche in campo bibliotecario sono stati condotti diversi esperimenti legati, in modo particolare, alle problematiche dell'interscambio bibliografico e alla necessità di superare le limitazioni che le tradizionali implementazioni dei formati MARC hanno progressivamente mostrato nel nuovo scenario della comunicazione globale e della interoperabilità dei dati che si accompagnano alla diffusione di Internet.

Pesano infatti, come un retaggio della storia ormai più che trentennale del MARC, i vincoli imposti dal formato che ne costituisce il supporto di base, l'ISO 2709, il quale, oltre a non essere immediatamente compatibile con il mondo Web, rende estremamente difficoltosa la manipolazione e la fruibilità stessa dei record generati [2]. Questi ultimi sono infatti costruiti sulla base di una netta separazione tra le informazioni bibliografiche vere e proprie, da un lato, e quelle relative al record nel suo complesso (*label*), alla lista dei campi e ai punti del record che ne segnano l'inizio (*directory*), dall'altro: non solo deve così essere demandata a un software apposito l'associazione tra i dati e le relative etichette di campo che ne garantiscono la leggibilità, ma, oltre a questo, ogni minima variazione alla lunghezza di un campo invalida tutti gli indici relativi ai campi successivi, il che obbliga a ricalcolare l'esatta posizione degli *entry point* dei dati, rendendo l'aggiornamento della *directory* estremamente complesso.

Tra i primi e più significativi tentativi che hanno ipotizzato la sostituzione dell'ISO 2709 con un supporto più idoneo è il progetto della Library of Congress che

GIOVANNA GRANATA, Dipartimento di Studi medievali e moderni, Università di Chieti, Chieti Scalo, e-mail g.granata@unich.it.

Per tutti i siti Web l'ultima consultazione è stata effettuata il 14 giugno 2003.

fin dal 1995 ha lavorato all'utilizzazione di SGML per costruire una DTD direttamente incentrata su USMARC [3]. A partire dal 2000, a seguito dei nuovi sviluppi sui linguaggi di *markup* maturati nell'ambito del W₃C, la DTD originaria è stata riformulata in XML [4], mentre in Francia il progetto BiblioML del Ministère de la culture et de la documentation ha realizzato la traduzione XML di UNIMARC [5]. In entrambi i casi il formato originario è stato interamente mappato non solo per quanto riguarda la sua struttura di base, ma anche per quanto riguarda tutti gli aspetti sintattici, cosicché nelle DTD relative sono stati elencati tutti i campi, con l'esplicitazione di quelli obbligatori e di quelli opzionali e, per ciascun campo, sono stati segnalati gli indicatori possibili e i diversi sottocampi, arrivando così a un livello di complessità estremamente elevato.

Un diverso approccio è quello rappresentato dal progetto MARC/Perl che, a partire dal 1999, ponendosi il problema di processare record MARC in ambiente Perl¹, ha sviluppato un modulo Marc::XML per la traduzione XML di MARC [6], ma, a differenza dei progetti della Library of Congress e di BiblioML, senza fare riferimento a una DTD esplicita e quindi evitando di definire a priori il contenuto dei campi. Altri software, realizzati in ambiente commerciale², hanno optato per questa soluzione che tuttavia non garantisce alcuna possibilità di validazione e di controllo sulla correttezza del formato, anche se, rispetto alla precedente, ha il vantaggio di essere di un'estrema semplicità.

Su una via intermedia si pone il tentativo sviluppato nell'ambito del progetto Open Archive Initiative (OAI), il cui interesse si appunta proprio sulla definizione di protocolli semplici per lo scambio di dati tra archivi diversi. In questo contesto il Virginia Tech Digital Library Research Laboratory (VT DLRL) è approdato all'elaborazione di un modulo MARC XML che si fonda su una DTD, questa volta esplicita e successivamente elaborata in XML Schema [7], in cui, tuttavia, si è scelto di definire genericamente solo "fixfiled", "varfield" e "subfield", rispettivamente per i campi a lunghezza fissa e variabile e per i sottocampi. In questo modo la robustezza del formato è garantita, anche se a un livello abbastanza generico, con riferimento solo alla presenza degli elementi astrattamente previsti ma, al contempo, eventuali modifiche alla specifica del formato originario non rendono necessarie altrettante modifiche dello schema. Dati i vantaggi di questo approccio, anche la Library of Congress più di recente si è orientata nella medesima direzione, accantonando il progetto iniziale, che era vincolato alla manutenzione di una DTD estremamente specifica, per la costruzione di un XML Schema semplificato in cui sono definiti soltanto "leader", "controlfield" e "datafield" [8].

Pur differenziandosi per quanto riguarda il livello di specificità adottata, tutte queste proposte sono incentrate sulla sostituzione con XML del solo supporto di base ISO 2709 e, anche se ne rendono possibile una più facile manipolabilità e leggibilità, non mettono in discussione il modello concettuale del record MARC. In una diversa direzione sembra invece andare la recente realizzazione, da parte della Library of Congress, del Metadata Object Description Schema (MODS) [9]: allo scopo di garantire una maggiore compatibilità con altri schemi di metadati, si è infatti rinunciato a tradurre in maniera integrale il formato nello XML Schema, producendo un nuovo *set* di elementi che incorporano solo alcuni dei

¹ Il Perl è un linguaggio di programmazione, tipicamente utilizzato per la scrittura dei moduli CGI nelle pagine Web.

² Per esempio il modulo MARC XML di Logos research systems.

campi originari, provvedendo a raggrupparli più opportunamente e sostituendo alle etichette numeriche l'equivalente linguistico. In tal modo, anziché trasportare il MARC nel nuovo ambiente, si è compiuto un primo passo verso il suo superamento.

Più deciso ancora, in questo senso, risulta il progetto XMLMARC condotto dallo staff della Lane Medical Library di Stanford che, sotto la guida di Dick R. Miller, ha in realtà portato allo sviluppo di una struttura dati del tutto diversa rispetto alle specifiche del MARC, severamente criticato per la sua ridondanza e inconsistenza e per la scarsa integrazione con il mondo del Web: ne deriva sostanzialmente un nuovo formato bibliografico, XOBIS (XML Organic Bibliography Information Schema), esplicitamente pensato per ottimizzare la gestione dell'informazione in contesto digitale [10-13].

In effetti l'introduzione di XML, sebbene comporti alcuni vantaggi rispetto ai formati originali, continua però a presentarsi con un basso valore aggiunto, nella misura in cui mantiene la loro struttura generale, continuando pertanto a riprodurre gli stessi vincoli. Uno dei più gravosi è, in particolare, quello che deriva dalla scarsa possibilità di strutturazione dei dati e che trova le sue ragioni nella stessa norma ISO 2709. Essa, infatti, prevede due soli livelli di articolazione dei dati, campi e sottocampi, evidentemente insufficienti quando si debbano rappresentare realtà bibliografiche più complesse, e ha per questo forzatamente orientato lo sviluppo delle diverse implementazioni nazionali verso un modello enumerativo, talora ridondante e, soprattutto, tendenzialmente "piatto". XML, invece, proprio per le sue potenzialità sintattiche nella gestione di strutture arboree fortemente annidate, potrebbe prestarsi meglio alla definizione di un nuovo formato dati, più articolato del MARC e, per questo, da un lato più coerente, dall'altro più rispondente alle esigenze di caratterizzazione dei dati bibliografici.

