2017

Music Social Tagging as a Validation Tool for the FRBR Conceptual Model

Kabel Nathan Stanwicks
University at Albany, State University of New York, kstanwicks@albany.edu

MUSIC SOCIAL TAGGING
AS A VALIDATION TOOL
FOR THE FRBR CONCEPTUAL MODEL

by

Kabel Nathan Stanwicks

A Thesis

Submitted to the University at Albany, State University of New York

in Partial Fulfillment of

the Requirements for the Degree of

Master of Science

College of Engineering and Applied Sciences

Department of Information Science

2017

Music Social Tagging as a Validation Tool for the FRBR Conceptual Model

by

Kabel Nathan Stanwicks

Table of Contents

Abstract

The International Federation of Library Associations developed its Functional Requirements for Bibliographic Records (FRBR) without performing user testing to ensure that the model would meet the needs of professionals and users. Analyzing user-generated social tags in relation to FRBR entities and attributes will help determine if the layperson describes objects in a manner that conforms to the FRBR conceptual model. Number one songs from the weekly Billboard Hot 100 charts from 1958 through 2013 were randomly sampled, tags associated with the sampled songs were pulled from the last.fm Web site, and tags were analyzed to determine their relation to FRBR entities and attributes. A percentage of tags map to FRBR entities, validating those entities' place in the model. User tags that do not map to a FRBR entities shed light on additional means for resource access and discovery, as well as potential issues to consider should the FRBR conceptual model be revised in the future.

Keywords: Functional Requirements for Bibliographic Records, social tagging, information organization, music

Introduction

The International Federation of Library Associations (IFLA) developed its Functional Requirements for Bibliographic Records (FRBR) conceptual model in 1998. Due to both time constraints and financial constraints, leading experts developed the model without performing user testing with library professionals or library users to ensure that the model would meet the needs of professionals and users alike. Nearly twenty years after FRBR was published, information professionals and scholars continue to struggle with the concepts presented in, and the implementation of, the FRBR conceptual model. Few researchers have sought to examine the degree to which the FRBR conceptual model aligns with users' mental models. Developing an understanding of how these intersect can aid in validating the FRBR model, designing search and discovery interfaces, and developing metadata standards that meet users' needs.

Can social tagging provide valuable insight into how the layperson categorizes Web resources and objects? Do user-generated tags conform to the FRBR conceptual model? Can social tagging assist information professionals in understanding ideas and concepts that people associate with, and use to describe, objects? Analyzing user-generated social tags in relation to FRBR entities and attributes can help determine if the layperson describes objects in a manner that conforms to the FRBR conceptual model. This study examines the extent to which user-generated tags for sound recordings map to the FRBR model, and with which FRBR attributes user-generated tags align. This study's results help validate portions of the FRBR conceptual model, suggest improvements to the FRBR model, and provide recommendations for improving music resource description.

