

Star-crossed lovers or heavenly match? ISBD and Linked Data - A Love Story

Mélanie Roche

Metadata Department, Bibliothèque nationale de France, Paris, France.

E-mail address: melanie.roche@bnf.fr



Copyright © 2017 by **Mélanie Roche**. This work is made available under the terms of the Creative Commons Attribution 4.0 International License:
<http://creativecommons.org/licenses/by/4.0>

Abstract:

This paper discusses how the ISBD Linked Data Study Group brought together ISBD, IFLA's worldwide-used bibliographic standard, with the technology of Linked Data at the basis of the Semantic Web. The road to this surprising union has not always been a smooth one, but the author is confident that it will be fruitful. The paper focuses on the methodology and achievements of the group, without leaving aside the challenges encountered — and more importantly, those to come. The stakes are high: at a time when both the bibliographic and the outside world are moving fast, it can become increasingly difficult to keep pace with the constant changes. Why bother with ancient standards when new conceptual models are being published every other year — or so it seems? How is interoperability still relevant today when new cataloguing codes are taking over the world? To put it bluntly: what is the point of bibliographic description in the digital world?

Keywords: ISBD, LoD, interoperability, standards, cataloguing

Introduction

It all began — as it often will — with a clashing of values. Stage left, a venerable 40-year-old, red-bound library standard, the beacon of a long-lived dream called Universal Bibliographic Control. Stage right, a brand-new technology that would lead the world from Web 2.0 to Web 3.0. Two communities, both equal in dignity: librarians and application developers, each one dotting over their most precious offspring. The stage is set, let the play begin.

ISBD and LoD — the International Standard for Bibliographic Description and Linked Open Data: a modern-day library version of Montagues and Capulets? Hopefully the story told in that paper will not be as tragic, but it will certainly prove as eventful.

1. The use for a go-between

One couldn't say that it was love at first sight between ISBD and Linked Data. In fact, when you think about it, those two seem to have very little in common. ISBD is one of IFLA's most

widely used standards. The reference where cataloguing is concerned, it touches upon librarians' core business – the description of all kinds of library material. It first came into being 40 years ago, as a means to implement Universal Bibliographic Control, an IFLA program that aimed at setting universal rules for the sharing of bibliographic records. It relies upon standardised punctuation, anything but obvious abbreviations, and elements dispatched in Areas numbered 1 to 8 that any cataloguer (they weren't called metadata librarians back then) could recite in their sleep. Like any standard, it has evolved through time, adapting through the manifold changes of the publishing industry. Once upon a time, each type of material had its own ISBD. Nowadays, since the publication of the consolidated edition in 2011, one single red-bound book binds them all. The 2011 edition also introduced an “Area 0” to dissociate the Content Form from the Media Type, confirming a paradigm shift that had initiated a few years ago in the bibliographic world.

Linked Data, on the other hand, is a relatively new technology. The term was introduced in 2006 by Tim Berners-Lee — the father of the Web himself — to describe “a method of publishing structured data so that it can be interlinked and become more useful through semantic queries”¹. The two important words here are *interlinked* and *structured*: indeed, linked data is interlinked because it relies on the use of URIs, that provide permanent addresses for objects and therefore lasting links between them. And it is structured because it uses a data model called Resource Description Framework (RDF)², that can express any statement about any resource in the simple syntactical form of *subject — object — predicate*. These statements are called triples, where the subject and object can either be URIs or nodes, and the predicate can either be a URI or a literal. Linked *Open Data* follows essentially the same requirements, but adds the use of open licenses and non-proprietary formats. Linked Open Data is often seen as the final stage of Linked Data.

To put it in a nutshell, ISBD used to be the cornerstone of UBC, while Linked Data is at the core of the Semantic Web. The differences between the two epitomise therefore two radically different approaches for quite a similar objective, that is the use and sharing of information (any kind of information for Linked Data, *versus* bibliographic information for ISBD). While ISBD firmly upholds the “top-bottom approach”, according to which standardisation can only come from the top and trickle down to the local, Linked Data is a relentless advocate for the “bottom-up approach” — meaning that global understanding arises from the interoperability of local data³.

ISBD revolves around the notion of “record”, that is to say a complete entity, a whole that has a beginning and an end. There is no such thing as a record in the Semantic Web: only linked data — emphasis on *data*. And there is no beginning or end to the linking of data. The shift from record to data, although a necessary one, is still an ongoing process in the community of librarians, if only for the practical reason that a record is easily perceived as being stored *somewhere* (either a drawer or a computer), whereas visualising exactly where *data* is stored can prove a little bit more tricky — and somewhat dizzying.