Rispetto al tentativo della Lane Medical Library, l'ipotesi che si intende verificare è quella di utilizzare il modello FRBR per la costruzione di questo nuovo formato.

FRBR

Lo studio dell'IFLA sui requisiti funzionali dei record bibliografici (FRBR) [14], dopo anni di attenzione alla prassi descrittiva implicita nella formulazione delle ISBD e dopo le soluzioni per molti versi compromissorie e non perfettamente coerenti imposte con i principi di Parigi, ha riportato la riflessione sulla catalogazione a un livello di maggiore astrattezza concettuale [15-17]: utilizzando i presupposti dell'analisi entità-relazione sviluppata nell'ambito della teoria dei data base, sono state definite le entità in gioco, le loro reciproche relazioni e i relativi attributi, con riferimento precipuo alle funzionalità logiche di ogni elemento³.

Originariamente pensato come puro strumento di analisi, FRBR si è da subito imposto come termine di riferimento obbligato per verificare la congruità dei formati bibliografici più diffusi: fin dall'inizio, non a caso, la sua mappabilità con UNIMARC è stata oggetto di verifica da parte degli estensori dello studio IFLA [14,

³ Le entità sono state raggruppate a seconda che riguardino i prodotti del lavoro intellettuale o artistico (gruppo 1) o che individuino i responsabili del contenuto e della produzione fisica del lavoro stesso (gruppo 2) o il suo soggetto (gruppo 3). Le entità principali, quelle del gruppo 1, sono distinte in Opera, Espressione, Manifestazione ed Item, con riferimento alla creazione intellettuale originale (Opera), alla sua realizzazione artistica e intellettuale (Espressione), alla realizzazione fisica di quest'ultima (Manifestazione) e ai singoli esemplari che ne riflettono la forma fisica (Item).

p. 117 ss.] e, più recentemente, è stata esaminata anche la compatibilità di USMARC con i requisiti funzionali in esso identificati [18]; inoltre ne è stato valutato l'impatto ai fini di una revisione delle ISBD [19] e si è dovuta misurare con il modello anche la discussione, in atto in diversi paesi, sull'aggiornamento dei rispettivi codici catalografici⁴, mentre sta crescendo l'interesse da parte di diverse comunità per la ricerca di una struttura comune in grado di rendere effettivo l'obiettivo dell'interoperabilità [24].

Oltre che come strumento di analisi, tuttavia, FRBR per la sua completezza è stato oggetto di attenzione anche per quanto riguarda la effettiva strutturazione dei dati, offrendo un nuovo paradigma di riferimento per l'insieme delle problematiche relative alla gestione e al recupero dell'informazione. In questo contesto sono stati realizzati alcuni progetti pilota, tra cui i più interessanti per l'accuratezza dell'analisi e l'ambizione degli obiettivi sono AustLit [25], finalizzato alla costruzione di un *gateway* bibliografico per la letteratura australiana e, in ambiente europeo, VisualCat che, specificatamente pensato per la catalogazione in contesti digitali, usa FRBR come ontologia più generale nel quadro di una struttura RDF finalizzata al trattamento di materiali diversi e quindi all'uso di sistemi diversi di metadati [26].

Le potenzialità di FRBR come formato bibliografico suggeriscono, però, di spingere l'impiego al di là della gestione interna dei dati e di utilizzarlo anche per il loro interscambio sostituendo in questo il MARC, le cui limitazioni nella strutturazione di realtà bibliografiche complesse potrebbero così essere superate. Apparentemente, in realtà, il modello relazionale che FRBR propone parrebbe poco funzionale a questo scopo. In astratto, infatti, l'ideale per l'esportazione dei dati è una struttura dati "piatta" e fortemente analitica: in questo modo, da un lato tutta l'informazione è disponibile contemporaneamente e può meglio essere controllata, dall'altro la granularità dei dati permette una migliore gestione e una più semplice mappatura. Il MARC sembrerebbe incarnare meglio questo modello e il suo successo più che trentennale ne è una conferma, ma alcune considerazioni possono mettere in discussione questa conclusione.

In primo luogo si deve osservare che la peculiarità della stessa informazione bibliografica impone di per sé una modellizzazione relazionale, laddove in particolare si tratta di gestire "descrizioni a più livelli", liste di autorità ecc. Non a caso lo stesso UNIMARC, come evoluzione più recente del MARC, ha introdotto il blocco 400 per la gestione dei link, rinunciando, in questo modo, a una delle principali caratteristiche del formato originario, l'autonomia che ogni record ha dagli altri. Per mantenere tale autonomia, che è un elemento evidentemente prezioso per agevolare l'interscambio, in linea di principio sarebbe necessario includere i dati dei record correlati uno nell'altro in maniera completa, ossia in modo da conservare l'intera struttura dei legami, ma ciò risulta di difficile gestione nelle attuali implementazioni⁵, anche nell'ipotesi di una loro traduzione in XML: il linguaggio, infatti, rende possibili molteplici livelli di annidamento, ma in questo caso ogni record che avesse legami con altri record, dovrebbe includere ciascuno di essi e questi a loro volta i loro correlati, in teoria senza limitazione alcuna, rendendo l'informazione estremamente ridondante. D'altra parte, l'utilizzo del semplice identificativo del record per esprimere il

4 Per quanto riguarda l'Italia, cfr. [20-23].

5 L'uso degli *embedded field*, non del tutto rispondente alle specifiche ISO 2709, è previsto dalla norma UNIMARC, ma chiaramente sconsigliato negli ultimi *update* a favore della tecnica standard.

legame non garantisce la presenza dello stesso tra i dati esportati e dunque rischia di comprometterne la completezza⁶.

Per quanto riguarda FRBR, invece, in primo luogo esso non presenta limiti strutturali all'articolazione dell'informazione, cosicché nell'ambito della stessa unità descrittiva possono essere correttamente rappresentate le diverse relazioni esistenti tra le entità come chiaramente indicano gli esempi dello studio IFLA⁷. In secondo luogo, diversamente dal modello MARC, la struttura granulare di FRBR permette di ipotizzare l'annidamento non di un intero record in un altro, ma dei singoli blocchi informativi corrispondenti alle diverse entità in gioco le quali nel loro complesso costituiscono il record bibliografico che, in tal modo, manterrebbe la sua autonomia.

