Thank you for coming to this webinar, where we will discuss FRBR as a foundation for RDA and library databases.
During the 1990s IFLA, the International Federation of Library Associations and Institutions, commissioned a new look at the bibliographic universe. The result was the document *Functional Requirements for Bibliographic Records*, or FRBR (published in 1998).

Text available in print or at:
http://www.ifla.org/en/publications/functiona-requirements-for-bibliographic-records
FRBR was joined by a companion volume, *Functional Requirements for Authority data*, or FRAD, published last year in 2009. FRAD is an expansion of FRBR and adds a number of entities not found in FRBR. There is also an extension of FRBR called *Functional Requirements for Subject Authority Data*, still under development. This presentation will be based on both FRBR and FRAD, but I will refer to the model as a whole as FRBR.

FRBR analyzes the bibliographic universe and divides it into a set of entities, such as persons, corporate bodies, concepts, works, and so forth.
FRBR

- A conceptual model of the bibliographic universe
- Based on the entity-relationship model developed for databases

FRBR is not a cataloging code. It is a conceptual model of the bibliographic universe based on a database modeling technique called ...
... “entity-relationship,” first introduced in the 1970s. This model is widely used in database design, but until recently hasn’t been used extensively in library databases. In this model a specific database universe is defined, and this universe is divided into specific entities linked by specific relationships.

An entity is something that can be distinctly identified within the context of the database. For example, a business database might define as entities “customers,” “employees,” “managers,” “stores,” “suppliers,” etc. A genealogical database might define as entities “persons,” “places,” “events.”

A relationship is an association between two or more entities. A business database might define a relationship between a particular store and an employee. A genealogical database might define a “father-child” relationship between a male person and his children.

In the model, entities and relationships are defined by attributes. An attribute is a characteristic that may identify instances of entities or relationships. For example, one of the attributes of a person is his or her birth date; other possible attributes for a person might be where he lives, his profession, his marital status, and so forth. Entity-relationship databases are designed with the entities, relationships, and attributes needed for the purpose of the database. A personnel database might need to define lots of attributes and relationships for persons (e.g., SSN, sex, marital status, position in the company, salary, etc.). A bibliographic database would not define all possible attributes and relationships for “person”, just those needed for the purposes of the database, such as name, possibly birth and death dates, relationship to works he/she created, etc. RDA, based on FRBR, defines entities, relationships, and attributes. Most of our cataloging under RDA will consist of describing the attributes of the different FRBR entities.
There are many different conventions of diagramming the entity-relationship model. One of the most basic methods is to use a rectangle for an entity, a diamond for a relationship, and an oval for an attribute. I have found this model to be the most convenient for describing an actual database, and so I will be using this diagramming technique during most of this presentation. FRBR’s diagramming model is a little different; I will show FRBR diagramming in a moment.
In this diagramming method entities, relationships, and attributes are linked by lines. The simplest version uses single lines, as shown here. To illustrate more complex situations other types of lines may be used, including lines with arrows at either end to show whether the relationship is reciprocal or not. In this presentation most of the diagrams will have the simple lines shown here. In the basic model both entities and relationships can have attributes, but in FRBR attributes have only been defined for entities, so no attributes for relationships will be shown in this presentation.
FRBR has two diagramming techniques, one for entity-relationship sets (i.e., the abstract model), and another for specific instances of entity-relationship. This slide illustrates the entity-relationship set diagramming technique. Entities are shown in rectangles as in the classic model, but relationships are simply shown by words next to the lines. In this illustration, “work”, “expression”, “manifestation”, and “item” are entities; “is realized through”, “is embodied in” and “is exemplified by” are relationships. Single arrows mean that only one instance of an entity can occur in the relationship; double arrows mean that more than one instance can occur. For example, look at the relationship between item and manifestation. In the FRBR model, a manifestation can be related to more than one item, but an item can be related to only one manifestation.
FRBR Diagramming