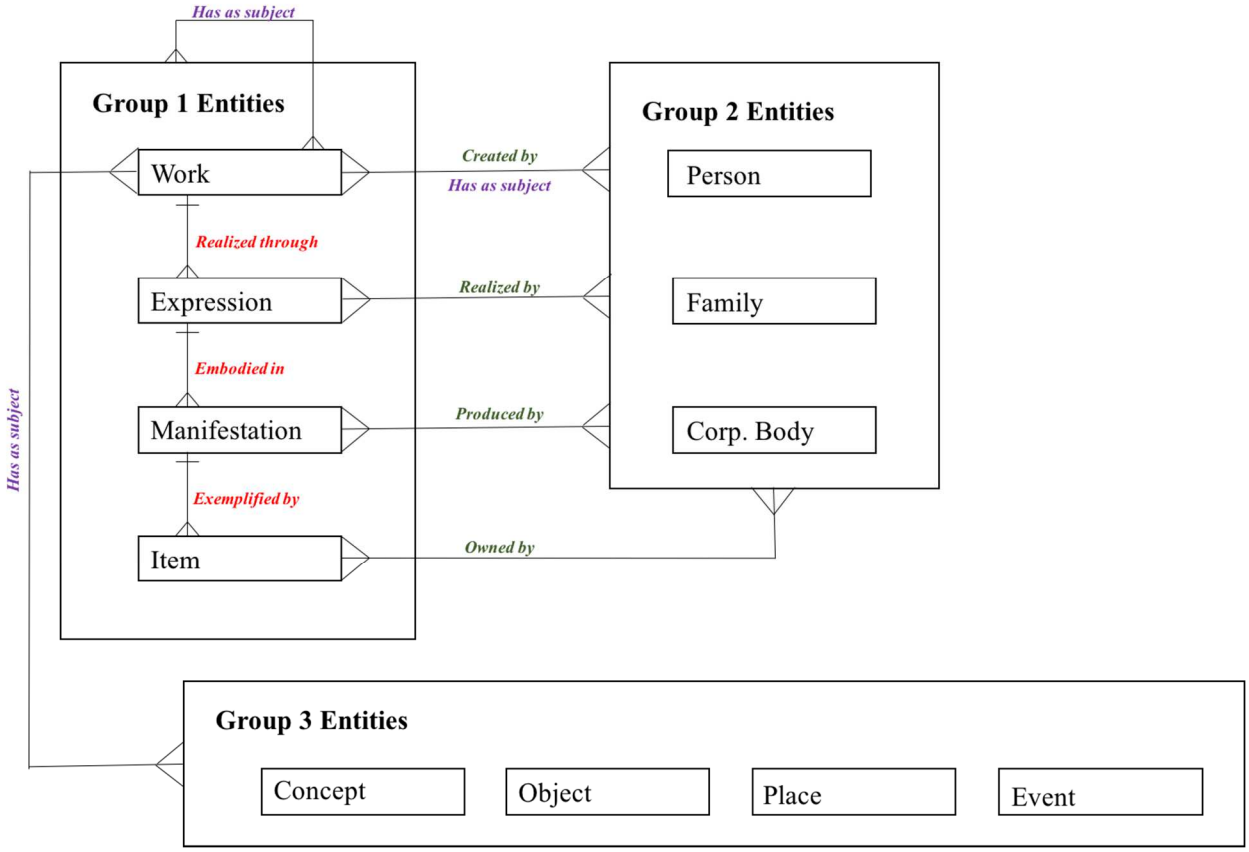
Literature Review

Functional Requirements for Bibliographic Records

The FRBR study's purpose—and subsequently FRBR's purpose—was "to produce a framework that would provide a clear, precisely-stated, and commonly-shared understanding of what it is that the bibliographic record aims to provide information about, and what it is that we expect the record to achieve in terms of answering user needs (IFLA Study Group on the Functional Requirements for Bibliographic Records [IFLA], 2008)." FRBR's two main objectives were to establish a framework to connect bibliographic record data to users' needs for those records, and to recommend basic functionality for bibliographic records that are created by national bibliographic organizations (IFLA, 2008). Tillett (2005) describes FRBR as "a conceptual model of the bibliographic universe." Tillett (2005) goes on to explain that FRBR is needed to improve users' abilities to navigate the bibliographic universe, and to provide data interoperability to libraries, museums, and other information organizations to enable these organizations to share their descriptive metadata. What came out of the FRBR study group is an entity-relationship model for what Albertsen and van Nuys (2005) refer to as real-world objects, as opposed to object descriptions. The FRBR model consists of three entity categories: Work, Expression, Manifestation, and Item entities are categorized as Group 1 entities; Name entities are categorized as Group 2 entities; and Subject entities are categorized as Group 3 entities (Figure 1). FRBR outlines four user tasks—find, identify, select, and obtain—for bibliographic records. Patton (2005) explains that there are four user tasks related to authorities records (i.e., name and subject entities) within FRBR—find, identify, contextualize, and justify. Le Boeuf (2005a) explains that Lubetzky and Ranganathan developed concepts that were FRBR-esque, and Heany and Serrai had object-oriented concepts that are also pre-cursors to FRBR; however,

the FRBR model has momentum where these other concepts did not because FRBR is backed by IFLA.

Figure 1: FRBR Entity-Relationship Model



Machine Readable Cataloging records, or MARC records, allowed libraries to separate some of the bibliographic record's components, and the FRBR model takes this even further by using an entity-relationship model (Howarth, 2012). Taylor (2007) explains that the FRBR model entities are divided into three groups, provides an overview of the FRBR model's entities and attributes, and includes detailed definitions and examples of each entity. Group 1 entities (i.e., Work, Expression, Manifestation, and Item) are the "products of intellectual or artistic endeavor;" Group 2 entities (i.e., Person, Corporate Body) are those "responsible for content, production, or custodianship of Group 1 entities;" and Group 3 entities (i.e., Group 1 and 2

entities, Concept, Object, Event, Place) are subjects for Group 1 entities (Taylor, 2007). The FRBR model's relationships link entities from different groups, different entities within the same group, and entities that are the same type to help users search for and discover resources in bibliographic databases (Žumer & Riesthuis, 2002). Albertsen and van Nuys (2005) explain that the Group 1 entity abstraction hierarchy focuses on a bibliographic set's common properties in that a Manifestation has common properties for all its items, an Expression has all the common properties of its Manifestations, and a Work has all the common intellectual properties for its Expressions. The FRBR model and uniform titles place the Work at the center of the bibliographic record with Expressions, Manifestations, and Items around the central Work (Le Boeuf, 2005b). Works can have multiple Expressions, and Expressions can have multiple Manifestations within the FRBR model (Albertsen & van Nuys, 2005).

The concept of the Work is central to the FRBR model; however, this concept is not original to FRBR. IFLA (2008) defines a work as "a distinct intellectual or artistic creation." Žumer and O'Neill (2012) look at the term's history in bibliographic resources, stating that Sir Thomas Hyde and Panizzi are among the earliest known users of the term "work," while Lubetzky and Smiraglia go to great extents to define the term. However, FRBR's definition and concept of work are vague and abstract (Le Boeuf, 2005b; Žumer & O'Neill, 2012). Le Boeuf (2005b) states that different cultures may view the distinction between one work and a new work differently because the FRBR final report's Work concept is abstract; therefore, national bibliographic conventions will be left to determine the boundaries between works, which may lead to variation in how different cultures handle the work concept. Žumer and O'Neill (2012) write that the varying definitions of the work concept also lead to confusion with aggregates—it

is not clear if the work is the combined grouping in an aggregate or remains the individual components. This is also an issue with individual songs that are compiled on a sound recording.

Some researchers have attempted to eliminate the ambiguity in FRBR's definition of a work. Smiraglia (2007a) posits that it is the Superwork that is essentially the FRBR Work entity because the Superwork is the concept that is then realized. Žumer and O'Neill (2012) believe that FRBR's Work definition can be improved by defining a work as "the smallest distinct and autonomous entity." Žumer and O'Neill (2012) also provide criteria for assessing a work's autonomy, stating that an autonomous work must meet at least one of the following: it has its own identity, it is created independently, it is able to stand alone, or it has distinct intellectual property rights. Le Boeuf (2005b) feels that FRBR's definition of a work is misleading, and highlights how the FRBR definition aligns with Umberto Eco's view of a work, "one 'thing' that remains 'almost' the 'same,' beyond all of the different ways it can be 'uttered.'" The author likens the work concept to a solar system with various iterations of a work gravitating around a sun, and interrelated solar systems working in harmony within a galaxy. This analogy illustrates the complex nature of the work concept.

