

A Faceted Data Model for Bibliographic Integration Between MARC and FRBR

Seungmin Lee*

Chungnam National University
Republic of Korea
E-mail: ableman@cnu.ac.kr

ABSTRACT

Although MACHine Readable Cataloging (MARC) and Functional Requirements for Bibliographic Records (FRBR) are currently the most broadly used bibliographic structures for generating bibliographic data in the library community, each has its own weaknesses in describing information resources in diverse media. If the MARC format could be implemented in a structure that reflects the multi-layered characteristics of FRBR, its use could address current problems and limitations in resource description. The purpose of this research is to propose an alternative approach that can integrate the heterogeneous bibliographic structures of MARC and FRBR through the applications of facet and facet analysis. The proposed faceted data model is expected to function as a conceptual structure that can mediate between MARC data elements and FRBR attributes in order to utilize these structures in a more reliable and comprehensive way.

Keywords: MARC, FRBR, Facet, Facet analysis, Integration

1. INTRODUCTION

To handle the increase of information resources in diverse media, many in the library community have relied on tools such as Anglo-American Cataloging Rules (AACR) and MACHine Readable Cataloging (MARC) in order to manage and organize these resources. Currently, the MARC format is

the most broadly used bibliographic standard for encoding and exchanging bibliographic data based on the descriptive rules provided by AACR. Although AACR and MARC format are suitable for more traditional resources such as books and printed materials, they may not be the most appropriate tools for describing new forms of resources, such as digital resources on the Web that are remotely

Open Access

Received date: January 12, 2013
Accepted date: February 23, 2013

*Corresponding Author: Seungmin Lee
Assistant Professor
Department of Library and Information Science
Chungnam National University, Republic of Korea
E-mail: ableman@cnu.ac.kr

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0/>). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

accessed. Because of its strict and rigid structure, the MARC format is limited in its ability to describe digital resources and cannot adequately represent the semantics and dynamic natures of those resources. MARC's ability to represent relationships among bibliographic entities with multi-layered characteristics is also problematic because of its linear and single-layered structure.

There is growing awareness of the need for a more flexible structure for bibliographic data which can handle a variety of resource media, represent the relationships among descriptive entities, and describe the multi-layered characteristics of digital resources. Currently, metadata, generally defined as data about data, is recognized as a powerful tool for generating bibliographic data and standardizing resource description. However, many communities have developed unique metadata standards to satisfy their own purposes. This tendency has led to a flood of heterogeneous purpose-specific metadata standards that can be only used in a specific community. It also leads to duplication of metadata records in different formats and inefficient use of existing records. Although those metadata standards are designed to provide standardized practices for resource description, they have failed to generate standardized resource description because of the lack of commonly accepted description rules.

To overcome these problems and to cope with the dynamic nature of new types of resources, the International Federation of Library Associations and Institutions (IFLA) proposed the Functional Requirements for Bibliographic Records (FRBR) model, which focuses on the organization of bibliographic elements and provides for multiple relationships among descriptive entities. Although FRBR can support the representation of multi-layered characteristics of resources, it also has several weaknesses as a bibliographic standard. For example, it does not provide sufficient data elements for resource description. And, because its strict hierarchical structure prescribes the relationships among entities and attributes, the predetermined relationships are also too rigid to provide the flexibility necessary to describe the dynamic nature of digital resources.

Many approaches have been explored in attempts to overcome these problems and to achieve interoperability between MARC and FRBR in order to combine the strengths of and complement the weaknesses of both MARC and FRBR. If the MARC format could be implemented in a structure capable of describing multi-layered characteristics, it could address current problems and limitations in resources description. These approaches are based on the approaches to metadata interoperability and focus on similarities between the two sets of descriptive elements. However, they have shown that many FRBR attributes do not have data elements that can be mapped directly to MARC and vice versa because of their unique characteristics.

The purpose of this paper is to propose an alternative approach that could interrelate between MARC data elements and FRBR attributes by constructing a conceptual data model. It provides a set of core bibliographic elements through the applications of facet analysis in order to conceptually integrate MARC data elements and FRBR attributes.

2. APPROACHES TO INTEROPERABILITY

2.1 Current Approaches to Interoperability

The efforts of the library community to achieve interoperability between bibliographic structures and their applications have made use of several different approaches, mostly applied to achieve metadata interoperability. Although these approaches are focused on establishing semantic relationships between the components of metadata standards, they have adopted different methods that reflect the aspects of the standards on which they focus.

