

To: Joint Steering Committee for Revision of AACR

FROM: Jennifer Bowen, ALA representative
John Attig, ALA representative designate

SUBJECT: Comments on RDA Scope and Structure

General comments

In its comments on the draft of RDA Chapters 6 and 7, ALA made the following recommendation about the RDA development process:

Adopt a top-down development approach. ... Determine a clear and explicit scope for RDA in terms of whom it will serve, what resources it will address, and what types of metadata it will produce.

The RDA Scope and Structure document is a direct response to this recommendation, and ALA thanks the JSC and the RDA Editor for preparing this important document. This is exactly the sort of fundamental document that ALA sees as the basis for the entire RDA development.

In its comments on the draft of RDA Chapters 6 and 7, ALA goes on to suggest ways in which such a fundamental document can be used as the basis for making decisions on the principles and objective of RDA, developing a set of general guidelines, and using all this as the foundation for the content of RDA. ALA still believes that this is the proper way in which to develop RDA content.

Although RDA Scope and Structure was distributed by the JSC as an informational document that defines “the framework for the development of RDA,” ALA considers that this document lies at the very foundations of the RDA project. We therefore believe that its content should be discussed, and questions related to its provisions should be decided in the high-level context of this document, and *then* implemented in the drafts of the RDA text. We would therefore like to take this opportunity to provide our comments on this document. We recommend that the JSC constituencies be asked to respond, either to the RDA Scope and Structure document itself, or to these comments.

Furthermore, ALA has from the beginning of the process commented on the difficulty of reviewing drafts of specific content without seeing either the Glossary, which defines many of the critical terms not contained in the text itself, or the conceptual explanations that are to be included in the General Introduction. We hope the intention is that the content of RDA Scope and Structure will contribute to both of these. We therefore urge that the document be maintained as decisions are made, and used as a means of developing content for the Glossary and Introduction and making this content available to reviewers of the RDA drafts.

Two general points may be made about the following comments:

1. While many of the provisions of RDA Scope and Structure have clearly governed the content of the RDA drafts, this is not always true. All too often, it is clear that a relationship to an external model has been discovered or invented after the fact; in some cases, there is no discernible impact of the stated features of the external standards on the text of RDA. This is the opposite of the process that ALA recommended: Begin by declaring the RDA schema (elements and attributes) based on whatever models are appropriate, then build the guidelines and instructions on that schema. RDA Scope and Structure makes it quite apparent that the basic content carried forward from AACR2 is being taken as a given, and that the scope, structure, and schema are being manipulated in convoluted ways in order to justify that basic content. This is particularly obvious to those who do not approach resource description through the AACR tradition; as these are a major intended audience for RDA, this is particularly unfortunate.
2. Some of the comments received by ALA suggest that RDA Scope and Structure misapplies some of the provisions of the external models. Those comments are included below to alert the JSC to these suggestions and to anticipate comments that may be made during the upcoming meetings with DCMI and IEEE-LOM representatives.

Introductory paragraph

The report of the outcomes of the October 2006 JSC meeting states: “The JSC affirmed the role of the IME ICC draft Statement of International Cataloguing Principles as the basis for the cataloguing principles used throughout RDA.” Yet there is no mention of the Statement in this document. Although the relationship of the Statement to specific details about the scope or structure of RDA would be difficult to identify, we would like to see this affirmation included in the introductory paragraph to this document.

Some members of the ALA community feel that there is a misalignment between the RDA concepts derived from the FRBR/FRAD models and the structure of RDA, and that it is therefore a mischaracterization to state that these serve as underlying conceptual models for RDA; the most that can be said, in their opinion, is that RDA is strongly influenced by FRBR and FRAD in certain specific ways.

1. Scope

Descriptions created by applying RDA need to support much more than just resource discovery. Metadata users also compare descriptions, verify citations, use electronic resources, etc. — and many user needs are for *information*, not *resources*. While

resource-discovery tasks form the core of RDA's scope, it would be unwise to ignore the importance of other user tasks, or to draw too absolute a line.

