RDA entity boundaries

Gordon Dunsire, RSC Technical Team Liaison Officer, 12 June 2020

Abstract

This paper proposes a development of the guidance and instructions for determining if an RDA metadata description set is referring to an entity being described or to a different entity. The criteria for making a distinction between two entities are essential for the creation of RDA metadata that are well-formed and meet the requirements of local applications.

The paper proposes the revision and relocation of existing Toolkit content, and poses questions on the coverage and granularity of Toolkit content on this topic.

Background

This proposal is a follow-up to the discussion paper on 'Work boundaries' discussed by the RDA Steering Committee at its meeting in October 2019 in Santiago, Chile.[1]

Following informal discussions early in 2020, the scope of the development was widened to include the boundaries of all RDA entities.

The proposal does not include a development of the guidance and instructions for work groups, but does include guidance on the use of work groups to collocate works with boundaries determined by different criteria applied by different agents who create metadata.

Discussion

An agent who creates metadata applies a set of criteria for determining if an entity being described already has a metadata description set or requires a new description. The set of criteria is termed an 'entity boundary'.

The criteria are specified as a difference in the values of one or more elements associated with a single entity.

The criteria fall into two categories:

1. Conformance with the semantic integrity of the type of entity. These criteria are mandatory for well-formed RDA data.
2. Conformance with "bibliographic and cultural conventions" (as noted in the LRM) that are specified for a particular application. These criteria are optional.

An entity boundary may include criteria in both categories.

Previous discussion has indicated the need for a specific guidance chapter on entity boundaries.

Recommendation 1: Remove the sections for describing a distinct entity from the guidance pages for describing a work, expression, and manifestation, and remove the sections for Transformation boundary of a work and a static work from the guidance on Diachronic works. These sections are located at the URLs:

---

[1] Work boundaries. Available at http://www.rda-rsc.org/sites/all/files/Work%20boundaries%20briefing%20paper.pdf

https://beta.rdatoolkit.org/en-US_ala-4d4d3f5b-8d94-3ee5-89d8-241a98366db4/section_ewz_sbp_vdb

https://beta.rdatoolkit.org/en-US_ala-f2747cbc-74d2-3131-a94b-e30effad9d09/section_qlp_33n_zdb

https://beta.rdatoolkit.org/en-US_ala-cf0b18a4-5a55-3358-94b0-2d4fb5449314/section_vrw_j25_xdb

https://beta.rdatoolkit.org/en-US_ala-443c463e-b075-3e6a-9377-eb230c6b8281/p_fvr_njh_vdb

https://beta.rdatoolkit.org/en-US_ala-443c463e-b075-3e6a-9377-eb230c6b8281/p_k3w_yv4_vdb

The guidance on *Transformation boundary of a diachronic work* is general and contextual and is best retained in the guidance for Diachronic works. The specific boundary criteria for diachronic work transformations are included in the instructions for *Boundary of a work* in Appendix 2. Appropriate cross-references will be added to the Toolkit.

Recommendation 2: Add a specific guidance chapter on entity boundaries. The proposed content is given in Appendix 1.

Recommendation 3: Add the term 'entity boundary' to the RDA Terms vocabulary, with the definition "A set of criteria that is applied by an agent who creates metadata to determine if a new entity is being described." Include the alternate term 'boundary' as a cross-reference in the Toolkit Glossary.

For consistency, boundary criteria should be presented in the Toolkit as a set of conditions associated with an instruction to treat the entity being described as a new entity that is different from the entity that is described by existing metadata. The instruction may be mandatory or optional.

Mandatory criteria

Most mandatory criteria are associated with the physical characteristics of an entity and are intrinsic to the type of entity.

The RDA entities that have physical boundaries are *Item*, *Manifestation*, *Person*, *Place*, and *Timespan*.

Physical characteristics may be associated with unique values of space and time. For example, different values of the places and dates of birth and death indicate that a person being described is not the same person that has already been described.

The physical characteristics that distinguish an item arise from modification of the item; the common physical characteristics of exemplars of a manifestation form the description of the manifestation, not the item.

A place is distinguished by its location; a timespan is distinguished by its beginning and ending.

Nomen has a mandatory boundary determined by the nomen string and the entity of which it is an appellation. Any difference in a nomen string or the entity to which it refers requires a description of a new nomen.

Optional criteria

Optional criteria for an entity boundary reflect cultural conventions and the requirements of an application for distinguishing instances of the same entity.

Optional criteria can be included in the boundaries of all RDA entities.

Boundary granularity

Only the broadest element applicable to a boundary criterion needs to be specified. By default, any instruction that applies to an element also applies to the subtypes of the element,

Differences in values that arise from the application of different recording methods or from granularity are ignored when determining a boundary.