Chan and Zeng (2006) have grouped current approaches to interoperability into three categories based on the level of relationships established between metadata standards: schema level, record level, and repository level. In schema level approaches, the focus is on the elements of a scheme that are independent of any application. Crosswalks, application profiles, and registries are methods included in this category. In record level approaches, metadata records are integrated through the mapping of elements. This level encompasses element mapping

and data integration methods. In repository level approaches, the objective is to map value strings associated with particular elements. Metadata repository and aggregation are representative methods in this third category.

Other researchers have contributed substantially to understanding metadata interoperability. Moen (2004) has argued that mechanisms for addressing metadata interoperability evince four main approaches: mapping, crosswalks, application profiles, and metadata registries. According to Moen, mapping is a process that identifies semantically equivalent elements in different standards, and crosswalks implement the basic rationale of element mapping, making mapping and crosswalks very similar approaches. Application profiles offer schema-level interoperability for sharing information about metadata standards in order to exchange and reuse elements. Metadata registries provide indexes to metadata terms and official definitions as well as to local variations and extensions that enable the reuse of existing elements.

Hodge (2005) has proposed major approaches to interoperability, including metadata frameworks, crosswalks, and metadata registries. The metadata framework approach integrates various standards into a single standardized scheme. It is a reference model which provides a conceptual structure into which other standards can be placed. A crosswalk matches the elements, semantics, and syntax of one standard with those of one or more other standards, where possible. A metadata registry is based on element mapping at the schema level. Hodge states that a metadata registry is a metadata database that stores terms and definitions of the components of metadata schemes and provides extensions of the terms in order to support reuse and exchange of elements.

Based on these approaches to metadata interoperability, several methods have been explored in attempts to integrate MARC and FRBR. Delsey (2002) mapped MARC21 data elements to FRBR attributes based on the direct element mapping approach. Aalberg (2005) refined FRBR attributes and created mapping tables to match FRBR elements to MARC data elements. However, these approaches have shown that many MARC data ele-

ments do not have equivalent FRBR attributes that can be directly mapped because they focus on similarities between the two sets of descriptive elements. In addition, these approaches attempted to map between MARC data elements and FRBR attributes without considering structural differences, although MARC has a single-layered structure and FRBR adopts a hierarchical structure with multi-layered attributes.

These difficulties in achieving interoperability between MARC and FRBR mainly result from the heterogeneity of bibliographic structures. Each structure has different representation of descriptive entities, different structural frameworks, and different levels of granularity from the other. For these reasons, current approaches have generally failed to achieve reliable interoperability between MARC and FRBR.

2.2 Methodology

MARC and FRBR are bibliographic systems with pre-determined structures. MARC can be considered as a bibliographic structure based on a standardized cataloging rule for resource descriptions. It contains more than 2,000 bibliographic entities under a strict and rigid structure. FRBR can also be considered as a cataloging rule with a conceptual structure for bibliographic description. However, it does not have specific entities but provide concepts that categorize bibliographic entities under the conceptual structure.

To achieve interoperability between these two heterogeneous structures and to fully utilize the advantages of each, it is necessary to adopt different approaches from those applied to metadata interoperability. Bibliographic entities conceptually represent general aspects of resources and consist of structures of bibliographic records, whereas metadata elements indicate specific aspects of information resources. For this reason, interoperability between MARC and FRBR may not utilize approaches based on direct mapping between elements in different metadata standards.

This research tried to construct a conceptual structure that would function as a data model which integrates descriptive aspects of information resources. This approach incorporates the strengths

of merging and mapping to make two heterogeneous bibliographic structures interoperable. In most cases, merging and mapping play important roles in integrating and achieving interoperability between heterogeneous structures.

Mapping can be defined as the process of establishing relationships between semantically equivalent elements in different structures (Kurth, Ruddy & Rupp, 2004). Thus, mapping refers to the process of associating elements of one set with elements of another set. In bibliographic description, mapping can provide conceptual connections among data elements in two or more bibliographic structures. These conceptual connections can establish simple relationships or associations among elements without any change or modification. Because the characteristics of the original elements are retained after mapping, the process can achieve interoperability between heterogeneous structures while retaining the unique characteristics of the original structures. However, mapping itself is not enough to achieve interoperability since it can only provide conceptual relationships at the data element level. It is necessary to establish relationships at the structure level to achieve full interoperability between two or more structures.

Merging generally takes two or more entities and reconstructs them into a single new entity (Tenant, 2004). Thus, if one entity is merged with another, they are combined to make a new structure. The merged entity does not retain the structures or characteristics of the original entities. Merging creates a new structure that is totally different from those of the original entities.