The music domain highlights many of the problems with the FRBR model's definition of the work, despite how FRBR addresses several bibliographic issues in the music domain. Le Boeuf (2005a) contends that the FRBR model's work concept is poorly defined because the FRBR definition focuses primarily on books; however, the FRBR model is intended to handle a wide variety of formats. A work's dependence or independence within the FRBR model is determined by whether it has a distinctive name or title, but this dependency test is problematic for larger works that consist of components that can stand on their own (Vellucci, 2007).

Vellucci (2007) further explores the idea of dependency by examining the interaction between music and text in songs. If a song's music and lyrics are written by different people, what is the primary work—the music or the text? Furthermore, does the combination of music and text form a unified work or aggregate works? These questions lack straightforward answers, and the answers most likely vary from song-to-song. Le Boeuf (2005b) questions how FRBR handles song lyric changes, "does the substitution of new lyrics in a preexisting song result in a new instance of the Work entity, or in a new instance of the Expression entity linked to the same Work entity?" The author cites the songs *America* and *God Save the King* as examples of songs that have the same music but different lyrics. FRBR has the potential to outline these relationships, which are not present in library catalogs, but the model requires clarity to do so. Miller and Le Boeuf (2005) explore a similar issue with mises-en-scene and choreographies and advocate for treating these two work forms as autonomous works, rather than taking the approach used for libretti and treating mises-en-scene and choreographies as second-tier works. The authors write, "Creations of 'mixed responsibility' can be considered as entities related to, but independent from, given source material, especially textual drama and preexisting musical Works (Le Boeuf, 2005)." Nicolas (2005) explains that Oral Tradition Work (OTW) does not fit within the FRBR work concept. The author explains that a work is defined in OTW as "content shared by all of its various versions and all of the performances that embody it." Nicolas (2005) goes on to explain that different OTW versions are part of the same work because versions must be versions of an original work. The author further posits that the FRBR model's work definition hinders OTW description because different versions of a work should fall under a common work for collocation purposes; however, the FRBR model treats OTW version relationships as derivative relationships.

Smiraglia (2007a) examines derivative bibliographic relationships, which the author defines as bibliographic relationships among works with a common progenitor. Examples provided include simultaneous editions, successive editions, predecessors, amplifications, extractions, accompanying materials, musical presentation, notational transcription, persistent work, translations, adaptations, and performances. The author explains that these types of works fall under the FRBR expression entity, and, therefore, are not considered works according to FRBR's work definition. The persistent works used as one example of derivative bibliographic relationships are defined as bestsellers that appear in new editions over time. In a sense, many of the other examples of derivative bibliographic relationships provided either result from, or lead to, a work's status as a persistent work. Smiraglia (2007a) also presents the term *instantiation*—the "realization of a phenomenon in time" and "a concrete exemplar of a work as it has appeared at a specific point in the lifetime of the work"—as a simplified way to examine derivative bibliographic relationships spun off from a work. Smiraglia (2007b) explains that instantiation networks form when information objects permeate a culture—more popular works are likely to be subject to instantiation.

Smiraglia (2007b) refers to derivative bibliographic relationships with a new term—the works phenomenon. The author asserts that previous methods for clustering and collocating works in library catalogs—like the uniform title—do not provide users with clear choices from which to select an appropriate instantiation of a work. Smiraglia (2007b) presents a study using persistent works to demonstrate the need for more robust clustering mechanisms. The author generated a sample of best-selling titles for 1900–1995 by taking the 10 best-selling fiction titles and 10 best-selling non-fiction titles for each year of this time frame from *Publisher's Weekly*. The author pulled bibliographic records for each work from OCLC, RLIN and the Web, analyzed

the records, labeled each record with derivative relationship types, and used SPSS to perform quantitative data analysis. The author categorized derivative relationships in this study using relationship types (i.e., simultaneous derivations, successive derivations, translations, amplifications, extractions, adaptations, performances) developed for one of his previous studies on derivative works and two new relationship types (i.e., predecessor, accompanying material) created specifically for this study. Smiraglia (2007b) found 100% of the fiction works were subject to derivation, and 97% of the non-fiction works were subject to derivation—only one work in the sample was not subject to derivation. The author also observed that the most popular best-sellers had larger derivative bibliographic networks. Smiraglia (2007b) concludes that an instantiation of works over time theory can be developed and used to design search interfaces for retrieval of works. The author claims that this study shows instantiation's increased prevalence over earlier studies; therefore, more complex uniform titles or FRBR-based linkages should be developed to improve search and discovery for these resources.

Bennett, Lavoie, and O'Neill (2003) studied the occurrence of derivative works in a bibliographic database by extracting and analyzing a sample of bibliographic records from WorldCat. The author refers to derivative works—or those works with multiple expressions—as complex works. The author also provides terms and definitions for two other types of works—elemental works are works with a single expression and single manifestation, and simple works are works with a single expression and multiple manifestations. The bibliographic data extracted for this study were divided into these three work categories. Records labeled complex works were further categorized into six subcategories—augmented, revised, collected/selected, multiple translations, multiple forms of expressions, and multiple forms of translations with multiple forms of expressions—with the author noting that these subcategories are not mutually

exclusive. This study found that these complex works, or derivative bibliographic relationships, are not common. Most sampled works (78%) were elemental works, and complex works accounted for a small number (6%) of the sampled works. A different study applied an algorithm to WorldCat records and found similar results with approximately 96% of records categorized as distinct works—works with a single expression and manifestation—synonymous with what Bennett et al.'s (2003) study referred to as elemental works (Hickey & O'Neill, 2005).

