

From Record Management to Data Management: RDA and New Application Models BIBFRAME, RIMMF, and OliSuite/WeCat

Mauro Guerrini & Tiziana Possemato

To cite this article: Mauro Guerrini & Tiziana Possemato (2016) From Record Management to Data Management: RDA and New Application Models BIBFRAME, RIMMF, and OliSuite/WeCat, Cataloging & Classification Quarterly, 54:3, 179-199, DOI: [10.1080/01639374.2016.1144667](https://doi.org/10.1080/01639374.2016.1144667)

To link to this article: <http://dx.doi.org/10.1080/01639374.2016.1144667>



Published online: 29 Feb 2016.



Submit your article to this journal [↗](#)



Article views: 352



View related articles [↗](#)



View Crossmark data [↗](#)

From Record Management to Data Management: RDA and New Application Models BIBFRAME, RIMMF, and OliSuite/WeCat

Mauro Guerrini^a and Tiziana Possemato^b

^aUniversity of Florence, Florence, Italy; ^b@Cult, Rome, Italy

ABSTRACT

The reflection provoked by RDA produced the awareness that the flat format of MARC 21 records is inadequate in expressing the relationships between bibliographic entities that the FRBR model and RDA standard consider fundamental. RIMMF and BIBFRAME indicate to software developers a way to think that is consistent with RDA. In Italy, @Cult, a software house and bibliographic agency working for Casalini Libri, has taken on the charge of following and facilitating the transition: OliSuite/WeCat provides an implementation of RDA that integrates vocabularies and ontologies already present in the Web by structuring the information in linked open data.

ARTICLE HISTORY

Received September 2015
Revised December 2015
Accepted January 2016

KEYWORDS

RDA; RIMMF; BIBFRAME;
OliSuite/WeCat; linked data;
Semantic Web; RDF

Library data in the Semantic Web

The tradition of publishing catalog data in electronic form is rooted in the history of libraries, more than in other communities. The MARC format was established precisely for the exchange of catalog metadata between different institutions; its full name—MACHINE Readable Cataloging—indicates the intent of the format to be read and interpreted by machines. This must not, however, mislead us to believe that libraries today are compliant with the requirements of current web paradigms, so that the details of their catalogs are *of the web* and no longer just *on the web*, where bibliographic and authority records may be used by programs and software. Libraries still have a long way to go for their bibliographic data to be really used in today's Semantic Web: some distance must be bridged between the reality of the data produced by cultural institutions and the techniques required to construct the Semantic Web. Libraries must upgrade not only their technology but also their conceptual framework.

The need for change has been voiced by libraries in various ways and on many occasions. The most significant of these calls for change, regarding the Semantic Web, is “On the Record,” the report of the Library of Congress Working Group on

CONTACT Mauro Guerrini ✉ mauro.guerrini@unifi.it 📍 University of Florence, SAGAS, Via S. Gallo 10, Florence 50129, Italy.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/wccq.

Published with license by Taylor & Francis. © Mauro Guerrini and Tiziana Possemato

the Future of Bibliographic Control,¹ first made public on January 9, 2008, which refers to the need to:

- transform textual description in data sets that can be used for processes and automatic processing by machines;
- make data elements uniquely identifiable within the informative context of the web;
- ensure data compatibility with technology and web standards;
- develop a more flexible, extensible metadata language, applicable in the whole context of the web.

The declaration of a move toward the Semantic Web is both clear and official, confirming what was announced as far back as 2005, with the publication of data schemas and ontologies in Resource Description Framework (RDF).

From metadata to significant data

Where do we begin to analyze the bibliographic record? Why is it now proving inadequate in the context of the Semantic Web? The history of library catalogs and archives shows bibliographic records being widely used since ancient times as descriptions of the resources collected by libraries and archives. These descriptions are comprised of metadata. In the conventional sense of the term, metadata is data regarding other data: in the descriptive record of a book or a film, for example, the metadata is the title, the name of the author or director, the year of publication, and so on. The main characteristics of metadata are:

- its artificial nature, built on the resource;
- its descriptive purposes;
- its usability by machines; metadata must be structured in such a way to be machine processed and, therefore, used by a computer.

The fields of MARC records that were used earlier to identify and document specific pieces of metadata respond fully to the first two characteristics: they are constructions connected with the object, or resource, to describe it, or rather to provide descriptive data for it.

However, in the Semantic Web, the first two characteristics are not sufficient. They must be accompanied by the third characteristic, which itself must be extended in meaning to embrace the sense of *comprehensibility to machines*. Metadata, in the context of the Semantic Web, is extended to become synonymous with *data that can be not only processed but also interpreted by a machine*. In this way the syntactic structures of metadata create the significant layer of the web that makes the meanings of the resources they describe and their relationships with other resources explicit to programs and software.

Bibliographic record versus RDF

The bibliographic record, a set of descriptive metadata for a resource, has always been at the center of libraries' attention and activities. The transition from paper catalogs to electronic catalogs, which increased dramatically with the introduction

of the MARC format in the mid-sixties, and the corresponding growth of interest in standards and information formatting (e.g., the International Standard Bibliographic Description [ISBD], from the end of the same decade), have amplified the centrality of the record. It has become the catalyst for a number of different services offered to users of various types of shared cataloging programs, inter-library loan and cooperative networks. The centrality of the record has, however, led to a series of drawbacks:

- The production of variant versions of standard records, to accommodate individual libraries' local needs, which are not always justifiable
- The preparation of records with high production costs, albeit very detailed, high-quality records
- The creation of "siloed resources" that for the most part are not usable outside the library world

In the Semantic Web, MARC records are considered a valuable source of information for the vast amount of *data* and extensive semantic encoding they contain. The set of metadata collected to describe a production of Shakespeare's *As You Like It*, in the example MARC record below, is rich in detail and every tag/field identifies an element's data type and its relation to the object described.

LC Control No.:	2009025024
LCCN Permalink:	http://lccn.loc.gov/2009025024
000	02384cam a2200445 a 450
001	15789102
005	20100204124659.0
008	090624s2009 enka b 000 0 eng
020	__ a 9780521519748 (hbk.)
020	__ a 0521519748 (hbk.)
020	__ a 9780521732505 (pbk.)
020	__ a 0521732506 (pbk.)
035	__ a (OCoLC)ocn416716652
040	__ a DLC c DLC d YDXCP d BWKUK d BWK d FDA d UKM d BTCTA d CEF d CDX d DLC
050	00 a PR2803.A2 b H35 2009
082	00 a 822.3/3 2 22
100	1_ a Shakespeare, William, d 1564–1616.
245	10 a As you like it / c edited by Michael Hattaway.
250	__ a Updated ed.
260	__ a Cambridge, UK; a New York : b Cambridge University Press, c 2009.
300	__ a xv, 240 p. : b ill.; c 24 cm.
490	0_ a New Cambridge Shakespeare
504	__ a Includes bibliographical references (p. 240).
505	0_ a Introduction. Journeys; Plays within the play; Theatrical genres; Pastoral; Counter-pastoral; The condition of the country; Politics; 'Between you and the women the play must please'; Gender; Nuptials; Sources; Date and occasion; Stage history; Recent critical and stage interpretations – Note on the text – List of characters – The play – Textual analysis – Appendixes : 1. An early court performance?; 2. Extracts from Shakespeare's principal source, Lodge's <i>Rosalind</i> ; 3. The songs.
650	_0 a Fathers and daughters v Drama.
650	_0 a Exiles v Drama.
600	10 a Shakespeare, William, d 1564–1616. t As you like it.
655	_7 a Pastoral drama. 2 gsafd
655	_7 a Comedies. 2 gsafd
700	1_ a Hattaway, Michael.