Bibliographic structures are usually heterogeneous structures with unique characteristics. This is the main reason that interoperability between different structures is obstructed. Although merging eliminates the heterogeneity of different structures and data elements, it can also eliminate the unique characteristics of each structure and data element by merging them into a new structure.

To mediate these limitations and problems, this research utilizes the strengths of both mapping and merging in order to integrate between two different bibliographic structures, instead of achieving interoperability between them. The integration process

may require clear and comprehensive criteria that can be used as conceptual spaces for merging data elements in different structures. Applications of facet analysis are applied to set up the criteria for integrating the two sets of bibliographic structures and constructing a faceted data model.

3. APPLICATION OF FACET ANALYSIS

3.1 Characteristics of Facet Analysis

A facet, in its simple meaning, is a conceptual categorization. It generally refers to a concept group, consisting of generic terms, used as a general manifestation of a compound subject to denote components of the subject (Ranganathan, 1962). In the library community, a facet is defined as “clearly defined, mutually exclusive, and collectively exhaustive aspects, properties, or characteristics of a class or specific subject” (Taylor, 1992). From different perspectives, a facet also refers to a partitioning of vocabulary or grouping of terms obtained by the division of a subject discipline into homogeneous or semantically cohesive categories (Svenonius, 2000).

The process of partitioning domain vocabulary and generating facets is often called facet analysis (or faceting). It is a mental process involving analysis of a subject into its facets based on a set of postulates and principles, resulting in a knowledge structure with clearly delineated semantic relationships between concepts. This structure provides a framework to accommodate various types of concepts along with syntax rules for their combination (Kumar, 1987).

Facet analysis derives two processes: analysis and synthesis. Ranganathan (1967) demonstrated facet analysis as “the process of breaking down subjects into their elemental concepts.” These concepts can be synthesized, which is the process of recombining concepts into subject strings or creating new compound terms.

Through the process of analysis and synthesis, facet analysis can be used as a tool to identify and represent relationships between concepts in a certain subject. Each facet contains a number of terms that will be considered to be conceptually equivalent (Harter, 1986). Based on the relationships of

concepts, the process also provides a framework of vocabulary with enough flexibility to include new subjects in it because it can synthesize or combine subjects with facets. Facet analysis also provides for the organization of concepts in modular hierarchies by separating unrelated or dissimilar concepts and grouping related or similar concepts. Thus, relevant concepts are identified by partitioning domain vocabulary into mutually exclusive facets (Priss & Jacob, 1998).

3.2 Types of Facets

Although the approaches of facet and facet analysis have many strengths in organizing and categorizing related concepts, it is difficult to define facets and to prescribe the semantic range of each facet. Because of the ambiguous or abstract semantics of facets, there are concepts which cannot be categorized exclusively into only one facet (Svenonius, 2000). Another problem is raised when two concepts with different attributes have the same label. Facet analysis does not provide a way to distinguish between these concepts, resulting in ambiguity of concept relationships.

To address these weaknesses and to support the organizing and categorizing process through facet and facet analysis, this research divides facets into two types according to their functions in bibliographic description: *class facet and property facet*.

Class facet refers to the fundamental attributes of every bibliographic entity in resource description. The semantic range of each class facet will be broad and abstract in order to encompass all the related concepts in resource description. It can also support the categorization process by encompassing entire concepts and categorizing related concepts with each other. Through these processes, class facet prescribes the semantic range of facet structures generated by facet analysis. Each class facet is located at the top of the structure and provides a semantic framework in which related concepts can be placed together.

Property facet can be considered as subclass of class facet. It is located under each of the related class facets and constructs hierarchical facet structure by specifying the semantic range of the class facets. It represents the category of the attributes of

concepts used in resource description. This facet also defines representative aspects, properties, and characteristics of a bibliographic entity.

When applied to MARC and FRBR, the limitations of facet and facet analysis can be complemented by the bibliographic structures of MARC and FRBR. These bibliographic structures have categorized entities based on their meanings in bibliographic descriptions with hierarchical structure. The semantic range of each entity is predefined with concrete structure. In addition, the label of each bibliographic entity is mutually exclusive with each other. Therefore, when applied to bibliographic structures, facet analysis can provide a clear facet structure. In addition, the class facets and property facets can be applied to each entity contained in both structures.

However, the applications of facet and facet analysis can provide integration of bibliographic entities in both structures instead of interoperability. Interoperability is conceptually based on the direct mapping of entities with the same or similar meaning. In contrast, the bibliographic structures are designed to provide bibliographic rules and standards to generate bibliographic data, although they contain concrete entities.