Authors look at the various types of component works found in the bibliographic universe. Le Boeuf (2005b) writes that a sum of works can also be a work. In other words, an aggregate of works can be considered a work in its own right. However, Vellucci (2007) highlights the ambiguity and disagreement surrounding the aggregate work concept—some consider an aggregate work a manifestation containing works, while others believe aggregation creates a new work because of the intellectual effort that goes into aggregating resources. Le Boeuf (2005b) takes Smith and Varzi's (1997) terminology for physical and human-generated geographic boundaries—*bona fide* and *fiat*, respectively—and applies these terms to musical works in exploring whole-part relationships and component works. Le Boeuf (2005b) refers to a complete symphony as a *bona fide* work, and the author refers to the symphony's individual movements as *fiat* works. This analogy attempts to highlight the boundaries within a work that composers and authors create to break up a larger work. Individual chapters within a book might provide another example for this analogy. Vellucci (2007) also examines whole-part relationships and aggregated works based on how the FRBR model handles dependencies. Vellucci (2007) contends, "it is obvious that the individual pieces in a musical anthology have an independent part/whole relationship to the anthology," and an individual piece in an anthology, "continues to stand as an independent work."

Le Boeuf (2005b) and Vellucci (2007) look at works and their component parts from two different perspectives by examining different-natured component works. Le Boeuf (2005b) claims that a component part of a work is a work, citing Wagner's Ring as an example. This example—Wagner's Ring—groups multiple complete operas into a larger work; a film trilogy or novel series works in the same way. Vellucci (2007) argues that larger musical works, such as operas and symphonies are integral units with component parts—the component parts being individual arias or movements in this case. However, the author also addresses how creating aggregate works out of these component parts leads to complications in works' structural relationships within the FRBR model. This demonstrates that context is important when examining component works, and a component part's status as a work can change based on context. A unit within a work might not be a work in and of itself when examining the larger work; however, this may change when looking at that same unit as part of an anthology or aggregate work.

Nicolas (2005) writes that FRBR's Expression entity is like the concept of versions, and Kilner (2005) recommends that the term 'version' replace the term 'expression' in the FRBR model. Kilner (2005) states that implementers view the Expression entity as the most difficult FRBR entity to implement. Nicolas (2005) explains that people have difficulties with the Expression entity because it is so flexible—not because of flaws—and advocates for implementers and users to sort out how the Expression entity is defined.

The problems with the Expression entity largely revolve around the degree to which expressions differ from the original work. The problems with the flexibility afforded to the Expression entity are illustrated by what the FRBR final report refers to as Variant Expressions. IFLA (2008) states that bibliographic distinctions between variant expressions is dependent on

the nature of the work. Therefore, the Expression entity lacks solid criterion for determining if a Work's Expressions constitute new, or separate, Expressions. On the other hand, IFLA (2008) indicates that a modified or revised text is considered a new Expression; however, the AustLit Gateway—a FRBRized database for Australian Literature—takes a different approach to this. AustLit Gateway does not necessarily view a minor text change to be considered a new Expression. An Expression is only considered new if the Expression's changes affect the Work's reception or meaning (Kilner, 2005). Jonsson (2005) argues that Expression records are good to have for clearly identifiable Variant Expressions, but these records are probably of little use to users.

Žumer & Riesthuis (2002) and Le Boeuf (2005b) criticize the FRBR model for its massive departure from traditional cataloging practices. Le Boeuf (2005b) explains that FRBR is a top-down structure, but library catalogers approach resource description from using a bottom-up approach. Žumer & Riesthuis (2002) expand on the difficulties experienced by librarians, and opine that the FRBR model is difficult for traditionally-trained cataloging librarians, citing issues with the model's terminology and entity definitions. Le Boeuf (2005a) questions whether users will find the FRBR model difficult to use, given the model's difficulties as perceived by librarians and practitioners.

Howarth (2012) summarizes the history of library catalog records—cataloging started as a manual process in card catalogs and microfilm catalogs, then computers brought about library automation and the introduction of the online public access catalog (OPAC) and later Web OPAC. The OPAC was thought to have "liberated" data from the card catalog (Howarth, 2012). Despite "liberating" catalog record data in certain ways, the OPAC and existing cataloging codes still left many constraints (e.g., main entry, quantity limits on subject terms) on catalog record

data. Implementing the FRBR model in library catalogs could remove many of these remaining constraints and further liberate catalog record data.

But what will it take to implement the FRBR model in library catalogs, and can this be accomplished with current cataloging practices and technology? Some believe that implementing the FRBR model is feasible, while others feel new approaches are needed for an effective implementation. Bennett et al. (2003) reason that implementing FRBR in library catalogs might not be burdensome because their study in applying the FRBR model to WorldCat found that approximately only 20% of WorldCat records would be candidates for FRBRizing; furthermore, they state algorithms can be used—and will be helpful—in implementing FRBR and FRBRizing catalog records. Gradmann (2005) argues that libraries cannot implement the FRBR model with current practices and should migrate to standard internet technologies for catalogs to implement FRBR. The author advocates for expressing FRBR using a Resource Description Framework Schema (RDFS) model and implementing Resource Description Framework (RDF) catalogs to align with the Semantic Web, and outlines the benefits to taking this approach when implementing the FRBR model. An rdfs:frbr model will allow library catalogs to become Web-transparent and reduce data redundancy, simplify cataloging work through inference-based models, provide libraries with Web-search exposure, and provide integrated library system vendor independence (Gradmann, 2005).

Cataloging standards enable efficient cataloging workflows and bibliographic record sharing (Hoffman, 2010; Howarth, 2012). Howarth (2012) explains that IFLA and national bibliographic agencies have had to adjust cataloging codes and standards when MARC and its national derivatives allowed libraries to share bibliographic records, and argues that cataloging code revisions and new standards are independent of cataloging technologies and systems.