Occorre inoltre tenere presente che l'effettiva funzionalità di un formato di interscambio dipende, in realtà, essenzialmente dalla struttura del data base di destinazione. È vero, infatti, che se esso stesso è "piatto", l'importazione di un file fortemente strutturato può essere complessa, perché richiede l'attraversamento ripetuto dello stesso per ricompattare i dati. Viceversa, però, la gestione di un record monolitico con un basso livello di articolazione interna in un data base fortemente strutturato rende necessaria la ricostruzione dei link impliciti (per esempio tra Espressione e Manifestazione), che sono invece direttamente individuabili in un record che indica esplicitamente le relazioni. Non a caso, date le difficoltà di tale operazione, l'uso massiccio del MARC come formato di comunicazione ne ha imposto l'uso anche come struttura di gestione dei dati, in contro tendenza rispetto allo sviluppo di data base di tipo relazionale cui si è assistito in settori diversi da quello bibliografico, e con le conseguenti limitazioni che ciò comporta [27]. Da questo punto di vista FRBR potrebbe prestarsi meglio come struttura di interscambio dati, non solo in quanto consente di esprimere correttamente e in modo naturale gli aspetti relazionali dell'informazione bibliografica, ma anche in quanto tendenzialmente più vicina alla struttura di gestione dei dati all'interno del data base.

Quello che si propone nel seguito è la costruzione di una DTD per FRBR, ma prima di esaminare i principali problemi che ciò pone, occorre sottolineare che essa sarà implementata in riferimento al modello astratto, così come questo è definito nello studio, pur nella consapevolezza che in vista della sua applicazione concreta al trattamento del materiale catalografico debbano preliminarmente essere effettuate alcune scelte operative. In particolare per come è costruito FRBR, il cui obiettivo primario non è propriamente quello di specificare un formato di gestione dati, è evidente che in esso non sono fissati limiti alle possibilità di interconnessione dei dati, mentre nell'ipotesi di un suo impiego pratico diventa necessario delimitare da una parte i "confini" del record, dall'altra i confini tra le entità stesse: a quale livello di complessità spingersi nel comprendere le diverse relazioni descritte dallo studio e dove fissare il limite per cui da espressioni diverse di una stessa opera si passa a opere tra loro diverse? Queste sono solo alcune delle domande che sono

⁶ Tale è per esempio la soluzione dell'*output* UNIMARC prodotto da SBN: nei campi del blocco 400 sono indicati i BID dei record correlati, accompagnati alle righe ISBD "compatte", cioè senza etichettatura di campo, cosicché per completare l'informazione è necessario leggere l'identificativo, cercare il record corrispondente nella base dati ed esportarlo autonomamente, associandolo in fase di importazione al record di partenza.

⁷ Come mostrano gli esempi dello studio IFLA [14], esso può descrivere senza difficoltà Opere che contengono altre Opere, con le relative Espressioni, Manifestazioni e Item.

state poste da chi ha verificato l'utilizzabilità del modello a fini catalografici [20]: a esse la proposta che segue non intende offrire una soluzione definitiva, affrontando preliminarmente soltanto l'analisi degli aspetti strutturali che sottostanno alla traduzione del modello stesso in XML. Eventuali modifiche alla DTD potranno essere discusse in seguito alla definizione di opportune modalità di impiego.

XML e FRBR

Il primo problema da affrontare deriva dal fatto che, a un esame superficiale, XML si presenta come un modello completamente gerarchico, apparentemente all'opposto di quello che è descritto nello studio IFLA: la specifica, infatti, impone che non sia possibile chiudere un campo senza aver prima chiuso tutti i suoi sottocampi, cosicché tra campi non può esistere altro che un rapporto di subordinazione secondo una struttura di tipo arboreo⁸.

A ogni elemento, tuttavia, possono essere associate una serie di meta-informazioni (gli attributi) che ne completano e arricchiscono il contenuto e, sebbene esse siano ovviamente specifiche per le singole strutture dati, la norma ne ha previste tre (ID, IDREF e IDREFS) che consentono di ampliare fortemente le potenzialità di XML in senso relazionale. L'attributo ID, infatti, indica l'identificativo univoco eventualmente associato a un campo, mentre IDREF e IDREFS contengono il riferimento a uno o più ID. In questo modo, benché l'originaria struttura arborea venga formalmente mantenuta intatta, l'insieme delle relazioni tra i singoli campi viene fortemente arricchito consentendo di rappresentare link anche tra elementi di rami diversi.

Poiché, dunque, l'apparente incompatibilità tra XML e FRBR viene così superata, si può ipotizzare l'implementazione di una DTD che, limitata alle sole relazioni principali del primo gruppo (Opera, Espressione, Manifestazione e Item), potrebbe dunque presentarsi nel modo seguente:

```

1. <!DOCTYPE frbr [
2. <!ELEMENT frbr (frbr_record)+ >
3. <!ELEMENT frbr_record (work | expression | manifestation | item)+ >
4. <!ELEMENT work (data)+ >
5. <!ATTLIST work
   a. id ID #REQUIRED
   b. realized_through IDREFS #IMPLIED
6. >
7. <!ELEMENT expression (data)+ >
8. <!ATTLIST expression
   a. id ID #REQUIRED
   b. embodied_in IDREFS #IMPLIED
   c. realization_of IDREF #IMPLIED
9. >
10. <!ELEMENT manifestation (data)+ >
11. <!ATTLIST manifestation
   a. id ID #REQUIRED
   b. exemplified_by IDREFS #IMPLIED
   c. embodiment_of IDREFS #IMPLIED

```

(1)

8 È proprio questa caratteristica che permette di verificare facilmente la completezza e la correttezza dei dati anche in casi di strutture fortemente articolate e complesse, rendendolo quindi particolarmente adatto al trasferimento di dati: la perdita di parte dei dati e la coerenza della loro struttura logica possono essere infatti immediatamente verificate.

```

12. >
13. <!ELEMENT item (data)+ >
14. <!ATTLIST item
    a. id ID #REQUIRED
    b. exemplar_of IDREF #IMPLIED
15. >
16. <!ELEMENT data (#PCDATA) >
17. ]>

```

Tale DTD prevede un documento composto da uno o più record FRBR (2.); ogni record è composto da una o più occorrenze degli elementi *work*, *manifestation*, *expression* e *item* (3.); ognuno di questi elementi è caratterizzato da informazioni specifiche, genericamente indicate con il campo *data* (4., 7., 10., 13.) e da una lista di attributi (5., 8., 11., 14.): l'identificativo dell'elemento (ID) e la lista dei link (*realized_through*, *embodied_in*, *exemplified_by* e i loro contrari), ciascuno dei quali contenente l'identificativo o gli identificativi delle unità correlate (IDREF/IDREFS).