By applying these different types of facets, semantic similarity between concepts can be clearly identified and the semantic relationships between concepts can be established in the facet structure. The constructed facet structure can support consistent extension of facets and facet analysis by representing concept relationships and functions as a faceted data model for integrating MARC and FRBR.

4. CONSTRUCTION OF FACETED DATA MODEL

4.1 Analysis of MARC and FRBR

The first step in the process of constructing faceted data model for integrating bibliographic structures involves the generation of facet vocabulary in different structures. This generation of vocabulary begins with analyzing and identifying bibliographic structures from which to extract concepts of elements and the concept relationships.

4.1.1 Analysis of MARC data elements and structure

MARC provides a standardized record structure for encoding and exchanging bibliographic data. As Moen and Benardino (2003) assert, MARC originated as a means to communicate bibliographic data about printed materials. However, it has evolved to address the representation of numerous bibliographic data types, including computer files, maps, serials, music, visual materials, and archival materials.

MARC consists of three main components in a bibliographic record: the leader, the directory, and the variable fields. These bibliographic components are enumerated in a predetermined structure. In addition, the MARC format uses a set of tags, indicators, delimiters, and subfield codes, which are applied to pre-determined MARC fields.

The MARC format is an analytical system with a linear structure that can fully describe bibliographic entities through application of almost 2,000 descriptive data elements. Although MARC analyzes data elements in detail, it simply enumerates those elements in a single-layered format that is pre-determined. This structural rigidity cannot fully support representation of resources with multi-layered bibliographic relationships. It is also problematic for MARC format to describe new types of digi-

tal resources because it was originated for traditional printed materials.

Another restriction on the MARC format is that its structure is based on the concept of main entry. Main entry (i.e., the 1XX field tag) relies on authorship of a work as the primary access point. With the advent of online catalogs, main entry is losing its importance because there are multiple possible access points other than author. The prescribed structure based on main entry also separates related data elements, thereby resulting in the possible duplication of data in the record.

4.1.2 Analysis of FRBR data elements and structure

FRBR provides a conceptual model for bibliographic records that defines logical relationships among bibliographic objects in terms of an entity-relationship model. FRBR defines the structure of a catalog record as a set of relationships among multiple entities, whereas MARC uses a linear and flat bibliographic structure. FRBR identifies three types of entities that are relevant to bibliographic objects: Group 1 (Work, Expression, Manifestation and Item); Group 2 (Person and Corporate body); and Group 3 (Concept, Object, Event and Place). As a further definition,

Table 1. Components of Directory

Variable Fields	Field tags	Descriptions
Control Field	00X	Control number, data and time, etc.
	0XX	Control information, numbers, codes
Data Fields	1XX	Main entry
	2XX	Titles, edition, imprint
	3XX	Physical description, etc.
	4XX	Series statements
	5XX	Notes
	6XX	Subject access fields
	7XX	Name, etc. added entries or series; linking
	8XX	Series added entries; holdings and locations
9XX	Reserved for local implementation	

Table 2. Entities Comprising FRBR Groups

Group	Entities	Attributes
Group 1	Work	Work title, form or genre, date, performance medium, intended audience, etc.
	Expression	expression title, form of the expression, language of the expression, type of score, scale of a map, etc.
	Manifestation	manifestation title, publisher, date of publication, form of carrier, dimensions, manifestation identifier (e.g. ISBN), terms of availability, etc.
	Item	location or call number, barcode, provenance, condition, access restrictions on an item, etc.
Group 2	Person	names, dates, titles or other designations, etc.
	Corporate body	name, number, place, date, other designation, etc.
Group 3	Concept	Term
	Object	Term
	Event	Term
	Place	Term

Group 1 comprises the products of intellectual or artistic endeavour that are named or described in bibliographic records: work, expression, manifestation, and item. Group 2 comprises those entities responsible for the intellectual or artistic content, the physical production and dissemination, or the custodianship of such products: person and corporate body. Group 3 comprises an additional set of entities that serve as the subjects of intellectual or artistic endeavour: concept, object, event, and place (IFLA Study Group on the Functional Requirements for Bibliographic Records, 1998).

Each entity in each of FRBR's three groups can be expanded by using attributes which can serve as a means for users to formulate queries about a particular object (IFLA Study Group on the Functional Requirements for BR, 1998). A total of 97 attributes are defined in FRBR in terms of the characteristics of an entity, rather than as specific data elements.

