What Does Giving Primacy to a Certain Entity Cause in a Conceptual Model for Cataloging?

Expression-Entity Dominant Model Revisited

Shoichi Taniguchi

Which entity is given primacy in a conceptual model for cataloging is an important issue in metadata interoperability. This study investigates the implications and consequences of giving primacy to different entities among models and the merit of the expression-entity dominant model. FRBR and four other models derived from FRBR that give primacy to different entities are examined. Several modeling issues, such as optionality or necessity of establishing entity instances, cardinality between entities, and treatment of titles and statements of responsibility that appear in a resource, are examined for each model and the results are compared.

The International Federation of Library Associations and Institutions (IFLA) Study Group on the Functional Requirements for Bibliographic Records developed a conceptual model for the bibliographic universe to be dealt with in cataloging.[1] This model—referred to as "FRBR model" here—was constructed with the entity-relationship modeling technique. Various other models for the entire bibliographic universe, or for a limited scope such as for musical resources, have also been proposed. FRBR and other models consist of multiple entities to represent a bibliographic resource in terms of entity-relationship modeling, or multiple classes in terms of the Resource Description Framework (RDF), a standard model for data interchange on the web. FRBR defines ten entities that include a group of four bibliographic entities, work, expression, manifestation, and item, to represent a bibliographic resource. These four entities all seem to be necessary to describe a resource from a theoretical viewpoint, but actually the work entity and/or the expression are not mandatory (i.e., can be omitted) in some cases when the implementation of the model is considered, whereas the manifestation entity is always required. The item entity is also omitted when item-specific information is not needed. This is a logical consequence inferred from the model.

However, few models declare, or even address, which entity (or class) is to be given primacy among bibliographic ones, while an individual model implicitly (and thus substantially) gives primacy to a certain entity. If dominant entities are

Shoichi Taniguchi (taniguchi@z2.keio .jp) is a Professor in the School of Library and Information Science, Keio Gijuku Daigaku, Tokyo, Japan.

different among models consisting of the same set of entities, the optionality or necessity of establishing an entity instance, assignment of some attributes to an entity, etc., will be different among those models, and finally, different metadata for the same resource will be created in accordance with the models. Therefore, whichever entity is given primacy within a model is an important issue for metadata interoperability.

Taniguchi recognized this point and introduced a viewpoint regarding which entity is given primacy among bibliographic entities in a model.[2] He outlined a model giving primacy to expression-level entity, i.e., an expression-dominant model, by indicating differences from the FRBR model, which deals with the manifestation as being dominant. The expression entity is defined in FRBR as "the intellectual or artistic realization of a work in the form of alpha-numeric, musical, or choreographic notation, sound, image, object, movement, etc., or any combination of such forms," while the manifestation is defined as "the physical embodiment of an expression of a work."[3]

The expression-dominant model intends: (1) to shift to a more content-oriented model from one based on a resource's physical features (i.e., manifestation-dominant model), and (2) to organize bibliographic resources primarily at the expression level, rather than at the work level. Both the expression and work entities bear the content aspect of a resource, but the expression is more stably grasped and identified than the work. The expression has "the form of alpha-numeric, musical, or choreographic notation, sound, image, object, movement, etc.," which can be objectively observed, and usually has a clue such as a title, statement of responsibility, etc. that identifies it or notifies a change to it.[4]

Research on the expression-dominant model is limited to Taniguchi's studies, and related research and projects have appeared thereafter, which will be described later.[5] In this paper, FRBR and four other models derived from FRBR that give primacy to different entities in the models are examined, such as the expression-dominant model, work-centric model, etc. For each model, several modeling issues, such as optionality or necessity of establishing entity instances, cardinality between entities, and treatment of titles and statements of responsibility that appear in a bibliographic resource, are examined and the results are compared between the models. The resultant differences among the models led to (1) the implications and consequences of giving primacy to different entities and (2) the merit of the expression-dominant model.

Research and Projects on the Expression-Dominant Model

The issue of which entity should be given primacy among bibliographic entities in a model has not previously been

examined except in studies by Taniguchi. No studies have attempted to examine the expression-dominant model (or its equivalent) as another choice of conceptual metadata model for the bibliographic universe. However, if the scope of exploration is extended beyond the conceptual models, some related research and projects on the expression-dominant model can be found.

FaBiO, the FRBR-aligned Bibliographic Ontology, imports the FRBR bibliographic entities as main "classes" in RDF vocabulary, and adds "properties" between them (i.e., relationships in the entity-relationship modeling), such as "hasManifestation" and "isManifestationOf" between work and manifestation, "hasPortrayal" and "isPortrayalOf" between work and item, etc., which are not defined in FRBR.[6] While FaBiO uses these FRBR classes, it places emphasis on the expression class by associating with the expression all the content description "properties" (i.e., attributes in the entity-relationship modeling) such as title of journal article, publication year, etc. FaBiO assigns only properties related to physical carrier and format to the manifestation class. It is a kind of expression-dominant model, although it does not address that modeling issue.

Another example is the Dublin Core Application Profile for Scholarly Works.[7] This application profile is based on FRBR; it defines the entities scholarlyWork (renamed from "work" in FRBR), expression, manifestation, and copy (renamed from "item"). However, it clearly shifts the focus to the expression entity. Title, description, identifier, date available, etc. are all associated with the expression, while only format, date modified, and publisher are associated with the manifestation. Currently, further studies are underway to represent complex real-world situations related to scholarly publications under the Common European Research Information Format (CERIF) development.[8]

Additionally, two studies conducted by Pisanski and Žumer revealed that users hold different views on the bibliographic universe, but generally have FRBR-like views.[9] Their studies also revealed that users generally seek bibliographic resources at the expression (not work) level or at the manifestation level, depending on their needs at the time, which coincides partly with the benefit of the expression-dominant model.

Google Scholar can be considered from a similar viewpoint. Search results in Google Scholar provide the title of a paper or report as well as a number indicating how many "versions" of the paper or report are available on the web. This number is linked to a list of the versions available for a paper or report. Google Scholar seems to try to collocate papers and reports at the expression-level while ignoring differences in file locations and formats, but it does not create detailed metadata for such resources. Web-scale discovery services implemented in libraries conduct a similar collocation to combine both print and digital editions of a resource.

Coyle argues that the expression-dominant model is an appropriate approach to organize resources in federated search systems that combine physical and digital versions of the same content resources.[10]

Models Giving Primacy to Different Entities

The FRBR model consists of the four entities to represent bibliographic resources: work, expression, manifestation, and item. The entity "item," "a single exemplar of a manifestation," is not considered in the current discussion.[11] Instances of the entity "item" are required for every resource to record location, condition, and/or other administrative data. However, the entity has no relation to the issue of which entity is given primacy in the model, except in cases where resources are unique, such as rare books and incunabula.

The following models are derived from FRBR by changing the entity to be given primacy:

- Model 1: Expression-dominant model, which was originally proposed by Taniguchi while referring to the FRBR's four bibliographic entities model.
- Model 2: Manifestation-dominant model, which is FRBR itself.
- Model 3: Work-centric model, which gives primacy to the work entity within the FRBR model's structure.
- Model 4: Model consisting of the two entities—the work entity and the combined expression-and-manifestation entity, where the latter entity is given primacy. It is a model blended from Models 1 and 2.
- Model 5: Model consisting of the two entities—the combined work-and-expression entity and the manifestation entity, where the former is dominant. It is blended from Models 1 and 3.

