IFLA Cataloguing Section

ISBD Review Group

ISBD/Linked Data Study Group
http://www.ifla.org/node/1795

Minutes of the ISBD Linked Data Study Group's Meeting
Paris, Bibliothèque nationale de France, December 9th – 10th 2014

Attendees: Mirna Willer, Gordon Dunsire, Françoise Leresche, Massimo Gentili Tedeschi, Michaël Jeulin (Tuesday), Mélanie Roche

Skype Attendance: Elena Escolano Rodriguez, Dorothy McGarry, Ricardo Muñoz Santos

Tuesday, December 09th 2014

Agenda : Guidelines for use of ISBD as Linked Data

Starting from a previous draft, the group discussed and wrote out the following sections of the document:

- Alignments and mappings;
- Using both constrained and unconstrained properties;
- Data presentation and layout;
- Glossary: the group aimed to define the word "mapping" but since no consensus could be achieved, the definition was left blank for the time being;
- Examples : by correcting an incorrect use of constrained properties, the exemples are more pedagogical. They illustrate the fact that complex cases with groupings cannot simply copy simple ones but have to use unconstrained properties.

The day ended with a Skype meeting during which the progress of the day's work was summarized to some of the members who could not physically attend the meeting. The revised version of the draft was sent to the rest of the group.

Below are the main points that have emerged from the group's discussions.

Distinction between past and future / transcription and recording / rules and input

- we need to distinguish between what is needed for the publication of **legacy data** (unbundling of al ISBD implicit sub-groupings) and input for the **future**. The difficulty lies in that LD is all about access while ISBD is all about identification: decomposing existing ISBD data into RDF data proves therefore to be a challenge.

- the above distinction really amount to distinguishing between **transcription** (rules) and **recording** (input). Currently, as Gordon pointed out, ISBD tends to confuse the two processes. Although Françoise agrees it is worth considering that ISBD should distinguish better between the two in its future developments, it will remain fundamentally limited to bibliographical description, ie identification through the conventional representation of pieces of information, and not with access. Sense will therefore still emerge from a prescribed order of information.

Distinction between gathering and sequencing

- there emerged the need for an **intermediary element** between "Label" and "Area" (for instance, "Group"), so as to manage sequencing, e.g. in the case of repeated aggragated elements.

- establish clear-cut distinctions between gathering and ordering: unconstrained properties allow for the **gathering**, while encoding schemes determine the **order** of information. The problem is, since the technical community disagrees on the use of encoding schemes, they do not exist yet and it will take a lot of human work to produce them (analysis of punctuation + creation of XSLT sheets).

Creation of an extension section

- creation of an **extension section** that would suggest extension mechanisms for the communities to use, provided that they tell us what they need (which is the stumbling stone, given the fast pace of the changes they are faced with).

- the section will also tackle the explicitation of all the implicit groupings of the ISBD stipulations, which have not yet been dealt with (so far the ISBD RG has only been concerned with the explicit groupings).

- this will have to wait for the revision of the ISBD. Meanwhile,we are leaving it up to individual communities to use our element sets, according to the following rule of thumb: a community dealing with anything other than a very simple ISBD record should use unconstrained properties.

Future trends

- need for **more collaboration** with other groups such as the cataloguing section, the professional committee or the study group on standards, but more urgently with the FRBR RG. There is currently no attempt at synchronizing FRBR and ISBD, which is not how a standard ought to be developed : FRBR being currently under revision, does ISBD need to wait for the consolidated model before it sets out its own revision? What we know so far is that FRBR Group 1 is most likely to remain unchanged, and since ISBD is primarily concerned with the description of the manifestation, it may be best to start on without further delay. One solution might be that once the revision of ISBD is undertaken, the chairperson of the FRBR consolidated group should attend meetings.

- **RDF is becoming extremely expensive**. In the future, maybe only international agencies such as IFLA will be able to afford "pure" RDF, while others (the BnF and the likes) may have to contend with Dublin core, being pressured as they are by commercial / semi-commercial instances (Google, schema.org, OCLC and so on) who will aim for something cheaper and more practical even though inferior in quality.

Actions

- Gordon Dunsire will investigate whether the DC architecture group initiative have worked on sequencing issues in the context of their database of use cases for application profiles at fine-grained levels (a project in which Europeana and EDM heavily involved).

- the production of an Application Profile is of seminal importance today: it should be easily understandable by humans, so that we actually know what is going on, and should include the elements discussed during this meeting. Gordon will produce the Application Profile in the form of a table (element – repeatable – mandatory). The xml file will come later.

- Françoise Leresche will provide specific examples in the dedicated section.

- **next step for the group**: read out the draft while the conversation is still in our minds → gather comments as quickly as we can => review of ISBD to RDA mapping (Gordon).

Wednesday, December 10th 2014

Agenda: ISBD → FRBR alignment

The group went through the existing alignment table and discussed every element that had raised comments. The comments were examined, discussed and the final decision was documented in Column T.

The group stopped right before the "Notes" area. Decisions were made regarding some of the notes, but given the general state of fatigue of the attendees, these cannot be seen as final and may be subject to further discussion.

Like the previous day, the day ended with a Skype meeting for some of the members who could not physically attend. In addition to an account of the progress of the work, this closing meeting provided the group with the opportunity to set up future work.