However, the semantics expressed in the record can only be understood by a human or by programs that recognize the MARC tagging. The *individual elements of information* (see below) become meaningless if the semantic designations of MARC are ignored or not recognized:

- xv, 240 p.
- Includes bibliographical references (p. 240).
- Cambridge (UK)
- New York
- Cambridge University Press, 2009.
- edited by Michael Hattaway.
- Updated ed.
- Fathers and daughters- Drama.

A MARC record searched on the traditional web loses much of its *meaning*. Search engines treat such information as *blocks of text* and therefore the data is not expressed individually and unequivocally. The entire record on the other hand, with its complexity and wealth, can be exchanged, that is, reused by those who transfer a copy of the same record to describe their copy of the same resource in their own systems. One drawback of this is that data, which may be re-usable in many contexts, must be repeated in each MARC record to maintain its integrity. In the case of a new edition of the same work, cataloging rules consider the new edition a new resource that requires the creation of its own new record, even though the new bibliographic record may differ from the previous version only in the “edition” field. Furthermore, records created for an online public access catalog (OPAC) in the MARC format were previously used almost exclusively by libraries. Even many parts of the publishing community do not use MARC to exchange data with libraries or related services.

In the context of the Semantic Web, *individual elements* can regain the meaning that the entire MARC record with tags gave them by transferring the tagged data into RDF statements with a subject, a predicate, and an object. RDF statements enable machines to understand the meaning, moving us from a traditional web to the Semantic Web.

As an example, in [Figure 1](#) you can see some of the original elements from a MARC record, translated into RDF statements: the *statement* in the RDF model is a brief sentence formed by a subject (in the example the resource named “As you like it”) a predicate (“has subject”) and an object (“Fathers and daughters – Drama”). The resource “As you like it” has different attributes (author, editors, publication date, etc.) expressed through triples of subjects, predicates, and objects. For example, the editor attribute can be stated as the triple: “As you like it” (subject), is edited by (predicate), Michael Hattaway (object). Such a statement is available for reuse in the Semantic Web and is especially useful when the statement incorporates Uniform Resource Identifiers (URIs). Each triple so produced (e.g., “As you like it edited by Michael Hattaway”) is reported in a language that the machine can understand: each element of the triple, that can be identified by a

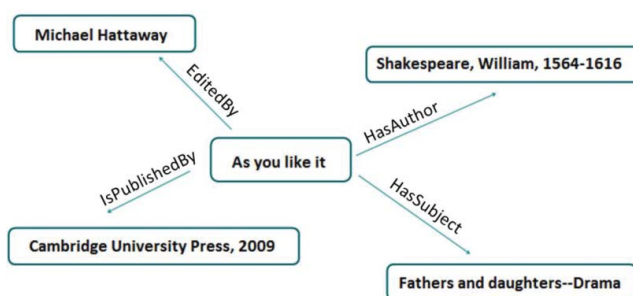


Figure 1. Elements and statements for bibliographic description on the web.

URI (or, in case of triple literal at least the subject and the predicate are URIs, while the object is a literal text) can be dereferenced by the machine: this means that the machine can read a definition expressed in the ontology for each entity used in the triple and understand its meaning.

Each element of the statement (subject – predicate – object) produced in this way can be recorded with a URI. The transfer of these statements in a way that can be understood by machines in the web transforms the data contained in the original MARC records into *integral parts of the Semantic Web*.

New cataloging models and linked data

The adoption of the logic and technology of the Semantic Web can have a different impact in various adoption contexts. Publishing open data or linked open data in an industrial context, for example, can mean registering details in Excel tables or converting databases in RDF, probably without changing the production processes of the data itself. In the case of libraries and archives the transition is generating a profound rethinking of current cataloging models, producing the re-establishment of principles and standards of cataloging regulations. The community based around libraries, such as library information system producers, is responding to this developmental need. Even the library community itself is making this transition, for example, the International Cataloguing Principles (ICP),² issued by International Federation of Library Associations and Institutions (IFLA) in 2009, intentionally talk about data rather than records.

Roy Tennant, in 2002, in an article published in *Library Journal* with the provocative title “MARC must Die,”³ marks the beginning of a new course of evolution. A wiki page dedicated to MARC was significantly (and ironically) changed from *MARC must die!* to *MARC is dead*.⁴

Resource Description and Access (RDA), initially conceived as an evolution of *Anglo-American Cataloguing Rules*, Second Edition (AACR2), opens up new prospects, manifesting itself as an international standard for assigning metadata to resources in the digital age. Guidelines (rather than cataloging rules) move within

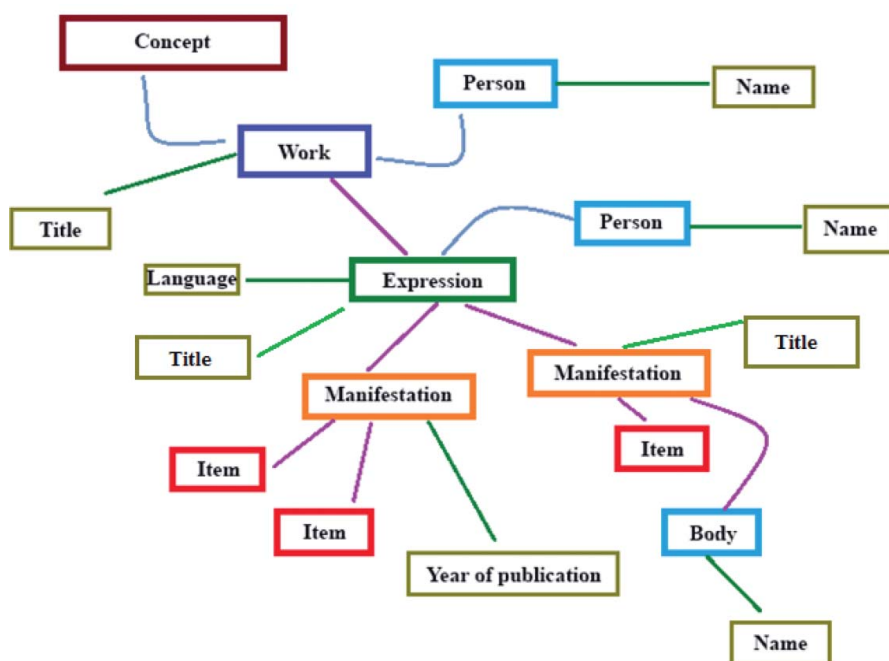


Figure 2. FRBR model proposed in RDA as a pattern for the development of rules in terms of linked data.

the philosophy of the Semantic Web and linked data. RDA has undertaken the publication of RDF representations of its terms for categories and relationships and provided guidance on the creation of linked data descriptions of new entity categories—persons, corporate bodies, families, and so on typically described in name authority files. Such representations can be reused in different contexts and for multiple purposes (see Figure 2).

The Figure 2 overview shows the entities, with some relationships and possible attributes. The image depicts in graphic form the resources in the Semantic Web, paying particular attention to explicit relationships between resources.⁵

The most significant and most important turning point for the cataloging community and for the future of universal bibliographic control has been the reflection provoked by RDA testing in 2010. The most consistent result was the awareness that the flat format of MARC 21 records is inadequate in expressing the *relationships* between bibliographic entities that the FRBR model and RDA standard consider fundamental. From this conclusion, in May 2011 the Library of Congress announced the *Bibliographic Framework Transition Initiative*, declaring in October 2011: “The new bibliographic framework project will be focused on the web environment, Linked Data principles and mechanisms, and the Resource Description Framework (RDF) as a basic data model.”⁶ That declaration from the Library of Congress ratified the transition from the traditional web to the Semantic Web with the official adoption of the technology connected to it.

