TO: Joint Steering Committee for Development of RDA

FROM: Kathy Glennan, ALA Representative

SUBJECT: Treatment of Subjects in RDA

RELATED: 6JSC/LC rep/3 and responses
                  6JSC/M/284-308 (at M.289.1)

Background

In May 2011, the Library of Congress Representative submitted a discussion paper (6JSC/LC rep/3, Chapters 12-16, 23, 33-37 (Group 3 entities and "subject")), on what might be done with the placeholder chapters relating to subjects in RDA. This paper was discussed by the Joint Steering Committee at its meeting in Glasgow, November 2011. To respond to this and other subject-related discussion papers and proposals relating to RDA, ALA formed the RDA SAC Subcommittee, a group of experts on subject analysis; this subcommittee provided the basis of ALA's comments on the 2011 LC discussion paper, and has continued thinking about this issue. The following set of recommendations originated with the subcommittee and has been endorsed by both ALA's Subject Analysis Committee and its Committee on Cataloging: Description and Access. ALA proposes a general strategy for dealing with subject entities, attributes, and relationships in RDA. If the JSC approves some version of such a strategy, ALA would be interested in preparing a specific proposal.

The LC discussion paper noted: "The general assumption of this discussion paper is that the JSC wishes to continue our policy of providing basic guidance within RDA for general libraries and others while referring out to specialist manuals, etc. (in this case other thesauri, subject heading lists, classification systems, etc.) for more specific instructions on form/structure of terminology, relationships among the values used as subject terms, and specifics for using any particular thesaurus or subject heading or classification list."

ALA strongly endorses this position, which forms the basis for our recommendations. Given the variety of subject systems in existence and use, we feel that RDA should provide basic guidance without imposing constraints, either upon the ability of individual subject systems to define their own structure and content, or upon the use of such subject systems by catalogers.

To this end, we offer the following recommendations for dealing with subject entities, attributes, and relationships in RDA.

Recommendations

1. **General approach to subjects in RDA:** ALA considered the question of the target audience for RDA guidelines and instructions relating to subjects. We believe that these should not be addressed either to developers of subject systems (i.e., to creators of approved subject structures and terms) or to catalogers applying subject systems. Both of these points of view are better addressed in the specifications of the individual subject systems.

ALA recommends that RDA include brief general guidelines, sufficient to allow subjects to be integrated into a general RDA data model. It would provide definitions of a few key entities, attributes, and relationships, thereby providing "hooks" that can be used to connect the RDA data model to the data models specified by individual subject systems. The goal would be to allow those applying RDA, including both catalogers and system designers, to understand subject concepts as part of the same conceptual model as the other features of RDA.

2. **Choice of model:** ALA recommends that RDA adopt the FRSAD approach, with a single subject entity, rather than the FRBR Group 3 entities that are the basis for the placeholder chapters currently included in RDA. The alignment of RDA with FRSAD should be explained in the introductory chapter of RDA, as a sub-section of RDA 0.3, "Conceptual Models Underlying RDA". *Rationale:* ALA believes that only the FRSAD model provides appropriate entities, attributes, and relationships for modeling subject treatment, while imposing minimal constraints upon the ability of subject systems to define their own structure and content. ALA acknowledges that the FRSAD Thema entity is defined as at a very high level of generality that subsumes all of the FRBR group 1, 2, and 3 entities — and that this does not correspond to the RDA model for these entities. However, we feel strongly that a single subject entity is preferable to the 11 FRBR entities, even though this single subject entity may not correspond exactly to the FRSAD Thema entity.

3. **Terminology:** ALA recommends that the FRSAD terms "Thema" and "Nomen" be replaced in RDA by the terms "Subject" and "Name of Subject". This modification of FRSAD terminology in RDA should be explained in RDA 0.3 and in the basic instructions for recording attributes of the Subject entity. *Rationale:* We find the use of Latin-based terminology in FRSAD to be inconsistent with the general principles of RDA and an unnecessary use of esoteric language.

4. **User tasks:** ALA recommends that the FRSAD user tasks be documented in RDA 0.0, "Purpose and Scope".

   The FRSAD user tasks are defined as follows:

   Find one or more subjects and/or their appellations, that correspond(s) to the user's stated criteria, using attributes and relationships.