General rules to keep in mind when aligning

- **be consistent**: the first rule by which we ought to abide when filling out the table is that of consistency. In similar cases, decision-making should always rely on the same premices.

- **read FRBR as it is** without anticipating its evolution: regardless of how old the definition of an element might be, and whether or not it has fallen out of usage, we will stick to the text without anticipating the revision of the definition.

- **favour direct mapping** so that the data speaks for itself. In theory, once an ISBD element has been identified as a subproperty of another ISBD element that has been properly mapped to he

corresponding FRBR element, this should be only a matter of semantics: any machine will infer that this subproperty matches the proper FRBR subproperty. But the semantic Web community prefers direct to indirect mappings, because it avoids the multiplication of triples into one gigantic soup.

- we should also keep in mind that aggregated statements are not necessarily superproperties of their components, and just because the elements are indented doesn't mean they are all subproperties: some are subproperties, some are components of an aggregated statement. The table should include a column with the URI and its relationship, so that we can keep track of what is a property, a subproperty, a compound etc.

- **look at the elements**, not the rules : what should prevail in the alignment is the formal definition of the element, then the stipulations if the definition is not clear enough, and lastly the examples (and only for the final decision-making). If the examples are inconsistent with the definition, then the definition still prevails. **Definition > stipulations > examples.**

- the definitions for "Specific Material Designation" and "Extent" given in the table are inconsistent with those in the ISBD glossary. The only way to map those elements is by assuming that the text of the ISBD is correct, and therefore revise the definitions in the OMR. If we decide to leave the definitions as they are, then the mapping cannot be made. The decision is up to the group.

Principles induced by the model

- FRBR being a model, **any alignment other than "<" should arouse suspicion** and be thoroughly examined. This is especially true for controlled vocabulary: since FRBR does not accommodate controlled vocabulary, ISBD elements using controlled vocabulary in ISBD will always align as "<".

- when more than one ISBD properties or subproperties can be matched to one single FRBR element, then they cannot be aligned as "=".

Creation of a 4th alignment

- it was decided to create a 4th alignment in addition to the existing "<", ">" and "=": **the symbol "C"**, standing for "Component", signals an aggregated statement (explicit in the ISBD but implicit in FRBR). It is a bent version of the "<" alignment, and could be paraphrased as "kind of narrower but not quite".

- "C" will alert whoever reads the mapping (FRBR or application developpers) on the compound nature of the ISBD element and the need to examine it very carefully: the alignment cannot be expressed in FRBR yet but may be next year, or it may be up to an individual application to interpret it using a particular syntax encoding scheme. Incidentally, it also allows us to be more accurate by stating what is actually going on.

- when a compound element in the ISBD has been identified as ">" than a FRBR element, there is no logical way to determine whether the elements which comprise the compound are indeed "<", ">" or = to the same element. The mapping therefore cannot be made (the option of a 5th symbol, "Ɔ", to signal a component greater than the FRBR element, has been discarded).

Distinction between content and carrier

- one the major issues discussed was, how do we account for variations in carrier that do not affect content (use of coloured paper, one particular state of an engraving, etc.)? to date, cataloguers are

told to make no difference between content and carrier, because it would be too much work for too small a benefit to the end users , but the argument is debatable as more and more users are interested in the presence of illustrations or colour for instance.

- at any rate, the decision should not be left to cataloguers, but should reflect the library collection's scope and purpose. It should also take into account whether the variation is intentional or not, although even chance or mistakes might become significant over time, and therefore might be worth recording as well.

- the present compromise is to consider that all works are aggregates (a text by an author + a preface by another author + illustrations, etc.). A more long-term solution might lie in **shifting focus from attributes to entities** : each entity being a different Expression, the record needs not say anything about them, but only state their presence and provide links.

Trends in other groups

- to date, ISBD has been more concerned with making the rules than thinking about the elements → even today in "RDA in France", there is still a long way to go before we really shift from "making a rule" to "thinking about the elements" (possibly a couple of years).

- the JSC is progressively shifting away from the idea that some sources of information are better than others: only transcribed elements need sources of information, but otherwise, when recording something, the source of information is the semantic web (a URI).

- the FRBR consolidation group has initiated the shift from attributes to entities but here again there is still a long way to go.

Actions

- although most of the definitions for FRBR elements match those in the OMR, some of them don't; Melanie Roche will replace them using the appropriate OMR wording.

- check the consistency of the decisions that were made in the alignment : Melanie Roche, under the supervision of Françoise Leresche

- **as a group**, decide whether we wish to change the definitions of "Specific Material Designation" and "Extent" in the OMR so as to put them in line with those in the ISBD glossary

- **next step for the group**: send out the alignment to the ISBD and FRBR Review Groups so that they can send their feedback. Before we do that, we will have to come to final decisions regarding the notes area → schedule :

- by **January 20th 2015**, each of us will have reviewed the commented "Notes" lines
- we try and come to a collective decision about the "Notes" elements, preferably in a future meeting, or possibly via email or Skype;
- we send the whole document to the ISBD and FRBR RG so their comments add up to ours.

Respectfully submitted,
Mélanie Roche (Bibliothèque nationale de France)
May 11th, 2014