An interesting testimony of the transition was the birth of the World Wide Web Consortium (W3C) Library Linked Data Incubator Group, whose purpose is to support and enhance global interoperability of library data on the web, by encouraging dialogue between experts of the Semantic Web.⁷ The report *Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services* (BIBFRAME),⁸ published by the Library of Congress on November 21, 2012, laid out the future.⁹

BIBFRAME

BIBFRAME is the result of a process of reflection on the functions of the catalog, the modernity of MARC and cataloging regulations, and of new types of resources in the bibliographic world. It represents the first step in defining a roadmap that is expected to propose a new bibliographic environment, which foresees the interconnection of entities generated by different sources and the enunciation of strategies and instruments capable of supporting the evolution in due course in the global communication circuit. It should, in particular, lead to the articulation of a new *bibliographic ecosystem*. The new instructions for the description and access to resources should be characterized by:

- a high level of analysis and *identification* of the data;
- an emphasis on *relationships*;
- widespread use of *controlled vocabularies*;
- an accurate, detailed description of resources;
- wide flexibility in the management of controlled items.

The model examines the evolution of cataloging standards and regulations and, therefore, makes a reflection on the role, the diffusion and the historical functions of MARC 21, and on the future of universal bibliographic control. Its theoretical development is based on the experience gained during the test phase of RDA. The possible conversion of bibliographic data into linked data is interpreted in terms of evolution rather than revolution. The considerations developed in the experimental phase of RDA are now deemed opportunities, if not necessities, for the use of the web as a model to express and connect information, according to the paradigm of the Semantic Web, that is to say through the structure of URIs and through the decentralization and division of the data. The expected results of the evolutionary process include the ability to achieve a simplification and optimization of descriptive processes for resources. That will allow, on the one hand, sharing of the description outside the library community, and on the other, the distribution of the task of attributing metadata at the international level among different cataloging agencies, in order to extend the coverage of bibliographic control and, subsequently, to better meet user requirements.

BIBFRAME offers an extremely simple data model, inspired by *Functional Requirements for Bibliographic Records* (FRBR) (but with significant differences); it aims to: distinguish between conceptual content and physical manifestations, that is, between the work and the instance of the work; identify an entity unambiguously;

highlight relationships between entities showing the nature of correlation. The model is structured in four main sections:

- *Work*: the conceptual essence of the cataloged resource, whose concept is similar to the *work* and the *expression* of FRBR
- *Instance*: an individual materialization of the work
- *Authority*: identification of the entities that have a relationship with the work or instance, for example, people, families, organizations, concepts, places
- *Annotations*: everything from the BIBFRAME model that decorates or enhances the resources with additional information, for example sample data, cover images, reviews, tables of contents, and so on

Figure 3 is extracted from BIBFRAME and represents the general structure of the model, with the work at the center of the graph, connected to the instance in which it occurs and the associated relationships.

At the center is the *work* with links to the *creator* and the *subject entry*, that is the descriptor expressing the main subject or theme; the work is connected to the

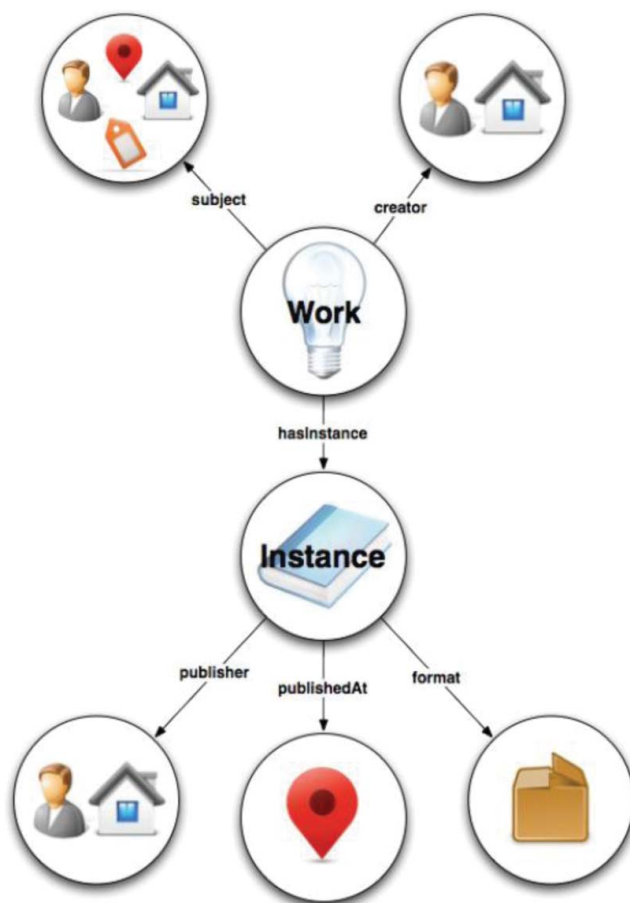


Figure 3. BIBFRAME: General structure of the model.

instance, with links to the publisher, the place and the format or mode of use: print edition, online edition, e-book, and so on.

To provide an example, Figure 4 shows a central work entitled *FRBR Report*. “IFLA Study Group” is the *creator*, and “Cataloging” the *subject describer*. *FRBR Report* has three instances, or three manifestations in FRBR terminology:

1. print version with cardboard cover;
2. pdf;
3. online version in HTML.