   Identify a subject and/or its appellation based on its attributes or relationships (i.e., to distinguish between two or more subjects or appellations with similar characteristics and to confirm that the appropriate subject or appellation has been found).

   Select a subject and/or its appellation appropriate to the user's needs (i.e., to choose or reject based on the user's requirements and needs).

   Explore relationships between subjects and/or their appellations (e.g., to explore relationships in order to understand the structure of a subject domain and its terminology).[1]

---

[1] *Functional Requirements for Subject Authority Data (FRSAD): a Conceptual Model.* (2010). IFLA Working Group on the Functional Requirements of Subject Authority Records (FRSAR). International Federation of Library

The JSC should consider extending the FRSAD user task *Explore* to FRBR Group 1 and Group 2 entities as well. This may be addressed in the FRBR Review Group's reconciliation of the FR models, results of which should be available by the November 2013 JSC meeting.

5. **Entities:** ALA recommends that there be only one subject entity in RDA: the FRSAD "Thema" entity — renamed "Subject" in RDA (see recommendation #3 above). Consistent with the treatment of names elsewhere in RDA, the FRSAD "Nomen" entity — renamed "Name of Subject" in RDA (see recommendation #3 above) — should be treated as an attribute of the **Subject** entity. The FRBR Group 3 entities (**Concept**, **Object**, **Event**, and **Place**) should not be treated as entities in RDA. ***Rationale:*** We believe that the simple FRSAD approach allows greater flexibility in allowing subject systems to define their own structure and content.

6. **The primary Subject relationship:** The FR models specify that the *has as subject* / *is subject of* relationship only exists between the **Work** entity and the **Thema** entity (see #13 below). ALA recommends that this specification be accepted provisionally. ***Discussion:*** We believe that there may be cases in which this is an inappropriate limitation, but we have difficulty developing convincing use cases. The most common instance cited is that of a change in subject scope between editions (expressions) of a work; the only way to avoid **Expression**-to-**Subject** relationships is to state arbitrarily that a change in subject scope always implies the existence of a new work. There may be other use cases that apply to the **Manifestation** and **Item** entities as well.

7. **Subject vs. genre/form:** The Subject entity as defined in FRSAD describes what a work is ***about***; it does not specify what a work ***is*** (i.e., its form or genre). Entities, attributes, and relationships for genre/form should be treated separately from subjects in RDA. ALA sees a critical need for RDA to deal with genre/form. We recommend that the JSC urge the FRBR Review Group to develop genre/form as an extension to the FR model. We also recommend that the JSC indicate its willingness to accept relevant proposals from JSC constituencies, regardless of whatever action is taken in the FR models.

8. **Subject chapters in RDA:** ALA recommends that most of the placeholder chapters dealing with subjects in RDA be eliminated. Instead, we recommend the following chapters:

> **Section 4. Recording Attributes of Subjects**
>
>> **Chapter 12. General Guidelines on Recording Attributes of Subjects**
>>
>> **Chapter 13. Identifying Subjects**
>> [These two chapters could probably be merged]
>
> **Section 7. Recording Relationships to Subjects**
>
>> **Chapter 23. General Guidelines on Recording the Subject of a Work**
>
> **Section 10. Recording Relationships between Subjects**
>
>> **Chapter 33. General Guidelines on Recording Relationships between Subjects**

Associations and Institutions. Available at: http://www.ifla.org/files/assets/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf (accessed 2013-01-21).

Appendix L. Relationship Designators: Relationships between Subjects
[There may not be any such designators; they would tend to be defined within each subject system.]

9. **Events:** ALA recommends that instructions covering events as subject headings (currently in placeholder Chapters 15 and 36) should be included in the general guidelines for recording attributes and relationships of subjects (Chapters 12/13 and 33 in the previous recommendation). Instructions for recording attributes and relationships of events as corporate bodies should remain in Chapters 11 and 32. ***Rationale:*** Not all subject systems will treat events in the same manner; RDA should not impose any constraints on how events should be treated as subjects.