The FRBR model focuses on the organization of data elements. It provides for multiple relationships among bibliographic entities by adopting a hierarchical structure. This hierarchical structure can clearly describe bibliographic relationships and deal with multi-layered characteristics of resources. In

contrast, MARC generally consists of manifestation-level and item-level information and the bibliographic elements are enumerated according to a linear structure. MARC also has bibliographic elements that correspond to work-level and expression-level elements in FRBR, but they are placed in fields related to authority files or uniform title. Therefore, MARC format is a mixture of work, expression, manifestation, and item information in bibliographic records within a linear structure that cannot express explicit relationships between entities.

FRBR enhances the retrieval of digital resources because it contains attributes which can be specific to digital resources, such as system requirements, file characteristics, mode of access, and access address, which MARC does not clearly provide. However, the FRBR model does not provide sufficient data elements to fully describe bibliographic entities, even though it can support the representation of multi-layered characteristics of information resources. Also, it has a pre-determined hierarchical structure, and the relationships among data elements are too rigid to provide the flexibility necessary when describing the dynamic nature of digital resources.

4.2 Generating Facet Vocabulary of Bibliographic Entities

Once bibliographic structures have been analyzed, the next step is to generate facet vocabulary by extracting elements from the two structures.

Each element has a unique meaning to represent a specific aspect of resources. However, the meaning is not just derived from the aspects of resources, but also from the context which is usually reflected in those structures. This context affects the semantics of elements and often causes the same element to be used in different ways. To address this contextual problem, the semantics of each element is considered along with the structural differences between the two structures when extracting elements.

The first step of the extraction process is the identification of superordinate elements in each structure. By identifying and comparing the elements with correspondents in another structure, commonly used superordinate elements in both structures can be considered as core elements because they share specific meanings in resource description. Then, all elements placed under superordinate elements in each structure are analyzed and extracted according to their semantics. These subordinate elements may have specific meanings to describe detailed aspects of resources. Thus, the types and quantities of these elements may vary across structures according to the levels of granularity. In addition, some elements may be used in more than one place in a structure with different labels. In this case, the extraction of subordinate elements only focuses on the semantics of the element based on the context, instead of the labels. By using the semantics as the criteria of element extraction, this duplicated use of the same element can be eliminated.

After these analyses, the extracted elements from the two structures are put together into the facet vocabulary and categorized according to their semantic similarities. This vocabulary reflects the semantic range of those structures and functions as the foundation of constructing a faceted data model.

Each shared meaning of superordinate elements

functions as a class facet that connects different structures through the related subordinate elements placed under each superordinate element. The subordinate elements serve as property facets because they represent specific attributes and characteristics of each bibliographic element in different structures.

4.3 Concept Group Identification

The basic strategy of categorizing extracted elements is to identify groups of related concepts that could be potential facets. Application of facet analysis converts extracted elements into concepts and creates a comprehensive set of candidate facets.

The elements in MARC are distributed under 10 categories represented by each field from 0XX to 9XX. Although these 10 fields encompass all the MARC elements, the meaning of each field is too broad to represent the specific aspects of each resource. The elements substantially used for resource description are the delimiters that specify the meaning of each field. Thus, this research analyzed the meaning of delimiters and extracted MARC elements based on the analyzed meaning. The elements in MARC were grouped into seven categories: Author, Title, Subject, Publication, Description, Identifier, and Format.

The structure of FRBR is extremely different from that of MARC because FRBR was originally designed as a data model focusing on the organization of bibliographic entities. Thus, the FRBR model does not provide sufficient elements to fully describe resources. However, FRBR also contains many elements for describing and managing resources within a hierarchical structure. FRBR also has 10 superordinate elements (entities) grouped into three from Group 1 to Group 3. However, the meanings of the elements are very broad and ambiguous. Thus, this research considered the attributes of each entity that specifies the aspects of a resource. The attributes in FRBR were also grouped into seven categories: Author, Title, Subject, Publication, Description, Identifier, and Format.

By comparing these superordinate elements extracted from both MARC and FRBR, this research creates seven class facets: Author, Title, Subject, Publication, Description, Identifier, and Format. In

Table 3. Example of Category of Facet (Class Facet *Format* only)

Class Facet	Property Facet
Format	Serials
	Musical works
	Image
	Cartographic
	Electronic resources
	Sound

the proposed data model structure, each class facet provides a space in which the related elements and attributes can be placed together according to their semantics.

Once class facets have been established, all other elements placed under each superordinate element in both structures are analyzed and facet analysis can extract the concept of each element. The extracted concepts are placed under each of the related class facets. These extracted concepts function as property facets which specify the semantic range of class facets.