Several criteria in the current beta Toolkit are worded in general terms without specifying any elements. An example is the Boundary of a manifestation given in Appendix 2. Other criteria reference one or more separate elements. An example that references multiple elements is included in the Boundary of an expression in Appendix 2.

Question 1: Should all boundary criteria be associated with one or more specific elements? That is, should the general criteria specify the elements to which they apply?

Boundary of an expression in Appendix 2 also illustrates another choice in the presentation of boundaries in the Toolkit. The element Expression: date of expression is presented separately from the long list of elements given in a subsequent condition; this reflects the current separation of boundary criteria.

Presenting boundary criteria as a set of separate conditions has the advantage of finer granularity that provides a flexible approach for policy statements and application profiles to select the specific criteria required. The statement or profile can simply indicate that a condition based on one or two elements is to be applied. If multiple elements are specified in a single optional condition, a policy statement or profile needs to provide further information if only a subset of the elements is used to determine a boundary.

On the other hand, presenting a finer granularity of boundary conditions increases the number of conditional options for the relevant entity.

Question 2: What level of granularity of boundary conditions should the Toolkit offer?

The basic choices are:

a) Use one condition to list all elements that are associated with optional boundary criteria, and leave it to policy statements and application profiles to refine the boundary.
b) Use one condition for each element that is associated with optional boundary criteria, to allow a simple selection by policy statement and application profile.
c) Use a combination of single-element and multiple-element conditions to reflect the presentation of an optional boundary in the current beta Toolkit.

Question 3: Are there any mandatory boundary conditions missing from Appendix 2?

Question 4: Are there any optional boundary conditions missing from Appendix 2?

Recommendation 4: Add a specific section on entity boundaries to the instructions for each entity. The proposed content is given in Appendix 2.

Recommendation 5: Add content to the guidance chapter on Representative expressions to indicate their use in determining work boundaries. The proposed content is given in Appendix 3, with the preceding and succeeding paragraphs.

Questions for discussion and recommendations

Question 1: Should all boundary criteria be associated with one or more specific elements? That is, should the general criteria specify the elements to which they apply?

Question 2: What level of granularity of boundary conditions should the Toolkit offer?

Question 3: Are there any mandatory boundary conditions missing from Appendix 2?

Question 4: Are there any optional boundary conditions missing from Appendix 2?

Recommendation 1: Remove the sections for describing a distinct entity from the guidance pages for describing a work, expression, and manifestation, and remove the section for Transformation boundary of a static work from the guidance on Diachronic works.

Recommendation 2: Add a specific guidance chapter on entity boundaries. The proposed content is given in Appendix 1.

Recommendation 3: Add the term 'entity boundary' to the RDA Terms vocabulary, with the definition "A set of criteria that is applied by an agent who creates metadata to determine if a new entity is being described." Include the alternate term 'boundary' as a cross-reference in the Toolkit Glossary.

Recommendation 4: Add a specific section on entity boundaries to the instructions for each entity. The proposed content is given in Appendix 2.

Recommendation 5: Add content to the guidance chapter on Representative expressions to indicate their use in determining work boundaries. The proposed content is given in Appendix 3, with the preceding and succeeding paragraphs.

Entity boundaries

An RDA *entity boundary* is a set of criteria that is applied by an agent who creates metadata to determine if a description of a new RDA entity is required.

The criteria are specified as a difference in the values of one or more elements associated with a single entity.

The criteria may indicate a new entity of the same type, or of a different type. For example, the criteria applied to an expression may determine that a new work is being described as well as a new expression.

In practice the values of an existing *metadata description set* are compared with the values that characterize the entity being described to determine if the existing metadata describe the same entity or a different entity. If it is determined that a different entity is being described, a new description is required.

Differences in values of an element are not treated as significant when they are the results of applying different recording methods, options within a recording method, or different levels of granularity. For example, the difference in the values "Nineteen-sixty" and "1960-Jan-01" for a related timespan is not sufficient to decide that they describe different entities.

The set of criteria is specific to an entity. Some criteria are required for conformance with well-formed RDA metadata and are determined by the semantic integrity of the type of entity. Some criteria are specified within an application and are determined according to "bibliographic and cultural conventions"(IFLA library reference model).

The semantic integrity of some entities is determined by characteristics that reflect the physical world. A significant difference in the values of one or more of the corresponding elements indicates that a different entity is being described and a new description is required. This applies to:

- Item
- Manifestation
- Person
- Place
- Timespan

The definition of Nomen requires that a nomen is distinguished by its nomen string and the entity of which it is an appellation.