10. **Places:** ALA recommends that the present content of Chapter 16, Identifying Places, be retained in RDA. However, we think that serious consideration should be given to moving this content to Chapter 11, Identifying Corporate Bodies. ***Rationale:*** RDA specifies how the names of certain places should be recorded. However, these instructions are limited to places as jurisdictions, and serve as the basis for naming governments and government corporate bodies. Moving these instructions to Chapter 11 would allow any instructions on places as subjects to be covered in the general guidelines for recording subject attributes and relationships (Chapters 12/13 and 33 in recommendation #9 above).

11. **Attributes of the Subject entity:** The FRSAD attributes of the **Thema** and **Nomen** entities, along with their definitions, are listed below.

    ALA recommends that (a) **Name of Subject** be treated as an attribute of the **Subject** entity, and that (b) the FRSAD attributes of **Nomen** be treated as attributes of the **Name of Subject** attribute. This would be the first case in which RDA defines attributes of an attribute, but it seems justified in this case. We also recommend that **Preferred Name of Subject** and **Variant Name of Subject** be defined as element sub-types of the **Name of Subject** element, in order to be consistent with the treatment of names elsewhere in RDA. These modifications have been included in the list below, along with our comments on each attribute.

    Subject [FRSAD 3.4 "Thema"]
    Definition: Any entity used as a subject of a *work*.

    – which has the following attributes/elements:

    Type of subject [FRSAD 4.1.1]
    Definition: The category to which a *thema* belongs in the context of a particular knowledge organisation system.
    Notes: These could be the FRBR Group 1 and Group 2 entities, and in certain implementations the FRBR Group 3 entities. However, other subject systems and implementations could require different values for this attribute.
    This should be considered for inclusion in RDA and for "core" status; if a subject system defines categories, the identification of these categories are necessary to the use of the system.
    This is data about data. It would presumably be used in mapping to encoding schemes.

Scope Note [FRSAD 4.1.2]
Definition: A text describing and/or defining the *thema* or specifying its scope within the particular subject authority system.
Notes: This is similar to the Cataloguer's Note element in RDA (5.9, 8.13, 24.8, 29.7). These are not core elements, but some of them are basic instructions.

Name of subject [FRSAD 3.5 "Nomen"]
Definition: Any sign or sequence of signs (alphanumeric characters, symbols, sound, etc.) that a *thema* is known by, referred to, or addressed as. Examples include "love," "∞," or "595.733".

– which has the following subtypes:

Preferred name of subject

Variant name of subject

Notes: FRSAD does not make the distinction between preferred and variant *nomens*, but for consistency with the rest of RDA, this distinction should be made. The preferred name should be a core element.
In the case of classification systems, the assumption is that the notation would be the preferred name of the subject, and the label or caption would be a variant name of the subject.

– and has the following attributes/subelements:

Type of Name of Subject [FRSAD 4.2.1]
Definition: Category to which the *nomen* belongs.
Notes: FRSAD states that "in addition to implementation-specific types, there are two important values": *identifier* and *controlled name*. Within RDA, Identifier for the Subject would be a separate element, and there might be instructions in RDA for constructing authorized access points representing subjects, similar to those for authorized access points representing names, although those instructions would need to be very general in nature in order to accommodate any subject system.
Other Type of Name of Subject values would be implementation-specific.

Scheme (LCSH, MeSH, AAT, etc.) [FRSAD 4.2.2]
Definition: The scheme in which the *nomen* is established, including value encoding schemes (subject heading lists, thesauri, classification systems, name authority lists, etc.) and syntax encoding schemes (standards for encoding dates, etc.).
Notes: This should be considered for inclusion in RDA and for "core" status; the identification of the particular subject scheme is necessary to the use of the system.
This is data about data.

Source consulted [FRSAD 4.2.3 "Reference source of nomen"]
Definition: The source in which the *nomen* is found. It may also be modeled as a relationship with the appropriate Group 1 entity.
Notes: Source consulted is a general RDA element applicable to any authorized access point. It should be considered for inclusion in RDA.
This is data about data.

Representation of name [FRSAD 4.2.4]
   Definition: The data type in which the *nomen* is expressed.
   Notes: Examples include *alphanumeric*, *sound*.
      There does not seem to be a comparable data element in RDA
   and it may not be necessary or appropriate to include this attribute
   in RDA.
      This is data about data.

Language of name [FRSAD 4.2.5]
   Definition: The language in which the *nomen* is expressed.
   Notes: This is data about data, and RDA has not (yet?) chosen to
      record such information for individual elements.