Models 3 to 5 were devised for this study while referring to the FRBR model. Model 3 was derived from FRBR by simply changing the dominant entity, whereas Models 4 and 5 were composed through the combination of multiple entities with given primacy. Models 1 to 5 will be examined in terms of several modeling issues to identify differences from each other. Those modeling issues are chosen as checkpoints that would reveal differences among the models.

Various other models can be found which reference FRBR or have similar multi-entity structures, such as BIB-FRAME and the "indecs" model. Although there seems to be overlap between the entities adopted by those models and the FRBR entities, slight (but significant in some cases) differences in entities' definitions seem to exist even if the same entity name is used. The BIBFRAME model, which is proposed in the Library of Congress' Bibliographic Framework Initiative, adopts the RDF class "work," whereas its

definition is different from FRBR's work, which will be discussed later.[12] Another example is the entity "expression" defined in the "indecs" metadata model, which is proposed primarily for e-commerce of content (intellectual property) in a network environment.[13] It is therefore complicated to analyze those models themselves in terms of which entity is dominant and to compare the resultant differences among the models. Instead, it is better to derive all possible models from FRBR as a base model and thus analyze those derived models from the same set of checkpoints. The resultant differences among the models lead to the implications and consequences of giving primacy to different entities. The draft FRBR-Library Reference Model (LRM), a consolidation of the FRBR, FRAD, and FRSAD conceptual models, adopts the four bibliographic entities—work, expression, manifestation, and item, whose basic structure is kept unchanged from FRBR.[14] The examination conducted in this study as it is will be applied to FRBR-LRM.

Incidentally, it might be theoretically possible to give primacy to all entities constituting a model, meaning that it is possible to deal with all entities equally. However, an individual model implicitly (and thus substantially) gives primacy to a certain entity. FRBR seems to not give primacy to any entities. The FRBR model as it is, however, substantially gives primacy to the manifestation entity, which will be examined later. The model is neither expression-dominant nor work-centric.

Model 1: Expression-Dominant Model

The purpose of giving primacy to the expression entity is to differentiate the content of a bibliographic resource from its physical carrier or format and to organize such resources at the expression level. The expression-dominant model proposed earlier is an example of this. Figure 1 shows the model at the instance level: one work instance, two expressions, and three manifestations, in addition to two instances of person, family, or corporate body (hereafter, PFC), and relationships between the instances. The word "instance" is used throughout this paper to distinguish an instance of an entity in a resource model from an entity type or class itself. Some principal attributes are also shown for the bibliographic entities.

a-1) Definitions of bibliographic entities, and the unit of establishing entity instances: The definitions of the entities are the same as those in FRBR. In comparison, there can be more than one criterion for the unit of establishing an expression instance within the expression entity (namely more than one criterion for determining the boundaries between one expression instance and another). The most granular one should be adopted while ignoring trivial variations. The latest amended version of FRBR states that "Minor changes, such as corrections of spelling and punctuation, etc., may be considered as variations within the

same expression."[15] Accordingly, an expression instance should be established, for example, at the level of the Japanese translation of Shakespeare's Hamlet by a person X, or that by a person X in year YYYY.

a-2) Optionality or necessity of creating bibliographic entity instances: Expression instance(s) are created for every resource being described; expression(s) are added to the model that represents a particular individual resource. This is a logical consequence deduced from the premise that the expression entity is chosen to be given primacy.[16] It can be represented by the minimum cardinality of the relationships between expression and other bibliographic entities, i.e. work and manifestation. If creating an expression instance for a resource is mandatory in the resource model creation, the minimum cardinality is 1 (not zero) on the expression side of the relationships between expression and other entities.

Figure 1. Model 1: Expression-Dominant Model

From the above, the manifestation is a kind of "weak" entity in this case. A "weak" entity is one that cannot be uniquely identified by its attributes alone and thus its existence is dependent on another entity, that is, the expression, which can exist without a work instance. Manifestation instances are depicted with double-lined rectangles in figure 1, which indicate that the entity is "weak."

Regarding creating an instance of the entity below the expression, i.e., a manifestation instance, there could be two possible interpretations. One is that a manifestation instance is required to represent a resource's physical aspects. The other is that the manifestation instance can be omitted in cases where no physical information on a resource is provided. This implies that only expression instance(s) are created for a resource. On the contrary, from a theoretical viewpoint, it is not necessary to create work instance(s) since the work entity is not dominant here and expression instance(s) can exist in themselves. In a practical situation, however, it is permitted to adopt a policy to create work instance(s) for every resource, if necessary. Additionally, developing a work instance is usually necessary to draw "subject" relationships to other entities such as concept, object, etc. defined in FRBR when it is suitable to represent the subject dealt with in a resource. If a work instance is not developed, associating an expression instance with such entities for subject
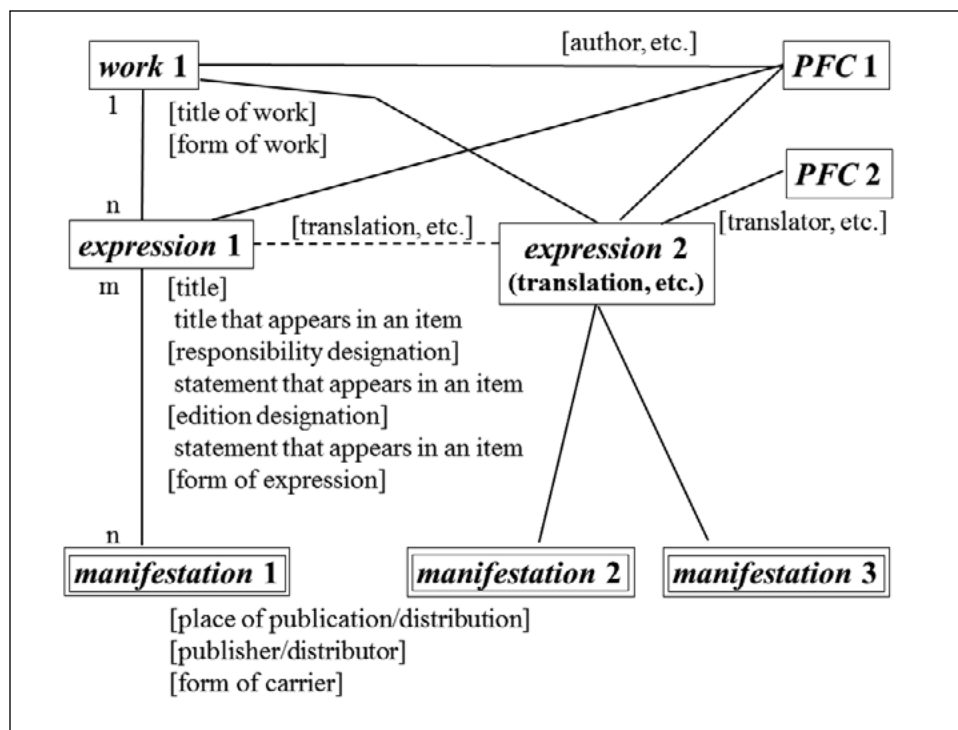
representation, instead, could be adopted as an expediency. Drawing "subject" relationships will not be considered further in this paper.