Lo schema ricalca fedelmente la proposta IFLA e potrebbe facilmente essere espanso per comprendere tutti gli elementi descrittivi e l'insieme delle relazioni in essa presenti. Ogni singolo record FRBR può così contenere l'intera informazione in modo da garantirne la completezza in sede di interscambio dati, ma, al tempo stesso, si articola in blocchi informativi (*work* ecc.) che, pur essendo correlati gli uni agli altri, mantengono però la loro autonomia e possono essere individuati come tali. Proprio questa autonomia, tuttavia, costituisce il rischio maggiore di questo approccio per quanto riguarda la robustezza del formato come supporto di esportazione: infatti è possibile incontrare un solo link errato per compromettere l'intera struttura logica, cosicché, da questo punto di vista, sarebbe opportuno limitare l'approccio relazionale il più possibile.

D'altra parte, benché gli aspetti relazionali della struttura di FRBR nello studio IFLA siano enfatizzati, in realtà solo una delle relazioni principali è strettamente tale e non potrebbe essere formulata diversamente, quella tra Espressione e Manifestazione. In questo caso, infatti, non solo per ogni Espressione ci possono essere più Manifestazioni, ma ognuna di queste può anche corrispondere a più Espressioni di Opere diverse. Viceversa tanto i rapporti tra Opera ed Espressione quanto quelli tra Manifestazione e Item possono essere visti come segmenti di una struttura gerarchica: non può infatti esistere un'Espressione senza un'Opera né una Copia senza una Manifestazione, ovvero, come chiaramente indica lo schema [14, p. 13, fig. 3.1], per ogni Espressione una sola è l'Opera di riferimento e per ogni Item una sola la Manifestazione. Di fatto lo schema potrebbe essere riformulato come segue, evidenziando il carattere di subordinazione o meno che assumono le relazioni tra le entità:



Data la loro natura gerarchica, in un modello di strutturazione dei dati basato su FRBR, i rapporti tra Opera ed Espressione come quelli tra Manifestazione e Item potrebbero essere gestiti anziché tramite link relazionali, come necessariamente dovrebbe invece essere per i rapporti tra Espressione e Manifestazione, semplicemente ricorrendo all’inserimento dei dati relativi a ogni entità subordinata (Espressione e Item) in forma annidata, “dentro” quelli dell’entità principale. Il record FRBR in questo approccio conterrebbe direttamente solo le entità Opera e Manifestazione ed è solo come elementi di queste ultime che troveremmo le altre due, secondo lo schema che segue:



Su questa base la DTD (1) dovrebbe essere così riformulata:

```

1. <!DOCTYPE frbr [
2. <!ELEMENT frbr (frbr_record)+ >
3. <!ELEMENT frbr_record (work | manifestation)+ >
4. <!ELEMENT work (data | expression)+ >
5. <!ELEMENT expression (data)+ >
6. <!ATTLIST expression
   a. id ID #REQUIRED
   b. embodied_in IDREFS #REQUIRED
7. >
8. <!ELEMENT manifestation (data | item)+ >
9. <!ATTLIST manifestation
   a. id ID #REQUIRED
   b. embodiment_of IDREFS #REQUIRED
10. >
11. <!ELEMENT item (data)+ >
12. <!ELEMENT data (#PCDATA) >
13. ]>
  
```

dove il rapporto tra Opera ed Espressione (4.) e quello tra Manifestazione e Item (8.) non è gestito tramite link, ma attraverso l’inclusione dei campi, mentre rimane invece espressa da un legame la relazione tra Espressione e Manifestazione (6b., 9b.). Tale approccio, che è molto più robusto perché il numero dei link è fortemente ridotto e l’annidamento dei dati permette un controllo più efficace, può essere assunto come base per la strutturazione di una DTD che prenda in considerazione l’intero modello FRBR.

La DTD per FRBR

Individuata la struttura di massima della DTD, per arrivare all’implementazione completa del modello proposto si tratta di esaminare più approfonditamente il rapporto IFLA allo scopo di definire le scelte più opportune in merito al trattamento delle relazioni non contemplate nello schema principale, ovvero quelle relative alle entità dello stesso tipo e alle entità del blocco autori e soggetti.

Per quanto riguarda le relazioni tra entità dello stesso tipo, occorre osservare che la loro introduzione porta una certa complicazione nella costruzione della DTD per-

ché, essendo ampia la casistica, oltre che la direzione del rapporto gerarchico implicato di superiorità o di subordinazione, se ne deve esplicitare anche la natura, a differenza di quanto avviene per le relazioni principali che sono invece definite univocamente⁹.

Una possibilità, che riproduce assai da vicino le esemplificazioni proposte da FRBR, è quella di trattare le entità correlate come elementi autonomi, indicando la direzione del legame e il tipo di relazione.

Per rappresentare questo modello occorre inserire, all'interno dell'entità di riferimento, un campo specifico (opera_correlata, espressione_correlata ecc.) nel quale, in forma di attributi, siano indicati l'identificativo della notizia da collegare e tutte le possibili relazioni in tutti i loro rapporti gerarchici.

Usando come esempio il caso del rapporto tra Opere, l'elemento `work` della riga 4. di (2) dovrebbe essere modificato con l'inclusione del campo `related_work`, come segue:

```

1. <!ELEMENT work (data|related_work|expression)+ >
2. <!ATTLIST work
   a. id ID #REQUIRED
3. >
4. <!ELEMENT related_work EMPTY>
5. <!ATTLIST related_work
   a. ref IDREF #REQUIRED
   b. relation_hierarchy (depends_on | required_by) #REQUIRED
   c. relation_type (part_of|successor|supplement|complement|
      summary|adaption|trasformation|imitation) #REQUIRED
6. >
```

In essa l'opera principale e l'opera correlata sono separate tra loro e interconnesse tramite link (2a., 5a.), mentre gli attributi del campo `related_work`, oltre all'identificativo della notizia da collegare, specificano la gerarchia (5b.) e la tipologia (5c.) del legame, proprio come nell'esempio dello studio IFLA [14, p. 66]:

w₁ Homer's *Iliad*
 Has a concordance: w₂ →

w₂ G.L. Prendergast *Prendergast's A complete concordance to the Iliad of Homer*
 ← Is a concordance to: w₁

Tale approccio, benché molto vicino alle esemplificazioni dello studio IFLA, moltiplica evidentemente l'uso degli ID e IDREF e contrasta pertanto con la soluzione generale che è stata proposta nella DTD (2). Più coerente rispetto a essa è una diversa soluzione che sfrutta le possibilità dell'annidamento per esprimere la gerarchia della relazione: in tal modo non solo sono ridotte le possibili tipologie di link, essendo l'inclusione stessa a definire la presenza di un legame "dall'alto - al basso", ma oltre a questo la relazione stessa può essere espressa come attributo dei blocchi infor-

