



Il software RIMMF: «RDA thinking in action»

Eleonora Belpassi

In due date successive, 30 gennaio e 26 giugno 2015, durante gli incontri dell'American Library Association (ALA) tenutisi a Chicago e San Francisco, hanno avuto luogo due iniziative comuni di sperimentazione del software RDA in Many Metadata Format (RIMMF) per la creazione di dati conformi alle linee guida Resource Description and Access (RDA).

Gli eventi – organizzato dal RDA Technical Development Team – hanno visto il contributo attivo di delegati di molteplici istituzioni culturali, che singolarmente o in piccoli gruppi hanno lavorato all'importazione ed arricchimento di dati principalmente incentrati su Jane Austen, le sue opere ed ogni risorsa da queste derivanti o ad esse riconducibili; senza vincoli di genere o di supporto.¹

La sperimentazione aveva come scopo immediato la creazione di una r-ball (RIMMF-Ball/RDF-Ball),² ovvero un nucleo di risorse correlate conformi allo standard di contenuto RDA, gestibili ed arricchibili con i relativi designatori di relazione utilizzando il software RIMMF, ed esportabili in formato Resource Description

¹ Jane-athon, “Creating linked data for Jane Austen and Blade Runner”, <http://www.rdatoolkit.org/janeathon>.

² <http://rballs.info/topics/p/jane/janethon1/index.html>.



Framework (RDF)³ in linea con le esigenze dei linked data (James Hennesly 2014).

Fin dal suo nascere, la rete ha messo le istituzioni culturali di fronte alla sfida della condivisione e di una gestione che integri e superi la dimensione locale e la tendenziale chiusura dei depositi catalografici. I dati di differente natura, registrati secondo differenti standard, conservati in separate banche dati, andrebbero riconvertiti entro parametri comuni, interconnessi e resi interoperabili, mantenendo però la ricchezza e la qualità che viene dalle specifiche tradizioni (Smith-Yoshimura 2011). Una risposta a questa esigenza di innovazione e condivisione è RDA, con cui la grande tradizione della catalogazione fa un passo avanti e si collega definitivamente all'era digitale (Bianchini e Guerrini 2015).

RDA è un insieme di linee guida in grado di definire “cosa” registrare di ogni possibile risorsa dell'universo culturale, avendo come riferimento il mondo della rete in cui i dati vengono assemblati *on the fly*, a seconda delle particolari esigenze degli utenti e dei gestori e quindi senza vincolarli ad uno schema di registrazione preordinato (Bianchini e Guerrini 2014).

Questa caratteristica viene definita “neutralità tecnologica”, che allude appunto alla mancanza di schemi prefissati di implementazione, per cui non viene fornita una precisa normativa per realizzare una specifica registrazione di questo modello di contenuto (Coyle 2014).

RIMMF è stato elaborato con l'intento di dare corpo all'esigenza di rendere concreto lo standard di contenuto RDA.

L'ultima versione del software, RIMMF3, sviluppato da The MARC of Quality (TMQ) è proposto in libera distribuzione per ambiente Windows con licenza Creative Commons.⁴

³ <https://www.w3.org/RDF>.

Il progetto prevede una serie di formati in ingresso (MARC21, MARCXML - da MARC21 e le registrazioni native in formato RIMMF txt); mentre i dati in uscita, oltre al formato nativo RIMMF txt, sono disponibili in MARC21, in modalità di visualizzazione OPAC, in XML e soprattutto in RDF utilizzabile per i linked data (Danskin 2015).

Il cuore funzionale di RIMMF è costituito da *template* strutturati in base alle entità FRBR/FRAD Gruppo 1 e 2 (Work, Expression, Manifestation, Item; Person, Family, Corporate Body). Ciascuno dei *template* presenti può essere editato individualmente, e compilandone i campi si creano automaticamente specifici e opportuni collegamenti (di primo e secondo livello) tra le entità designate. Si viene così a costituire un set di *template* (WEMI) più i riferimenti ad uno o più “agent” (Person, Family, Corporate Body).