a-3) Cardinalities between bibliographic entities: According to FRBR, the maximum cardinality of the relationship between work and expression is one-to-many; each work has one or more expressions that realize it and each expression realizes just one work. Contrary to this, that relationship's cardinality could be changed to many-to-many, which means that one or more work instances can be developed for a single expression instance in a resource model when necessary. It is valid if we adopt an interpretation that creating work instances depends on cataloging codes and various cultures or national groups—FRBR itself points out this issue—and consequently different works could be recognized for a single expression depending on such codes or others.[17] The cardinality of the relationship between expression and manifestation, in contrast, is many-to-many in this model, which is the same as FRBR.

b) Relationships to PFCs: PFCs have relationships with bibliographic entities to represent "responsibility" relationships, such as "is created by," "is realized by," etc. A creator, for example, an author of a textual work or a composer of a musical work, is linked to the work and expression instances that are created and realized by the creator. A reviser, translator, etc. who revises or translates an expression, or a performer of a musical work, is associated only with

expression instances. Figure 1 shows that the entity instance PFC 1 is linked to work 1 and expressions 1 and 2, while PFC 2 is linked only to the expression 2. If developing work instance(s) is optional and thus can be omitted, relationships between PFC and the expression are required to be represented in the resource model.

c) Treatment of titles and statements of responsibility that appear in a resource: Titles and statements of responsibility that appear in a resource should be associated with the expression entity in the expression-dominant model. This was noted earlier.[18] It implies that such titles and statements of responsibility can be handled as the attribute values of the title and responsibility designation of an expression instance without any problem. Such titles and others in a resource are reasonably abstracted to those at the expression level. They are used as external clues to the identity of expressions; the same title and statement of responsibility indicate the sameness of texts, images, or sounds, even with trivial variations, such as corrections of spelling and punctuation in texts, etc. Conversely, resources comprising the same expression rarely have different titles and statements of responsibility, except in cases of re-publication among different publishers, for example. Likewise, edition statements found in a resource are attributed to the expression when those statements represent the state of text, image, etc., such as "revised edition" and "Japanese translated edition." If statements are related to differences in form and format, they are attributed to the manifestation. The expression entity therefore has these attributes plus those defined in FRBR, such as form of expression, date of expression, language of expression, etc.

The manifestation entity provides the attributes about a resource's physical carrier and format, and its publication, production, and distribution. Titles and others that are attributed to the expression are not usually associated with the manifestation. A manifestation instance in principle does not have a title, statement of responsibility, etc. in a resource model.

d-1) Treatment of aggregate resources with collective titles: An earlier study of modeling of component parts in the expression-dominant model addressed two types of component, "document part" and "content part," which are physically an independent component and a dependent component, respectively.[19] The present study introduces a different viewpoint: whether an aggregate resource has its own collective title. A component part here is a "content part," which is not physically independent of its host.

When an aggregate host has a collective title and individual components within the host have their own titles, (1) an expression instance (and also a work, if appropriate) can be developed in a resource model for an individual component, and (2) the title of a component is associated with the expression instance in the model. Of course, an expression instance (and also a work) for the host resource is developed separately

in a resource model and should represent whole/part relationships to the instances for the components. Additionally, a manifestation instance for the host resource is developed in a resource model to represent the host's physical characteristics. Manifestation instances for individual components are not developed since components here lack physical characteristics except the location of a component within the host. Developing work instances for components and their host in a resource model depends on the policy described earlier.

d-2) Treatment of aggregate resources lacking collective titles: When a host resource lacks a collective title addressing the entire resource, expression instances (and also works, if appropriate) for individual components with their titles are developed and linked to the same manifestation instance for the host resource. Thereby, one manifestation instance can accommodate more than one expression in this case, representing many-to-one cardinality of the "is embodied in" relationship between expression and manifestation. Titles and statements of responsibility that appear in such a host, which are the combination of individual titles and statements of responsibility of the components, are associated with the manifestation—this is an exceptional case in the expression-dominant model.

An expression (and also work) instance for such a host resource lacking its own title is not usually created in this model. In a practical situation, it would be possible to conveniently create an expression (and also work) instance for such a resource with a devised title.

e) Treatment of abridgement, revision, translation, etc.: Abridgement, revision, translation, etc. result in different expression instances from original expressions. Those expressions with such relationships can be linked; the expressions 1 and 2 in figure 1 are linked with a dotted line, representing such a relationship. They are also linked to the same work instance from which abridgement, etc. originate if the work is developed. Collocation of expression instances at the work level, as a result, is attained.

f) Treatment of resources with equivalent content but different physical characteristics: Different manifestation instances are created for resources with equivalent content but different physical characteristics, such as various carriers or formats. In the model, such manifestations are linked to the same expression corresponding to that content. Collocation of manifestation instances at the expression level for such resources is properly attained.

g) Other issues: Developing work instances remains an issue in this model, while those instances are needed to properly represent the "responsibility" relationship to PFCs and the "subject" relationship to other entities. This issue cannot be solved by a theoretical discussion.

An appendix is provided to illustrate an example that is consistent with the expression-dominant model. For this illustration, the following expedients are adopted: (1)

using existing MARC21 biblio-graphic records; (2) transferring the data elements of the MARC records to the attributes of the bibliographic entities; (3) supply-ing data values to nearly manda-tory attributes if no data value is found in the MARC records—they are preceded by "+"; and (4) indicating relationships between bibliographic entities and PFCs under the former entities. MARC bibliographic records with LC control numbers 97001449, 80017667 and 88036703, rep-resenting a family of books, Margaret Maxwell's *Handbook for AACR2* and name author-ity records corresponding to the two persons (LC name author-ity control numbers 80017667 and 95028779), are used here. The resulting set of instances is one work, three expressions, and three manifestations; each expression has one manifestation

Figure 2. Model 2: Manifestation-Dominant Model

in this case. If there is a digital version of any of the books, only a new manifestation is added and linked to the proper expression. The two PFC instances are briefly illustrated.
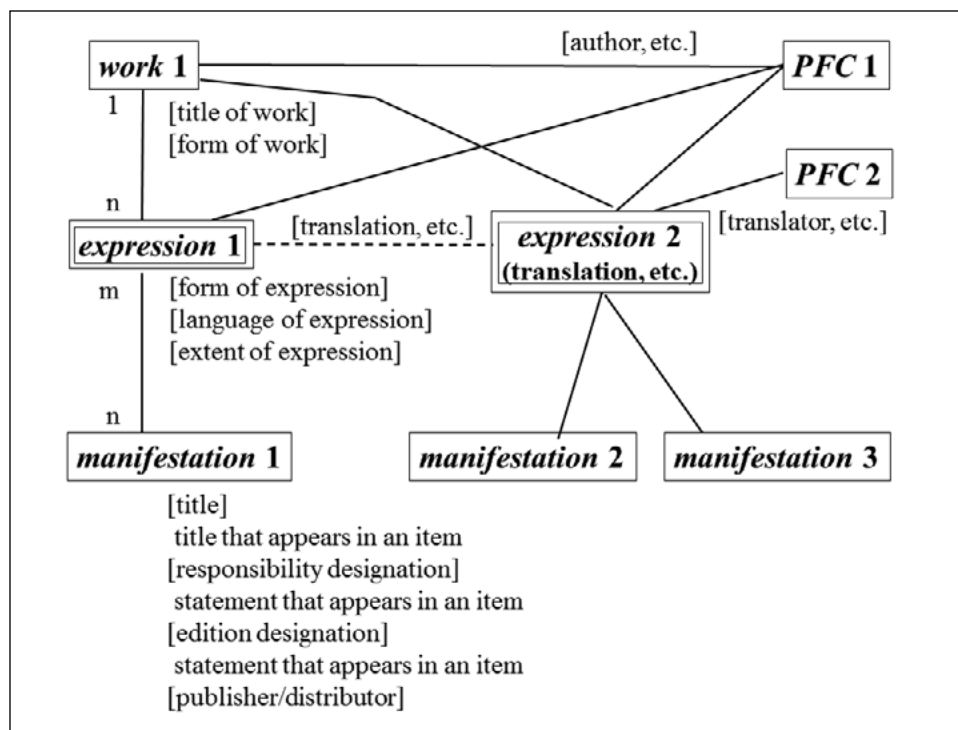
Model 2: Manifestation-Dominant Model

a) Model 2 is the FRBR model itself. Figure 2 shows the model at the instance level with some major attributes. Developing a manifestation instance is mandatory for every resource, since the manifestation entity is given primacy. FRBR intends that creation of a work instance is mandatory, but it does not provide any rationale. The expression may be a "weak" entity, depending on the work, and instance cre-ation for the expression is mandatory or optional, depending on the policy on work instance creation and on relationships between them.