⁹ Ad esempio per il rapporto tra opere, le relazioni previste sono: *Successor*, *Supplement*, *Complement*, *Summarization*, *Adaptation*, *Transformation*, *Imitation*, ognuna di queste in rapporto gerarchico di superiorità o di subordinazione: un'opera può essere supplemento di una seconda opera o avere essa stessa un supplemento. Invece per quanto riguarda il rapporto tra Opera ed Espressione, esso è definito solo dalla relazione *Is realized through*.

mativi inclusi invece che come campo autonomo, secondo lo schema che segue:

w₁ Homer's *Iliad*

w₂ G.L. Prendergast's *A complete concordance to the Iliad of Homer* (tipo di relazione: Supplement)

Su questa base il frammento di DTD (2), riga 4., può dunque essere così riformulato:

1. <!ELEMENT work (**work**|data|expression)+ >
2. <!ATTLIST work
 - a. relation (part_of|successor|supplement|complement|summary|adaption|trasformation|imitation) #IMPLIED
3. >

Le due Opere correlate sono una dentro l'altra (1.), in modo che sia implicitamente definita la gerarchia della relazione. In 2a. sono indicati in forma di attributi dell'Opera le possibili tipologie di legame evitando il ricorso al campo *related_work*.

Questo tipo di soluzione, estesa a Espressione, Manifestazione e Item, permette di gestire la maggior parte delle relazioni aggiuntive tra le entità del primo gruppo descritte dallo studio: quelle tra Opera e Opera (5.3.1) [14, p. 65], tra Espressioni della stessa Opera (5.3.2) [14, p. 71], tra Manifestazione e Manifestazione (5.3.4) [14, p. 79] tra Item della stessa Manifestazione (5.3.6) [14, p. 79], oltre che l'insieme delle relazioni parte/tutto (5.3.1.1, 5.3.2.1, 5.3.4.1, 5.3.6.1), secondo lo schema che segue:

* Opera A

- Opera B subordinata ad A: supplemento, riassunto, adattamento ecc. (5.3.1)
 - Espressione 1B (1 di Opera B)
 -
- Espressione 1A (1 di Opera A)
 - Espressione 1.1A (1 subordinata a 1A): traduzione, arrangiamento ecc. (3.5.2)
 -
-

* Manifestazione a di 1A

- Manifestazione b subordinata ad a: riproduzione ecc. (5.3.4)
 - Copia 1 di b
 -
- Copia 1 di a
 - Copia 2 di a subordinata a Copia 1: riproduzione, facsimile ecc. (3.5.6)
 -

Per contro essa non è in grado di supportare le relazioni tra Espressioni di opere diverse (5.3.2), tra Manifestazioni e Item di altre manifestazioni (5.3.5) e tra Item di manifestazioni diverse (5.3.6), in quanto esse non sono in rapporto gerarchico diretto e quindi non possono essere rese ponendo gli elementi in forma annidata¹⁰.

¹⁰ Anche la relazione parte/tutto a livello di manifestazione potrebbe comportare lo stesso problema nel caso le singole parti facciano riferimento a opere diverse, per esempio, come indica lo studio, quando si ha un film e la sua colonna sonora. Tuttavia la DTD ipotizzata già a livello di struttura principale gestisce le relazioni tra Espressione e Manifestazione tramite link.

Nel primo caso, in realtà, si ha a che fare con relazioni analoghe a quelle presenti tra Opera e Opera [14, p. 72] e, non a caso, lo studio suggerisce di trattarle a questo livello, trasformandole in relazioni tra le Espressioni di un'opera e un'Opera diversa, data anche la difficoltà che spesso si presenta nell'identificare lo specifico testo utilizzato come base per la nuova espressione [14, p. 74]. Da questo punto di vista tali relazioni possono essere ricondotte allo schema proposto osservando che in esso l'espressione di un'opera (Espressione 1 di Opera B), a sua volta annidata all'interno di un'altra opera (Opera A), è anch'essa necessariamente subordinata a quest'ultima, secondo la medesima relazione che le opere hanno tra loro.

Nel caso delle concordanze omeriche viste sopra, per esempio, otterremo un record di questo tipo:

w₁ Homer's *Iliad*

w₂ G.L. Prendergast, *A complete concordance to the Iliad of Homer* (tipo di relazione: Supplement)

e₁ testo originale delle concordanze di G.L. Prendergast

e₂ nuova edizione delle concordanze, rivista e ampliata da Benedetto Marzullo

Per quanto riguarda il rapporto tra Manifestazione e Item di manifestazioni diverse e quello tra Item e Item di manifestazioni diverse, invece, la situazione prospettata dallo studio è, in un caso, quella in cui un'intera edizione sia esemplata su una singola copia di un'altra edizione [14, p. 78], nell'altro, quella in cui una particolare copia sia rilegata insieme con un'altra, magari di un'opera diversa [14, p. 80]. Per entrambe le situazioni nella DTD l'uso di legami relazionali sarebbe preferibile rispetto alla ipotesi dell'annidamento che porterebbe a una complicazione eccessiva della struttura. Tuttavia il vero problema da valutare è l'opportunità stessa di includere nello schema le due relazioni in quanto esse comporterebbero l'esportazione di una massa di dati non strettamente necessari, ampliando il record eccessivamente. In effetti esso verrebbe a includere non solo la catena che, secondo lo schema esemplificativo proposto, conduce da Opera A a Item 2 di a, ma anche quella che dalla copia stessa si sviluppa nella direzione di altre Manifestazioni (o eventualmente di altre Opere) non necessariamente pertinenti rispetto alle esigenze dell'esportazione. Per evitare questo problema, la gestione ottimale delle due relazioni descritte potrebbe piuttosto essere quella, più tradizionale, di ricorrere a note in campi di tipo descrittivo.

La soluzione di legami relazionali è invece più adatta per quanto riguarda le entità degli altri due gruppi: autori e soggetti; sebbene, infatti, a rigore esse possano essere rappresentate anche attraverso il meccanismo dell'annidamento, tuttavia, data la loro complessità e poiché in genere provengono da archivi di autorità, è opportuno che siano gestite come realtà descrittive autonome.

A tale proposito lo studio IFLA definisce un solo tipo di relazione tra ciascuna delle quattro entità principali e quelle del secondo gruppo: "created by" a livello di Opera, "realized by" a livello di Espressione, "produced by" a livello di Manifestazione e "owned by" a livello di Item, mentre un'unica relazione, "has as subject", interessa le entità del terzo gruppo e la sola Opera.