Questo set può essere visualizzato e gestito anche come r-tree, ovvero un diagramma ad albero espandibile. Le ramificazioni possibili mostreranno anche in questa visualizzazione ogni nuova relazione che verrà associata alle varie entità.

Per creare dei record, oltre alla compilazione ex novo del set di *template* WEMI, è possibile importare dati bibliografici o registrazioni d'autorità nei formati MARC 21 e MARCXML.

⁴ È possibile scaricare gratuitamente l'applicativo dal sito del produttore e utilizzarlo liberamente per usi non commerciali, <http://www.marcofquality.com/wiki/rimmf3/doku.php>.



Figura 1 Esempio di r-tree (relationship tree) a partire dell'opera *Se questo è un uomo* di Primo Levi

Il programma dispone di un tool di ricerca e importazione che si collega alle risorse catalogafiche della Library of Congress, a VIAF e alle informazioni di Wikipedia. Permette ricerche combinate nel database LC per file bibliografici e di autorità (vengono forniti solo i primi 10 risultati per ogni ricerca). I risultati, prima dell'importazione, possono essere controllati in un visualizzatore MARC per valutarne la pertinenza.

Effettuata la scelta viene offerta l'opportunità di accettare o scartare le importazioni. Infatti si apre una finestra di dialogo che mostra la struttura r-tree della risorsa. Possono essere accettate o scartate le singole entità spuntando la relativa casella. Questo permette di rifiutare entità già presenti nel proprio *Entity index*.

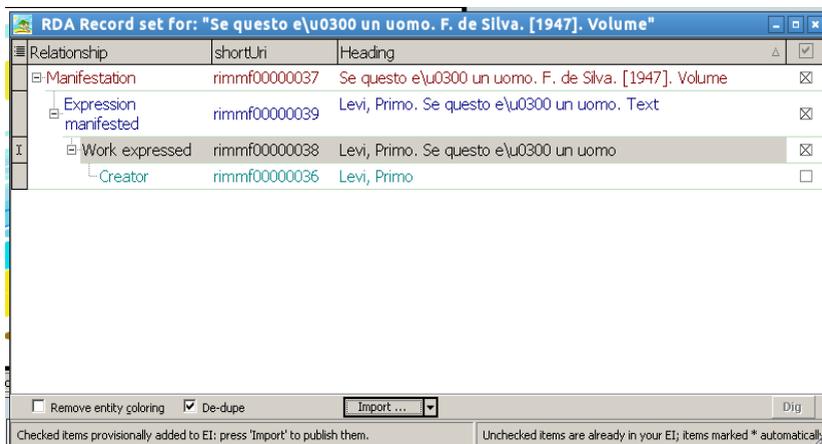


Figura 2 Maschera di visualizzazione di un esempio di set da importare con checkbox da spuntare.

I dati importati, vengono elaborati dal programma e inseriti nei *template* WEM più eventuali *agent*, a seconda della ricchezza del file MARC scelto. Si è specificato WEM e non WEMI, perché nel caso dell'importazione dei dati di risorse bibliografiche non vengono mai importati i dati relativi all'Item. Questa scelta è probabilmente dettata dal fatto che si tratta dell'oggetto fisico presente nella sede relativa, con le sue annotazioni specifiche (codici di catalogazione, molteplicità di copie, stato di conservazione, ecc.).

Un altro modo per importare dati è quello di recuperarli direttamente dai cataloghi online delle biblioteche che mettono a disposizione file scaricabili in MARC 21 o nel relativo MARCXML (il software RIMMF non riconosce UNIMARC, quindi per utilizzarlo sarebbe necessario un software di conversione UNIMARC → MARC 21).

Nel caso dell'interrogazione diretta dei cataloghi online, la ricerca e la scelta avvengono secondo le procedure offerte dalle diverse interfacce proposte dagli OPAC delle istituzioni prescelte.

Una volta scaricato il file, è sufficiente trascinarlo nella finestra di apertura di RIMMF per avviare l'importazione.