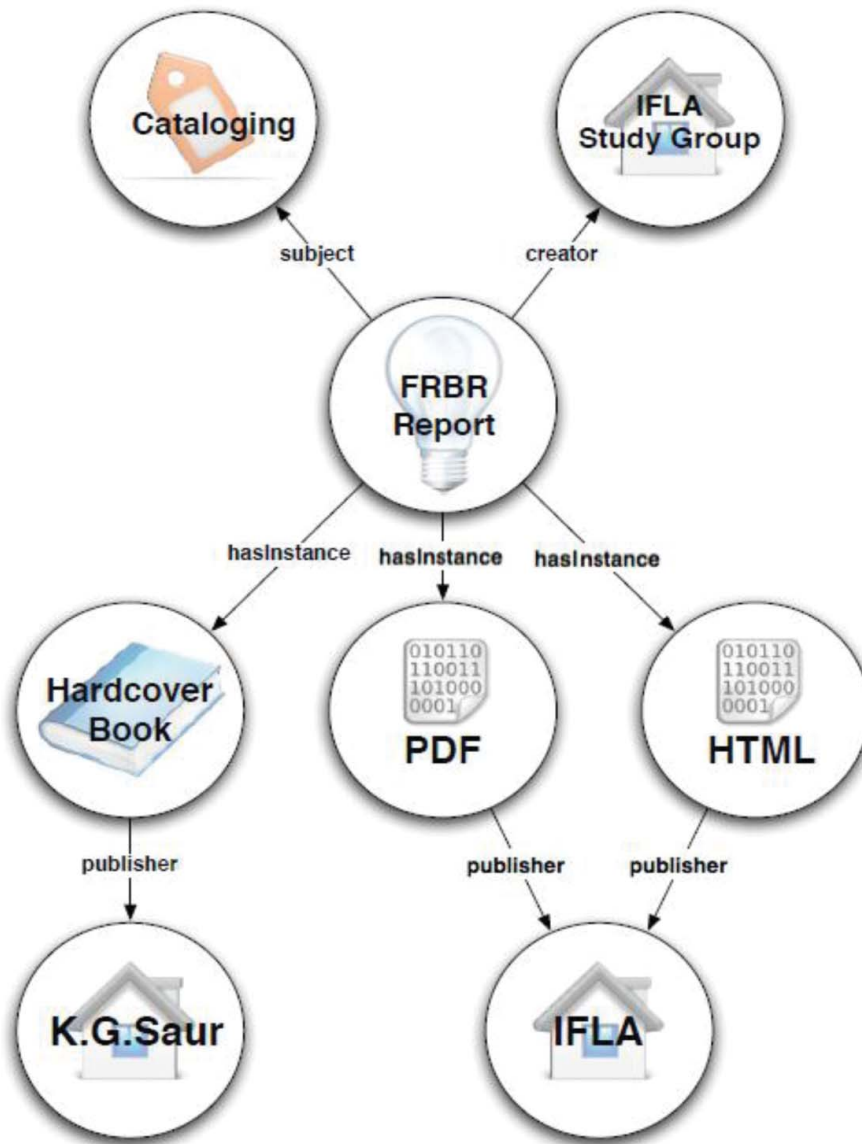


Figure 4. BIBFRAME: Example of the work *FRBR Report*.

The print edition with rigid cover is published by KG Saur, while the Portable Document Format (PDF) file and HTML are both published by IFLA.

BIBFRAME indicates to the library community and software vendors a path of change, centered on the information requirements of the user, data reuse in different contexts (interoperability), and the need for machines to be part of the process of creation and distribution of information. The project plans to assemble the data of the current “MARC resource” in a new logical architecture, enabling the cooperative assignment of metadata in a more self-contained descriptive level. The translation of the bibliographic records in linked data is, therefore, intended as a basis for starting a debate for discussion within the information community. The commitment to using the web as a place in which to express and connect information. BIBFRAME, therefore, is proposed as a reference model, according to which library catalogs will be able to truly enter the web; a web that is evolving from a network of documents to a network of data linked by qualified relationships.

Semantic Web technologies for the publication of data are already widespread in the world of cultural heritage. The same cannot be said of the methods and tools used for cataloging resources, for which we have no indications or examples of what these new systems for data creation will have to guarantee. BIBFRAME defines, however, some general guidelines that the new generation Integrated Library Systems (ILSs) have to follow, including:

- the prospect that they will be oriented towards a resource-based architecture (i.e., where the focus of description is a single BIBFRAME or FRBR logical entity rather than several entities compounded as a single object);
- the possibility that each resource thus identified or created can be linked to other resources, as in the entity-relationship FRBR model;
- the possibility that in the association of attributes to each resource the cataloger can draw information from the web, available according to the requirements of linked data, without having to create new attributes each time;
- the use of controlled lists of terms and, therefore, access in the construction phase of the entity to vocabularies and the lists of terms published for example on the Open Metadata Registry or on Vocab.org;
- the possibility to define in the system's configuration a logic for the creation of a URI, so that each new resource to be created is automatically attributed a URI.

All this expresses the general framework in which new ILSs must be rethought and written.

ILSs for the next generation

The passage from record to discrete statements linked by identifiers in the tools for mediation between the user and the bibliographic world is a change that no operator in the library sector can ignore. The new generation of ILSs will take into account the suggestions from linked data, proposing a new

model for creation, sharing and data representation. The development of data processing tools focuses on certain fundamental aspects, in particular:

- a greater division of single aspects (elements of description) in structuring information;
- the use of controlled vocabularies, ontologies, and authority files, to share a common language and extend such sharing to machines as much as possible;
- a clear identification of objects and entities, through identification codes and unique qualifying attributes;
- the creation of as wide a network as possible of connections between different entities, to ensure that each entity is connected to other entities with different specified relationship terms;
- a strengthening of the processes of discovery and, therefore, of sharing, of data in order to involve different operators who were traditionally excluded from these processes, such as publishers, who more than others are in contact with the authors, that is with those who produce information (primary sources).

These principles have been set as the starting points for the formulation of RDA. The most urgent challenge for manufacturers of new cataloging systems is therefore to integrate the principles of linked data in the construction of user-friendly systems.¹⁰

RDA in Many Metadata Formats (RIMMF)

RDA in Many Metadata Formats (RIMMF), developed by the company The MARC of Quality (TMQ),¹¹ is a prime example of an RDA-oriented system. The beta version was made available in January 2015 and can be downloaded for Windows, with Creative Commons licenses (license for free use, not for commercial use). RIMMF is an illustrative and educational tool for data visualization, which was created to help catalogers think in terms of RDA and, therefore, in FRBR, rather than in AACR2 and MARC. RIMMF is not, therefore, an ILS or a cataloging module, because the data created cannot yet be used for products, despite being exportable in various formats (RDF, XML, MARC). It can be considered more as a prototype of how an interface using RDA could be. Its value lies in the objective of simplifying the change of approach to cataloging, transforming it from the production process of a record to the process of identification and attribution of metadata to FRBR entities (Work, Expression, Manifestation, Item, Person, Family, Corporate Body, Concept, Object, Event, Place). The change is evident in the selection of the model to be used, in relation to the resource we describe: the focus on FRBR entities is immediately clear.

Figure 5 shows a sample entry webpage from RIMMF where descriptive data for a manifestation are provided. RIMMF workpages allow a cataloger to enter data for the Work, Expression, Manifestation, and Item (WEMI) entities associate with a particular resource.

Element Label	Text	RDA Rule	AAP
Manifestation			
RIMMF Identifier *	atcault00000013	no rule	
Composite Key *	The name of the rose, Harcourt Brace, 1994, Volume		
Title		2.3	
Title Proper	The name of the rose	2.3.2	<input checked="" type="checkbox"/>
Statement of Responsibility		2.4	
Statement of Responsibility Relating to Title Proper	Umberto Eco	2.4.2	
Statement of Responsibility Relating to Title Proper	translated from the Italian by William Weaver	2.4.3	
Edition Statement		2.5	
Designation of Edition	1st Harvest, ed	2.5.2	<input type="checkbox"/>
Publication Statement		2.8	
Place of Publication	San Diego	2.8.2	<input type="checkbox"/>
Publisher's Name	Harcourt Brace	2.8.4	<input checked="" type="checkbox"/>
Date of Publication	1994	2.8.6	<input checked="" type="checkbox"/>
Series Statement		2.12	
Title Proper of Series	A Harvest book	2.12.2	<input type="checkbox"/>
Title Proper of Series	Harvest in translation	2.12.2	<input type="checkbox"/>
Mode of Issuance	single unit	2.13	<input type="checkbox"/>
Identifier for the Manifestation	LCCN 94013818	2.15	
Identifier for the Manifestation	ISBN 0156001314	2.15	
Identifier for the Manifestation	http://ecw.loc.gov/94013818	2.15	
Note On Manifestation	Including the author's postscript.	2.17	
Note On Manifestation	A Helen and Kurt Wolff book	2.17	
Note On Manifestation	Includes English translation of: Il postume a il nome della rosa	2.17	
Media Type	unmediated	3.2	<input type="checkbox"/>
Carrier Type	volume	3.3	<input checked="" type="checkbox"/>
Extent statement *	536 p	3.4	<input type="checkbox"/>
Number of units *	536	3.4.1.3	
Type of Unit (Text) *	pages	3.4.5	<input type="checkbox"/>
Dimensions	21 cm	3.5	<input type="checkbox"/>
Uniform Resource Locator	http://www.loc.gov/casbi/description/har/94013818.html	4.6	
Uniform Resource Locator	http://www.loc.gov/casbi/description/har/94013818-b.html	4.6	
Expression Manifested	Eco, Umberto. Nome della rosa. English - atcault00000014-	17.1+	

Figure 5. RIMMF's page for the entry of data for a *Manifestation* entity.