Per rappresentare tale situazione in forma relazionale, la DTD, pur rimanendo gerarchica quanto alla sua struttura principale, dovrà essere modificata sulla base di un modello abbastanza vicino a quello esaminato per la (1). Come elementi di

ognuna delle entità del primo gruppo si dovrà prevedere un campo con il quale esprimere la relazione con le entità del secondo gruppo: `created_by`, `realized_by`, `produced_by` e `owned_by`, mentre in forma di attributi rispettivamente di ciascuna entità del primo gruppo e di ciascuno dei campi esprimenti la relazione saranno indicati l'ID e l'IDREF degli elementi da correlare; in più per la sola Opera sarà necessario inserire anche il campo `has_as_subject`, anch'esso con IDREF come attributo per garantire il legame tra l'Opera stessa e il Soggetto. Ovviamente le entità del secondo e del terzo gruppo devono essere definite a monte come elementi del record FRBR, cosicché, nel complesso, la DTD proposta (2) dovrebbe essere integrata e riformulata come segue:

```

1. <!DOCTYPE frbr [
2. <!ELEMENT frbr (frbr_record)+ >
3. <!ELEMENT frbr_record (work | manifestation | responsibility |
  subject)+ >
4. <!ELEMENT work (work|data|expression|created_by|has_as_subject)+ >
5. <!ATTLIST work
  a. id ID #REQUIRED
  b. relation (part_of|successor|supplement|complement|summary|
    adaption|trasformation|imitation) #IMPLIED
  c. >
6. <!ELEMENT created_by (#PCDATA) >
7. <!ATTLIST created_by ref IDREF #IMPLIED>
8. <!ELEMENT has_as_subject (#PCDATA) >
9. <!ATTLIST has_as_subject ref IDREF #IMPLIED>
10. ... ..
11. <!ELEMENT responsibility (person|corporate_body)>
  a. <!ATTLIST responsibility
  b. id ID #IMPLIED
  c. >
12. ... ..
13. ]>

```

dove l'attributo IDREF dell'elemento `created_by` (7.) permette il legame con l'ID dell'elemento `responsibility` (11b.), a sua volta specificato come `person` o come `corporate_body`. Inoltre, come è evidente dall'uso di `#PCDATA`, i campi che descrivono le relazioni (6., 8.), nonostante il legame, possono essere comunque gestiti come campi di dati, anziché come campi vuoti (`#EMPTY`): in questo modo il contenuto della relazione può essere presente nel record anche in forma testuale, permettendo di utilizzarne le informazioni per verificare il legame stesso, o anche per sostituirlo, nel caso esse non siano in partenza strutturate in *authority file*.

Una volta affrontate le problematiche che il modello IFLA pone a livello di struttura generale, occorre valutare alcune possibili scelte per quanto riguarda il trattamento del campo `data`: infatti, come si è visto analizzando le DTD proposte per la traduzione del MARC in XML, una delle principali differenze che esse presentano è costituita dalle modalità di gestione dei *tag* di partenza, ora indicati in maniera generica, semplicemente come *field*, ora esplicitati ed enumerati come veri e propri elementi.

La prima soluzione porta a DTD meno potenti per quanto riguarda il controllo sintattico, ma certamente più semplici e soprattutto tali da non dover essere modificate ogni volta che lo è il modello di riferimento. Questo problema è in effetti abbastanza cruciale per il MARC che, data la sua natura enumerativa, si è sviluppato nel tempo attraverso un continuo proliferare di campi per rendere possibile il trattamento di nuovi materiali e di nuove tipologie di supporto. Diverso è da questo punto di vista il caso di FRBR che, proprio nella misura in cui si propone di fissare i requisiti funzionali del record bibliografico, affrontandone l'analisi a partire dalla considerazione di un modello teorico, ha in sé le premesse per una definizione più fondata e stabile degli elementi in gioco. Per tali ragioni la soluzione di una DTD più complessa, ma in grado di garantire il controllo sintattico attraverso l'esplicitazione di quelli che lo studio IFLA definisce come attributi delle diverse entità, può essere preferibile, suggerendo di articolare il campo "data" in sotto campi, dopo averlo differenziato, a seconda dell'elemento in cui si trova incluso, in *work_data*, *expression_data* ecc.

Il frammento di DTD relativo all'Opera diventerebbe a questo punto:

1. <!ELEMENT work (work|work_data|expression|created_by|has_as_subject)+ >
2. <!ATTLIST work
 - a. id ID #REQUIRED
 - b. relation (part_of|successor|supplement|complement|summary|adaption|trasformation|imitation) #IMPLIED
3. >
4. <!ELEMENT work_data (title|form|...|equinox)+ >
5. <!ELEMENT created_by (#PCDATA) >
6. <!ATTLIST created_by ref IDREF #IMPLIED>
7. <!ELEMENT has_as_subject (#PCDATA) >
8. <!ATTLIST has_as_subject ref IDREF #IMPLIED>

Tale soluzione garantisce la possibilità di verificare la correttezza dei record, imponendo vincoli ferrei al suo contenuto; proprio per questo, tuttavia, potrebbe anche comportare la perdita di informazioni, nel caso il formato debba veicolare dati non originariamente strutturati secondo il modello FRBR, ma derivati da sistemi diversi. A tal scopo potrebbe essere utile inserire anche un campo *extra_data* in grado di accogliere tali informazioni, senza dover per questo intaccare l'intrinseca coerenza dell'insieme di campi indicati dallo studio. A differenza dei campi *data*, i campi *extra_data* non possono che essere trattati in maniera generica, eventualmente utilizzando l'attributo *field_name* per indicare il nome del campo originario.

Conclusioni

Per concludere si propone un esempio di strutturazione di un record FRBR secondo le scelte discusse per la articolazione della DTD.

Si tratta delle *Avventure di Pinocchio* (w_1) di Carlo Collodi, di cui sono considerate due diverse Espressioni, l'edizione critica italiana (e_1 , a cura di O. Castellani Polidori) e la sua traduzione in inglese (e_2 , a cura di N. J. Perella).

A sua volta l'edizione critica italiana (e_1) si materializza nella pubblicazione della Fondazione nazionale Carlo Collodi (m_1), ma anche in quella di University of California Press (m_2) che è contemporaneamente anche materializzazione di e_2 , in quanto contiene la traduzione inglese e il testo originale a fronte.

w₁ *Avventure di Pinocchio* di C. Collodi**e₁** edizione critica italiana a cura di Ornella Castellani Pollidori**m₁** *Le avventure di Pinocchio* / Carlo Collodi. - Ed. critica / a cura di Ornella Castellani

Pollidori. - Pescia: Fondazione nazionale Carlo Collodi, 1983.

m₂ *The adventures of Pinocchio: story of a puppet* / Carlo Collodi; translated, with an introductory essay and notes by Nicolas J. Perella. - Berkeley: University of California Press, 1986. (Testo italiano e inglese a fronte)**e₂** traduzione dell'edizione critica italiana a cura di Nicolas J. Perella**m₂** *The adventures of Pinocchio: story of a puppet* / Carlo Collodi; translated, with an introductory essay and notes by Nicolas J. Perella. - Berkeley: University of California Press, 1986.

Il documento XML relativo è il seguente: dentro l'Opera, qualificata dall'identificativo **w₁** (A), sono inclusi i campi **work_data** (A1.), **expression** (A3.) e **created_by** (A5.) che esprimono in forma di link il legame con l'autore.