b) A creator (author, composer, etc.) is associated in this model with the work and expression instances created and realized by the creator. A PFC that revises, translates, etc. an expression, or performs a musical work, is associated only with the expressions that the PFC realized. These are based on the premise that work and expression instances are properly developed in the model, but this is not assured as described above.

c) Titles and statements of responsibility appearing in or on a resource are associated with the manifestation entity, as FRBR describes. The model does not associate statement of

responsibility with the expression. Although FRBR defines the attribute "title of the expression," its position and treat-ment are vague; FRBR-LRM does not adopt such an attri-bute anymore. In figure 2, the expression lacks an attribute for title.

d-1) When an aggregate host resource has a collective title, (1) work (and expression) instances are developed in the resource model for individual components within the host; (2) the title of a component is associated with the work (not the expression) instance for the component; (3) a work instance (and an expression) for the host is developed; and (4) whole/part relationships between the component works and the host work (and between the component expressions and the host expression) are developed.

It is readily accepted that, even in the manifestation-dominant model, the title of a component, which appears along with the collective title of the host resource, is associ-ated with the work instance for the component. Because no manifestation instance is usually developed in a resource model for an individual component, we regard titles that appear in a resource but represent components within the resource as titles for the component works without any hesitation.

d-2) When an aggregate host lacks a collective title, work (and expression) instances for individual components with their titles are developed in the resource model, and these instances are linked to the same manifestation
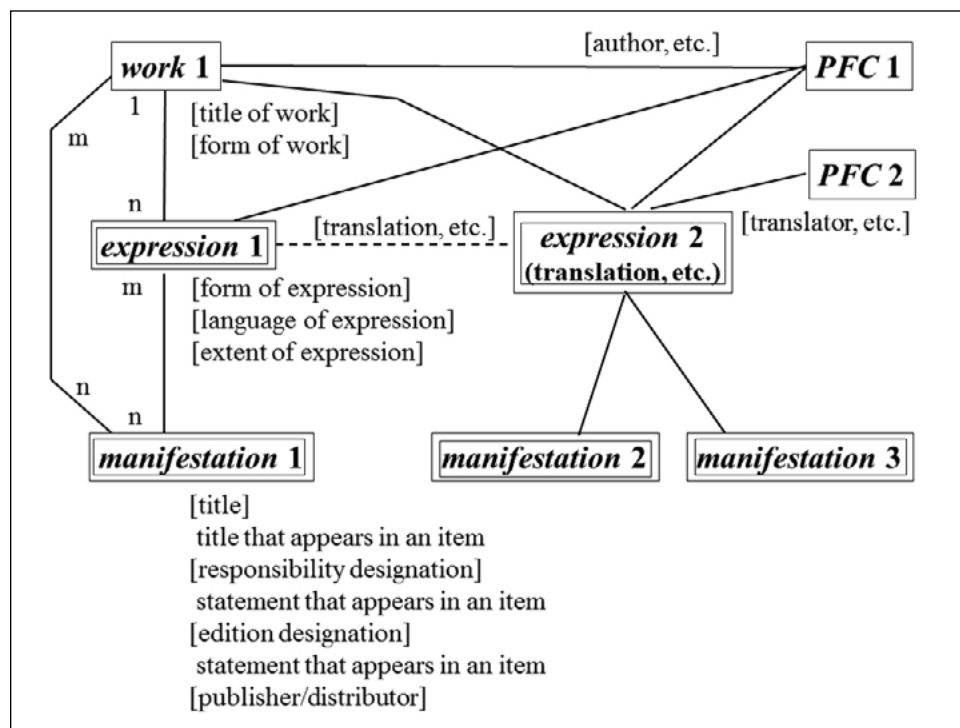
Figure 3. Model 3: Work-Centric Model

the resource's physical aspect is emphasized at that level. In contrast, treatment of the work and expression is uncertain. Whether work and/or expression instances are developed in a model for every resource, or for what cases those instances are developed, is unclear. In particular, treatment of the expression entity in this model is ambiguous while in the cases of the above b) and d) to f), expression instances take important roles.

Model 3: Work-Centric Model

a-1) In this model, the work entity is dominant among the entities while the definitions of the entities are the same as those in FRBR. Figure 3 shows this model at the instance level.

The model adopted by the Indiana University Variations project is similar to this model. The model in Variations2 focuses on recorded classical music and consists of the entities "work," "instantiation," "container," and "media object," which basically correspond to work, expression, manifestation, and item, respectively, in FRBR.[20] The Variations model, however, is work-centric, and hence "the Variations model does not re-use Instantiations on multiple Containers, whereas, according to FRBR, the same performance issued multiple times would be modeled as one Expression appearing on multiple Manifestations."[21] This means that the entity "instantiation" (i.e., being equivalent to the expression) is "weak" and dependent on the work. Variations3, the latest project, adopts a modified version of FRBR but is still work-centric.[22]

a-2) Only work instances are mandatory, whereas expressions and manifestations are optional and dependent on their corresponding works; that is, the expression and manifestation are "weak" entities. A manifestation instance is usually developed in a resource model to represent the physical aspect of a resource. Creating work instances for collected works including compilations, assembled collections, etc. in the resource model is an important issue involved in this model; how do we deal with such resources and develop work instances in a stable manner?

a-3) The relationship between work and expression and between expression and manifestation are the same as those in Models 1 and 2. The relationship between work and manifestation is newly introduced, of which cardinality

instance for the host. The title of the manifestation for a host in such a case is the combination of individual titles of the components. It is unclear whether developing a work instance (and an expression) corresponds to the host.

e) Revision, translation, etc. create different expression instances from those upon which the revision, etc. is based—this is the same as the treatment in Model 1. Those expressions are associated with the same work from which revision, translation, etc. originate, if the work instance is developed in the model. Figure 2 depicts such an expression-to-expression relationship with a dotted line.

f) Resources with equivalent content but different physical characteristics require the development of different manifestation instances for individual resources in resource models. These manifestations are linked to the same expression corresponding to that content, if the expression is developed in the model. However, expression instance creation is unclear in this model as previously noted. If those manifestations are linked to the same work embracing that content, instead of an expression, they are intermingled with other manifestations like revision, translation, etc. under the same work. Collocation of manifestation instances at the expression level is not attained.