- **cb₁** Kelmscott Press
  - is the producer of →
  - ← has a producer
    - **m₁** the 1891 publication of *Poems by the Way* by William Morris
    - **m₂** the 1892 publication of *The Recuyell of the Historyes of Troye* by Raoul Lefevre.
    - **m₃** the 1896 publication of *The Works of Geoffrey Chaucer*

This slide illustrates FRBR’s diagramming technique when it wants to show specific instances of an entity. In this illustration, the corporate body entity “Kelmscott Press” has a specific relationship (producer) to three manifestation entities, *Poems by the way*, *The Recuyell of the Historyes of Troye*, and *The works of Geoffrey Chaucer*.

I have illustrated the two FRBR diagramming techniques here in order to help you understand FRBR itself when you have a look at it. However, as mentioned, I find the traditional Entity-Relationship diagramming technique to be a bit clearer in a graphic presentation, and so I will be using it in the rest of this presentation.
To create a good entity-relationship database, careful planning is required. When you begin to design an entity-relationship database, one of the first things you need to do is define the entities and relationships. You need to define every entity and every relationship that is important to your particular database. On the other hand, one type of entity should not overlap with another. This careful planning process is exactly what the authors of FRBR, FRAD, and FRSAD have done, and it has taken years. The authors of FRBR thought about what types of things should be defined as entities in our bibliographic universe, and this is the result.

The entities in the FRBR model are divided into three groups. The first is defined as “the products of intellectual or artistic endeavor,” and the entities in this group are listed on this slide.

Work: “a distinct intellectual or artistic creation”
Expression: “intellectual or artistic realization of a work in the form of alpha-numeric, musical, or choreographic notation, sound, image,” etc.
Manifestation: “the physical embodiment of an expression of a work”
Item: “a single instance of a manifestation”—in other words, a copy
Here are some specific examples of these abstract entities.

All of these entities have attributes. Attributes are characteristics that would be necessary to describe each entity, to distinguish it from other entities of the same type.

**Work**: (what distinguishes it from other works?) title ("Gone with the wind"), form (novel), date of composition (pre-1936), etc. (music attributes include key, medium of performance)

**Expression**: (what distinguishes it from other expressions?) form (not literary form—physical form, i.e., text on paper, cassette, compact disc, electronic); date (in the case of the German translation, the date it was composed); language (1st expression = English; there may be other English expressions; language of the German translation is German, etc.); extent (e.g., the number of words, the duration of a recording, etc.)

**Manifestation**: (what distinguishes it from other manifestations?) title (i.e. exactly what is printed on the title page), statement of resp. (ditto), edition/issue designation, place of publication, publisher, date of publication, etc.

**Item**: (what distinguishes it from other items?) item identifier (e.g. barcode, possibly call number); provenance (who has owned the item?); marks/inscriptions; condition (missing its cover, etc.); access restrictions; location of the item (where is it vs. others?)

**NOTE**: Much of RDA consists of instructions for naming and describing the attributes of the entities. It looks to a situation where we might have a separate record or description for each entity, and so tells us what we should include in the descriptions of those entities—i.e., we’ll need to enter the attributes of the entities into our records.
This is how the primary entities are related to each other (read from slide)
To get more concrete, here are some specific instances of the FRBR work entity with various types of relationships. In this slide the novel is shown to have a relationship to two other works: a derivative relationship with the movie, and a descriptive relationship with the work *Vanity Fair* and *Gone with the wind: a critical comparison*. In a real database the work *Gone with the wind* (the novel) would have relationships with many other works; and it would also potentially have relationships with instances of all the other types of FRBR entities. The database becomes like a web, with dozens, hundreds, or even thousands of relationship links between any one entity and other entities. In this slide you can see further relationships between the movie and a work that describes it, and the work *Vanity Fair* and the work that also describes *Gone with the wind*. In a FRBR/RDA based database each of these work entities would have separate descriptions that the user could examine.
The second group of entities are those responsible for creating intellectual or artistic content in Group 1 entities. FRBR/FRAD have defined three: Person, Family, Corporate body. Note “family” is new to descriptive cataloging rules, and was added to the FRBR model through FRAD. RDA incorporates guidelines for describing instances of the family entity. This should be helpful to us as we move forward.