L'importazione di file da fonti diverse dalla LC e da alcune biblioteche dell'area anglosassone (Cornell University Library Catalog, National Library of Scotland, Trinity College Dublin Library Catalogue etc.) necessita di qualche intervento di revisione dei dati.

L'importazione dei dati MARC deve essere comunque rifinita e verificata. Vanno specificati in alcuni casi i corretti designatori di relazione RDA (ad esempio: *creator* → *author, composer, film maker* etc.; *contributor* → *translator, conductor, editor...*), vanno controllate le date riferite a *Work*, che riporta quasi sempre la data dell'edizione, e vanno specificate le eventuali fonti consultate per le correzioni.

Le possibilità di *editing* non si limitano all'arricchimento e alla precisazione dei campi RDA, ma possono essere gestiti gli *Authorized Access Point*, scegliendo le componenti che vanno a formarli nei *template Work* ed *Expression*. Possono essere aggiunte fonti consultate (*source consulted*) e arricchite le voci degli specifici *template* riferiti a opere di editoria piuttosto che produzioni multimediali, registrazioni musicali, o altro.

Ogni singola sezione e sotto-sezione RDA può essere aggiunta al *template* in rapporto alla sua relazione con il *Work, Expression, Manifestation, Item, Person, Family, Corporate Body*.

Salvando il *template* vengono salvate solo le voci compilate, ovvero alla sua riapertura saranno presenti solo i campi contenenti dati.

I *template* che compongono il set WEMI + *Agent* sono collegati tramite link incrociati di primo e secondo livello; possono essere stabiliti o eliminati link tra entità diverse, rappresentanti possibili relazioni che le legano (es. Traduttore ↔ opera tradotta; sceneggiatura ↔ opera adattata; opera musicale ↔

arrangiamento; et al.), che comunque devono essere presenti nel proprio entity index e definiti da designatori di relazione in RDA.

Aperto l'r-tree di un'entità compare una finestra in cui l'entità di partenza WEMI o *Agent* fa da radice all'albero delle relazioni collegate. Per espandere i nodi sottostanti, oltre ad aprire le relazioni dirette, si utilizza l'opzione 'dig', fino ad includere tutta la rete delle relazioni possibili all'interno del set.

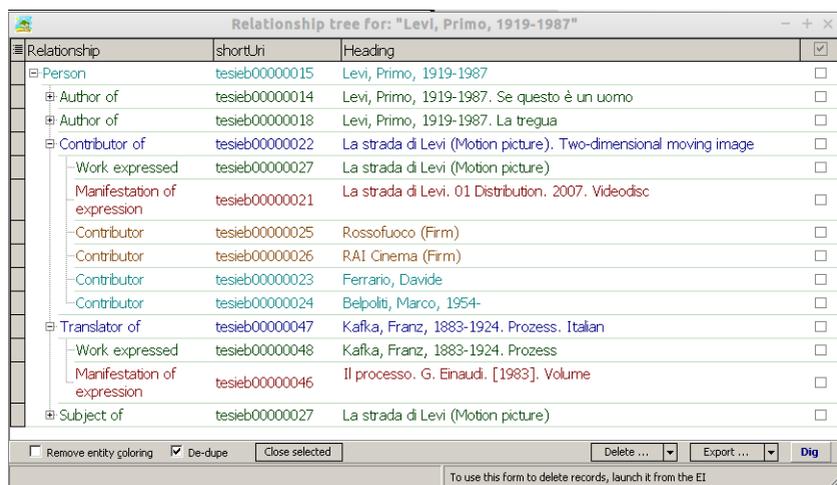


Figura 3 Esempio di r-tree con lo sviluppo di alcuni nodi

Cliccando su ciascuna voce dell'albero si accede al *template* corrispondente. Come d'altra parte all'interno dei *template*, cliccando sui link presenti, si aprono quelli collegati.

Sono presenti anche *hyperlink* alle fonti consultate (*source consulted*) e alle basi di dati citate nelle registrazioni importate (se presenti).