g) This model focuses on the manifestation, which includes both the resource's content and the physical characteristics. However, those two aspects (or characteristics) are not separable at the manifestation level; rather,

is many-to-many. This relationship is needed when an expression is omitted but the physical aspect of a resource is recorded with a manifestation. In figure 3, the relationship between work 1 and manifestation 1 is depicted, while relationships from work 1 to manifestations 2 and 3 can be also depicted.

b) The relationships between work and PFC and between expression and PFC are equivalent to those in Models 1 and 2. However, creating in a resource model an expression instance is optional and thus drawing the relationship between expression and PFC depends on the existence of expression instances.

c) Titles and statements of responsibility that appear in or on a resource are associated with the manifestation entity, in the same manner as that in Model 2. In comparison, it is generally difficult to abstract directly a title that appears in a resource as such to a title of the work since a work covers more than one language/script edition and abridged/revised/translated edition.

d) For aggregate resources with collective titles, the patterns described in Model 2 are valid for this model, although expression instance creation in a resource model is unclear here. Assignment of attributes in this model is also the same as that in Model 2.

e) The treatment of abridgement, revision, etc. in Models 1 and 2 is also applied in this model, while developing expression instances in a resource model is not assured.

f) The treatment of resources with equivalent content but different physical characteristics in this work-centric model is the same as those in Models 1 and 2, while developing expression instances is not clear in this model. Manifestations with different physical characteristics are linked to the same expression or work corresponding to that content; of course, the relationship with the expression is dependent on the existence of the expression instance. Collocation of manifestation instances at the expression level is attained only when necessary expressions and proper expression-to-manifestation relationships are developed.

g) Developing expression instances is an unresolved issue in this model. Both the expression and manifestation are "weak" entities and dependent on the work. A manifestation instance is needed to record a resource's physical
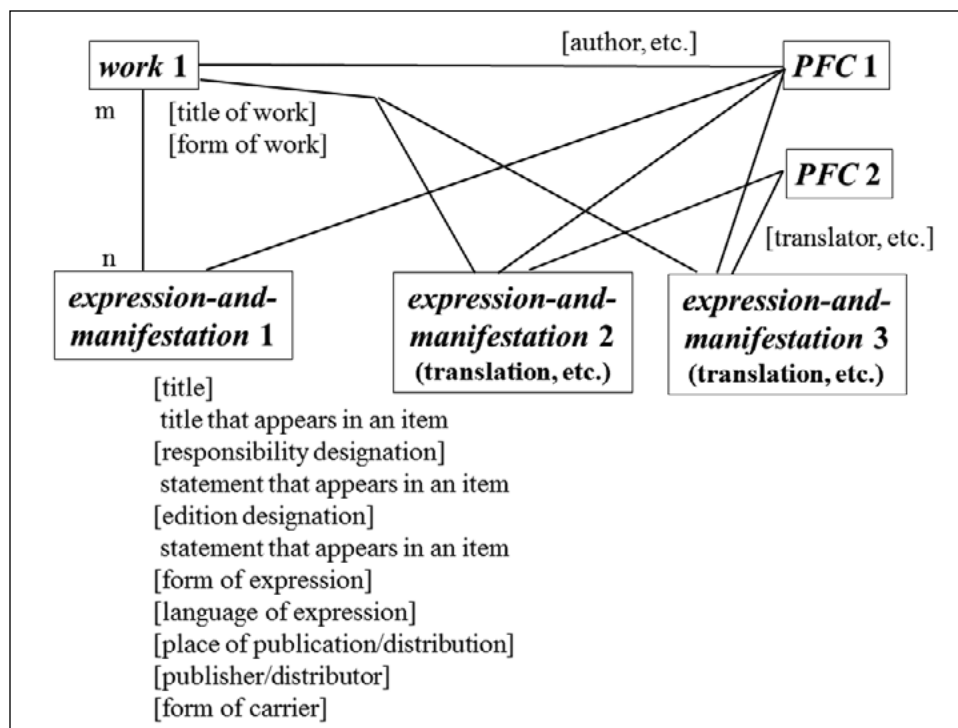

Figure 4. Model giving primacy to expression-and-manifestation

aspect. However, the treatment of the expression is not stable. It is also questionable whether all resources, such as compilations and assembled collections, can be properly managed at the work level.

Model 4: Model Giving Primacy to Expression-and-Manifestation

a) Model 4 is made up of two entities: the work and the combined entity of expression and manifestation in FRBR. The expression-and-manifestation (hereafter E-M) entity is given primacy in this model. Figure 4 depicts this model at the instance level. If the dominant entity is changed from the E-M to the work, the resultant model will be similar to Model 3 with minor differences. Hence, this section discusses the model in which the E-M entity is dominant.

An E-M entity instance is established on a unit of smaller original entity; that is, the unit of manifestation in usual cases, but that of expression in some cases. This model is similar to Model 2, being manifestation-dominant, in this respect. Additionally, whether a work instance is required is not clear, the same as in Model 2. An E-M instance is required for every resource. The cardinality of the relationship between work and E-M is either one-to-many or many-to-many, depending on the policy or interpretation of works, as described in Model 1.

This model seems to be similar to that implemented in

conventional cataloging practice; a uniform title authority record corresponds to a work instance and a bibliographic record corresponds to an E-M instance. FRBR, which is Model 2 in this paper, reflects conventional cataloging practice, but Model 4 would be more similar to it because an E-M instance is close to what a conventional bibliographic record represents.

b) The relationship between work and PFC is equivalent to that in Models 1 to 3. However, expression (e.g., text, sound, etc.) is embedded in the combined E-M entity, and thus the relationship between E-M and PFC is also developed in a resource model for representing the "responsibility" relationship. In Figure 4, the instance PFC 1, which is a creator, is associated with the E-M instances 1 to 3. PFC 2, which is a translator, etc., is linked to E-Ms 2 and 3.

c) Titles and statements of responsibility appearing in or on a resource are associated with the E-M entity. The E-M, being the resultant entity from the entities integration, has both attributes related to the expression—such as form and language of expression—and those related to the manifestation—such as place of publication/distribution, date of publication/distribution, form of carrier, etc.

d) An E-M instance (and a work) is developed in a resource model for an individual component and its host resource, when the aggregate host and its individual components have their own titles. The component's title is associated with the E-M for the component, of which the unit is in accordance with the unit of expression, which is smaller than that of manifestation in such a case. Whole/part relationships between the E-M instances (and between the work instances) can be developed in the model. When an aggregate host lacks a collective title addressing the entire resource, the same treatment is applied as that for a host having its collective title.

e) For cases of abridgement, revision, etc., different E-M instances from those upon which the abridgement, etc. was based are created in this model. Those instances are associated with the same work from which the abridgement, etc. originates.

f) Equivalent content with different physical characteristics causes different E-M instances for individual resources in the model. These instances are linked to the same work corresponding to that content. However, they are intermingled with other E-Ms like abridgement, revision, etc. under the same work. These two groups cannot be differentiated from each other based on their relationship to the work.

g) Collocating of instances at the expression level cannot be attained as described above. The model shows partially the characteristics of being manifestation-dominant. Collocation at the work level, in contrast, is attained if necessary work instances and their relationships to corresponding E-Ms are created in the model. "Responsibility" relationships between bibliographic entities (e.g., E-M and work)

and PFC may be complicated; it is not clear which E-M, work, or both is needed to represent such a relationship in a given case.