¹ https://en.wikipedia.org/wiki/Linked_data. Consulted on May 25, 2017.

² <https://www.w3.org/RDF/>. Consulted on May 25, 2017.

³ For a fully detailed presentation of the “bottom-up” approach, see Gordon Dunsire and Mirna Willer, *The local in the global: universal bibliographic control from the bottom up*, World Library and Information Congress: 80th IFLA General Conference And Assembly, 16-22 August 2014, Lyon, France. Online: <http://library.ifla.org/817/1/086-dunsire-en.pdf>. Consulted May 25, 2017.

However, ISBD and LoD share more than they knew. First, they both emphasise the importance of *structured* data: the ISBD uses meaningful punctuation to structure data, while Linked Data uses the grammatical construction of the triple. Meaning arises from structure. Second, both are fierce advocates of “the common good” — and do not shy from the pomposity of the word either, since the definition of Linked Data uses the word *useful*. To be sure, ISBD can be purchased like any publishing product, but it is also freely available online in eleven languages, because the idea behind it is that the standard serves a whole community of users. Likewise, the very name of Linked Open Data suggests that the sense of community prevails over that of business — although here also it can prove quite a lucrative one. Sharing — be it bibliographic records or Knowledge pure and simple — is ultimately the superior motive for the both of them. And it is worth mentioning that in both cases, the amount of work it requires often lies upon enthusiastic volunteers that dedicate much of their free time for this idea they have of the common good.

In spite of their many differences and everyone telling them they should fight each other to the death, ISBD and LoD needed therefore some kind of Friar Laurent to determine whether it was ancestral hate or indeed burgeoning love that they were feeling. This Friar Laurent came in the guise of a bunch of information professionals who refused to believe in this feud and instead chose to envision what a brilliant outcome the encounter of ISBD and Linked Data could bring forth.

2. The ISBD Linked Data Study Group

An emanation of the ISBD Review Group within the IFLA Cataloguing Section, the ISBD Linked Data Study Group was officially created in 2013. Previously known as the ISBD/XML Study Group, its aim since 2008 has been “to promote interoperability and foster the reuse/retrieval of bibliographic data in the Semantic Web”⁴ — in other words, to bring together ISBD and Linked Open Data. So far it has been chaired by Elena Escolano, Françoise Leresche, María Violeta Bertolini, and as of August 2016 the author of this paper.

The Study Group usually meets once a year in August, during the IFLA international conference, for a business meeting where we update our annual action list and plan the work ahead for the year to come. Throughout the year, most of the work is done remotely *via* emails, shared documents or videoconferencing. As this is not always sufficient, we also plan for “mid-term” meetings in the Spring to settle the last pending issues. All the minutes and agenda for the group's annual and mid-term meetings can be found online on the Study Group's webpage. This page also provides useful links to our publications and related papers.

Although English is our working language, the group comprises of native English, Italian, German, Spanish, and French speakers, all of whom are quite sensitive to translation issues. That is why we extended the general *Guidelines for translations of the IFLA namespaces in RDF*⁵ to the more ISBD-specific *Guidelines for translations of the IFLA ISBD namespaces in RDF*⁶, first published in 2015 and updated in 2017. This illustrates how the group fits into the

⁴ <https://www.ifla.org/node/1795>. Consulted on May 25, 2017.

⁵ <https://www.ifla.org/node/5353>. Consulted on May 25, 2017.

⁶ <https://www.ifla.org/publications/node/10835>. At the time this article was written, the online version is the 2015 version.

global IFLA structure: most of the members are either liaisons or members of LIDATEC, the technical committee in charge of IFLA namespaces⁷. Inclusion within the larger structure of IFLA is key, and it has always been clear to us that the ISBD LD SG is only one link in the chain of IFLA standards; that is why we have established lasting liaisons with the FRBR Review Group, also within the Cataloguing Section.

The Study Group's position, however, is not always an easy or indeed a comfortable one: finding ISBD a place in the Semantic Web eventually leads to questioning its very position in the bibliographic world, once a very stable and closed world which is now as fast-changing and expanding as the "real world". For instance, RDA (Resource Description and Access) was introduced in 2010: this content code for the description of library resources has been embedded in the Semantic Web from the very beginning. Bringing ISBD and Linked Data together also means therefore bringing ISBD together with its sister standards within the bibliographic universe, without the one taking advantage on the other. It is a dangerous road that leads to this union, on which one has to tread carefully if they wish to travel safely.

