

Moving forward with RDA

—Based on the presentation given at the National Library of China, October 26, 2015

Chris Oliver

University of Ottawa Library

When RDA, *Resource Description and Access*, was published in June 2010, it was published as a “first release”, the starting point for a standard that would continue to be developed according to defined goals and objectives defined both in the 2005-2009 Strategic Plan and the introductory chapter of the standard. The first version of RDA had attained minimum goals to signal a change in direction, a new approach to resource description and access, with sufficient content for release and initial implementation. The concept of “first release” was not explicitly stated either at the RDA Toolkit website or at the website of the Joint Steering Committee for Development of RDA (JSC), but it was generally understood in the community that had worked on RDA development. The phrase appears in many JSC approved documents, such as the [FAQ](#), presentations by JSC members, and JSC documents such as [Issues deferred until after the first release of RDA \(5/JSC/Sec/6/Rev\)](#). There were also several “placeholder” chapters, chapters that were empty with just one phrase: To be developed after the initial release of RDA. RDA was not incomplete when it was published in 2010. It was a complete standard that superseded its predecessor AACR2, and went beyond AACR2. But it was also a snapshot of the standard at that point of development. Its release did not signal the end of work, but the start of new phase.

In the 2010 version of RDA, one can see the groundwork for the fundamental change in direction that RDA represents, in contrast to AACR2. RDA was designed as a standard for the digital environment, as a standard with broader scope, and one based on an explicit theoretical framework through its alignment with the FRBR family of conceptual models.¹ These aspects are delineated and present in the 2010 version of RDA. But these aspects are also further developed and brought to a more advanced level as RDA develops and changes each year. The 2015 version of RDA is more in line with the fundamental goals and objectives than the 2010 version. And RDA will continue to develop, as technological requirements advance and as the standard becomes more completely integrated in the international cataloguing community.

Right from the start of RDA development, the standard was seen as a content standard, a standard for the data that is recorded, not a standard related to encoding or display. RDA defines the elements required for description and access and gives instructions on formulating

¹ The FRBR family of conceptual models refers to the three conceptual models developed under the auspices of IFLA (International Federation of Library Associations and Institutions): [FRBR](#) (Functional Requirements for Bibliographic Records); [FRAD](#) (Functional Requirements for Authority Data); [FRSAD](#) (Functional Requirements for Subject Authority Data).

the data that is recorded in each element. Instead of long text strings, data is recorded in clearly defined, distinct and precisely labelled elements. RDA is not tied to any single encoding schema or style for displaying data. Whether the 2010 or the 2015 version, the essence of RDA is a set of data elements to record significant bibliographic data about entities and to record the relationships of entities to each other.

RDA was designed with an eye to the future. The granularity and precision of data was intended to allow for large-scale machine-actionable processing of data, as well as enabling the use of this data in new technological environments, especially the semantic web of linked data. RDA places a heavy emphasis on the importance of recording relationships precisely, and in a post MARC environment, relationships become the key to meaningful data. RDA was also designed with an inherent flexibility, anticipating environments when there would be different ways of putting data together. However, we have had to implement RDA in a MARC environment and this future-oriented aspect of RDA has often been obscured. Especially during phase 1 of implementation, when cataloguers were anxious about the transition, a very rigid interpretation of RDA was presented, with everyone following one set of options and one way of encoding the data.

An example of RDA's flexibility can be seen in chapters 6, 9, 10 and 11. These chapters all have a recurring instruction: "record ...identifying attributes as separate elements, as parts of access points, or as both." These chapters discuss the data elements for the entities work, expression, person, family, and corporate body. At the end of each chapter, there is a section on constructing the authorized access points. RDA focuses on collecting data about each entity. Then, one may use some of these data elements to construct the authorized access point. But the main thrust is collecting data about the entity.

In a MARC environment, authorized access points are pre-constructed and must follow one particular order, especially in an environment of shared authority files. However, with the fluidity of web environments and new technologies, this rigidity may just be a passing phase. Tom Delsey, the architect of RDA, had a different vision. In an e-mail to me in 2010, he reiterated this vision:

The definition of separate elements means that data elements that have traditionally been incorporated into access points with a pre-set structure do not necessarily have to be stored and presented in that structure when they are recorded as RDA elements. There will be scope in the future for customizing the display of elements that were previously "locked-in" to a pre-set "one-size-fits-all" access point structure, to better serve the needs of a particular user community, or to construct access points "on the fly", using automated processes to produce a results set display that responds more directly to the user's search query.