Model 5: Model Giving Primacy to Work-and-Expression

a) Model 5 consists of two entities: a combined entity of work and expression and the manifestation, where the former is given primacy (see figure 5). A work-and-expression (hereafter W-E) instance is usually established for a smaller unit, namely, that of the expression, not the work, and creating that instance is mandatory. A manifestation is also required for every resource, while the manifestation is a "weak" entity dependent on W-E. The cardinality of the relationship between W-E and manifestation is many-to-many.

If we were to give primacy to the manifestation among these two entities, the resultant model would be substantially equivalent to Model 2, i.e., the manifestation-dominant model. This section therefore deals with the model giving primacy to the W-E entity.

Meanwhile, the distinction between the above two entities in this model is similar to that between "Work" and 'Instance" in the BIBFRAME model. BIBFRAME's "Work" and "Instance," which are defined as RDF classes, correspond to the combined W-E and the manifestation, respectively.[23] However, BIBFRAME seems to adopt a policy that does not give primacy to either class (i.e., entity), since it is intended to be used to accept a wide variety of metadata, including metadata based on any model, i.e., a "Work"-dominant model and an "Instance"-dominant one.

b) Both the relationship representing "creation" and that representing "realization," like revision and translation, are drawn between W-E and PFC. These two relationships are not differentiated with the associated instances, without relationship designators. Figure 5 depicts these relationships between W-E 1 and PFCs 1 and 2.

c) It is not clear whether titles and statements of responsibility that appear in or on a resource are associated with the W-E or the manifestation; both would be possible. If the first choice is adopted, the resultant model will be similar to Model 1. In contrast, if the second choice is adopted, the resultant model will be similar to Models 2 and 4. The W-E entity puts together attributes associated with the work and the expression in FRBR as a result of the entities integration.

d) When an aggregate host has a collective title, (1) a W-E instance is developed in a resource model for an individual component within the host; (2) the title of a component is associated with that instance for the component; and (3) a W-E instance for the host is developed separately and has whole/part relationships to the W-Es for the components. In comparison, when a host lacks a collective title for the entire resource, there are two scenarios. One is that

W-Es for individual components with their own titles are developed in a resource model, and these instances are linked to the same manifestation for the host resource. No W-E for the host is of course created. The other is that just one W-E, in addition to one manifestation, is developed in a resource model for the host, with its title being recorded is the combination of individual titles of the components. No instance for a component is created in this scenario.

e) For the cases of abridgement, revision, etc., different W-Es from the instances based on for abridgement, etc. are created in this model. Hence, collocation at the work level, which aggregates all expressions under a certain work such as original expression, derivative ones, etc., is not attained in this model. Introducing another upper-level

Figure 5. Model 5: Model giving primacy to work-and-expression

entity like "super-work" is needed to attain that collocation, but this results in a similar model to Model 1, i.e., the expression-dominant model.

f) Resources with equivalent content but different physical characteristics result in different manifestations for individual resources. These instances are linked to the same W-E corresponding to that content.

g) There remain some unclear or unresolved issues such as: c) treatment of titles and statements of responsibility that appear in a resource and d) treatment of aggregate resources, in particular, that lack have collective titles.

Discussion

The results of examining Models 1 to 5 are summarized below.

a-2) Optionality or necessity of creating bibliographic entity instances: Creating in a resource model an entity instance at the level that is given primacy is mandatory as a logical consequence of giving primacy to a certain entity. An expression instance is required in the expression-dominant model, while a manifestation is required when the manifestation entity is dominant. Other entities below the dominant one—if we understand multi-entity models in a hierarchical manner—are in principle "weak," the existence of which is dependent on another entity. Meanwhile, regardless of which
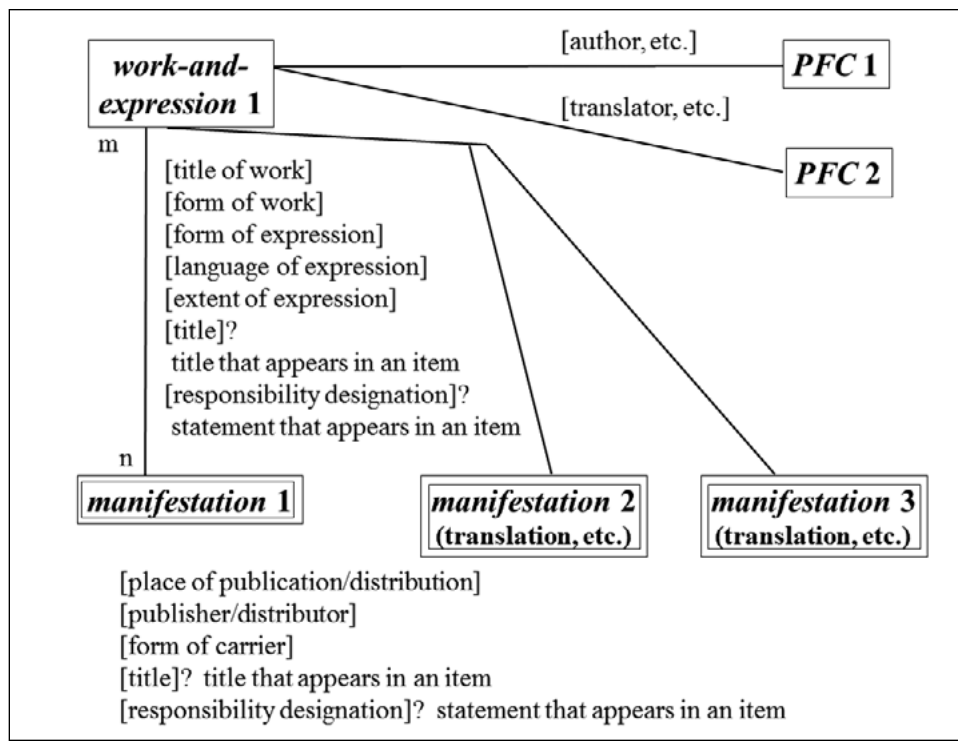
entity is dominant, the manifestation entity (or its equivalent in derivative models) is required to describe a bibliographic resource's physical characteristics. Creating a manifestation instance (or its equivalent) in a resource model therefore is mandatory except in cases where a resource's physical characteristics do not need to be recorded.

a-3) Cardinalities between bibliographic entities: Changing a dominant entity in a model causes no change in the cardinalities of relationships between bibliographic entities. The cardinalities of those relationships are many-to-many, except that between work and expression, which is one-to-many in FRBR but still debatable. These are also valid even in the derivative models, i.e., Models 4 and 5.

b) Relationships to PFCs: PFCs, which are responsible for a resource's intellectual content, are associated with the work and the expression (or their equivalents). When there are both work and expression entities, creators and other secondary contributors for the content are differentiated with the linked entities. In the expression-dominant model, these two are properly differentiated. If there is only either work or expression in a model, creators and other contributors for the content are not differentiated with the linked entities; we need another mechanism to differentiate them, such as relationship designators adopted in RDA (Resource Description and Access).[24] Changing a dominant entity in a model influences the extent of the requirement of an entity instance and thus places a constraint on the relationships

between PFCs and bibliographic entities.

c) Treatment of titles and statements of responsibility that appear in or on a resource: Titles and other information that appear in a resource are in principle associated with the dominant entity in a model. Exceptions are the models that give primacy to the work entity or its equivalent, i.e., Models 3 and 5. In these models, there is a gap between the treatment of such titles and the titles for components within a resource, which is described in d-1) and d-2) below.

d-1) Treatment of aggregate resources with collective titles: Expression and/or work instances (or their equivalents) are developed in a model for individual components within an aggregate host. They, or one of them, usually correspond to the dominant entity in a model. Concurrently, an instance for the host is developed in a model at the dominant entity level and the levels below it. Whole/part relationships between the instances for components and that for the host are developed at the same entity level, such as expression-to-expression and work-to-work relationships.