The three entities here are defined much as we would expect:
- Person: “an individual or persona established or adopted by an individual or group”
- Family: “Two or more persons related by birth, marriage, adoption, or similar legal status, or who otherwise present themselves as a family.”
- Corporate body: “an organization or group of individuals and/or organizations acting as a unit”
Here are some concrete instances of these abstract entities.

Attributes have been defined for each of these entities.

For person, these include: name, dates, title, other designation, gender, place of birth, place of residence, language of person, field of activity. For example, Margaret Mitchell has a name, she has a gender, dates, and so on, as do Claude Debussy and George W. Bush. The fact that the attributes of each of these are different is what distinguishes them from one another.

For corporate body, attributes include: name, number (e.g. for meetings), place associated with the corporate body, date associated with the corporate body, type of corporate body, language of the corporate body, its field of activity.

For family, attributes include: name, type of family (e.g. clan, dynasty, etc.), dates of family, places associated with family, history of family.

The RDA chapters dealing with these entities tell us how to record the attributes of the entities. Again, RDA is looking toward a database structure where we would have a separate entity record or description for each person, corporate body, or family, and within that entity record we would describe the entity, or in other words, record its attributes. We would create such a record only once for each entity rather than our current practice in MARC of repeating much information every time we create a new bibliographic record. Entity records would be linked to other FRBR entity records, e.g., the entity record for Margaret Mitchell would be linked to the work record for *Gone with the wind* via a “creator” relationship link. Any entity record can be linked to any other entity record as appropriate.
Here we have various relationships between entities. Margaret Mitchell has a creator relationship with Gone with the Wind. The work has a “realized through” relationship with the two expressions; and notice the expression entities also have a relationship to each other. Finally, the German expression has a “translated by” relationship with another person entity, Martin Beheim-Schwarzbach.
Group 3 entities are entities that can be subjects of works, expressions, manifestations, or items. Any of the entities can be the subject of a work—for example a person entity, from group 2, could be the subject of a biography.

The “new” entities in Group 3 are:

- Concept: “an abstract notion or idea”
- Object: “a material thing”
- Event: “an action or occurrence”
- Place: “a location”
Here are some concrete examples of Group 3 entities. All the examples are LCSH or currently formed name access points.

Attributes of concept include: term of concept, type of concept
Attributes of object include: type of object, date of production, physical medium, place of production, etc.
Attributes of event include: Date associated with event, place associated with event
Attributes of place include: term for the place

The RDA guidelines for recording attributes of concept and objects are not yet written and do not appear with the first publication. Guidelines for recording attributes of events and jurisdictional places are in the current publication.
Here are a few group three relationships to the work Gone with the wind. In order to keep on familiar ground I’ve used LCSH style subject strings. However, note that with current LCSH practice there is a mixture of FRBR entities when a subject string is subdivided. For example, Georgia is a FRBR place entity, History and fiction are concept entities, and Civil war is an event entity. When combined as a string the whole thing would probably be regarded as a concept. If the strings were broken up (faceted) it would be possible to be more precise about the entity. This needs thinking out. Another point to note: in RDA fictitious characters are grouped within the person entity, not the concept entity, so Scarlett O’Hara is labeled a “person” in this example. [Pause to look at the diagram.]
FRBR/RDA Attributes

- FRBR, a *model*, defines attributes, but does not tell us how to record the data
- RDA, a *cataloging code*, defines attributes and does tell us how to record the data

FRBR defines a set of attributes for each entity in the model. Because it is not a cataloging code, FRBR does not define how the information is to be recorded. For example, “name of person” is one of the attributes of the “person” entity in FRBR. FRBR defines this attribute: “The name of a *person* is the word, character, or group of words and/or characters by which the *person* is known”, and points out that a person may be known by more than one name, and that libraries normally select one of the names as a uniform heading. But it does not tell us how to form the data to be recorded in this element, and if we are one of the libraries that wants to select one as a uniform heading, it does not tell us how to make that choice. That is the province of a cataloging code, such as RDA. RDA also defines attributes for entities, but because it is a cataloging code it also informs us how to record the data and in the case of the “name of person” attribute, it tells us how to choose between competing forms.
For this webinar I will take one FRBR entity, “person”, and demonstrate how this works out in RDA.