The present scope statement addresses the needs of users of descriptions created by applying RDA. It is equally important to address the needs of users of the content standard itself. This has been a matter of controversy since the inception of the RDA project. ALA feels that this is an important part of the scope of RDA and should be addressed in this document.

ALA also feels that it is inappropriate to claim that RDA is directed to the entire resource description community, when we are constantly being told that its provisions do not meet the needs of significant parts of that community. We suggest that it be acknowledged that RDA comes from the tradition of library descriptive cataloging and that it is primarily (but not exclusively) designed to meet the needs of that community, while making an effort to be relevant to a broader range of resource description activities. This is a complex matter that deserves careful consideration. The following statement is offered as a possible starting point:

RDA is primarily intended for immediate use in the library cataloguing community. RDA may also be useful for other resource description activities within libraries and other information management organizations. RDA also participates in the ongoing global discourse on descriptive metadata.

1.1 Definitions

There are definitions throughout section 1 of the document. Section 1.1 seems to be definitions of "Basic Concepts."

In this section and elsewhere, this document uses many concepts and terms from identified external models. However, the document does not always indicate **how** these terms and concepts relate to the corresponding RDA terms and concepts. It is helpful when the document refers explicitly to an external model at points in the document where that model is used; footnote 5 is a particularly good example for its specificity. In other cases, it is not always clear where the external definitions are used exactly, where they are used or extended to build RDA definitions, and where the RDA definitions employ external terms in ways quite different from (and sometimes even in conflict with) their use in the external models.

There are numerous differences in the terminology used in RDA and each of the models referenced. The selective use of terms from different sources in this document is confusing. It might be worth constructing a table that relates RDA, FRBR/FRAD, DCAM, and <indec> terms. It is not clear that there is virtue in using terminology from other models, often unfamiliar to most RDA users, in defining RDA concepts; it also gives the impression — sometimes inaccurately — that RDA is in fact using these terms

in a way that is compatible with these models. On the other hand, there is often no good reason why RDA cannot use a term established in an existing model, where the meaning is the same.

In too many cases, definitions in this document contain wording that is descriptive rather than definitional. For example, “either tangible or intangible in nature” in the definition of *resource* adds nothing to the definition of the term; such descriptive information should be kept separate from the formal definition.

Resource

This definition is narrower than the one in the DC Abstract Model. Either RDA should accept the DCAM meaning, or change the term to *bibliographic resource* to remove the conflict. Also, consider adopting some of the DCAM language: either “any information object that has identity” (current) or “any information object that might be identified” (proposed). In general, we support closer alignment of definitions with the external models.

Resource discovery

In the spirit of more closely aligning definitions, we question the need to change the FRBR term “entities” to “resources.”

Descriptive data

Access point control data

The DC Abstract Model uses *property/value pair* to cover what RDA would characterize as both the element and its content; it uses *value representation* to cover the content alone. It is not clear whether “data” in these terms includes both the elements and their content, but we suspect that it does. There are other cases in this document in which “value representation” is used when clearly both element and content are meant.

The operational definition of *controlled access point* should not be a footnote. As we argue elsewhere, this is an important concept in RDA and deserves to be defined formally, either here or (perhaps) in section 1.6.

1.2 Descriptive data

1.3 Access point control data

ALA considered the request from the IFLA Cataloguing Section to add an acknowledgement of “the contributions of the ISBDs to identification of data elements and the structure of RDA chapters.” Within ALA, support for adding this acknowledgement to the Scope and Structure document was mixed, with some members preferring that the document be more focused upon the future rather than on the historical underpinnings of RDA. On the other hand, other ALA members feel that maintaining ties with the past is not a problem, so long as it is clear that we are attempting to reconceive how that past should be modified to meet the needs of the future.