The data listed for each entity are formulated as attributes of the resource and taken directly from the set of RDA elements, also available as an RDF element set on the Open Metadata Registry (<http://rdvocab.info/>). Each attribute is associated with the corresponding instruction of the RDA Toolkit, which requires authentication for access: in this way the use of an essential working tool for the new generation of catalogers is simplified. The functions currently available in RIMMF are:

- the creation of new models of cataloging;
- the creation of new catalog records;
- the importation of data from external databases;
- the creation of relationships between entities belonging to a single database;
- the display of the most simplified hierarchical structure (tree) created by

WEMI relationships that link the various entities describing a resource.

The possibility to view or import RIMMF data from external databases is of particular interest. External databases can include bibliographic databases and OPACs

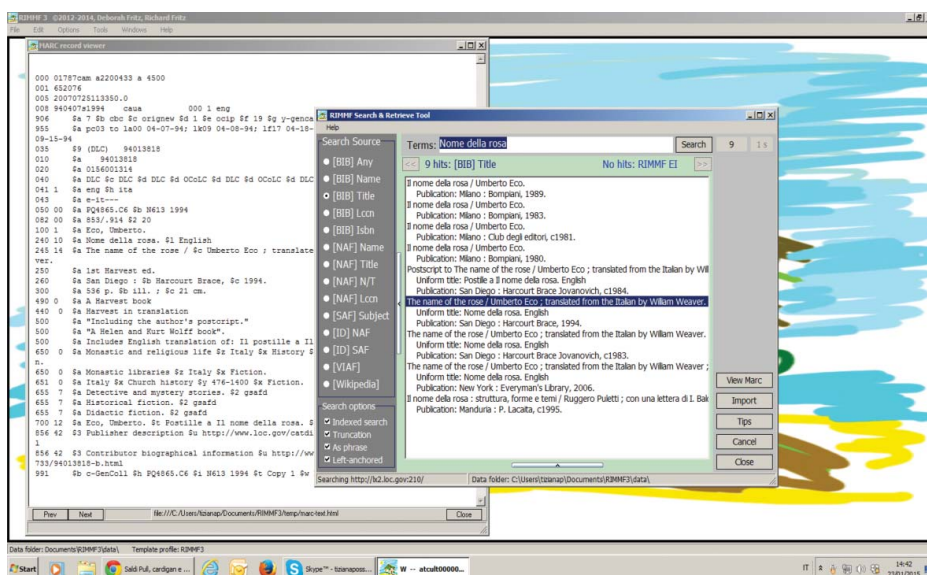


Figure 6. RIMMF's search function to find data in external databases.

(e.g., the Library of Congress), authority files (e.g., Virtual International Authority File [VIAF]), encyclopedias, and online directories (e.g., Wikipedia). Figures 6 and 7 show how RIMMF enables searching and importing. The view/import mechanism has a dual purpose:

1. to facilitate the sharing of data, to avoid creating from scratch data that already exist on other databases;
2. to allow for the creation of external links as early as the data structuring phase, bringing forward in the process one of the most important tasks in data structure.

Figure 8 shows the visualization of data in a hierarchical structure: this provides a clear and outlined overview of the complex network of relationships that can be built between different resources.

RIMMF3, available from January 2015, offers a series of new functions, such as:

- an element set based on the RDA Registry that includes application support for the many hundreds of new relationships (and their inverses);
- creation of the `rimmf.com` namespace to support all elements added by RIMMF to support RDA;
- improved MARC mapping with many different ways to get MARC data into RDA;
- support for the export of RDA entity records as RDF linked data.

OliSuite/WeCat

Next-generation software must necessarily take into account the RDA guidelines. Such software in the future will include tools that will help the transition from an

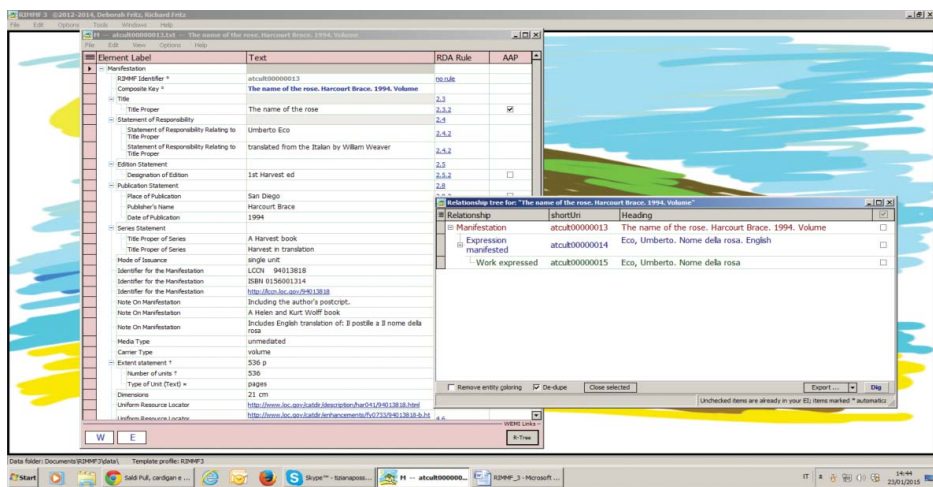


Figure 7. RIMMF's screen for importing records in MARCXML from the Library of Congress, mapped and visualized according to the structure of RDA.

era, lasting over 50 years, dominated by data structured in MARC format, to a new era, which has still not been perfectly delineated and established. However, BIBFRAME leads us to believe that the library sector will go in the direction of linked open data and express in its *language* an evolution toward new models of attributing metadata. The first effort that the new software already is making is precisely the adaptation of language, with the introduction of a terminology that has so far been absent from applications, but proposed by ICP, namely, the introduction of labels such as *authorized access point* as well as *identification of the work* or *identification of the expression* to replace *uniform title*, and so on. This new language is evidence of a notable change in the structure of the model for attributing

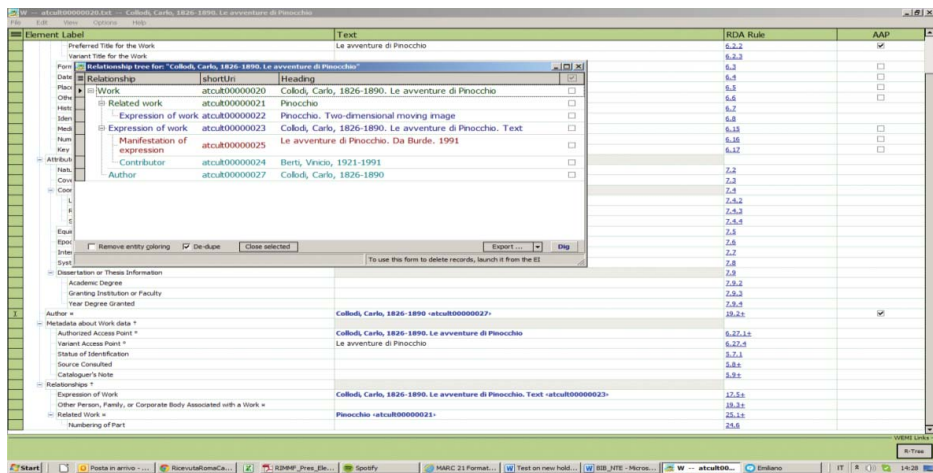


Figure 8. RIMMF's hierarchical visualization of the relationships between different entities.

metadata.¹² These changes of language in the interfaces of next-generation software are not yet fully supported by MARC 21 updates. MARC 21, for example, still retains the outdated terminology *Uniform Title* for the tags 130 and 240.