For example, if the author of my work is Geoffrey Brown, and this is a common name, there are likely to be many Geoffrey Browns. The current requirement is to differentiate firstly by dates:

Brown, Geoffrey, 1926-
Brown, Geoffrey, 1935-
Brown, Geoffrey, 1941-
Brown, Geoffrey, 1954-
Brown, Geoffrey, 1960-
Brown, Geoffrey, 1963-
Brown, Geoffrey, 1964-

Does a user necessarily know the date of birth of the author they want to read? If a student wants a book by their professor, would they know which Geoffrey Brown to choose? Perhaps in the future, a user can ask for the persons to be differentiated by another piece of data that is more meaningful for them, such as profession or occupation:

Brown, Geoffrey (Sailor)
Brown, Geoffrey (Psychologist)
Brown, Geoffrey (Banker)
Brown, Geoffrey (Mechanical engineer)
Brown, Geoffrey (Computer scientist)
Brown, Geoffrey (Biologist)
Brown, Geoffrey (Author)

The biology student would quickly be able to identify and select the Geoffrey Brown they want. The sections on constructing access points are very much written for the current environment, but the rest of those chapters already envision a more flexible future environment.

The cataloguing world is already moving towards the idea of flexibility and away from “one size fits all”. An example would be [VIAF](#), the Virtual International Authority File. One identifier leads to a variety of different ways to record the name of the person, etc. By bringing together the authority files from different countries around the world, VIAF presents a new model where there is not one preferred form, but a number of preferred forms of the name, depending on the geographical context, language, etc.

Confucius, 0551?-0479? av. J.-C.
Confucius
Kongzi, 551-479 v.Chr.
Confucij, ca. 551-479 a.C.
Confucius, 551-479 a.C.
Kong, Qiu v551-v479
孔子, 552-479 B.C.
Kongzi
Konfucius, 551-479 f.Kr

(from VIAF)

During RDA development, there was a deliberate move to define data elements rather than use strings and paragraphs as in AACR2. We often look at RDA through the lens of the linear text as presented on the RDA tab in the Toolkit, organized according to the essential steps required to describe a resource and give access to it. On the Tools tab of the Toolkit, there is an alternative view, RDA as a type of data dictionary. There, the RDA instructions are clearly structured as an element set, with the elements and sub-elements organized according to the FRBR model.

Each RDA element was designed to be precise and unambiguous, to contain only one kind of data, and to use controlled vocabulary wherever possible. Thus, in the future, any data element has the potential to be usable to index the data, to search and retrieve, as well as create meaningful displays. Also, with precisely parsed data, any element can be used by both a human and a computer. The more precise and granular the data, the more reliably the data can be manipulated by computers. In 2010, one saw the beginnings of this approach. However, as communities work with RDA, new opportunities are seen to make RDA data more machine-actionable. There is currently a discussion paper proposing a major step in breaking down the data for extent and dimensions into much finer, more precise data chunks that can be better manipulated by machines in the future.

The discussion paper is called [Machine-Actionable Data Elements for Measurements, Extent of the Carrier, Pagination and Foliation, Dimensions, Extent of the Content, and Duration – Discussion Paper \(2015\) \(6JSC/ALA/Discussion/5\)](#). At this point, the paper proposes the option of text strings alone or text strings and a more granular set of measurement sub-elements. But perhaps in the future there would just be the more granular set of measurement sub-elements. Greater granularity might also allow for easier manipulations if one wishes to present data using different numbering or language terms.

For example:

string	measurement sub-elements	
1 volume	measurement type	carrier extent units
	measurement unit	volume
	measurement quantity	1
approximately 600 slides	measurement type	carrier extent units
	measurement unit	slides
	measurement quantity	600
	measurement qualifier	approximately
25 x 35 cm	measurement type	height
	measurement unit	cm
	measurement quantity	25

measurement type	width
measurement unit	cm
measurement quantity	35

No decisions have been made yet, but it is interesting to see that this discussion has gone through several iterations and is an ongoing subject at JSC meetings.