4.4 Construction of a Faceted Data Model

Construction of a faceted data model was based on the generated facet vocabulary. Class facets derived from superordinate elements can have a broad range of semantics and may not be specific enough to represent particular characteristics of a resource. The concepts derived from subordinate elements function as property facets and are nested under class facets derived from superordinate elements. These property facets divide the broad meaning of the class facet into more specific units of semantics and establish semantic relationships between class facets and property facets. Property facets related to the same class facet were subsequently categorized based on their functional similarities.

Following the categorization of extracted concepts into class facets and property facets, a hierarchical structure of a faceted data model can be con-

structed. Class facets, which are commonly used in both structures, are placed at the top of the data model. Each class facet contains elements equivocal or similar to property facets from both MARC and FRBR. The class facet consists of property facets, which functions to specify attributes or characteristics of the class facet. This group of facets prescribes the semantic concepts of different bibliographic structures. In addition, the group of facets describes the semantics of elements in both structures by specifying the function of the element under a specific context.

Based on semantically distinct categories of facets, a faceted data model was constructed, which categorizes constituent facets according to their functional roles. The structure of the faceted data model consists of three components: facet group (class facet and property facet), extracted concept, and bibliographic elements in both structures. The facet group has its own hierarchical structure of relationships between class facets and property facets. This group of facets has specific relationships with other components of the model which indicates how elements in bibliographic structures are conceptually connected with facets. By this step, the duplication of elements can be eliminated. In addition, the problems of the separation of same or similar elements in different places in each structure can be also addressed. For example, both 245 and 246 fields represent a title of a work, but those fields are separated in the MARC system. FRBR also separates the attributes related to a title of a work into three places: title of work, title of expression, title of the manifestation. Through the use of class facets and property facets, these separated and distributed elements can be grouped together in regard to their semantics.

In the proposed faceted data model, there are seven class facets at the top of the structure and 21 property facets under basic facets according to the semantic similarities between them (see Table 5).

These facets can be assigned to each element in MARC and FRBR. Through these facets, both MARC and FRBR elements can be integrated and interoperable with each other.

The faceted data model is not intended to describe specific information resources but to provide

Table 4. Example of Facet Vocabulary (Class Facet *Format* only)

Class Facet	MARC	FRBR	Property Facet
Format	254, 255, 256 310, 321 352 342, 343 362 856	Work Expression Manifestation	Serials: publication frequency
			Serials: publication status
			Musical works: performance medium
			Musical works: score type
			Musical works: key
			Image: technique
			Image: color
			Cartographic: coordinates
			Cartographic: scale
			Cartographic: technique
			Electronic resource: system
			Electronic resource: access
			Sound: technique
			Sound: reproduction
Sound: playing			

Table 5. Class Facets and Property Facets of the Proposed Faceted Data Model

Class Facets	Property Facets
Author	Person, Corporate Body, Meeting
Title	Title Statement, Series Statement
Subject	Classification Number, Keyword
Description	Edition, Summary, Representation
Identifier	Identifier
Publication	Publisher
Format	Serials, Musical Work, Cartographic Work, Computer File, Image, Microform, Electronic Resource, Sound Recording, Other Formats

a set of core bibliographic elements. Facets in the model can be connected to both the MARC and FRBR because these facets were extracted from the elements contained in the MARC and FRBR systems. If any facet in the data model can be connected with any of the corresponding MARC data elements and FRBR entities/attributes, a user can utilize MARC for detailed descriptive elements and FRBR for representation of bibliographic relationships.

4.5 Implementation of a Faceted Data Model

Elements of bibliographic structures are used to describe specific aspects of a resource and derive their values from the resource being described. However, the values associated with a resource are not values of the element per se, but of the concept that is the core meaning of the element. In this sense, an element is the representation of a concept. The format of an element can be changed

Table 6. Example of the Interoperable Elements on Faceted Data Model (Class Facet Title Only)

Class Facet	Property Facet	MARC Element		FRBR Element	
Title	Title statement	245 \$a	Title statement	Work	Title of work
				Expression	Title of expression
				Manifestation	Title of the manifestation
	Title proper	246 \$a	Title proper	Work	Title of work
		246 \$b	Remainder of title	Expression	Title of expression
				Manifestation	Title of the manifestation
	Uniform title	130 \$a	Uniform title		
		240 \$a	Uniform title	Work	Title of work
		730 \$a	Uniform title		

according to the context of the associated bibliographic structure, as reflected in its structure and syntax, while the concept, which is the translated meaning of the element, is not changed regardless of differing contexts.