Oltre a questi ci sono i link diretti tra i singoli campi dei *template* e l'*element set* del *RDA Toolkit* corrispondente. È anche prevista l'impostazione della modalità di accesso al *Toolkit* stesso.

È importante notare che i dati contenuti nella directory radice dell'*Entity Index*, non sono vincolati ai singoli *template* e ad una organizzazione “a schede”, ma vengono composti secondo le esigenze di editing o visualizzazione, mantenendo i legami di relazione e la strutturazione ad albero (è possibile infatti l'esportazione dei dati in XML e RDF).

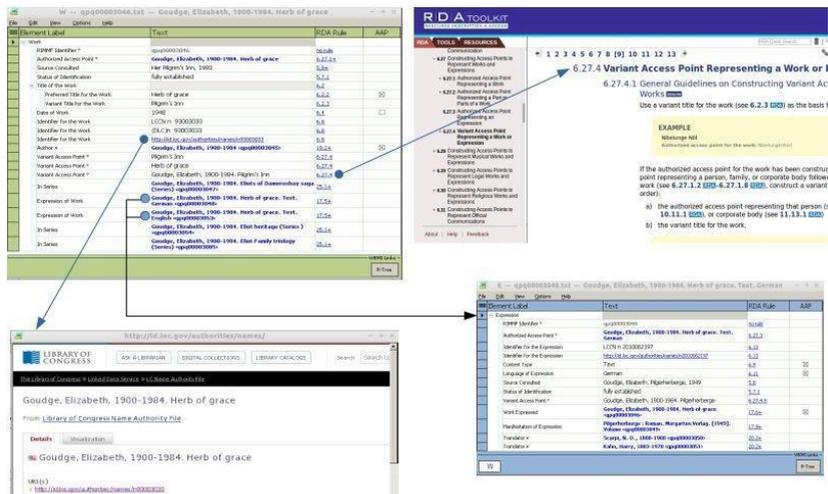


Figura 4 Visualizzazione schematica di alcuni esempi di link a partire dal *template Work*.

Editing dei *template*

I record importati nei *template* di RIMMF abbiamo già più volte detto che necessitano di correzioni e integrazioni, e soprattutto di interventi per definire adeguatamente le relazioni secondo le linee guida RDA.

Quando si compie una revisione nelle schede è necessario avere presente alcuni riferimenti generali. Eccone una lista a partire dai

core-element illustrati per FRBR da cui derivano i *template* di RIMMF.

Work	Expression
Title of the work	Form of expression
Form of the work	Date of the expression
Date of the work	Language of expression
Other distinguishing characteristic	Other distinguishing characteristic
Intended termination	Medium of performance
Intended audience	Context for the expression
Context of the work	Extent of the expression
Others...	Summarization of content
	Others...
Manifestation	Item
Title of manifestation	Item identifier
Statement of responsibility	Provenience of the item
Edition/Issue designation	Marks/Inscriptions
Place of publication/distribution	Exhibiton history
Publisher/distributor	Condition of the item
Date of publication/distribution	Treatment
Series statement	Access restriction of the item
Form of carrier	Fingerprint
Extent of carrier	
Physical medium	
Others	

Tabella 1: Lista di *core-element*

Naturalmente le revisioni avvengono in base alle voci disponibili o a quelle di cui si intendono arricchire i *template*.

Alcuni esempi di interventi effettuati durante la sperimentazione del software:

- Nel caso di aggiunta di componenti all'*access point* è necessario un doppio click sul campo AAP per aggiornarne la dicitura.
- Controllo della data dell'Opera (in mancanza di specifici Authority file riguardanti il *Work*, la data viene attribuita dal software, utilizzando quella della *manifestation*, con evidenti possibili errori).
- Inserimento della *source consulted* da cui è stata reperita una eventuale informazione (ad esempio un link alla pagina Wikipedia relativa)
- Cancellazione di *variant access point* presenti, se identici al *preferred title*.
- Controllo dei designatori delle relazioni dell'opera rispetto agli eventuali *agent*.
- Specificare i *relationship designators* (per es. *creator* → *author*; *contributor* → *translator* ...).