However, the FRBR model has the potential to enhance libraries' abilities to share bibliographic data and records on an international scale not yet seen. Tillett (2005) writes about the FRBR model's potential impact on resource description and access, stating, "We are moving towards an era when we have the ability to share and re-use bibliographic descriptions created anywhere in the world and to tie the bibliographic descriptions with real access, so users can obtain the resources they want." Just as MARC led to major code revisions, FRBR requires code revisions and collaboration on a much greater scale to reap all the FRBR model's benefits.

Zeng and Žumer (2010) examine the Functional Requirements for Subject Authority Data (FRSAD) model in relation to other knowledge organization models to examine FRSAD's potential for international data sharing in the future. The authors provide an overview of the FRSAR working group's work. The FRSAR working group looked at enhancing the FRBR Group 3 entities, considered other approaches and models for providing subject access, and focused on developing a new conceptual model to address works' "aboutness" within the FRBR model. The authors explain that the working group's model allows FRBR Group 1, 2, or 3 entities to have a subject relationship with the work. The Simple Knowledge Organization System Reference (SKOS) uses labels to represent concepts, and Zeng and Žumer (2010) assert that the FRSAD thema and nomen entities are in line with this SKOS concept. The authors write that OWL Web Ontology Language provides comprehensive semantic relationships between the concept of themas that FRSAD tries to convey. The DCMI Abstract Model employs a one-to-one principle for describing objects, and Zeng and Žumer (2010) state that FRSAD allows thema to be independent of nomen as does the DCMI Abstract Model. Zeng and Žumer (2010) believe that the FRSAD conceptual model will facilitate international subject authority data sharing

because it takes into consideration the constructs of other knowledge organization models and standards.

Some see the FRBR model's potential for clustering search results by the model's entity relationships as a boon to library catalogs' usability. This demonstrates FRBR's user-focused nature, despite the absence of user studies that went into developing the FRBR model. OPACs display bibliographic records in unorganized search results lists, which do not convey the relationships among materials in the search results (Carlyle, 1999; Le Boeuf, 2005a; Svenonius, 2000). Svenonius (2000) believes that FRBR can provide a solution to this issue. Carlyle (1999) explains that research shows that clustered search results may help users identify needed resources, and researchers have proposed clustering by subject or work to assist users in navigating search results lists. Carlyle (1999) also points out that Lubetzky advocated that library catalogs should show relationships between a work's versions; however, organized displays still have not been incorporated into OPAC design. Arastoopoor et al. (2011) rationalize that digital libraries need meaningful search results displays that can show how materials relate to larger bibliographic families because digital materials cannot show relationships through physical collocation like materials on a shelf; therefore, a FRBR-based search results display must "…employ an in-depth hierarchical structure based on specifications of the bibliographic family" to be meaningful to users. FRBR is poised to accomplish this since it is a hierarchical entity-relationship model.

Ercegovac (2006) looks at science fiction materials in four collections, bibliographic relationships in a sample of entities within these collections, and how to express bibliographic entities and relationships using current cataloging standards. The author claims that library catalog users have difficulty reviewing long displays of search results; therefore, reorganizing

catalog search results might make it easier for users to sift through lengthy search results without burying relevant results. To accomplish this, library catalogs could break up large results lists into smaller, meaningful groupings by taking advantage of data available in MARC records. The author suggests system design should incorporate different views to support different user communities. This study examined FRBR Group 1 entities as defined in the FRBR study, Tillett's taxonomy of content relationships, and Smiraglia's modification of Tillett's work. The author used descriptive survey methodology to examine catalog records for Abbott's *Flatland* in three research libraries' catalogs—UCLA, University of California-Melvyl, and Library of Congress—and OCLC WorldCat. The catalogs were searched in March 2003 and again in April 2004 using an author-title search. The Library of Congress' FRBR display tool was used on records retrieved from the Library of Congress' catalog. 86 records found across the different catalogs were exported, coded to include currently used MARC tags, and augmented with new fields (e.g., parent) to aid analysis. The parent field produced results that showed expressions and manifestations produced by the parent records. The author concludes that user interfaces need to provide guidance in searching and browsing materials based on explicitly expressed relationships, which can then be applied to user interfaces in global digital libraries.

Hoffman (2010) references Cutter's user-focused principles related to cataloging, which state that the user's convenience is more important than the cataloger's convenience. To what extent do library cataloging codes and standards place this focus on the user? The FRBR model is intended to be user-focused, but how was it developed to ensure the user's convenience? Hoffman (2010) writes that cataloging has been a standards-based practice but has recently shifted its focus from standards to users; however, the author argues that, "Catalogers are told to meet users' needs, but they are never told how to do it."

Requesting catalogers to focus on users is, in part, complicated by the lack of user studies completed when crafting cataloging codes and standards. Hufford (1991) reviews the history of cataloging codes and points out the lack of empirical testing in developing these codes, and suggests, "Though authors of past codes had the catalog user in mind when creating their codes, there is no evidence of an empirical study of use of data elements in the bibliographic record affecting any of these codes." Multiple authors argue that cataloging standards are not user-focused because most cataloging research does not focus on users (Borgman, 1996; Carlyle, 1999; Hoffman, 2010). Carlyle (1999) also addresses the lack of research into user behavior, and explains that this research void has led to the design of systems that are not in line with user perceptions. Hoffman (2010) explains that librarians and researchers want cataloging codes to be user-focused, but these codes fall short of that goal, are criticized for their biases, and do not adequately reflect their users' cultures. FRBR is cited as another model that is claimed to be user-centered, despite not studying users while developing the model (Hoffman, 2010; Madison, 2000).