The JSC has also created the RDA namespace and has published the element set and controlled vocabularies at the [RDA Registry](#). The RDA Registry provides the infrastructure to make the RDA element set and the RDA controlled vocabularies accessible on the web, suitable for a linked data environment. The data is available in the computer readable formats expected by those working in this environment, and the data is openly accessible so that it can be used not only by the library community but also by other cultural heritage communities. When RDA elements set are used, whether whole or in part, it helps build towards better interoperability of data.

RDA is a set of practical instructions. It is currently used in many environments as a replacement for AACR2 or other cataloguing standards while libraries continue to work in a traditional MARC environment. But RDA is also part of the groundwork for moving the data that libraries have created and continue to create into new technological environments and ensuring that that data can be used effectively in the future.

RDA was also designed as a standard with a broader scope, in two senses: in the sense of connecting with other cultural heritage communities beyond libraries, such as archives and museums; also in the sense of breaking down the Anglo-American limitations of its predecessor and making it easier to apply RDA in different language and cultural contexts. In the 2010 release, at RDA 0.11, there is the statement: RDA is designed for use in an international context. The statement remains and signals an important intention. RDA's predecessor, AACR2, was translated into many languages and used in many countries beyond the four author countries, Australia, Canada, Great Britain and the United States. However, AACR2 was written with an "Anglo-American" perspective, favouring one particular set of parameters – language, script, numbering, calendar, culture, etc. Most of the pre-2010 development of RDA was done by the four author countries, countries that share the same language and similar cultures. Canada, a bilingual country, was often the sole voice reminding the group to be aware of limitations to instructions that work only in English. The intention to internationalize was real, but putting it into practice takes time as well as feedback from other communities. As the group of RDA implementers broadens, it provides new opportunities to identify the "Anglo-American" remnants and eliminate them.

One successful example was the reversal of instructions about initial articles when recording titles of works. In the 2012 April update to RDA, a change to the original instructions was

implemented, arising from a [proposal](#) from the Deutsche Nationalbibliothek. As they translated RDA, they identified that the main instruction in RDA at 6.2.1.7 was a carry-over of “Anglo-American” practice. It was going to cause major problems for their implementation of RDA because of the role of initial articles in an inflected language such as German. The main instruction said to omit the initial article; there was also an alternative instruction to record the title of the work with the initial article. The Deutsche Nationalbibliothek proposed that the two instructions should be reversed, thus eliminating an “Anglo-American” bias in RDA. The change was accepted and implemented.

The JSC has undertaken two initiatives to support and expand internationalization: the establishment of the Translations Working Group and the proposed new governance structure for JSC and the Committee of principles that oversaw the development of RDA.

As the number of RDA translations grows, the JSC issued a document outlining a policy for translations ([Translation Policy for RDA and RDA Toolkit 6JSC/Policy/6](#)). In this document, the JSC commits to creating a Translations Working Group because in the past few years the feedback from translation teams has been enlightening and led to positive changes in the standard. From the section on establishing this new working group:

The JSC is keen to get input from translators in regard to how the standard might be improved, especially in regard to phrasing or word choices that are problematic for non-English speakers and in regard to instructions that may reflect cultural bias. To facilitate this input the JSC has formed a working group for translations that can supply reports on language issues and help the JSC to identify solutions to the problems that arise. Translations groups are expected to provide this sort of input to the JSC and may be invited to supply a representative to the working group.

(Translation Policy for RDA and RDA Toolkit 6JSC/Policy/6)

An interesting aspect of the structure of the Toolkit is the way that the translations have been integrated into the Toolkit. There is functionality so that one can switch from one language version to the other, either for the whole text, or at a particular point in the text, allowing one to compare the way each translation expresses the text for a particular instruction. One can also compare the texts between language versions with split screens. The texts of the different translations are interlinked through the use of identifiers for each paragraph; those identifiers also carry information about the language version and the content version (i.e., in line with which update). The translations are integrated and presented as multiple facets of the same work, in the same place, within a single subscription, which seems to affirm the importance of the translations as key documents, as opposed to treating them as separate, ancillary documents.