Caratteristica qualificante del software è la possibilità di creare collegamenti navigabili tra le risorse presenti nell'*Entity Index*, costituendo così quella che viene definita una *r-ball* (nella designazione la *r-* è collegabile a RDA, RIMMF, RDF e Resource).

La possibilità di attuare la scelta del nodo centrale intorno al quale sviluppare la *r-ball* e i conseguenti arricchimenti, permette di individuare intenzionalmente un riferimento intorno al quale articolare la costituzione del nucleo dei dati in uscita. Non si tratta quindi di una semplice trasformazione di dati presenti in deposito, esportati in formato RDF.

La *r-ball*, una volta arricchita dai collegamenti che si sono creati all'interno dell'*Entity Index*, contiene un dataset interrogabile, sia

internamente al programma RIMMF come un database – tramite una maschera di interrogazione presente in fondo alla finestra dello stesso *Entity Index* – sia soprattutto, dopo la sua esportazione in RDF, e il caricamento in un RDF *triplestore*, disponibile alle interrogazioni online con un linguaggio come SPARQL, predisponendo adeguate interfacce. Questo apre alle potenzialità di mettere a disposizione i dati elaborati nelle r-ball, per un utilizzo in rete, e questi dati possono essere generati da iniziative specifiche e mirate, anche da istituzioni di medie e piccole dimensioni, in possesso del patrimonio delle registrazioni MARC che RIMMF è in grado di convertire e riutilizzare esportandoli poi anche in RDF.

Conclusioni

I dati delle risorse che RIMMF mette insieme in un modo che non è presente nelle diverse fonti dalle quali si sono acquisiti, rende anche possibile interventi di arricchimento e sviluppo di collegamenti inclusivi, permettendo di costituire una rete di relazioni che associano caratteristiche dei dati bibliografici e di quelli di autorità, all'interno dei quali si possono stabilire anche relazioni di derivazione, di affinità, di associazione, di contributo, tanto da potere sviluppare delle reti ampliabili e non predeterminate.

In questa prospettiva i dati delle risorse arricchite e integrate strutturalmente possono essere esportati creando la RDF-ball e così raggiungere la rete, mantenendo la loro derivazione qualitativa dalle istituzioni culturali che le hanno prodotte, e arricchendosi di tutte quelle relazioni che attraverso questo processo di riqualificazione possono essere aggiunte da contesti culturali istituzionali o privati e dalle comunità degli studiosi.

RIMMF fornisce anche l'occasione per confrontarsi con le problematiche della realizzazione delle registrazioni secondo le

linee guida RDA,⁵ per non dimenticare che uno dei suoi scopi dichiarati all'origine era quello didattico.

Nella pratica dell'utilizzo, le definizioni delle linee guida vengono esplorate più approfonditamente, incrociandosi con la contestualizzazione attuata da RIMMF, in una reciproca verifica. Infatti, dalla sperimentazione degli incontri ALA di Chicago e San Francisco, sono emerse considerazioni operative, soprattutto relativamente ai designatori di relazione RDA, portate poi all'attenzione dello specifico gruppo di lavoro JSC (tra gli stessi organizzatori del Jane-athon ci sono membri del nuovo Joint Steering Committee for Development of RDA).

Nel novembre 2006, la Biblioteca del Congresso ha convocato un gruppo di lavoro per esaminare il futuro del controllo bibliografico nel 21° secolo. Il gruppo di lavoro ha interpretato questo compito nel modo più ampio, esaminando le tendenze e le prassi del momento, gli sviluppi nuovi ed emergenti e la crescente gamma di partecipanti nell'ambiente in continua evoluzione della produzione, distribuzione e uso della conoscenza. Sono state individuate delle aree di lavoro tematiche:

- Utilizzatori e modi d'uso dei dati bibliografici;
- Strutture e standard per il controllo bibliografico;
- Economia e organizzazione del controllo bibliografico.