L'Espressione inclusa dentro l'Opera, l'edizione critica **e₁**, contiene a sua volta al suo interno il campo **expression_data** (A3a.), i campi che esprimono il legame con le Manifestazioni (**m₁** ed **m₂**) contenenti il testo italiano (A3c.-d.) e con l'autore (A3e.), nonché la seconda Espressione (**e₂**), cioè la traduzione inglese, a essa subordinata gerarchicamente (A3f.) secondo la modalità con cui si è deciso di rappresentare la relazione *Expression to Expression* di FRBR 3.5.2.

Al rigo B. si chiude l'Opera **w₁**, ma subito dopo se ne apre una seconda **w₂** (C1.-6.): oltre alle *Avventure di Pinocchio*, infatti, l'edizione inglese contiene, a cura dello stesso traduttore Nicolas J. Perella, anche un saggio introduttivo e delle note che sono rappresentati come opera a sé stante con un proprio autore (C5.) e legata alle *Avventure di Pinocchio* (**w₁**) da una relazione di soggetto (C6.).

Seguono quindi le Manifestazioni (E.-H.) entrambe articolate nei campi **manifestation_data** (E1., G1.), con i relativi sotto campi, e **embodiment_of** (E3., G3.-5.) che contengono gli identificativi delle Espressioni cui sono connesse. Chiude il documento il blocco relativo agli autori (I.-N.).

I. <?xml version="1.0"?>

II. <!DOCTYPE frbr PUBLIC "http://www.unich.it/~granata/frbr/frbr.dtd">

III. <frbr>

IV. <frbr_record id="r1">

<!-- work 1 -->

A. <work id="w1">

1. <work_data>

a. <title>Le avventure di Pinocchio</title>

2. </work_data>

3. <expression id="e1">

a. <expression_data>

i. <language>italian</language>

ii. <other_distinguishing_characteristic>Edizione critica

iii.</other_distinguishing_characteristic>

b. </expression_data>

c. <embodied_in ref="m1"/>

d. <embodied_in ref="m2"/>

e. <realized_by ref="auth2">Castellani, Ornella</realized_by>

```

        f. <expression id="e2" relation="translation">
            i. <expression_data>
                *. <language>english</language>
            ii. </expression_data>
            iii.<embodied_in ref="m2"/>
            iv. <realized_by ref="auth3">Perella, Nicolas J.</realized_by>
        g. </expression>
    4. </expression>
    5. <created_by ref="auth1">Collodi, Carlo</created_by>
B. </work>
<! -- work 2 -->
C. <work id="w2">
    1. <work_data>
        a. <title>Introductory essay and notes</title>
    2. </work_data>
    3. <expression id="e3">
        a. <expression_data>
            i. <language>english</language>
        b. </expression_data>
        c. <embodied_in ref="m2"/>
    4. </expression>
    5. <created_by ref="auth3">Perella, Nicolas J.</created_by>
    6. <has_as_subject ref="w1"/>
D. </work>
<! -- manifestation 1 -- >
E. <manifestation id="m1">
    1. <manifestation_data>
        a. <title>Le avventure di Pinocchio</title>
        b. <statement_of_responsability>Carlo Collodi</statement_of_responsability>
        c. <edition>Ed. critica</edition>
        d. <edition_statement_of_responsability>a cura di Ornella Castellani Pollidori</edition_statement_of_responsability>
        e. </edition_statement_of_responsability>
        f. <place_of_publication>Pescia</place_of_publication>
        g. <publisher>Fondazione nazionale Carlo Collodi</publisher>
        h. <date>1983</date>
        i. <extent_of_carrier>LXXXIV, 281 p. : ill.</extent_of_carrier>
        j. <dimensions_of_carrier>30 cm</dimensions_of_carrier>
    2. </manifestation_data>
    3. <embodiment_of ref="e1"/>
F. </manifestation>
<! -- manifestation 2 -- >
G. <manifestation id="m2">
    1. <manifestation_data>
        a. <title>The adventures of Pinocchio : story of a puppet</title>
        b. <statement_of_responsability>Carlo Collodi ; translated, with an introductory essay and notes by Nicolas J. Perella</statement_of_responsability>
        c. <place_of_publication>Berkeley</place_of_publication>
        d. <publisher>University of California Press</publisher>
        e. <date>c1986</date>

```

```

        f. <extent_of_carrier>VIII, 498 p. : ill.</extent_of_carrier>
        g. <dimensions_of_carrier>21 cm</dimensions_of_carrier>
    2. </manifestation_data>
    3. <embodiment_of_ref="e1"/>
    4. <embodiment_of_ref="e2"/>
H. </manifestation>
<! -- authors -- >
I. <responsability id="auth1">
    1. <person>
        a. <name>Collodi, Carlo</name>
    2. </person>
J. </responsability>
K. <responsability id="auth2">
    1. <person>
        a. <name>Castellani, Ornella</name>
    2. </person>
L. </responsability>
M. <responsability id="auth3">
    1. <person>
        a. <name>Perella, Nicolas J.</name>
    2. </person>
N. </responsability>
V. </frbr_record>
VI. </frbr>

```

RIFERIMENTI BIBLIOGRAFICI

- [1] W3C. *Extensible Markup Language (XML) 1.0*. 2. ed., W3C Recommendation 6 October 2000. <<http://www.w3c.org/TR/REC-xml>>.
- [2] Giovanna Granata. *XML e formati bibliografici*. «Bollettino AIB», 40 (2000), n. 2, p. 181-191.
- [3] Library of Congress. Network development and MARC standards office. *MARC SGML Archive*. <<http://lcweb.loc.gov/marc/marcsgmlarchive.html>>.
- [4] Library of Congress. Network development and MARC standards office. *MARC XML DTD*. <<http://www.loc.gov/marc/marcxml.html#marcdtd>>.
- [5] France. Ministère de la culture et de la communication. *BiblioML*. <<http://www.culture.fr/BiblioML>>.
- [6] MARC::XML - A subclass of MARC.pm to provide XML support. <<http://marcpm.sourceforge.net/dep/marc::xml/MARC-XML.html>>.
- [7] Virginia tech digital library research laboratory. *MARC XML*. <<http://www.dlib.vt.edu/projects/OAI/marcxml/marcxml.html>>.
- [8] Library of Congress. Network development and MARC standards office. *MARC 21 XML Schema*. <<http://www.loc.gov/standards/marcxml>>.
- [9] Library of Congress. Network development and MARC standards office. *Metadata Object Description Schema (MODS)*. <<http://www.loc.gov/standards/mods>>.
- [10] Dick R. Miller. *XML and MARC: a choice or replacement?*. Testo presentato al Joint MARBI/CC:DA meeting presentation at ALA, July 10, 2000.
Slides: http://xmlmarc.stanford.edu/ALA_2000_Chicago/index.htm.
Testo: http://xmlmarc.stanford.edu/ALA_2000.htm.