Practitioners are making efforts to meet users' needs while adhering to standards, but researchers need to make conscious efforts to study users when revising cataloging codes and standards. Hoffman (2010) examined how cataloging units balance standards with users' needs and found that they achieve this by following cataloging standards with the users in mind because the standards do not meet the users' needs. Hufford (1991) reasons that catalog use studies are necessary when revising cataloging codes. However, this has not happened despite warnings of this nature over the years. Carlyle (1999) states that library catalog search result displays should be generated around users' needs.

Riva and Oliver (2012) write that RDA is in alignment with FRBR and FRAD; however, RDA is not an implementation of these two conceptual models. The authors examine how RDA converges with, and diverges from, the FRBR and FRAD conceptual models. They explain that the RDA and FRAD bibliographic entities' fundamental concepts are the same, and RDA's attributes can be traced to both the FRBR and FRAD conceptual models. They state that FRBR and FRAD's Group 1 entity relationships are organized slightly differently, and RDA draws its Group 1 entity relationships from both models. The authors indicate that user tasks for bibliographic data are essentially the same—obtain access to a resource—under FRBR and RDA, as well as FRAD and RDA. However, RDA provides simpler definitions than FRBR and FRAD. The authors explain that RDA does not define users, FRBR addresses users as anyone using bibliographic records, and FRAD defines two sets of users—those creating authority data and those using the data. The authors confirm that RDA's bibliographic entities line up with those in both FRBR and FRAD; however, there are 5 FRAD elements—names by which entities are known, identifiers, controlled access points, rules, and agency—that do not appear in FRBR. They also maintain that attributes among RDA, FRBR, and FRAD are similar, but RDA offers a higher level of granularity in its attributes. Riva and Oliver (2012) conclude that RDA mirrors FRBR and FRAD in the models' entities, attributes, relationships, and user tasks; furthermore, RDA pulls together the two models where FRBR and FRAD diverge in their treatments of Group 1 entity relationships.

Hider and Liu (2013) explore how RDA elements support the FRBR model's user tasks. The authors explain that the RDA, FRBR, and FRAD models address certain user tasks that users should be capable of completing with a catalog, and RDA's user tasks are grounded in those of FRBR and FRAD. They also state that RDA allows for the use of 463 elements and sub-

elements in bibliographic records and 59 elements in authority records; however, it is not yet clear the extent to which the required RDA elements support the user tasks defined in RDA, FRBR, and FRAD. The authors posted a questionnaire survey—querying users about typical bibliographic tasks performed in the catalog—to the library's Web site, and survey results were used to select representative user tasks. The authors used existing AACR cataloging records, which contained both required and optional RDA elements, in this study; they used think-aloud protocol to determine which elements participants used while searching the catalog. A total of 20 participants—6 expert catalog users from the library staff and LIS faculty were invited to participate, and 14 students recruited through flyers in the library—took part in the study. 10 tasks were given to the non-expert users and 20 tasks were given to the expert users. The authors used TechSmith's Morae software to capture the participants' actions in the catalog and their spoken explanations for what they were doing and thinking. The authors analyzed how frequently participants used particular elements in their searches and with which user tasks these elements were associated. The authors coded identification and selection tasks together, and they recorded both deliberate use and unintentional element use. Records used for this study were mapped to RDA because RDA elements were used for the study's data analysis. The authors collected 182 survey responses from the library's Web site, and specific item, subject and author/organization searches were ranked highly in importance. Respondents also ranked the usefulness of RDA-required elements and some non-required elements. Participants in Hider and Liu's user study completed 260 bibliographic tasks. The authors used 37 RDA elements throughout the user study, and they recorded the title proper as the most used element. The authors found that many RDA elements were not used, which they reported was not surprising. The authors found that element use between expert and non-expert users did not vary

significantly. The authors concluded that some non-required elements may be just as important as some required elements in fulfilling user tasks, and they noted that some elements are more useful to librarians and others are more useful to users.

Pisanski and Žumer (2010a; 2010b) performed a two-part study to determine if the FRBR model aligns with users' mental models. Pisanski and Žumer (2010a) posed the question, "Do mental models of individuals resemble the FRBR conceptual model of [the] bibliographic universe?" The authors provided participants with cards containing FRBR entities and asked participants to create at least three groupings with the cards. The authors performed cluster analysis of the participants' card sorts using the Ward method. Participants were provided two sets of cards—each set corresponded to a different Work—and the authors asked the participants to develop a derivation chain for the FRBR entities contained on their cards. Pisanski and Žumer (2010a) reported that participants were confused by the tasks they were asked to perform, participants' card sorts did not contain any overlap, and none of the participants' card sorts aligned with the FRBR model. Participants grouped expressions together more frequently than they did manifestations. The authors also developed concepts for the entities and asked participants to develop concept maps using one of the card sets from the sorting exercise—14/30 participants generated a FRBR work-expression-manifestation-item concept map. Pisanski and Žumer (2010a) concluded that study participants did not have a shared mental model of the bibliographic universe. The authors found that participants' models did not overlap, but some models were FRBR-like in nature, and participants tended to group the original Expression closely with the Work itself.

Pisanski and Žumer's (2010b) second study in mental models presented participants with eleven pairs of publications and asked questions related to each pair's similarity and

substitutability—similarity and substitutability were found to be closely related. They based their methodology in-part on Carlyle's (2001) research design. The authors asked participants to describe the books with which they were presented to determine if participants described materials in terms related to FRBR entities. The authors looked at elements that participants used to describe the books in this study and participants' rationale for associating specific elements with specific books. The authors address results related to FRBR aggregates, citing that participants paid attention to the manifestation's main work and did not recognize additional works. The authors note that participants found pairings of books and movies based on those books to be most dissimilar; and participants cited language, medium, and contents as elements that make materials not substitutable. Pisanski and Žumer (2010a; 2010b) conclude that FRBR is a viable conceptual model for cataloging, but a careful implementation is necessary because individuals' mental models differ from one another's and the FRBR model.