Il gruppo di lavoro ha previsto, per il controllo bibliografico, un futuro che sarà collaborativo, decentrato, di portata internazionale, e *web-based*. La realizzazione di questo futuro avverrà in collaborazione con il settore privato e con la collaborazione attiva degli utenti della biblioteca. I dati saranno

⁵ RDA Toolkit. "Examples of full RDA records (JSC) – Bibliographic Records". http://www.rdatoolkit.org/sites/default/files/6jsc_rda_complete_examples_bibliographic_apr0913_rev.pdf.

raccolti da più fonti, il cambiamento avverrà rapidamente e il controllo bibliografico sarà dinamico, non statico.

Nel maggio 2011 la stessa Library of Congress ha annunciato la Bibliographic Framework Transition Initiative (BIBFRAME): il nuovo progetto di struttura bibliografica focalizzato sull'ambiente Web, secondo i principi e le tecnologie Linked Data, e il modello dati RDF. Viene così sancita la linea di sviluppo conforme al web semantico e l'adozione ufficiale della tecnologia ad esso collegata.

Il 21 novembre 2012 ancora la Library of Congress pubblica il report *Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services* (BIBFRAME),⁶ dando l'avvio a ciò che si prospetta come una nuova era per lo sviluppo della catalogazione.

RIMMF e altre simili iniziative che cercano di riqualificare i dati presenti in MARC21, UNIMARC, e altri standard presenti nei depositi istituzionali, dovranno anche confrontarsi con il nuovo cambiamento in corso, che trasformerà i formati storici a cui ci stiamo riferendo in un nuovo standard di registrazione: BIBFRAME (Bibliographic Framework Initiative) (Shieh 2013).

BIBFRAME andrà a sostituire MARC ed è predisposto per avere molte delle caratteristiche che stiamo affrontando qui: la sua totale integrazione con le esigenze della rete, il web semantico e i linked data; e offrirà dei tool di lavoro, rendendo concreta questa ipotesi di cambiamento (Baker, Coyle e Petiya 2014). Quindi dopo la trasformazione concettuale interverrà una trasformazione operativa di cui si dovrà prendere atto nel corso del suo sviluppo.

RIMMF si colloca in questa fase di transizione: la sperimentazione di RDA ha fatto crescere l'esigenza di strumenti operativi immediatamente utilizzabili per vedere attuate registrazioni di dati conformi alle sue linee guida. D'altra parte

⁶ BIBFRAME. "Documentation", Library of Congress. <http://www.loc.gov/bibframe>.

l'evidente inadeguatezza di MARC al contesto della rete, e il suo futuro abbandono, creeranno presumibilmente un processo lungo e laborioso (anche per la necessaria conversione della mole dei dati diffusi in questo formato).

Da queste esigenze – e proprio da specialisti della gestione di MARC – è partita l'iniziativa di RIMMF: effettuare registrazioni strutturate secondo il modello FRBR e le definizioni delle linee guida RDA, in modo da permettere anche di utilizzare le registrazioni MARC degli archivi correnti e sia in grado di confluire nei futuri repository delle triple RDF.

La particolarità che caratterizza RIMMF è la scelta di un riferimento attorno al quale aggregare il nucleo dei dati (che restano pur sempre riutilizzabili in altri modi senza esigenze di riconversione). Il web semantico sognato da Berners-Lee è una rete infinita di rimandi reciprocamente definentisi: RIMMF dà la possibilità di costituire la sua r-ball con un criterio di scelta deliberato.

Da un primo nucleo di dati, infatti, saremo in grado di costruire, attraverso opportuni collegamenti instaurati tra gli elementi della descrizione delle risorse, una r-ball molto grande, costituita da risorse associate, così da avere tutte le risorse connesse, i collegamenti alle autorità, etc. e la possibilità di visualizzare il prodotto della nostra catalogazione in vari formati.

RIMMF permette di mantenere il controllo del processo di strutturazione della rete delle risorse costituite, mantenendo l'identità del catalogatore e la consapevolezza dell'importanza del ruolo che possono occupare le biblioteche ed altre istituzioni culturali come agenti attivi nel Web.