Implementation of the faceted data model is based on the contextual instantiations of facets in facet vocabulary that represent the semantic functions of elements in resource description. Thus, the faceted data model incorporates three primary components: the elements, the facets, and the functions of the facets. Each of the components is key to implementation of the faceted data model. Each component is connected to other components by specific relationships. A class facet identifies the core and broad meaning of an element, providing the semantics for the element. The property facet stands in for the roles of each element in resource description.

Each class facet, property facet, and elements in both structures are closely related with each other. Property facet represents a specific part of the semantics of class facet and the meanings of elements are converted to property facets based on the prescribed semantic range of related class facets. Both property facets and the meanings of elements consist of the semantic range of the superordinate class facet.

In spite of their different functions, the class facet and property facet serve as containers that can hold similar yet heterogeneous concepts of elements

which comprise the class facet. Using these two types of facets, the different functions of each facet and their relationships can be specified and differences in semantic ranges can be mediated.

Based on the semantically distinct categories illustrated in Figure 2, a faceted data model was constructed that categorizes constituent facets according to their functional roles. The facet structure consists of three components: a group of facets (class facet and property facet), roles of facets, and meanings of elements. Class facet represents the primary concept that subsumes related subordinate concepts represented by elements in both MARC and FRBR. Property facet contains concepts that describe specific aspects of the class facet and functions as subordinates of the class facet. This set of facets has its own hierarchical structure of relationships between facets. Property facet represents a specific part of the semantics of class facet and a group of facets (class facet-property facet pair) clarifies the semantic range of bibliographic structures, which is determined by the meaning of each element in both structures.

The faceted data model is not intended to be used for any actual resource description, but to provide a set of facets that can relate elements from different bibliographic structures. Facets, which represent elements from different structures in a context-independent manner, can semantically connect the elements in both the MARC and FRBR because the facet is extracted from the elements contained in

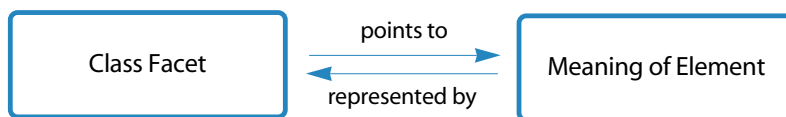


Fig. 1 Relationship Between Class Facet and Corresponding Meaning of Element

those structures. This process is optimized to make elements in a bibliographic structure from different structures interoperable by identifying and linking elements with the same concept, and thus to establish semantic relationships between those structures. Ideally, the faceted data model generates a conceptual structure that can connect elements from different bibliographic structures on the basis of semantic, syntactic, and structural similarities by identifying the conceptual orientation of seemingly disparate elements within a single, context-independent data model.

5. CONCLUSION

This research has constructed a faceted data model for integrating heterogeneous bibliographic structures such as MARC and FRBR. This model is not intended to describe any specific information resources but to provide a set of facets used in bibli-

ographic description. A set of semantics of elements in different structures was categorized according to their semantic similarities through the application of facet analysis.

This research specified two types of facet: class facet and property facet. Class facet was placed at the top of the structure and displays the framework of the faceted data model. Property facet, nested under each class facet, specifies the semantic range of the class facets.

These different types of facets are assigned to each element in both bibliographic structures based on the semantics of those elements. Through these assigned facets, the elements with the same facets can be connected with each other. This connection integrates both MARC and FRBR into the proposed faceted data model.

The faceted structure of the proposed data model also provides the capability of integrating semantic and structural interoperability by taking into account the contextual differences between MARC