Criteria specified for different applications may vary in the choice of elements, the range of variation in values of each element, and the choice of an existing description set for comparing values.

Criteria may be specified by an *application profile*.

The specification of criteria may be supported by a *vocabulary encoding scheme* for RDA Entity: category of RDA entity or an appropriate *entity subtype*. Each term or concept in a scheme may be used to define the boundary elements and values of the application that uses the scheme. If an application allows only one value to be assigned to each entity, a change in the value of a category becomes the criterion for a new description.

For general guidelines on application profiles, see Guidance: Application profiles.

For general guidance on data conformance, see Guidance: Introduction to RDA. Data conformance.

Boundary of an agent

The relative boundary of the entity is determined by bibliographic and cultural conventions.

Boundary of a collective agent

The relative boundary of the entity is determined by bibliographic and cultural conventions.

Boundary of a corporate body

The relative boundary of the entity is determined by bibliographic and cultural conventions.

Boundary of an expression

The relative boundary of the entity is determined by bibliographic and cultural conventions.

CONDITION

There is a significant difference in the values of elements of an expression from the values of a *representative expression*.

> OPTION
>
> Record a new instance of Expression.

CONDITION

There is a significant difference in a value of Expression: creator agent of expression or an element subtype

> OPTION
>
> Record a new instance of Expression.

CONDITION

There is a significant difference in a value of Expression: date of expression.

> OPTION
>
> Record a new instance of Expression.

CONDITION

There is a significant difference in a value of one or more of

- Expression: aspect ratio
- Expression: aspect ratio designation
- Expression: capture information or an element subtype
- Expression: colour
- Expression: content type
- Expression: designation of version
- Expression: duration
- Expression: extent of expression
- Expression: form of notation or an element subtype
- Expression: intended audience of expression
- Expression: interactivity mode
- Expression: key of expression
- Expression: language of expression
- Expression: prime meridian
- Expression: relief type
- Expression: scale
- Expression: script

> OPTION
>
> Record a new instance of Expression.

CONDITION

An expression is an *aggregating expression.*

There is a significant difference in the values of Expression: <u>aggregates</u>.

OPTION

Record a new instance of <u>Expression</u>.

Boundary of an item

The absolute boundary of the entity is determined by physical characteristics of the entity.

CONDITION

There is a significant difference in a value of one or more of

- Item: modifier agent or an element subtype
- Item: accompanied by item or an element subtype

Record a new instance of Item.

Boundary of an item

Boundary of a manifestation

The absolute boundary of the entity is determined by physical characteristics of the entity.

The characteristics of a manifestation are derived from the characteristics that are common to all items that exemplify the manifestation.

CONDITION

An item that has not been modified has values of characteristics that differ from those of a manifestation.

Record a new instance of Manifestation.

CONDITION

A manifestation has values of relationship, identifying, or carrier characteristics that differ from another manifestation.

> OPTION
>
> Record a new instance of Manifestation.

Boundary of a nomen

The absolute boundary of an instance is determined by its nomen string and the entity of which it is an appellation.

CONDITION

There is a significant difference in a value of one or more of

- Nomen: nomen string
- Nomen: appellation of RDA entity of or an element subtype

Record a new instance of Nomen.

Boundary of a person

The absolute boundary of the entity is determined by physical characteristics of the entity.

CONDITION

There is a significant difference in a value of one or more of

- Person: date of birth
- Person: date of death
- Person: place of birth
- Person: place of death

Record a new instance of Person.

CONDITION

There is a significant difference in a value of Person: field of activity of person.

OPTION

Record a new instance of Person.

CONDITION

There is a significant difference in a value of Person: period of activity of person.

OPTION

Record a new instance of Person.

Boundary of a place

The absolute boundary of the entity is determined by physical characteristics of the entity.

CONDITION

There is a significant difference in a value of Place: location.

Record a new instance of Place.

Boundary of a place

Boundary of a timespan

The absolute boundary of the entity is determined by physical characteristics of the entity.

CONDITION

There is a significant difference in a value of one or more of

- Timespan: beginning
- Timespan: ending

Record a new instance of Timespan.

Boundary of a work

The relative boundary of the entity is determined by bibliographic and cultural conventions.

A *static work* may be realized in two or more expressions that differ in content without changing the intellectual or artistic conception of the work.

An expression of a static work may be embodied in two or more manifestations without changing the content. A manifestation that carries the same content embodies the same expression and work.

A difference in content between two expressions is necessary but insufficient as an indicator that each expression realizes a different work. For example, an expression that is a translation of another expression is not usually described as a realization of a different work. Instead, both expressions are described as realizations of the same work.