In Italy, @Cult,¹³ a software house and bibliographic agency working for Casalini Libri (and by Casalini for the Library of Congress), has taken on the charge of following and facilitating the transition from an environment based on records to an environment based on linked data statements. As an established manufacturer and distributor of ILSs in the world of libraries and archives (i.e., Amicus, which later evolved into OliSuite), @Cult has assimilated in the cataloging module named WeCat, part of the OliSuite ILS, the updates 9–20 of MARC 21 to address RDA, especially as it relates to content, media, and support (types and characteristics), attributes for names and resources, and all types of relationships. The WeCat module is a piece of software that, while still registering data in MARC 21, allows data to be structured according to RDA guidelines in a way, therefore, that is oriented towards linked open data (LOD). There are several features that express this orientation. The software offers specific codes for FRBR entities as can be seen already in the definition of search indexes for the construction of queries as shown in Figure 9. The attribution of metadata starts with the selection of models that offer catalogers the attributes envisaged by FRBR for each specific entity.

OliSuite/WeCat has enhanced its authority control functions addressing the need, common to RDA and the principles of linked open data, to describe the entity that is being analyzed with a high number of attributes, thus creating an object that is perfectly identifiable and reusable in many areas. The possibility to link work or expression descriptions and attributes (formerly relegated to authority records) in more efficient ways with manifestation descriptions is a brand new

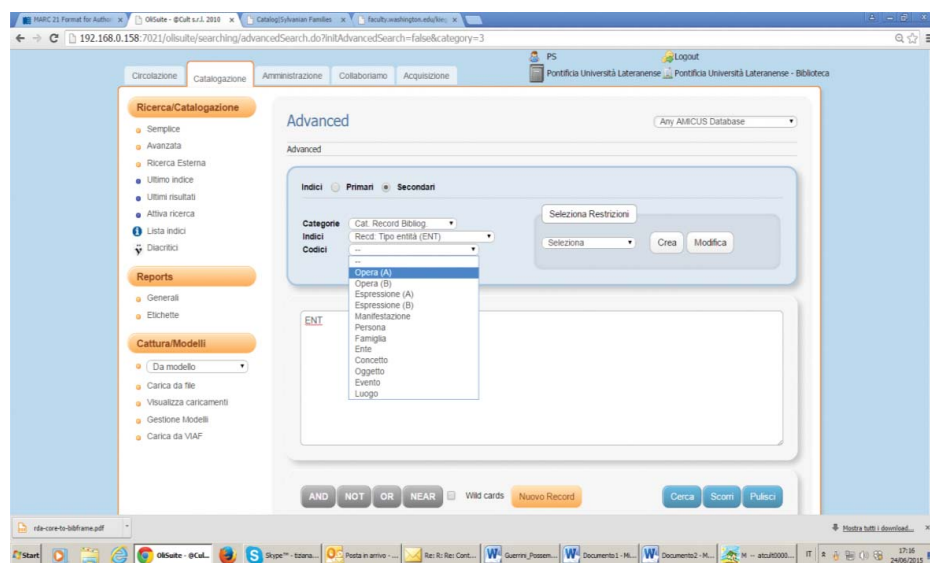


Figure 9. OliSuite/WeCat: The of search indexes for FRBR entities.

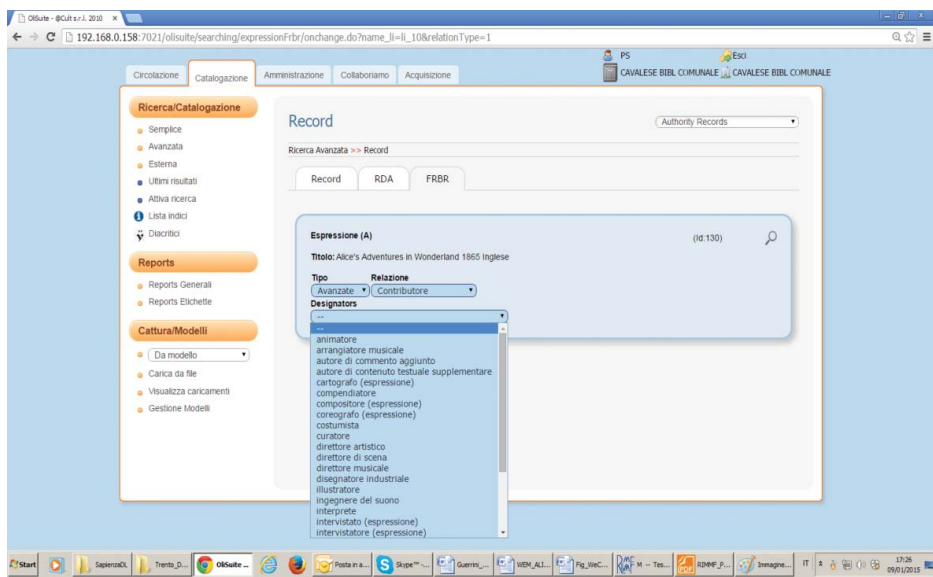


Figure 10. OliSuite/WeCat: Creation of relationships between an expression and a person (in the role of contributor).

element, introduced to allow the creation of a network of connections that constitutes the heart of both RDA and linked open data. Relationships between entities foreseen by OliSuite/WeCat reflect the different types of relationships foreseen by RDA appendices:

- I—Relationships between a resource and people, families, and organizations (corporate bodies) associated with the resource (e.g., see Figure 10),
- J—Relationships between Works, Expressions, Manifestations, and Items (e.g., see Figure 11), and
- K—Relationships between people, families, and organizations (corporate bodies) (e.g., see Figure 12)

of the RDA Toolkit, are supported in the function for the creation of relationships between resources.

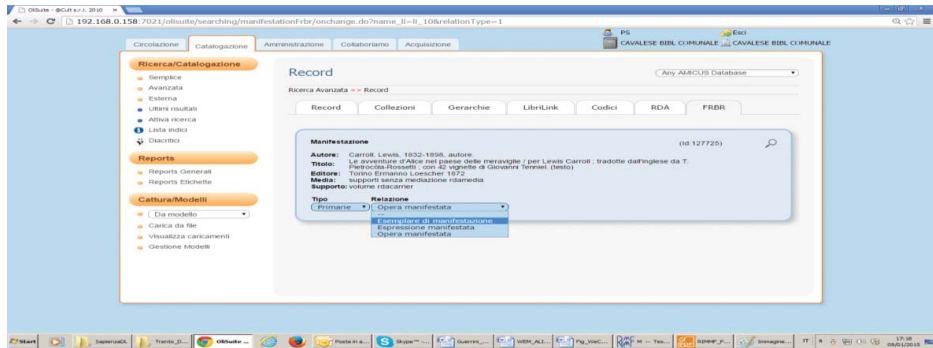


Figure 11. OliSuite/WeCat: Creation of relationships between a manifestation and a different entity of FRBR group 1.