FRBR and FRAD define several attributes of “person”. These are mostly taken directly into RDA, where they are worked out more fully, and in some cases refined with subelements. In RDA the act of recording attributes of an entity is referred to as “identifying” so RDA chapter 9 is called “Identifying Persons”. Chapter 9 is within a larger RDA section called “Recording attributes.” Within each “entity” chapter the guidelines begin with a scope note and general guidelines defining the entity, as you can see here. The central portion of each chapter consists of subsections detailing each attribute of the entity, including a definition of the attribute (also referred to as an “element” in RDA) and guidelines for recording the information. Finally, at the very end of each chapter, is a subsection detailing how to construct an access point to represent the entity. I will not draw these notes here. It is extremely important that you understand this structure or you will get confused. The central subsections of the chapter do not necessarily have anything to do with the access point. They are intended to give guidance for recording attributes, not constructing an access point.
How might this work in the real world? “Margaret Mitchell” is an instance of the “person” entity. Many of the RDA-defined attributes apply to her. How would we work RDA out in a FRBR-based database?

The first RDA attribute is “name of person”. RDA works this out into two subelements, “preferred name” and “variant name.” RDA instructs us, for all of these forms, to invert; and it instructs us to choose the commonly known form as the preferred name. In this demonstration I will record the preferred name and one variant name. The preferred name is based on usage. Variant names can come from any source. Regarding Mrs. John Robert Marsh, in RDA, as in AACR2, if a married person is identified only by a partner’s name, the term of address is considered an integral part of the name, and hence is recorded as part of the name attribute, as here.

In RDA a fuller form of name is not required to “fill out” elements already found in a preferred or variant name. Because Mitchell had a middle name, the fuller form of her forename is “Margaret Munnerlyn”.

“Date associated with the person” is an element which RDA has subdivided into subelements: date of birth, date of death, and period of activity. Note the RDA element only calls for the year to be recorded in most cases. I am giving the format RDA prescribes if two persons with the same name are born or died in the same year. We will see that this data is recorded slightly differently in MARC.

The attribute “gender” in RDA can be recorded either “male” “female”, or “unknown”. If none of these terms is appropriate, another may be provided by the cataloger.

Place of birth and language of the person are other possible attributes to record. RDA prescribes the forms shown.

There are many other elements and subelements called for in RDA. Note that the only a few are core or required. Preferred name and dates are core. Remember this doesn’t mean the dates necessarily have to be part of the access point. It just means they need to be recorded as an element of the RDA record if known. There are a few other elements that are core if needed to distinguish one person from another.
These same RDA elements translate into MARC in this way.

There is no discrete field in current MARC for recording the preferred name. The only place to record it is as a part of the authorized access point in 100. Subfield $a and $c (not present in this example) contains the preferred name prescribed by RDA. Note “1900-1949” is not part of the preferred name, but a cataloger addition to the access point.

Similarly there is no discrete place to record the variant name. Until MARC expands to allow this, it has to be recorded as part of a variant access point, in $a and $c of a 4XX field.

There is also no discrete place to record the fuller form of name element. Until MARC expands to include this element, it is recorded in $q of the access point fields, 1XX, 4XX or 5XX.

The date subelements are recorded in field 046; subfield f is date of birth, subfield g is date of death. Note the MARC formatting is different from the formatting called for in RDA--MARC calls for YYYYMMDD.

Place elements associated with a person are recorded in 370; $a is the place of birth.

375 contains the gender element. It is recorded in MARC exactly as RDA calls for it.