- [11] Dick R. Miller. *XML: libraries' strategic opportunity*. «Library journal. NetConnect supplement», summer 2000, p. 16-19. <<http://elane.stanford.edu/laneauth/LJ>>.
- [12] Chr. Johnson Bruce. *XML and MARC: which is "right"?* «Cataloging & classification quarterly», 31 (2001), n. 1, p. 81-90.
- [13] Stanford University. Lane medical library. *Medlane*. <<http://laneweb.stanford.edu:2380/wiki/medlane>>.
- [14] International federation of library associations and institutions. Study group on the functional requirements for bibliographic records. *Functional requirements for bibliographic records: final report*. München: Saur, 1998. Trad. it.: *Requisiti funzionali per i record bibliografici: rapporto conclusivo*. Roma: ICCU, 2000.
- [15] Pino Buizza. *Dai Principi di Parigi a FRBR*. «Bibliotime», 5 (2002), n. 1. <<http://www.spbo.unibo.it/bibliotime/num-v-1/buizza.htm>>.
- [16] *Seminario FRBR: Functional requirements for bibliographic records = Requisiti funzionali per record bibliografici, Firenze, 27-28 gennaio 2000*, atti a cura di Mauro Guerrini. Roma: AIB, 2000.
- [17] Carlo Ghilli – Mauro Guerrini. *Introduzione a FRBR: Functional requirements for bibliographic records = Requisiti funzionali per record bibliografici*. Milano: Bibliografica, 2001.
- [18] Tom Delsey. *Functional analysis of the MARC 21: bibliographic and holding format*. <<http://www.loc.gov/marc/marc-functional-analysis/home.html>>.
- [19] Patrick Le Boeuf. *L'impact di modèle FRBR sur les révisions à venir des ISBD: un défi pour la Section de Catalogage de l'IFLA*. In: *67th IFLA council and general conference: Boston, August 16-25, 2001*. <<http://www.ifla.org/IV/ifla67/papers/095-152af.pdf>>
- [20] Istituto centrale per il catalogo unico. Commissione RICA. *L'applicazione del modello FRBR ai cataloghi: problemi generali e di impiego normativo, 22 ottobre 2001*. <<http://www.iccu.sbn.it/rica-frbr.doc>>.
- [21] Isa De Pinedo - Alberto Petrucciani. *FRBR and the revision of the Italian Author Cataloging Rules (RICA)*. In: *ELAG 2002: semantic web and libraries. 26. Library system seminar, Rome, 17-19 April 2002*. <<http://www.ifnet.it/elag2002/papers/pap5.html>>.
- [22] Isa De Pinedo. *Prospettive nell'applicazione degli FRBR alla revisione delle RICA*. <<http://www.iccu.sbn.it/DePinedo.doc>>. In: *Catalogazione e controllo di autorità: giornate di studio: Roma, Istituto della enciclopedia italiana, 21 e 22 novembre 2002*. <<http://www.iccu.sbn.it/ricaaf.html>>.
- [23] Alberto Petrucciani. *Struttura delle norme di scelta dell'intestazione: le RICA e i nuovi modelli di analisi*. <<http://www.iccu.sbn.it/PetruccianiRICA.doc>>. In: *Catalogazione e controllo di autorità: giornate di studio: Roma, Istituto della enciclopedia italiana, 21 e 22 novembre 2002*. <<http://www.iccu.sbn.it/ricaaf.html>>.
- [24] David Bearman. *A common model to support interoperable metadata: progress report on reconciling metadata requirements from the Dublin Core and INDECS/DOI Communities*. «D-Lib magazine», 5 (1999), n. 1. <<http://www.dlib.org/dlib/january99/bearman/01bearman.html>>.
- [25] *AustLit: Australian literature Gateway*. <<http://www.austlit.edu.au>>.
- [26] Paul Hendrik Jørgensen. *Applicazioni pratiche di FRBR e RDF nella catalogazione, nel controllo di autorità e nella ricerca delle relazioni tra le risorse elettroniche*. In: *Le risorse elettroniche: definizione, selezione e catalogazione: atti del Convegno internazionale, Roma, 26-28 novembre 2001*, a cura di Mauro Guerrini, con la collaborazione di Stefano Gambari e Lucia Sardo. Milano: Editrice Bibliografica, 2002.
- [27] Cinzia Bucchioni – Serafina Spinelli. *Qualche riflessione su FRBR*. «Bibliotime», 3 (2000), n. 1. <<http://www.spbo.unibo.it/bibliotime/num-iii-1/buccspin.htm>>.

A DTD XML for FRBR

by Giovanna Granata

W3C's recent introduction of the eXtensible Markup Language (XML) and the luck that it has encountered from its very first appearance have given rise to a multitude of projects in fields that differ considerably from one another. Apart from the structuring of documents, the sector for which it was originally invented, due to its syntactic potential, the language is by now quite widespread for numerous other uses, from the planning of data bases based on the use of DTD (Document Type Definition) and XML tables for the layout of their internal structure, to the verification of the completeness and correctness of the files given through their functionality of control and the convalidation of their structure, even going so far as the definition of logical data transmission formats that use XML as a simple structural support: the possibility of dividing information according to complex plans, strictly encased thanks to the use of beginning and end markers, does in fact make it possible to represent large classes of data in a relatively simple manner.

It is precisely with reference to this last aspect, that of data transmission, that various experiments have been carried out in the library field linked especially to problems of bibliographic interexchange and to the necessity of overcoming the limitations that the traditional uses of MARC formats have gradually shown in the new scenario of global communications and the interoperability of data that accompanies the diffusion of Internet.

Actually, although it does have some advantages with respect to the original formats, the introduction of XML does continue to appear with a low added value, to the extent that it maintains their original structure, thus continuing to reproduce the same limits. The hypothesis that it is intended to check in the article is that of using the FRBR model for the construction of this new format. What is proposed is the construction of a DTD for FRBR, implemented in reference to an abstract model, even while aware that some operational choices have to be made prior to its concrete application in the treatment of catalographic material. In particular, because of the way in which FRBR is structured, its main objective not being that of specifying a format for data management, it is clear that no limits for the possibility of data interconnection are fixed in it, while in the hypothesis of its practical application it becomes necessary to define on the one hand the "boundaries" of the record, and on the other the boundaries between the entities themselves: at what level of complexity should we arrive in understanding the different relationships described by the study and where should we fix the limit at which from different expressions of a same work we pass to works that differ from one another? These are only some of the questions posed by those who checked the usability of the model for cataloguing reasons: the proposal presented does not intend offering a definitive solution to these, dealing initially only with the analysis of the structural aspects which are subordinate to the translation of the model itself into XML. Possible modifications to the DTD can be discussed later.

GIOVANNA GRANATA, Dipartimento di Studi medievali e moderni, Università di Chieti, Chieti Scalo, e-mail g.granata@unich.it.