L'opportunità che RIMMF offre per sperimentare concretamente le linee guida RDA è (per adesso) libera e gratuita, ciò significa che chiunque e qualsiasi istituzione può sperimentare le nuove linee guida di catalogazione. Mentre nel caso di BIBFRAME il

sistema (tuttora in evoluzione) è creato da un'istituzione specifica, la Library of Congress con partner privati di alto profilo, RIMMF offre la possibilità, una volta scaricato, di essere liberamente sperimentato potendo contattare gli sviluppatori, in un'ottica collaborativa.

Quello che ci si può attendere dai depositari delle tradizioni culturali con le loro eredità di conoscenze e competenze, le loro garanzie di affidabilità e controllo critico, è una aperta e consapevole accettazione della sfida tecnologica della rete e dell'orizzonte del web semantico a cui potranno apportare un contributo qualitativo degno della loro storia.

Bibliografia

- Baker, Thomas, Karen Coyle, and Sean Petiya. 2014. "Multi-Entity Models of Resource Description in the Semantic Web: A Comparison of FRBR, RDA and BIBFRAME." *Library Hi Tech*.
- Bianchini, Carlo, and Mauro Guerrini. 2014. *Introduzione a RDA - Linee Guida per Rappresentare E Scoprire Le Risorse*. Editrice Bibliografica. Milano.
- . 2015. "RDA: Resource Description and Access: The New Standard for Metadata and Resource Discovery in the Digital Age." *JLIS.it*. <http://leo.cineca.it/index.php/jlis/article/view/10963>.
- Coyle, Karen. 2014. "Multi-Entity Models of Resource Description in the Semantic Web: A Comparison of FRBR, RDA, and BIBFRAME." *Library Hi Tech*. <http://kcoyle.net/LHTv32n4preprint.pdf>.

E. Belpassi, *Il software applicativo RIMMF...*

- Danskin, Alan. 2015. "FRBR unMARCed: RDA Cataloguing with RIMMF." CIG Linked Data, February. <http://www.cilip.org.uk/cataloguing-indexing-group/presentations/linked-data-what-cataloguers-need-know-2015>.
- Hennelly, James. 2014. "Fun with Dick and Jane (and RDA): Creating Linked Data for Jane Austen and Blade Runner." *RDA Toolkit*. <http://www.rdatoolkit.org/janeathon>.
- Shieh, Jackie. 2013. "A Trasformative Opportunity: BIBFRAME at George Washington University, an Early Experimenter." *Evolution of Bibliographic Data Exchange*. http://www.niso.org/apps/group_public/download.php/11939/IP_Shieh_GWU_BIBFRAME_isqv25no4.pdf.
- Smith-Yoshimura. 2011. "Social Metadata for Libraries, Archives, and Museums." *OCLC Research*. <http://www.oclc.org/research/publications/library/2012/2012-01r.html>.

ELEONORA BEMPASSI, Università di Firenze. eleonora.belpassi@virgilio.it.

Belpassi Eleonora. Il software RIMMF: «RDA thinking in action». *JLIS.it*. Vol. 7, n. 1 (January 2016): Art: #11495. DOI: [10.4403/jlis.it-11495](https://doi.org/10.4403/jlis.it-11495).

ABSTRACT: RIMMF software is grew out of the need to visualize and realize records according to the RDA guidelines. The article describes the software structure and features in the creation of a r-ball, that is a small database populated by recordings of bibliographic and authority resources enriched by relationships between and among entities involved. At first, it is introduced the need that led to RIMMF outcome, then starts the software functional analysis. With a description of the main steps of the r-ball building, emphasizing the issues raised. The results highlights some critical aspects, but above all the wide scope of possible developments that open the Cultural Heritage Institutions horizon to the web prospective. Conclusions display the RDF-linked-data development of the RIMMF incoming future.

KEYWORDS: BIBFRAME; Cataloguing; R-ball; RDA; RIMMF.

Submitted: 2015-08-30

Accepted: 2015-09-23

Published: 2016-01-15