Pisanski and Žumer's (2012) user study seeks to verify the FRBR model. The authors used prominent bibliographic structures from their previous study on mental models, and presented these structures along with a list of bibliographic entities to 120 participants in this study. Study participants were asked to select the pre-determined structure that worked best with the bibliographic examples, or develop new structures if they felt that the structures presented did not work. The authors report that just over half (64 out of 120) of the participants selected the FRBR structure as an appropriate grouping for the bibliographic entities with which they were presented. The authors performed a chi-squared test to examine the probability that participants would equally select all structures presented, but this proved false, as the FRBR structure was chosen more frequently than the other structures. The authors conclude that FRBR provides a

good model, and—as with their previous study—they did not encounter an alternative model to FRBR based on their results.

Zavalina (2007) states that the Institute for Museuam and Library Services (IMLS) Digital Collections and Content project developed a metadata schema and created a collection registry of digital collections—including images, text, physical objects, sound files, interactive resources, moving images, and datasets—funded through the IMLS National Leadership Grant since 1998 and the Library Services and Technology Act since 2006. The author argues that the Gateway to Education Materials subject scheme is inadequate for indexing the registry. This paper examines similarities in user keywords pulled from the registry's transaction logs with the GEM, LCSH, and AAT controlled vocabularies. The author seeks to answer what is the degree of semantic similarity between users' keywords and controlled vocabulary terms, how does the FRBR set of entities compare to data in user searches, and what is the ratio of subject and known-item searches in the registry? Records were pulled from the transaction logs for this study, and the author manually extracted users' keyword search query strings from the transaction log. The author categorized user queries into seven FRBR search categories—work, person, corporate body, concept, object, event, and place. The author asserts that it was not possible to delineate precisely what users were searching for from the transaction logs; therefore, FRBR Expression, Manifestation, and Item entities were not examined in this study. The author mapped the user queries to subject terms in the three controlled vocabularies examined. The author was the only coder for this study, and states it was therefore not necessary to calculate intercoder reliability rates. Zavalina (2007) found nearly 75% of user searches fell within the object, concept, place, or individual categories, 67% of searches were known-item searches for particular collection titles, and unique search terms were used 79% of the time. In comparing

search terms to the three controlled vocabularies, the author found 2.2% of the queries matched unique search terms from GEM, 76.2% of terms matched LCSH, and 26.3% of search terms matched terms from the AAT. It should be noted that this study and its findings lack validity because content analysis was not completed by multiple coders with intercoder reliability testing performed across multiple coders' analyses.

Zavalina (2012) cites the quality of subject metadata, the application of controlled vocabularies in information systems, and subject heading structures as the main problems associated with subject access to information resources. The author states that cataloging principles— from Lubetzky's revision of Cutter's work through AACR2—have not addressed subject access, and the author contends that RDA, which is steeped in the FRBR model, finally attempts to address subject access but still falls short. The author explains that FRSAD introduces the thema entity, which is a superclass of all entities and can be the subject of a work, whereas FRBR Group 3 entities, which focus on subject, have limited definitions. The author thinks that relationships between FRBR entities other than works are potentially useful in signifying subjects, but are not included in the FRBR model. The author states that RDA diverges from the FRBR model in that it allows relationships within the FRBR Group 3 entities, but RDA does not prescribe how these relationships function.

Zavalina (2012) recommends examining how users perform subject searching to validate FRBR as an effective model. Zavalina's (2012) study analyzes searches from the Opening History digital library by comparing the searches with FRBR subject entities and relationships. This study pulled and analyzed one year's worth of individual search queries from the search log in light of the FRBR Group 1 Work entity. Two coders performed mapping, and intercoder reliability was calculated at 93.36% in Zavalina's (2012) study—making this a valid study,

whereas Zavalina's (2007) study lacked this validity. Zavalina (2012) reports that place and object—both FRBR Group 3 entities—constituted the top user search categories observed in this study, and the highest rates of entity co-occurrence in searches related to FRBR Group 3 subject entities. Zavalina (2012) concludes that FRBR-based search query categorization demonstrates the importance of subject searching, this study's results show the importance of relationships between subject entities, and the results can provide justification for developing these relationships within the RDA and FRBR constructs.

Bowen (2005) describes a user interface that was developed by University of Rochester staff to provide better access to a video collection and a music CD collection. This Web-based search tool used data from MARC records but allowed users to bypass challenging searches in the library's online catalog. MARC relator codes were used to generate FRBR Group 2 entities that could be searched in this user interface, which allowed users to search for videos by director and sound recordings by performer. Some catalog maintenance was required ahead of time to insert relator codes in records that were missing them. The user interface returns search results using a SQL script written in Perl to query the online catalog and then return relevant results to the Web user interface. The interface also provides language and genre browse options. The University of Rochester performed usability testing, which demonstrated users' difficulties in navigating the online catalog and databases. The library developed an interface using FRBR relationships to improve usability, make it easier for novice library users to find quality resources, and provide a tool for library staff to perform FRBR usability testing. The author concludes that FRBR has the potential to improve access to library resources, and states that catalog records already contain key data needed to implement FRBR capabilities; however,

vendors still need to redesign discovery interfaces to provide users with FRBR-inspired capabilities.