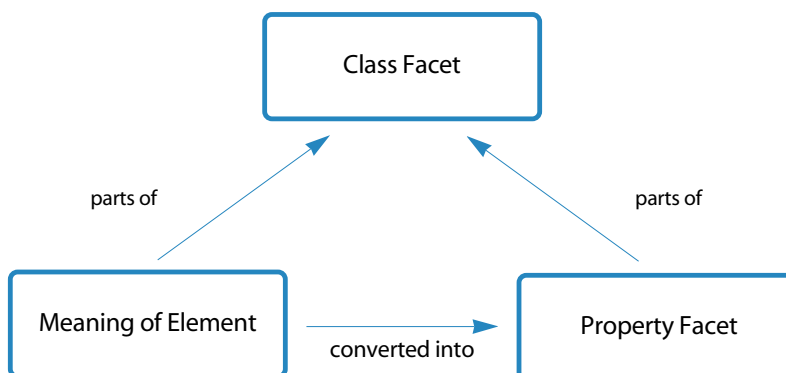


Fig. 2 Relationships Across Different Types of Facets and Meaning of Element

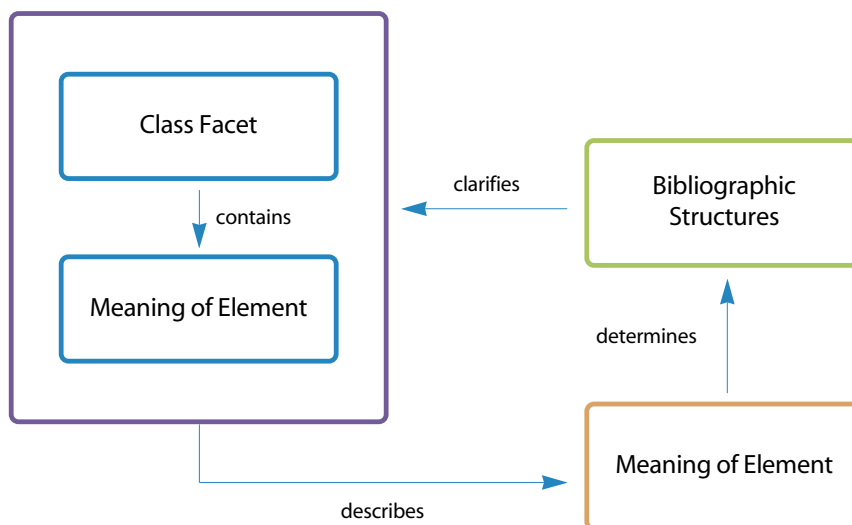


Fig. 3 Conceptual Structure of the Faceted Data Model

and FRBR. Thus, integration of heterogeneous bibliographic structures into this faceted data model can provide an alternative approach to utilize these structures in a more reliable and comprehensive way by reflecting the semantics of and the relationships between elements used in bibliographic structures.

REFERENCES

- Aalberg, T. (2005). *From MARC to FRBR: A case study in the use of the FRBR model on the BIBSYS database* [PowerPoint slides]. Retrieved from www.fla.fi/frbr05/aalberg2BIBSYSfrbrized.pdf
- Chan, L. M., & Zeng, M. L. (2006). Metadata interoperability and standardization - a study of methodology, Part I: Achieving interoperability at the schema level. *D-Lib Magazine*, 12(6). Retrieved from <http://www.dlib.org/dlib/june06/chan/06chan.html>
- Delsey, T. (1998). *The logical structure of the Anglo-American Cataloguing Rules-Part I, Drafted for the Joint Steering Committee for Revision of AACR*. Retrieved from www.rda-jsc.org/docs/aacr.pdf
- Harter, S. P. (1986). *Online information retrieval: Concepts, principles, and techniques*. New York: Academic Press.
- Hodge, G. (2005). Metadata for electronic information resources: From variety to interoperability. *Information Services & Use*, 25(1), 35-45.
- IFLA Study Group (1998). *Functional requirements for bibliographic records, final report*. IFLA Study Group on the Functional Requirements for Bibliographic Records.
- Kumar, P. S. G. (1987). *Introduction to colon classification*. Nagpur, India: Dattsons.
- Kurth, M., Ruddy, D., & Rupp, N. (2004). Repurposing MARC metadata: Using digital project experience to develop a metadata management design. *Library Hi Tech*, 22(2), 153-165.
- Moen, W. E. (2004). Metadata interaction, integration, and interoperability [PowerPoint slides]. In *National Information Standards Organization (NISO) Workshop: Metadata Practices on the Cutting Edge*, Washington, DC.
- Moen, W. E. & Bernardino, P. (2003). Assessing metadata utilization: An analysis of MARC content designation use. In *Proceedings of International*

- Conference on Dublin Core and Metadata Applications, DC-2003*, Seattle, WA, (pp. 171-180). Retrieved December 12, 2012 from <http://dcpapers.dublincore.org/ojs/pubs/article/view/745/741>.
- Priss, U., & Jacob, E. K. (1998). A graphical interface for faceted thesaurus design. In *Proceedings of the 9th ASIS SIG/CR Classification Research Workshop*, (pp. 107-118). Silver Spring, MD: American Society for Information Science.
- Ranganathan, S. R. (1962). *Elements of library of classification*. Bangalore: Sarada Ranganathan Endowment for Library Science.
- Ranganathan, S. R. (1967). *Prolegomena to library classification*. New York Asia Publishing House.
- Svenonius, E. (2000). *The intellectual foundation of information organization*. (3rd ed.) Cambridge, MA: MIT Press.
- Taylor, A. G. (1992). *The organization of information*. Englewood, CO: Libraries Unlimited.
- Tennant, R. (2004). A bibliographic metadata infrastructure for the 21st century. *Library Hi Tech*, 22(2), 175-181.