As the ISBD data elements are acknowledged as one of the primary influences on FRBR, and RDA acknowledges the FRBR model as the source of RDA elements, the statement by the IFLA Cataloguing Section is certainly true, as far as it goes. In fact, RDA has a very complex relationship to various features of the ISBD, which the JSC seems to be redefining in ways that are central to the scope of RDA. The move of ISBD presentation elements out of the text of RDA and into an appendix is one example of this, as is the limited, selective support for some of the ISBD structure through the definition of sub-elements. This redefined relationship with ISBD needs to be explained more fully, either somewhere in this document (as some would prefer) or elsewhere (such as in the RDA General Introduction).

ALA suggests that the apparent restriction of RDA elements to FRBR/FRAD attributes and relationships may not be absolute. The FRBR/FRAD models clearly form the basis for the elements in RDA. FRBR has been out for a decade, during which time many have pointed out attributes that were inappropriately not included in FRBR or FRAD, including some AACR2/ISBD elements. In fact, there are some elements in RDA — the three elements derived from the RDA/ONIX Framework come to mind — that are not to be found in FRBR. The JSC has never been shy about extending the boundaries of its underlying standards, and should not in this case. The statements in the first paragraphs of these two sections should be qualified as “generally” true.

The statement (1.3, 2nd paragraph) that certain attributes are covered on a selective basis should be explained.

The statements on user tasks relating to resource management and data management may need to be qualified, particularly regarding the latter. Experience has shown that metadata about metadata (meta-metadata?) are crucial to the effective use, sharing, and processing of that metadata. It should be included within RDA to the extent possible. In fact, we can think of at least one element already in RDA — *Issue, part, or iteration used as the basis for the description* — that falls into this category. Although we have not seen the chapter on “Other information used in access point control,” we suspect that there are some examples of meta-metadata that are candidates for inclusion there as well.

ALA believes that subject relationships should not be ruled completely out of scope. We understand that the JSC has agreed that some mention of subject access will be included in RDA. We would urge that this decision be reflected in the Scope and Structure document. We support the inclusion of the concept of subject access within RDA, with appropriate references to relevant standards. We also suggest that RDA recognize that access points formulated according to Part B of RDA might be appropriate for use as subject access points, as well as primary and secondary [descriptive] access points.

Section 1.3 uses the term *controlled access point* in relation to FRAD. In the 2005 draft, this term was simply *access point*; we would appreciate confirmation that this is a deliberate change in FRAD.

ALA disagrees that “relationships between controlled access points, as defined in FRAD, are also out of scope” (final paragraph of 1.3). The *parallel language relationship*, for example, is crucial in the successful navigation of today’s multilingual and multiscrypt information universe.

1.4 Elements

ALA is not convinced of the conceptual value of treating relationships as elements. We prefer that the data elements used to express the relationships be treated as elements, but not the relationships themselves. In addition, in our comments on the draft of Chapter 6 (Related resources), we objected to the characterization of relationships as mandatory or optional, as opposed to the elements that express the relationships.

We do not believe that it is accurate to state that every RDA element corresponds to an attribute or relationship in FRBR or FRAD, or that the scope of the elements always matches that of their FRBR/FRAD counterparts. If nothing else, it is necessary to include RDA element sub-types and sub-elements. If the JSC wishes this statement to be true, then some changes will need to be made — beginning with the new elements based on the RDA/ONIX Framework. On the other hand, we suggest that only a general correspondence can or needs to be asserted.

The distinction between elements, element sub-types, and sub-elements is of fundamental importance in providing a way of capturing the structural relationships between related attributes. However, the explanation given here is not particularly clear or accurate. The use of the term “sub-class” to define “sub-type” is not helpful to understanding this concept, and the assertion that “the defined scope of the element sub-type is co-extensive with the defined scope of the element” is not accurate — or else the inclusion of a Scope rule for each sub-type (see section 2.3) is completely redundant; we suspect that it is not always true.