An exception is Model 2, i.e., FRBR, in which those instances can be developed in the model for components and their host, but they are not both instances at the dominant entity level. This indicates that proper treatment of aggregate resources having collective titles is not assured in Model 2.

d-2) Treatment of aggregate resources lacking collective titles: The same treatment of components as that described in d-1) is applied in every model. In contrast, a manifestation (or its equivalent) is developed in a model for a host, regardless of whether the manifestation entity is dominant. Additionally, "embodiment" relationships are developed in a model between the instances for components, which are expressions and/or works (or their equivalents), and the instance for the host, which is a manifestation (or its equivalent).

e) Treatment of abridgement, revision, translation, etc.: Such resources create in a model different expressions (or their equivalents) from the expression upon which the abridgement, etc. were based. This is independent from the issue as to which entity is dominant. However, if the expression (or its equivalent) is not dominant in a model, it is not assured that proper instances are fully developed for such resources.

f) Treatment of resources with equivalent content but different physical characteristics: Different manifestations (or their equivalents) are created in a model for such individual resources. This is independent from the question as to which entity is dominant. Those manifestations are linked to the same expression or work (or their equivalent) corresponding to that content. As a result, collocation of manifestation instances at the expression level for such resources is properly attained. However, if those manifestations are linked to the same work (not an expression), they are intermingled with other manifestations like revision, translation,

etc. under the same work. These two groups cannot be differentiated from each other based on their relationships to the work.

It is worth noting how the user tasks that FRBR defines are related to the discussion in this paper. FRBR defines the four user tasks: find, identify, select, and obtain, and each task is further divided into "find work," "find expression," etc. User tasks related to the dominant entity in a model have a key position in the sequence of user actions performed by users. Users begin their "find" tasks with the dominant entity in most cases, and that entity is necessarily "identified" or "selected" in the action sequence. In the expression-dominant model (i.e., Model 1), a series of tasks thought to be the mainstream begins with the task "find expression" and then "identify expression" or "select expression." After that, one or more manifestation instances that are linked to each of those expression instances are "identified" or "selected" by the user as appropriate. Subsequent tasks (e.g., "identify or select manifestation," and "obtain item") are then performed in turn. The task "find manifestation" is subordinate to the mainstream. The reason is that sufficient data (i.e., attributes) to accomplish the tasks (including the "find" task) are assigned to the expression entity in this model, whereas the manifestation entity does not have such data. The task "find work" is also a possible action that users take first, but the completion of that task is dependent on the comprehensive development of work instances.

In contrast, in the FRBR model (i.e., Model 2), the task "find manifestation" would be carried out first; the manifestation entity has a solid basis for its accomplishment. The tasks "find work" and "find expression" would be less frequently performed, since (1) it is not clear whether work and expression instances exist in all resources and (2) attributes associated with the entities and used as clues to find them are very restricted; this is particularly true of the task "find expression." The details have been discussed in prior studies by Taniguchi. Similar discussions apply to Models 3 to 5.

Conclusion

Five models including FRBR were examined in terms of several modeling issues, such as optionality or necessity of establishing entity instances, cardinalities between entities, and treatment of titles and statements of responsibility that appear in a resource. Those models consist of FRBR entities or their derivatives and give primacy to different entities among the models. The following implications and consequences of giving primacy to different entities were confirmed.

The direct consequence of giving primacy to a certain entity in a model is (1) an instance of the entity that is given primacy is created for every resource and (2) titles and

statements of responsibility that appear in a resource are associated with the dominant entity, with some exceptions. In the expression-dominant model, expression instance(s) are created for every resource, and titles and other information that appear in or on a resource are associated with the expression entity. These have already been confirmed in prior studies.

These two issues have an impact on (1) drawing relationships between PFCs responsible for a resource's intellectual content and bibliographic entities, namely, work and/or expression (or their equivalents); (2) treatment of aggregate resources and possible resultant collective titles; (3) treatment of abridgement, revision, translation, etc.; and (4) treatment of resources with equivalent content but different physical characteristics.

The expression-dominant model makes it possible to effectively address these issues. Creators and other secondary contributors of the content are differentiated by "responsibility" relationships with the linked entities, that is, either the work or the expression. Component parts within an aggregate resource are represented by expressions and works. Their host resource is represented by a work, an expression, and a manifestation when the host has its collective title, or with only a manifestation when the host lacks a collective title. Abridgement, etc. and resources with equivalent content with different physical characteristics are properly represented by the expression and manifestation entities. Collocation of instances at both the work level and the expression level are fully attained.

These are the characteristics and merits of the expression-dominant model, leading to consistency with a content-oriented model, neither a physical features-oriented nor work (i.e., more abstract construct)-oriented model. The other models, including FRBR, were confirmed as unsuitable for content-oriented in this study. A tendency to separate content from physical features will increase, and thus the same expression will increasingly appear in various formats and carriers. Additionally, most users will move toward a more content-oriented model; users often search for a specific expression (e.g., text in a certain language) and select the manifestation (e.g., a printed book, e-book, or audio file) linked to the expression in accordance with their choice. To handle this situation, studies should begin with a theoretical examination of possible models. The study conducted in this paper reexamined the expression-dominant model as one possibility. Instead, it might be possible to deal with some issues of content-oriented metadata creation at the level of metadata application profiles, or cataloging guidelines and instructions that are subsequent to the modeling and form the cataloging practice; however, this does not lead to a fundamental solution.

It is true that, even if a certain model is selected, its implementation varies depending on application profiles, cataloging guidelines and instructions. For example, if the FRBR model is adopted, multiple application profiles and cataloging guidelines and instructions like RDA can be developed. Even for RDA, some implementation scenarios (i.e., metadata schema) are proposed. This implies that a model in itself does not prescribe the metadata structure and cataloging practice that accord with the model. In some cases, hence, the same metadata records could result from following different models. However, models prescribe the whole framework of and essential points on metadata. Examination at the level of application profiles and others does not provide a fundamental solution.

This study is the first step toward content-oriented metadata creation. For the modeling, further examination of the models in terms of specific resource types, that is, by limiting resource types, is needed in the next step of this study. Another examination of the models by converting the same set of actual extant data to those proper to individual models would be worthwhile to confirm differences among the models. Of course, an examination of metadata schema and cataloging guidelines and instructions that are consistent with the adopted model is also needed to reach the stage of practical application of content-oriented metadata creation.

References

1. IFLA Study Group on the Functional Requirements for Bibliographic Records, *Functional Requirements for Bibliographic Records. Final Report* (München: K.G. Saur, 1998), accessed October 27, 2016, http://www.ifla.org/files/assets /cataloguing/frbr/frbr.pdf; *Functional Requirements for Bibliographic Records. Final Report, As Amended and Corrected Through February 2009* (International Federation of Library Associations and Institutions, 2009), accessed October 27, 2016, http://www.ifla.org/files/assets/cataloguing/frbr /frbr_2008.pdf.
2. Shoichi Taniguchi, "A Conceptual Model Giving Primacy to Expression-level Bibliographic Entity in Cataloging," *Journal of Documentation* 58, no. 4 (2002): 363–82.
3. IFLA Study Group, *Functional Requirements for Bibliographic Record* (1998), 18, 20.
4. Ibid., 18.
5. Taniguchi, "A Conceptual Model Giving Primacy"; Shoichi Taniguchi, "Conceptual Modeling of Component Parts of Bibliographic Resources in Cataloging," *Journal of Documentation* 59, no. 6 (2003): 692–708.
6. Silvio Peroni, "The Semantic Publishing and Referencing Ontologies," in *Semantic Web Technologies and Legal Scholarly Publishing* (Springer, 2016), 121–37; Silvio Peroni and David Shotton, "FaBiO and CiTO: Ontologies for Describing Bibliographic Resources and Citations," *Web Semantics* 17 (2012): 33–43.