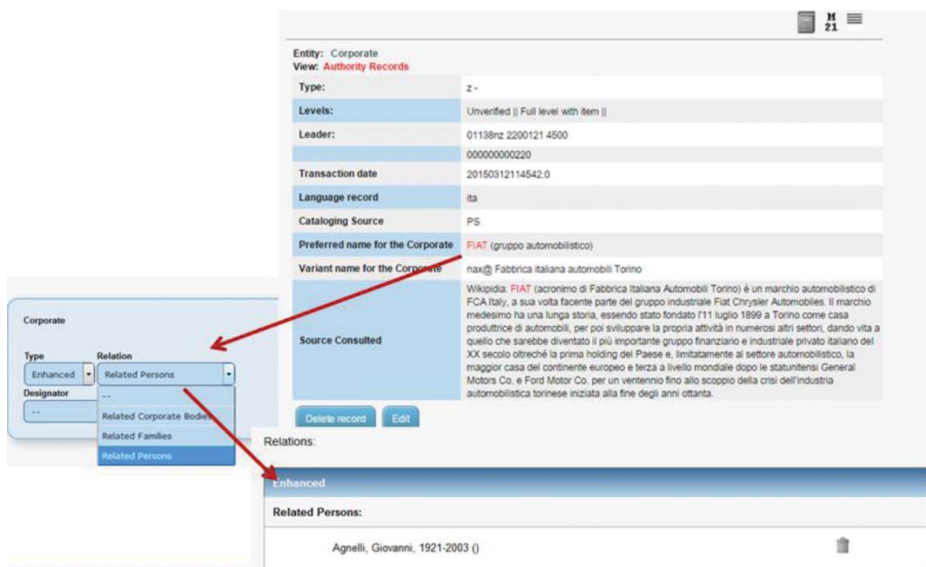


Figure 12. OliSuite/WeCat: Creation of relationships between a corporate body (*FIAT* - auto group) and a person (*Giovanni Agnelli*).

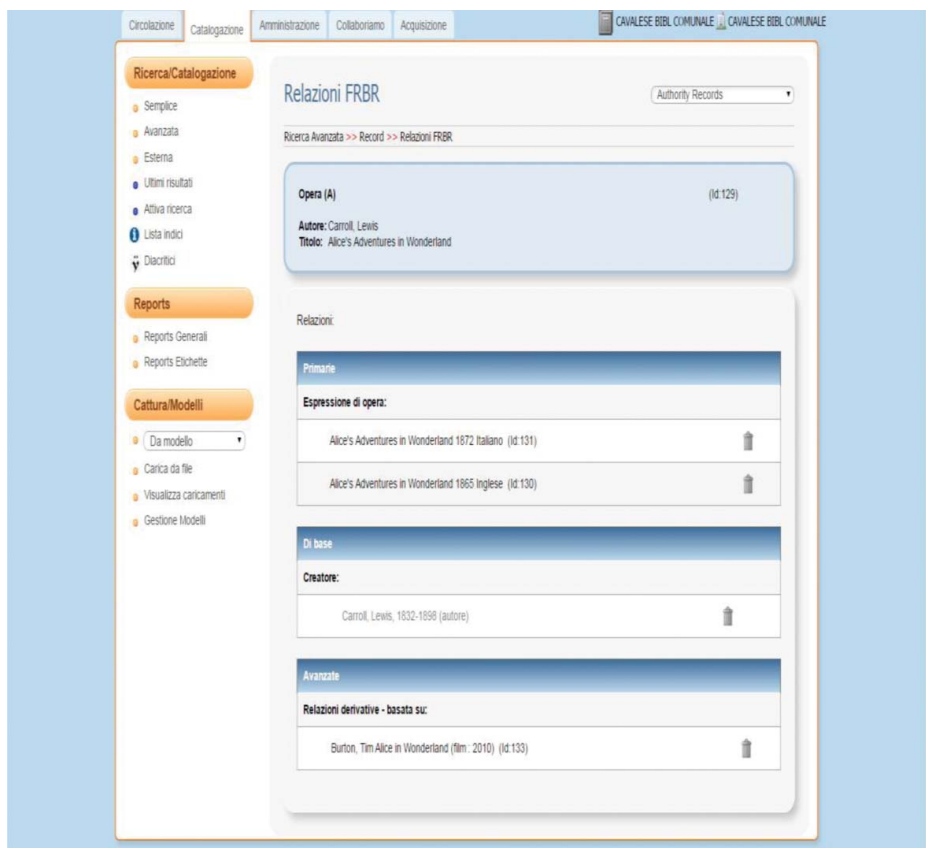


Figure 13. OliSuite/WeCat: The work *Alice's Adventures in Wonderland* with its relationships.

All lists of terms used and proposed by OliSuite/WeCat come from controlled vocabularies, most of which are published in RDF on the Open Metadata Registry. The metadata attribution software shares the purpose of using a language that is formulated in data sets and in ontologies that can be understood by machines. Each precisely identified entity, therefore, becomes a node of a more or less wide-spread network, which is depicted graphically to allow the cataloger to maintain control of the network developed from a point of observation. See the example in Figure 13 of relationships between a work (*Alice's Adventures in Wonderland*), its expressions (the 1872 Italian version and the 1865 English version), its creator (Lewis Carroll), and derivative works (the Tim Burton film in 2010).

In order to uniquely identify the resources, which are common to RDA and linked open data, OliSuite/WeCat offers the search function to capture identifiers used for the very same entity in other datasets and external sites (see Figure 14), to prepare the interlinking mechanism, which allows data to acquire the fifth star foreseen by Tim Berners-Lee's Semantic Web ratings system.¹⁴

A further new element of OliSuite/WeCat is the function that links registration to a framework of data conversion in RDF, using different ontologies, the result of an important European project. The project, entitled Automatic publication under LInked dAta paradigm of library Data (ALIADA), has been carried out by a group of Italian, Spanish, and Hungarian organizations: @Cult (Italian), ScanBit (Spanish), Tecnalía (Spanish), ARTIUM—Centro Museo Vasco de Arte Contemporáneo (Spanish), and Szépművészeti Múzeum (Hungarian). More information about the ALIADA project can be found at <http://www.aliada-project.eu/>. ALIADA intends to create a framework for the production and publication of data with the linked open data paradigm. Version 1.0 is available on GitHub, for the open source community. This OliSuite/WeCat connection between a piece of cataloging software and a data converter for linked open data has an important purpose: to simplify the update functions of a dataset

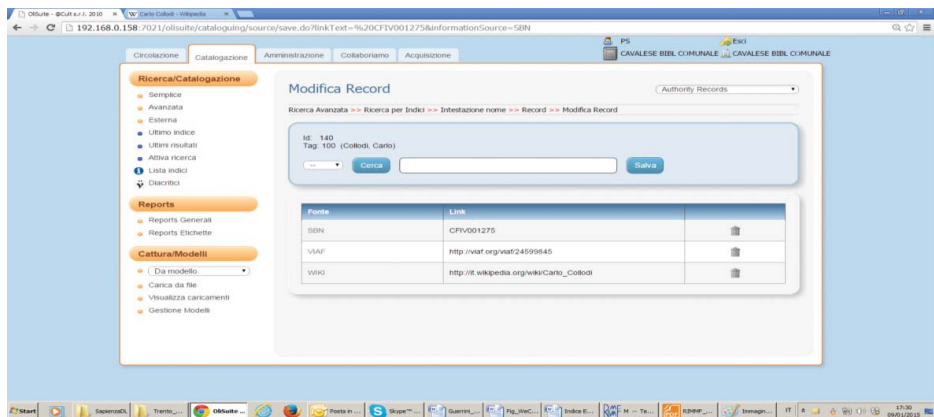


Figure 14. OliSuite/WeCat: Search and capture from the identifiers of the person entity *Collodi, Carlo* in three different data sources, SBN, VIAF, and Wikipedia.

in RDF that is already published on the web, allowing those who create the data to guarantee a real-time update of the published data.