1.5 Value representations

Value strings are one category of value representations. Sections 1.5 and 1.6 are closely related, but the relationship is not particularly clear. Consider combining into a single section on “Values,” with subsections that explain the relationship between *values*, *value representations*, and *value strings*.

ALA is not clear how these categories are applied within RDA — with the possible exception of the categories of element-types used to express relationships, which is not clearly related to the list of categories that precedes it. We suspect that this section represents an after-the-fact attempt to relate RDA to the <indec> Metadata Framework.

If that is the intent, then the concepts need to be much clearer, and they need to be applied to the structure and content of RDA; we do not see that this has yet been done.

We believe that the list of categories may misrepresent the <indec> Metadata Framework. These are categories of <indec> *attributes*, **not** <indec> *values*; *value representations* in the DC Abstract Model is the equivalent of <indec> *values*. These <indec> attributes are syntactic categories, *values* and *value representations* are semantic categories. The categories listed in this section might be characterized as attributes of a resource, but **not** as categories of value representations. [Note: We have not been able to access the <indec> Metadata Framework online to confirm this.]

The use of the term “quality” by <indec> is unfortunate; it may be misinterpreted by those unfamiliar with the model as an assertion that these attributes measure the quality of the resource being described — which is clearly not how the term is intended.

The second main paragraph and the definition of *description* misuse the term *value representations*, as defined in the DC Abstract Model. An entity does not have a value representation; only values have representations. The correct term in the definition of *description* should be “property/value pairs” which comprises both the element and its content. In the paragraph above, these are more categories of elements used to express relationships than categories of value representation.

1.6 Value strings

The three categories given are not “forms” of a value string; one cannot tell by looking at a string whether it is transcribed or not. These are categories of value string.

Transcription is an attribute of a value string that identifies its relationship to a source of data. In the DC Abstract Model, the distinction between structured and unstructured strings is based on particular syntax encoding schemes and in RDA, on the specific guidelines for recording the string. The distinction between transcription, on the one hand, and structured/unstructured is an important one because the two are not mutually exclusive.

The DC liaison to CC:DA informs us that “in DCAM, strings are simple things without inherent structure (there was such a thing as a “structured value” in the old days, but we’ve pretty much abandoned those as a bad idea.” This is an example of how this document uses concepts from DC (and other models) in the course of introducing complexities that are quite foreign to those models. ALA is concerned that if RDA is not going to respect the spirit of its underlying models, then it should be more specific in describing its relation to those models.

The use of the <indec> attributes in the definitions of *structured* and *unstructured strings* does not particularly aid in understanding these concepts.

Another category of value string that might be considered by the JSC is *controlled vocabularies*. ALA would like to see recognition of controlled vocabularies as a distinct category of value representation, as well as reference to controlled vocabularies at relevant places in the text of RDA. The distinction between *structured* and *unstructured* seems to deal with syntactic features, while *controlled vocabularies* deals with semantic features of the value strings.

1.8 Record structure

File structure, record structure, encoding and presentation are all distinct concepts; both encoding and presentation are independent of the **structure** of either individual records or of files of records. This section does not in fact address file structure — although the inclusion of the RDA Implementation Scenarios within this document might be a good idea — nor does it address record structure. The mappings and other information to be presented in the appendices deal either with encoding schemas or with presentation features, and we are not sure that it is accurate to characterize what we intend to do with ISBD presentation features as a “mapping.”

In the second sentence, “value representations” should be “property/value pairs” because this is a case that involves both the element and its content.

2. Structure

ALA feels that the scope of Parts A and B should be more directly related to the FRBR Group 1 and Group 2 entities respectively. Part A covers description of and access to bibliographic entities/resources, while Part B covers control of names of persons, etc. The introductory statements to section 2 and its subsections could make this clearer, and the titles of the parts might focus on this distinction: “Bibliographic resources” and “Agents” has been suggested as a succinct way of doing this.