One of the means to address this issue has been to publish the ISBD element set in the Open Metadata Registry (OMR)⁸, "a fundamental piece of technical infrastructure for the Semantic Web"⁹ maintained by Metadata Management Associates that allows the publication and mapping of distinct metadata schemas, schemes and Application Profiles so as to "support the key goals of metadata discovery, reuse, standardization and interoperability locally and globally"¹⁰. Several bibliographic standards and models are declared in the OMR, such as the RDA element set, the Unimarc element set and the FRBR element set. That is a way of making sure that no standard is left aside, because it is true that union is strength. It so happens that so far the spectrum of our alliances has "naturally" fallen within the realm of bibliographic description (a difficult habit to break for cataloguers), but it does not have to be: we would be more than happy to hear from other Study Groups, Task Groups or whatever their unit is called, who also strive to expose their standard into the Semantic Web.

3. Achievements and challenges ahead

The Study Group's first work consisted in publishing various alignments and maps, the use of which is to bridge the gap between cataloguers and application developers. The difference between an alignment and a map is that "an alignment is a match between one or more elements in one schema to one or more elements in another schema. The match does not have to be exact; it can be equivalent, broader, or narrower. A mapping is an expression in RDF of a single alignment between two elements. A map is a set of mappings between the two schemas."¹¹. In other words, the alignment is most useful to a human eye, while the map targets Web applications.

⁷ <https://www.ifla.org/node/9428>. Consulted on May 25, 2017.

⁸ <http://metadataregistry.org/schema/show/id/25.html>. Consulted May 25, 2017.

⁹ <https://www.w3.org/2001/sw/wiki/OpenMetadataRegistry>. Consulted May 25, 2017.

¹⁰ *Ibid.*

¹¹ *Guidelines for Use of ISBD as Linked Data*. Online: <https://www.ifla.org/node/10834/>. Consulted May 25, 2017.

The first alignment matched the RDA and the ISBD element sets. It was prepared by Gordon Dunsire and the group in 2011, and updated in 2014 following a later release of RDA. This document took the form of a table, in which all the ISBD elements were matched as narrower than (<), equal (=) or broader (>) than their corresponding RDA elements. The definitions of both element sets were taken from the OMR. Other alignments followed, such as the alignment of the Content Form and Media Time (the aforementioned Area 0 of ISBD) with the RDA/ONIX Framework. Finally, the same method was used to align the ISBD element set with the FRBR element set, achieving an even more complete harmonisation of IFLA bibliographic standards and models in the Semantic Web. A first version was published in 2016, and should be updated this year in Wroclaw.

What we learn from these alignments is that, quite predictably, there hardly ever is a perfect equivalence between two element sets. Harmonisation, however, can be achieved inasmuch as it focuses not on “identity of content” but on functional interoperability, that is to say “that by which “records valid under one of the standards should be capable of being mapped to either of the other standards. It is recognised that some issues will take longer to resolve than others and a few issues may prove to be irreconcilable, but steps can be taken to limit the impact of such differences.””.¹²

In 2016, the group achieved one of its primary objectives by publishing the *Guidelines for Use of ISBD as Linked Data*¹³ – a document providing guidance and examples for developers who wish to use ISBD namespaces in the Linked Data environment. One of the requirements for this document was to give concrete examples, which was done by extracting “real-life records” from Spanish, Italian and French online catalogs and translating them into RDF triples (see figure 1 below). Five examples were taken to take into account the manifold choices and possibilities with which one is presented when expressing ISBD in Linked Data.

¹² https://www.ifla.org/files/assets/cataloguing/isbd/OtherDocumentation/isbd2rda_alignment_v3_1.pdf.

¹³ <https://www.ifla.org/node/10834/>. Consulted May 25, 2017.