In the new [2015-2020 strategic plan](#), the first priority is to make RDA an internationally recognised standard, followed by the priority to increase the adoption of RDA internationally. These goals are supported by proposals for a new governance structure which makes radical

changes to broaden the participation in RDA development and secure a broad range of perspectives during the decision-making processes. The Joint Steering Committee and the Committee of Principals have been renamed the RDA Steering Committee (RSC) and the RDA Board. There will be a mixture of expert and regional representation. Regional representation will be based on the UN areas rather than the original author countries. For the RDA Board, it will include, among others, representatives from national institutions, one from each of these six areas: Africa, Latin America and The Caribbean, North America, Asia, Europe, Oceania. The RSC will include, among others, one regional representative from each of these same six areas as well. In addition, the RSC will also include a representative from the Translations Working Group. There is a credible effort to support internationalization with a matching governance structure that distributes decision-making among a broader group of international representatives.

One of the defining characteristics of RDA is that it is a practical standard based on a robust theoretical framework. During the development process, RDA was explicitly aligned with FRBR and FRAD. The introductory chapter of RDA announces the standard's alignment with the IFLA conceptual models. The evidence is seen in the very structure of the document, the choice of vocabulary, the key role of user tasks, the move to understanding bibliographic data in terms of entities, attributes and relationships, and the focus on the role of relationships. With the April 2015 update, the alignment with the FRBR family of conceptual models was completed with the addition of the alignment with FRSAD. In practical terms, it does not mean the addition of a set of entities and attributes to define "subject" or "thema", with corresponding instructions on recording that data. Instead, the approach was to stay at a high level of generality. Subjects can be categorized in many different ways, depending on language, culture, discipline, specialization of a collection of resources, etc. The decision was taken not to define the subject elements, but mainly to add some text to chapter 23, on subject relationships.

In the 2010 version, RDA acknowledged the importance of the subject relationship, in a simple way: recording a subject relationship was considered "core" and included in 0.6 (core elements), and, at that time, there was no text in chapter 23. In 2015, the alignment with the FRBR family of models was expanded to include FRSAD. The section on the alignment was switched from 0.3 to 0.2, putting it into an even more significant position within the introductory chapter of RDA. Text was added to chapter 23, defining terminology, functional objectives and principles, core elements, the instructions for recording the subject relationship and relationship designators (following the pattern established in other chapters). Appendix M was also added with controlled vocabulary for the relationship designators for subject relationships.

The text in chapter 23 introduces a new notion, the *identifiable subject system*, which refers to "a standard for subject access points and/or classification numbers used by the agency creating the data." This creates a way to link to outside subject schemas or controlled vocabularies

without predetermining the choice of one particular perspective in terms of how subject access is organized or what terminology is used. This is another example of the way in which RDA is designed not to be “one size fits all” but to accommodate a range of needs and to promote flexibility within a standardized framework.

The alignment with the three models that form the FRBR family of conceptual models marks a point in the completion of unfinished tasks. However, it will not be the end point in terms of making changes to maintain alignment with the FRBR family of conceptual models. The IFLA FRBR Review Group is in the process of consolidating the three models into one. The process of consolidation has required remodelling to resolve points where the models conflict or are inconsistent with each other. The remodelling process is also introducing some streamlining in the consolidated model, deprecating some entities, and introducing some new entities. The consolidated model follows the spirit of the three former models and does not represent a radical break. However, the consolidated model may have some effect on the RDA elements and the structure of RDA. It will not create changes that radically affect the data recorded though it may change how we understand certain areas of data. The model is under development and is expected to undergo world-wide review in 2016. Until the model is published, one cannot be certain about its impact on RDA. The JSC (or RSC, as they are now known) are aware of the work on the consolidated model and are in close communication with the FRBR Review Group. The consolidated model is seen by JSC as one of several possible sources of impact on the content and structure of RDA. In early 2015, the JSC decided to adopt a working principle called “[the pragmatic approach](#).” if a proposal for revision is made in an area likely to be affected by external changes, RDA revisions in that area will be deferred or referred to a working group looking at the complete impact of that particular external change.

The standard in 2015 is not the same as the standard issued in the first release. There have been many changes to advance RDA and to make it better aligned with the vision and ongoing goals for the standard. The changes have not deflected RDA away from its original defining characteristics: a cataloguing standard for the digital environment; a standard aiming to connect with the broader cultural heritage community and also with the broader international cataloguing community; a standard based on a robust theoretical framework through its alignment with the IFLA conceptual models. In the developments since the first release, one notes a continuing commitment to be a flexible and adaptable standard, open to other communities. RDA will keep changing both in response to technological advances and, as importantly, in response to the needs of different linguistic and cultural communities as the goal of internationalization is actively pursued.