2.1 Part A – Description

The structure of Part A is not clearly stated in the introductory paragraph. The decision to organize Part A primarily by the FRBR user tasks which the data elements support leaves some significant gaps. For example, identification is confined to manifestations and items; there are data elements that are identified in FRBR as primarily supporting the identification of works and expressions, yet there seems to be no place for them in this structure. More fundamentally, chapters 6 and 7 do not seem to be related to the FRBR user tasks in any way.

ALA would have preferred that some single organizing principle could be applied to Part A in a manner similar to the organization of Part B according to the FRBR Group 2 (and Group 1) entities. It is not clear that such an organizing principle could be fitted to Part A without significant reorganization, but we do feel that the present structure is too complex and inconsistent to be readily understood by catalogers.

2.2 Part B – Access point control

There are data elements that apply to all the entities covered in Part B; some of these are included in the chapter on Other information used in access point control. ALA would like to see these common elements addressed *before* the chapters on individual entities, either as part of the initial chapter in Part B or in a new chapter covering common elements.

Access points representing places

Places occupy an ambiguous position in the FRBR model, functioning clearly as Group 3 entities (geographic coverage as a subject relationship), but in more subtle ways as Group 2 entities. Place *as such* cannot be “responsible for the intellectual or artistic content, the physical production and dissemination, or the custodianship of [Group 1] entities.” In AACR2, the rules for place names function almost exclusively as jurisdiction names to be used as or as part of names of corporate bodies or as qualifiers. If it is the intention of RDA to limit its guidelines on place names in this way, then it would be more appropriate for this chapter to be merged into the chapters on “Access points representing corporate bodies” and “Access points representing works, expressions, manifestations, and items.” On the other hand, if other uses of controlled access points for places are anticipated, that should be stated in the scope of the chapter on “Access points representing places.” Among other implications, this would make it clear whether non-jurisdictional places (such as geographic features) are included in the scope of RDA.

Other information used in access point control

ALA questions the need for this chapter as described. To the extent that the elements in question apply to all the entities covered in Part B, they should be addressed in a general chapter dealing with common elements. To the extent that they are particular to one or more of the entities, they should be included in the appropriate chapter covering that entity. Our assumption is that almost the entire chapter falls into the former category.

2.3 Format

ALA questions whether the section on notes should be placed in the section on the format of the guidelines. The most important issue addressed is when content is characterized as a note an element, as opposed to being treated as an element in its own right. This discussion should be elsewhere, perhaps as part of the discussion in 1.6 of structured and unstructured value strings. The *format* of guidelines for notes — and also those relating

to sub-elements and element sub-types — seems inappropriate for inclusion in this high-level document.

2.5 Appendices

Capitalization

ALA remains unconvinced that consistency of capitalization is required to support any of the FRBR user tasks.

Abbreviations

ALA is persuaded by the JSC's decision that the use of most abbreviations is a barrier to User Convenience. ALA will consider whether this principle should be applied to the specification of abbreviations in Part B of RDA, in spite of the resulting need to change many existing access points (and thus violating the principle of Consistency). There is a strong sentiment in ALA that the principle of User Convenience should be of primary importance in decisions regarding the content of RDA.

Initial Articles

ALA is not certain that RDA guidelines should require catalogers to identify initial articles in order to omit them from controlled access points. We urge the JSC to consider this issue in its discussions on the content of Part B. If RDA does *not* call for omission of initial articles, then their treatment becomes a matter either of encoding or presentation, and can be handled in the relevant appendices by referring to existing lists of initial articles in relevant encoding and presentation standards.

Record structures for descriptive data

Record structures for access point control data

As argued above, the content of these appendices is not really the structure of the records that contain RDA content, but rather the **separate** issues of the application of different **encoding standards** and of different **presentation standards** to RDA content. We recommend that the titles and scopes of these appendices be modified to make this distinction.