The OliSuite/WeCat, born in the European context, has an international dimension, and is aimed to worldwide libraries; its development plan, devised by @Cult and shared with a large number of users, is oriented in two directions:

1. make the system gradually, but ever more significantly, independent of the MARC format, with an “object-oriented” metadating system that identifies and describes elements and not records, and is, therefore, closer to the world of the Semantic Web; and
2. port the system as open source, to extend the interest about the system to a wider community of experts, fans, and supporters.

Conclusions

The diffusion of RDA, the consolidation of BIBFRAME, the Semantic Web, and linked open data oblige information managers (agencies) to reflect deeply on the future of ILS software and the practice of assigning metadata. The awareness that we face a new generation of systems is now mature and increasingly widespread. Some doubt remains regarding how to achieve the transition from the old to the new dimension of information management: how can millions of existing pieces of data (structured, e.g., in MARC) be saved, while maintaining equally effective services built with this data? The time has arrived to invest in this development, but not before some of the questions that remain open have been resolved. These include:

- the difference between the FRBR model, absorbed for example by RIMMF and RDA, and the BIBFRAME model, which is watched with interest even so, particularly in relation to activities of data conversion to RDF;
- the future of MARC 21, for which, to date, there is no official declaration of its demise; indeed, the MARC Advisory Committee continues to work on the development and maintenance of the format. This means that thousands of MARC records are added every day to the already impressive mass of data that currently exists in this format. It will still be necessary to deal with these data in the near future.
- the still immature condition of BIBFRAME and, in particular, of the vocabulary produced by the model¹⁵: in data conversion processes, it is difficult to use a tool that is under construction. The first user tests are indubitably encouraging, but it will certainly be necessary to face continuous updates of the procedures for the conversion of data from MARC format to BIBFRAME, until these are stabilized.

To try to be active, nonetheless, and to take real steps forward in terms of this fundamental passage from the record to the data, we must start from the key concepts expressed by RDA: *identify* and *describe*, *link*, *represent*:

- *Identify and describe*: the current stage of data processing is still possible with conventional systems, which structure the information (records) in MARC. However, it is essential to implement everything already provided by RDA; in particular, the linguistic change produced by the FRBR model and the use of unique identifiers, such as URIs, from the construction of the data. The application of RDA, even in a traditional MARC context, therefore, allows for data to be arranged in such a way that, described with its attributes, data can then be more easily linked to other data and then represented in the many existing modalities.
- *Link*: the action of linking expresses the key concept of open linked data. The construction of links is accomplished mainly through the use of controlled vocabularies and ontologies, structured according to the RDF model: using terms taken from vocabularies constructed for the web to describe a resource (e.g., type of content, format, the intended instrument for a musical composition, medium) increases the chance of machines understanding this data, by enhancing the capacity of sharing the same information. Providing mechanisms in an ILS such as dictionaries, word lists, and ontologies available on the web, as well as the very structuring of the data, allows us to express the relationship between the resource type (e.g., book) and its attribute (e.g., title) in terms of the relationship between subject–predicate– object, creating a first, fundamental nucleus (the triple) of a larger network of bibliographic data.
- *Represent*: the publication of data for use, a function traditionally assigned to the OPAC and expressed in HTML pages, evolves into a variety of options in relation to areas of use and the users to which it is addressed. The general reference is still the FRBR model. Every other mode of representation is possible and desirable:
 - *Content*: all information traditionally expressed in an OPAC can be merged and enriched with information from thematic portals, online encyclopedias, specialized sites, and so on;
 - *Disciplinary contexts*: all disciplines are now available and no longer limited to traditional silos of libraries, archives, museums, publishers, and so on;
 - *Usage*: all uses are available, including use by scientific sites or sites with integrated commercial functions.

For information managers, the challenge is to harmonize tradition with innovation, the old with the new, in an enlightened manner, thus creating a new system of information management that is both advanced and economically sustainable. Representing information becomes in this way a varied and changing mode that follows the principle of user convenience.

The application of RDA is the first step in achieving this change. RIMMF and BIBFRAME Editor indicate to software developers a way to think that is consistent with RDA and open linked data systems environment. OliSuite/WeCat provides an implementation of RDA that integrates vocabularies and

ontologies already present in the web by structuring the information in the linked open data. The evolution is still in progress, but much has been completed, facilitated by the decision to follow the evolution of RDA from the beginning and to use all its instructions.

Notes

1. Library of Congress. Working Group on the Future of Bibliographic Control, "On the Record," January 9, 2008, <http://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf>
2. IFLA, "IFLA Cataloguing Principles: Statement of International Cataloguing Principles (ICP) and its glossary," eds. Barbara Tillett and Ana Lupe Cristán (München: K. G. Saur, 2009), <http://www.ifla.org/publications/ifla-series-on-bibliographic-control-37>; for commentary, see Mauro Guerrini, Giuliano Genetasio, "I Principi internazionali di catalogazione (ICP)," *Universo bibliografico e teoria catalografica all'inizio del 21. secolo* (Milano: Editrice Bibliografica, 2012).
3. Roy Tennant, "MARC Must Die," *Library Journal*, October 15, 2002, <http://lj.libraryjournal.com/2002/10/archives/marc-must-die/>
4. MARC must die!: Community portal, http://marc-must-die.info/index.php/Main_Page. The MARC Advisory Committee continues its operations developing and maintaining the MARC 21 format.
5. Part of the overview was used by Barbara B. Tillett in the seminar on RDA organized by Florence University's biennial Masters course in Cataloging, with Casalini Libri, ICCU (Linked Heritage) and AIB (Italian Library Association. Section Toscana), March 18–20, 2013.
6. Library of Congress, "Bibliographic Framework Transition Initiatives," October 2011, <http://www.loc.gov/marc/transition/news/framework-103111.html>
7. W3C Incubator Group, "Library Linked Data Incubator Group Final Report," October 25, 2011, <http://www.w3.org/2005/Incubator/lld/XGR-lld-20111025>
8. Library of Congress, "Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services," November 21, 2012, <https://www.loc.gov/bibframe/pdf/marclid-report-11-21-2012.pdf>
9. Library of Congress, "Bibliographic Framework as a Web of Data."
10. Cf. Barbara B. Tillett, "RDA and the Semantic Web, Linked Data Environment," *Jlis.it* 4, no. 1 (2013), <http://leo.cineca.it/index.php/jlis/article/view/6303/7876>
11. The MARC of Quality. Providing quality training, software, and database services to help you create better MARC and RDA/MARC records (August 1, 2015), <http://www.marcofquality.com/>
12. Cf. Mauro Guerrini, with Giuliano Genetasio, "I Principi internazionali di catalogazione (ICP)," 171–216.
13. @Cult is a software house located in Rome, Italy, focused on information management systems. More information about the company is available at <http://www.atcult.it/> (accessed January 26, 2016).
14. Cf. Tim Berners Lee, "Linked Data," last change, June 18, 2009, <http://www.w3.org/Design/Issues/LinkedData.html>
15. Library of Congress, "Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services," November 21, 2012, <http://bibframe.org/vocab/>