A *representative expression* may be used to establish the values that are compared to determine if an expression that is being described is a realization of a new work.

For general guidelines on representative expressions, see Guidance: Representative expressions.

A common appellation may be used to collocate distinct works with common characteristics as a *work group*. This device can be used to harmonize the results of applying different work boundaries by different agents who create the metadata.

CONDITION

There is a significant difference in a value of Work: subject or an element subtype.

>   OPTION

>   Record a new instance of Work.

CONDITION

There is a significant difference in a value of Work: related agent of work or an element subtype.

>   OPTION

>   Record a new instance of Work.

CONDITION

There is a significant difference in the values of elements of an expression from the values of a *representative expression*.

>   OPTION

>   Record a new instance of Work.

CONDITION

There is a significant difference in a value of Work: expression of work for an expression whose content differs from another expression of the work as the result of a change in genre or literary form.

>   OPTION

>   Record a new instance of Work.

OPTION

Relate the new work to the other work using Work: transformation by genre of

CONDITION

There is a significant difference in a value of Work: expression of work for an expression whose content differs from another expression of the work as the result of a change in target audience.

OPTION

Record a new instance of Work.

OPTION

Relate the new work to the other work using Work: transformation by audience of

CONDITION

There is a significant difference in a value of Work: expression of work for an expression whose content differs from another expression of the work as the result of a change in editorial policy.

OPTION

Record a new instance of Work.

OPTION

Relate the new work to the other work using Work: transformation by policy of

CONDITION

There is a significant difference in a value of Work: expression of work for an expression whose content differs from another expression of the work as the result of a change in style.

OPTION

Record a new instance of Work.

OPTION

Relate the new work to the other work using Work: transformation by style of

CONDITION

There is a significant difference in a value of Work: extension plan that is a change from *integrating determinate plan* or *integrating indeterminate plan* to *successive determinate plan* or *successive indeterminate plan*.

OPTION

Relate the new work to the other work using Work: sequential version of

CONDITION

There is a significant difference in a value of Work: extension plan that is a change from *static plan* to *successive determinate plan* or *successive indeterminate plan*.

OPTION

Relate the new work to the other work using Work: serializedl version of

CONDITION

There is a significant difference in a value of Work: extension plan that is a change from *integrating determinate plan* or *integrating indeterminate plan* to *static plan*.

OPTION

Relate the new work to the other work using Work: static version of

CONDITION

A work is an *aggregating work.*

There is a significant difference in the values of Expression: aggregates of an expression from the values of the *aggregating expression*.

OPTION

Record a new instance of Work.

CONDITION

A work is an *diachronic work.*

There is a significant difference in the values of elements of an expression from the values of an expression that realizes a work.

OPTION

Record a new instance of Work.

CONDITION

A work is a *diachronic work.*

There is a significant difference in a value of Manifestation: carrier type for a manifestation that embodies a work.

OPTION

Record a new instance of Work.

OPTION

Relate the new work to the other work using Work: transformation by policy of

CONDITION

A work is a *diachronic work.*

There is a significant difference in a value of one or more characteristics of the plan for embodiment over time.

OPTION

Record a new instance of Work.

OPTION

Relate the new work to the other work using Work: <u>transformation by policy of</u>

CONDITION

A work is a *diachronic work.*

A value of Work: <u>preferred title of work</u> is in a language and script that divides text into words.

There is a significant difference in a value of Work: <u>preferred title of work</u> that is a change in the first five words that follow an initial article, if any.

OPTION

Record a new instance of <u>Work</u>.

OPTION

Relate the new work to the other work using Work: <u>transformation by policy of</u>

CONDITION

A work is a *diachronic work.*

There is a significant difference in a value of Work: <u>preferred title of work</u> that is a change in meaning of the title or an indication of a different subject.

OPTION

Record a new instance of <u>Work</u>.

OPTION

Relate the new work to the other work using Work: <u>transformation by policy of</u>

CONDITION

A work is a *diachronic work.*

There is a significant difference in a value of Work: <u>preferred title of work</u> that is a change in a name of a corporate body.

OPTION

Record a new instance of <u>Work</u>.

OPTION

Relate the new work to the other work using Work: <u>transformation by policy of</u>

Representative expressions

…

These characteristics are useful for describing and distinguishing a work. The values of a representative expression element can be used in the description of a work by copying them from a representative expression to the corresponding work element, although strictly speaking these elements concern expression characteristics rather than work characteristics.

A *representative expression* may be used to establish the values that are compared to determine if an expression that is being described is a realization of a new work.

For instructions on determining the distinction between different works, see Work: <u>Boundary of a work</u>.

Any expression can be used as a representative expression.

…