The language of the person element is recorded in 377. In current MARC practice, languages are recorded as the MARC language code in subfield $a, which is repeatable if the person used more than one language.

To complete this record we would include notes for the sources we consulted to find the information. RDA instructions for this are in chapter 8; for the most part they would be recorded in 670 fields, just as NACO practice has been until now.
Now let’s create a record for a work. The attributes, or elements, of the work entity are found in RDA chapter 6, “Identifying works and expressions.” Elements or attributes of work include title, form, date of the work (date the work was created), other distinguishing characteristic, and special guidelines for particular kinds of works like musical works. There are more elements and subelements defining work in RDA than there were for person, so I can’t fit them all on the screen, but these are enough to get us started. Let’s create a record for *Gone with the wind* using some of these elements.

One thing to note before we start: “Author” or “Creator” is *not* an attribute of work in RDA or FRBR. The author is a person, family, or corporate body, and as such is a separate entity with a relationship to the work. In a pure RDA work record the author’s name would not appear in the record, but the record instead would be linked to the record for the author. This is a dramatic difference from current MARC practice.
“The title of the work” element in RDA is subdivided into subelements “preferred title for the work” and “variant title for the work.” The preferred title is chosen as you would expect: choose the title by which the work is commonly known. I am not aware of any variant titles for this particular work, so this element isn’t recorded.

Gone with the wind does have a form. It is a novel.

We are instructed in RDA to record the date of the work, defined as the earliest date associated with the work. If we don’t know the complete history of the work, which we usually won’t, we can record the date the first expression of the work; this date in turn may be the first time the expression was published in a manifestation. In this case that would be 1936.

History of the work is another possible element, as is summarization of the content (RDA 7.10)
As it currently exists MARC has a hard time accommodating a pure RDA work record because the preferred and variant title elements do not have discrete MARC fields, just as we noted that preferred and variant personal name do not, and therefore are only recorded as part of the authorized access point in 1XX or 4XX fields. In this hypothetical example I am going to code the preferred title in 190. Some of the other elements are defined only for the bibliographic format (e.g. summary = 520). I am going to make up a 320 field for summary in this hypothetical example. As currently designed MARC doesn’t work well, but with some tweaking it could work.

As with the person entity, associated dates are recorded in 046. $k is date created.

Form of work is recorded in 380.

The history of the work element is not mapped in the RDA mappings, but it seems to me it fits within the definition of the MARC 678 authorities format.

The point of all this is that it might be possible to reshape MARC to allow it to contain RDA data and also reside in an entity-relationship database. Please note that all these elements describe the work. If we were cataloging in an entity-relationship environment we would only have to record this information one time. For example, we currently record the summary in each and every bibliographic record for this work, at least if we want to be consistent. In an entity-relationship environment we would only need to record it once, in the work record.

Work Entity Attributes (MARC--Hypothetical)

| 046  | $k 1936 |
| 190  | $a Gone with the wind [hypothetical] |
| 320  | $a Scarlett O’Hara, the spoiled daughter of a well-to-do plantation owner, must overcome the challenges facing the South during and after the Civil War [hypothetical] |
| 380  | $a Novel |
| 678  | $a Romantic novel first published in May 1936; it won the Pulitzer Prize in 1937. |
This is what the work record for *Gone with the wind* might look like under the current MARC field capabilities and the policies during the RDA national test.

Preferred title is recorded in $t$ of the 1XX field as part of the authorized access point because there is no discrete place to record the element.

There is no place to record the summary in the MARC authority format, so for the moment it will continue to be recorded redundantly in every bibliographic record for this work.

Form is recorded in 380.

History of the work presumably could be recorded in 678. I don't believe 678 has yet appeared in any work record, but this is how it might look.
Here's how these MARC records might fit into an ER database based on FRBR. The work entity description is linked to the person entity description by a “created by” relationship.