Example 1a: using fine granularity ISBD properties

```
@prefix ex:      <http://example.com/> .
@prefix isbd:   <http://iflastandards.info/ns/isbd/elements/> .
@prefix isbdcf: <http://iflastandards.info/ns/isbd/terms/contentform/> .
@prefix isbdcqss: <http://iflastandards.info/ns/isbd/terms/contentqualifica
tion/ sensoryspecfication/> .
@prefix isbdmt: <http://iflastandards.info/ns/isbd/terms/mediatype/> .
@prefix skos:   <http://www.w3.org/2004/02/skos/core#> .

ex:1
  isbd:P1001    isbdcf:T1009 ;
  isbd:P1002    isbdcqss:T1005 ;
  isbd:P1003    isbdmt:T1010 ;
  isbd:P1004    "El alcalde de Zalamea" ;
  isbd:P1007    "Calderón de la Barca" ;
  isbd:P1008    "14ª ed., 1ª en esta presentación" ;
  isbd:P1010    "edición de José María Ruano de la Haza" ;
  isbd:P1016    "Madrid" ;
  isbd:P1017    "Espasa" ;
  isbd:P1018    "2012" ;
  isbd:P1022    "179 p."@es ;
  isbd:P1024    "19 cm" ;
  isbd:P1026    "Austral" ;
  isbd:P1031    "50" ;
  isbd:P1032    "B 425-2012" ;
  isbd:P1042    "Bibliografía: p. [55]-63"@es ;
  isbd:P1154    "ISBN 978-84-670-3934-4" ;
  isbd:P1164    "Colección austral (1987)" .

isbdcf:T1009
  skos:prefLabel "Text"@en ;
  skos:prefLabel "Texto"@es .

isbdcqss:T1005
  skos:prefLabel "visual"@en ;
  skos:prefLabel "visual"@es .

isbdmt:T1010
  skos:prefLabel "unmediated"@en ;
  skos:prefLabel "sin mediación"@es .
```

Figure 1 : Example 1a in the *Guidelines for Use of ISBD as Linked Data*.

In the following years, the ISBD LD SG will have to tackle many challenges, the least of which is not the oncoming revision of the ISBD. The world (bibliographic and otherwise) has changed a lot since the 1970 and even the 2011 version of the standard: new models have emerged, that were in their turn reviewed to take a much higher-level approach of the bibliographic universe. The FR family of conceptual models (FRBR, FRAD and FRSAD¹⁴) was consolidated into the newborn Library Reference Model (LRM)¹⁵. The reviewed version of the International Cataloguing Principles was officially endorsed by the Committee on Standards earlier this year¹⁶. How will ISBD fit into this new-looking world? How can it build upon past alliances to enhance consistency and sense in the world of IFLA standards? We can reasonably argue that ISBD provides the fine granularity of information that could turn into an extension of the LRM. But before this work is completed, many steps have to be taken.

¹⁴ FRBR: Functional Requirements for Bibliographic Records; FRAD : Functional Requirements for Authority Data; FRSAD: Functional Requirements for Subject Authority Data.

¹⁵ <https://www.ifla.org/publications/node/11412>. Consulted July 3, 2017.

¹⁶ <https://www.ifla.org/publications/node/11015>. Consulted June 12, 2017.

In 2017, a task group was commissioned within the Study Group to produce an alignment between the ISBD element set and the Library Reference Model element set — although the latter has not yet been declared in the OMR. The document will very likely be discussed and validated during this year's WLIC in Wrocław. One of the many issues it raised was that of the relationships: the LRM is much more relational than FRBR used to be. ISBD is very little concerned with relationships: it mainly deals with a resource being described “book in hand”, as the colloquial phrase goes.

Conclusion

The work of the ISBD Linked Data Study Group has shown that despite their differences, ISBD and Linked Data were actually meant to be. What started like a clashing of values ended up in a series of alignments between data. What sweeter end could there be to the ears of modern-day librarians?

Publishing an IFLA standard in the Semantic Web is more than a question of visibility: it is one of the reasons librarians and libraries are still relevant today. Indeed, the publication itself is actually the easy part; along with it there comes a never-ending work of documentation, maintenance and update. The ISBD Linked Data Study Group realised this a long time ago. One of the next steps for the group would be to know exactly how the alignments, guidelines etc. are used by their intended audience, in order to serve our users better (both professional and otherwise). Expanding beyond the world of bibliographic description would also contribute to this “common good” that is so dear to our community.

The key to the Study Group’s (so far) successful enterprise has been inclusion and dialogue. The fact of being clearly identified within a Review Group has been extremely useful, all the more so for the 2017 edition of the WLIC, when our business meeting was assigned a spot in the program for the first time ever — and so was the ISBD Review Group. Maintaining dialogue between standards and establishing liaisons with other IFLA units is instrumental: where linked data is concerned, union really *is* strength.

Acknowledgments

The author of this article would like to thank all the members of the ISBD Linked Data Study Group, past and present, for their hard work, their boundless knowledge, and their affectionate support. María Violeta Bertolini, Vincent Boulet, Gordon Dunsire, Elena Escolano Rodríguez, Massimo Gentili-Tedeschi, Françoise Leresche, Dorothy McGarry, Clément Oury, Ricardo Santos Muñoz, Mirna Willer: knowing you, working with you, learning from you has made me feel like I have found a professional family to which I will always be proud to say I belong.