Script of name [FRSAD 4.2.6]
   Definition: The script in which the *nomen* is expressed.
   Notes: This is data about data, and RDA has not (yet?) chosen to
      record such information for individual elements.

Script conversion [FRSAD 4.2.7]
   Definition: The rule, system, or standard used to render the *nomen* in
      a different representation.
   Notes: This is data about data, and RDA has not chosen to record such
      information for individual elements.

Form of name (e.g., full, abbreviated, formula, etc.) [FRSAD 4.2.8]
   Definition: Any additional information that helps to interpret the
      nomen.
   Notes: This is data about data, and RDA has not chosen to record such
      information for individual elements.

Time validity of name [FRSAD 4.2.9]
   Definition: The time period in which the *nomen* is/was used or is/was
      valid with a subject vocabulary system.
   Notes: This is data about data, and RDA has not (yet?) chosen to
      record such information for individual elements.

Audience [FRSAD 4.2.10]
   Definition: The community or user group for which the *nomen* is the
      preferred form.
   Notes: There is nothing comparable in RDA. This is probably data
      about data, and will be very much implementation-specific. This is
      probably not a candidate for inclusion in RDA.

Status of identification [FRSAD 4.2.11 "Status of *nomen*"]
   Definition: The status of the *nomen* in a subject authority system.
   Notes: Status of identification is a general RDA element applicable to
      any authorized access point. It should be considered for inclusion
      in RDA.
         This is data about data.

ALA recommends that the JSC identify which of the FRSAD attributes should be included in RDA. We tentatively recommend that **Type of Subject**, **Preferred Name of Subject**, and **Scheme** be included in RDA and identified as Core elements. We recommend that **Variant Name of Subject**, **Identifier for the Subject, Scope Note**, **Source Consulted**, and **Status of**

Identification be included in RDA as non-Core elements; these are generally included in other sections of RDA.

Finally, we note that many of the attributes of **Nomen** in FRSAD fall under "data about data". We do not believe that RDA can ultimately avoid dealing with such data, and urges that the JSC take this opportunity to do further work on integrating data about data into the RDA data model.

12. **Access points:** ALA notes that RDA chapters on identifying entities include a section on constructing authorized and variant access points, and that this might be appropriate in the case of the **Subject** entity. There are some faceted subject systems that specify the creation of pre-coordinated subject strings made up of individual facets; in such a case, the whole string would be an access point, and the individual facets would be preferred names. However, this practice is not universal, and it may not be appropriate to build this distinction into RDA. We leave this as a question for the JSC to consider.

13. **Relationships:** The FRSAD relationships, along with their definitions, are listed below, along with our recommendations.

> ***Work to thema*** [FRSAD 5.1]
> *Work* has as subject *thema* / *Thema* is subject of *work*.
> *Notes:* This is the primary subject relationship that corresponds to the RDA placeholder chapter 23. This should be considered for inclusion in RDA. The content of such a chapter might include general principles of subject analysis.

> ***Thema to nomen*** [FRSAD 5.2]
> *Thema* has appellation *nomen* / *Nomen* is appellation of *thema*.
> *Notes:* If Name of Subject is treated as an attribute of the Subject entity in RDA (as suggested above), it would not need to be treated as a relationship.

> ***Thema to thema*** [FRSAD 5.3]
> Includes hierarchical and associative relationships between subject entities.
> *Notes:* The distinction between *thema*-to-*thema* relationships and *nomen*-to-*nomen* relationships is not one that appears clearly in RDA. Most RDA relationships are between entities, although the concept of preferred and variant names might be considered as comparable. It is not clear whether this needs to be incorporated into RDA.

> ***Nomen to nomen*** [FRSAD 5.4]
> Includes equivalence and whole-part relationships between names of subjects. The former would include variant names.
> *Notes:* See previous note.

ALA recommends that the **thema to thema** relationships be covered in a highly general way in RDA. As in other situations, this should be done in a way that does not impose undue constraints on the ability of subject systems to define their own structure and content. We considered suggesting that generic thesaurus-based relationships (broader term, narrower term, related term) be included, but even this may be imposing undue constraints on the ability of individual subject systems to define their own structure and content.