PAUSE HERE TO LET PEOPLE LOOK.
FRBR/FRAD define tasks that the user wants to accomplish when he or she approaches the library’s database. For FRBR these are:

1. The user needs to find materials relevant to his or her needs. A user begins this process by doing a search in the database.
2. Once a search is executed, users need to identify the resource, that is, they must confirm that the resource corresponds to what they were looking for.
3. If more than one resource corresponds to the search, they then need to select the resource most appropriate to their needs.
4. Finally, once the user has selected a resource he or she wants to obtain it.

FRAD adds the following tasks:

1. Contextualize. This means the user needs be able to place the entity he or she is seeing into context. This was apparently meant to apply mainly to creators and users of authority data, but it seems a useful task to keep in mind for all our users.
2. Justify means to document one’s reasons for choosing a name or term on which a controlled access point is based. This task is probably undertaken only by creators of authority data.

So we can partially judge FRBR and RDA by how well they promote these user tasks. Hopefully, by moving to an entity-relationship database structure we will be making it easier for our users (which include ourselves) to do this. Let’s see how an ER database might behave.
Let's suppose a user approaches the database looking for the name Margaret Mitchell. Since there are more than one, he or she might be given a list of names to choose from. There might be better ways to disambiguate these persons' names, but at the moment we're using dates, so let's go with them. The user chooses Mitchell, Margaret (Margaret Munnerlyn), 1900-1949.
The user might now find his/her way to the person record for Margaret Mitchell, which could be opened to reveal the details about Margaret Mitchell (the attributes we recorded in an earlier slide). In this way the user is identifying the person, one of the FRBR user tasks, confirming that this particular Margaret Mitchell is the right one. The display might also show related entities, as here. Relationships with other entities is another very good way to identify, perhaps better than opening the record and reading about Margaret Mitchell. The user would likely say, “Yes, the Margaret Mitchell who wrote Gone with the wind is the one I want.” Seeing this display the user might say “love story behind Gone with the wind? That looks interesting.” Clicking on that work record might lead to a display such as this:
From here the user could *contextualize* (the author of the work is Marianne Walker; it has other subjects aside from Margaret Mitchell), and choose between English, Russian or Japanese expressions. Or he or she could decide this isn’t really what was wanted, and go back to Margaret Mitchell.
The user decides he wants to read Gone with the wind.
Note that all entities in the database that are related to the work Gone with the wind are clearly shown. This includes related persons, works, subjects, and expressions. This contrasts with the current environment where the results of a search usually appear in a jumbled display with little showing how results are related to each other. Here, the user can choose any of the entities and follow the web to other related entities. In this case let's assume the user chooses an audio expression. Note in this simplified diagram there is only one audio expression; in reality there would be a separate expression for each performance. Also in this and following slides for the sake of space I've included attributes and relationships of certain entities right in the box, which differs from normal diagramming.
This expression is the one read by Linda Stephens. It exists in at least three manifestations. Reviewing the FRBR user tasks, the user has found and identified a resource; now he must select the one that meets his needs. Looking at the choices, he realizes he has neither a cassette player nor a CD player; so he selects the Playaway version.
The user discovers that a copy is conveniently available; he goes to the Media Center and obtains the Playaway version of this *Gone with the wind* audiobook.

There are surely better ways to display this information, but the point is to show how a user might navigate through an entity-relationship database.
Point 1: good for the cataloger

Point 2: good for the user

Point 3: Every piece of the FRBR study was based on perceived user needs to find, identify, select, and obtain library materials. RDA, following FRBR, also claims to be designed with the user in mind. It will be up to us in the coming months and years as we become familiar with RDA and begin to use it to decide whether it lives up to its promise; and where it doesn’t, to contribute to improving it.

FRBR/Entity-Relationship structure

- Why is an ER structure useful?
  - Only need to record attributes for an entity once rather than multiple times as in current MARC environment
  - Ease of navigation (“FRBRization”)
  - FRBR designed with user in mind
FRBR as a Foundation for RDA

Thank you!

Robert L. Maxwell
robert_maxwell@byu.edu