7. Julie Allinson, "Describing Scholarly Works with Dublin Core: A Functional Approach," *Library Trends* 57, no. 2 (2008): 221-243; Julie Allinson, Pete Johnston, and Andy Powell, "A Dublin Core Application Profile for Scholarly Works," *Ariadne* 50 (2007), accessed October 27, 2016, http://www.ariadne.ac.uk/issue50/allinson-et-al.

8. Jan Dvo ák, Barbora Drobíková, and Andrea Bollini, "Publication Metadata in CERIF: Inspiration by FRBR," *Procedia Computer Science* 33 (2014): 47–54.

9. Jan Pisanski and Maja Žumer, "Mental Models of the Bibliographic Universe. Part 1: Mental Models of Descriptions," *Journal of Documentation* 66, no. 5 (2010): 643–67; Jan Pisanski and Maja Žumer, "Mental Models of the Bibliographic Universe. Part 2: Comparison Task and Conclusions," *Journal of Documentation* 66, no. 5 (2010): 668–80.

10. Karen Coyle, *FRBR, Before and After: A Look at Our Bibliographic Models* (Chicago: American Library Association, 2016), 19–20.

11. IFLA Study Group, *Functional Requirements for Bibliographic Record* (1998), 23.

12. Library of Congress, "Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services," (2012), accessed October 27, 2016, http://www.loc.gov/bibframe/pdf/marcld-report-11-21-2012.pdf.

13. Godfrey Rust and Mark Bide, *The <indecs> Metadata Framework: Principles, Model and Data Dictionary* (WP1a-006-2.0), (2000), accessed February 25, 2017, http://www.doi.org/topics/indecs/indecs_framework_2000.pdf.

14. Pat Riva, Patrick Le Boeuf, and Maja Žumer, *FRBR-Library Reference Model, Draft for World-Wide Review* (2016-02-21), accessed February 25, 2017, http://www.ifla.org/files/assets/cataloguing/frbr-lrm/frbr-lrm_20160225.pdf.

15. IFLA Study Group, *Functional Requirements for Bibliographic Record* (2009), 20.

16. Taniguchi, "A Conceptual Model Giving Primacy."

17. IFLA Study Group, *Functional Requirements for Bibliographic Record* (1998), 17.

18. Taniguchi, "A Conceptual Model Giving Primacy."

19. Taniguchi, "Conceptual Modeling of Component Parts."

20. Indiana University Digital Music Library Project, "IU Digital Music Library Data Model Specification V2" (2003), accessed October 27, 2016, http://www.dml.indiana.edu/pdf/DML-DataModel-V2.pdf.

21. Jenn Riley, Casey Mullin, and Caitlin Hunter, "Automatically Batch Loading Metadata from MARC into a Work-based Metadata Model for Music," *Cataloging & Classification Quarterly* 47, no. 6 (2009): 519–43.

22. Indiana University Digital Music Library Project, "Definition of a FRBR-based Metadata Model for the Indiana University Variations3 Project" (2007), accessed October 27, 2016, http://www.dlib.indiana.edu/projects/variations3/docs/v3FRBRreport.pdf.

23. Library of Congress, "Bibliographic Framework as a Web of Data."

24. Joint Steering Committee for Development of RDA, *RDA: Resource Description and Access. 2014 Revision* (Chicago: American Library Association, 2016).

Appendix. An Example of a Set of Instances in Line with the Expression-Dominant Model

[work instance 1]
+title of the work: Handbook for AACR2
+date of the work: 1980-
082 00 |a 025.3/2 |2 21
630 00 |a Anglo-American cataloguing rules |x Handbooks, manuals, etc.
650 _0 |a Descriptive cataloging |x Rules |x Handbooks, manuals, etc.
is created by: 100 1_ |a Maxwell, Margaret F., |d 1927-
is realized through: <expression instance 1>
is realized through: <expression instance 2>
is realized through: <expression instance 3>

[expression instance 1]
245 10 |a Handbook for AACR2 : |b explaining and illustrating Anglo-American cataloguing rules, second edition / |c by Margaret F. Maxwell.
008 …s1980 …eng
504 __ |a Includes bibliographical references and index.

is realized by: 100 1_ |a Maxwell, Margaret F., |d 1927-
is embodied in: <manifestation instance 1>

[manifestation instance 1]
260 __ |a Chicago : |b American Library Association, |c 1980.
300 __ |a xi, 463 p. ; |c 24 cm.
020 __ |a 0838903010 (pbk.) : |c $8.00 (est.)

[expression instance 2]
245 10 |a Handbook for AACR2, 1988 revision : |b explaining and illustrating the Anglo-American cataloguing rules / |c by Margaret Maxwell ; with a new chapter by Judith A. Carter.
008 …s1989 …eng
504 __ |a Includes bibliographical references and index.
500 __ |a Rev. ed. of: Handbook for AACR2.
is realized by: 100 1_ |a Maxwell, Margaret F., |d 1927-
is realized by: 700 1_ |a Carter, Judith A.
is embodied in: <manifestation instance 2>

[manifestation instance 2]
260 __ |a Chicago : |b American Library Association, |c 1989.
300 __ |a ix, 436 p. : |b ill. ; |c 26 cm.
020 __ |a 0838905056 (alk. paper)

[expression instance 3]
245 10 |a Maxwell's handbook for AACR2R : |b explaining
and illustrating the Anglo-American cataloguing rules and
the 1993 amendments / |c Robert L. Maxwell with Margaret
F. Maxwell.
246 30 |a Handbook for AACR2R
008 …s1997 …eng
504 __ |a Includes bibliographical references and index.
500 __ |a Rev. ed. of: Handbook for AACR2, 1988 revision /
by Margaret Maxwell.
is realized by: 100 1_ |a Maxwell, Robert L., |d 1957-
is realized by: 700 1_ |a Maxwell, Margaret F., |d 1927-
is embodied in: <manifestation instance 3>

[manifestation instance 3]
260 __ |a Chicago, IL : |b American Library Association, |c
1997.
300 __ |a xii, 522 p. : |b ill. ; |c 26 cm.
020 __ |a 0838907040 (alk. paper)

[PFC instance 1]
100 1_ |a Maxwell, Margaret F., |d 1927-
046 __ |f 19270909
372 __ |a Library science |2 naf
375 __ |a female
377 __ |a eng
400 1_ |a Maxwell, Margaret Finlayson, |d 1927-
400 1_ |a Maxwell, Margaret, |d 1927-
670 __ |a Her Shaping a library, 1973.

[PFC instance 2]
100 1_ |a Maxwell, Robert L., |d 1957-
046 __ |f 19571214
372 __ |a Library science |a Cataloging |a Printing--History
|2 lcsh
375 __ |a male
377 __ |a eng
378 __ |q Robert LeGrand
400 1_ |a Maxwell, Bob, |d 1957-