Pisanski and Žumer (2007) examine OCLC FictionFinder and LibraryLabs FRBR implementations. The authors explain that the LibraryLabs prototype was developed by the National Library of Australia and searches a portion of the 16 million bibliographic records in Australia's national union catalog. They state that the prototype's interface provides a link to a FRBR-grouping view from within a bibliographic record's display—the FRBR grouping is not displayed with the initial search results. The authors claim that LibraryLab's FRBR results presentation allows users to identify relevant results in less time. They state that superworks relationships are included in this prototype, but inheritance is not presented in the display. They also explain that LibraryLab's manifestation groupings are incomplete, which may lead users to a false sense of completeness in their searching. Pisanski and Žumer explain that OCLC FictionFinder provides access to 2.8 million fiction works from OCLC WorldCat; however, groupings and displays do not fully follow the FRBR model, and works results are ranked based on the number of copies held by libraries as indicated in WorldCat. The authors state that FictionFinder focuses on the Work and does not provide easy access to individual editions, it lacks options to search by Manifestation criteria, and search results lists are not adequately reduced when limiters are applied. They argue that this does not facilitate OPAC use based on user tasks, which is something that the FRBR model is supposed to address. The authors conclude that neither LibraryLabs nor FictionFinder implements the FRBR model as conceived by IFLA, and both interfaces produce inappropriate results due to poor-quality cataloging records. They assert that FictionFinder provides users with a better presentation than does LibraryLabs, and both interfaces are better than traditional OPACs.

Zhang and Salaba's (2012) IMLS-funded project evaluates existing FRBR-based catalogs, performs user studies on the FRBR catalogs, and develops and evaluates a FRBR prototype catalog. OCLC FictionFinder, WorldCat.org, and Libraries Australia were evaluated in this study. Zhang and Salaba (2012) reported that participants were successful finding works and sets of works across all three platforms, participants experienced difficulty finding expressions and manifestations; participants experienced difficulty identifying manifestations based on publisher data, and participants were not successful obtaining items. The authors state that participants found some of the FRBR-based catalog enhancements useful. Participants liked the easy-to-use designs, links to related items, flexible sorting options, objects' availability and holdings information, single search boxes, and easily understandable language. However, participants felt that grouping and displaying search results within categories, as well as providing links to related works, would enhance the catalogs.

Zhang and Salaba's (2012) project team designed a FRBR catalog prototype. They used users' evaluations from the three FRBR catalogs studied, and gathered input from a new group of study participants to implement FRBR features and develop a user interface for their prototype. Zhang and Salaba (2012) asked participants questions about the library catalog design based on FRBR layouts and illustrations. They found that 72% of participants had a positive impression of the prototype catalog's results display, and 28% of participants indicated that the results display needed improvements. However, the authors report that suggested improvements were not related to the FRBR implementation. They state that a majority of participants felt the prototype catalog would be more helpful than traditional catalogs, participants found it was easy to navigate, and participants understood the FRBR search results display.

Zhang and Salaba's (2012) final study evaluated the FRBR prototype catalog developed for this project. They recruited participants to perform title, author, and subject searches in both the prototype catalog and a traditional OPAC; and they gave participants searches to find, identify, and then select a work, an expression, and a manifestation. The authors report that participants preferred the FRBR prototype catalog over the traditional OPAC, and the participants were successful in completing the FRBR-based search tasks within the prototype catalog. Zhang and Salaba (2012) conclude that user studies are helpful in designing and developing FRBR-based systems, and these studies' results indicate that FRBRized displays support users' search tasks.

Metadata standards have been developed alongside bibliographic standards in response to the Web developing over time, and crosswalks have been developed to facilitate data sharing across metadata standards (Howarth, 2012). Researchers have examined how existing metadata standards align with the FRBR model to further explore interoperability between metadata and bibliographic standards, and have worked to develop FRBR-based metadata models. Arastoopoor, Fattahi, and Parirokh (2011) developed a mapping matrix to determine how UNIMARC, Metadata Object Description Schema (MODS), Encoded Archival Description (EAD), and Dublin Core (DC) metadata standards map to the FRBR model, as well as the extent to which each of these metadata standards meets the requirements outlined by the FRBR model. The study's results show that UNIMARC is good for Manifestation entities; MODS is good for Work entities; EAD has elements suitable for Work, Expression, Manifestation, and Item entities; DC has some elements suitable for Work, Expression, Manifestation, and Item entities; but none of these metadata standards fully maps to the FRBR model. Howarth (2012) discusses how the FRBR study group and the International Committee for Documentation (CIDOC) group

explored potential shared-perspectives for their two models. The CIDOC Conceptual Reference Model (CRM) is an object-oriented metadata model developed for the museum and cultural heritage communities, whereas FRBR is an entity-relationship model developed for bibliographic data. The FRBR-CIDOC group developed an object-oriented version of the FRBR model—referred to as FRBRoo.

Chen and Chen (2004) look into the feasibility of applying the FRBR model to the metadata standard used for the National Palace Museum in Taipei's Chinese painting and calligraphy collections. The authors state one can develop a metadata system by applying Heaney's object-oriented approach—formulated in 1995—with the FRBR model. The authors examine how to potentially implement a FRBR metadata model for the Chinese painting and calligraphy collection and other museum digital libraries. The National Palace Museum used the Categories for the Description of Works of Art (CDWA) metadata standard, and the authors compare this format to the FRBR model. They conclude that FRBR can provide a good framework for clarifying metadata elements and relationships; however, it does not adequately support management and workflow metadata.

Riley (2011) provides an overview of Indiana University's Variations projects. Variations Digital Music Library began as a streaming audio service at Indiana University in 1996. Variations2 was a grant-funded project to explore digital library system architecture, metadata, and usability, among other things. The author explains that Variations2 uses a work-based metadata model that resembles—but is slightly different from—FRBR, and Variations3 examined making the Variations software available to other institutions and developing a sustainable work-based metadata model. The author states that the Variations3 project concluded that music metadata models must be FRBR-based to meet user needs. The author also provides a

brief overview of the IMLS-funded Variations/FRBR (or V/FRBR) project, which was a response to the LC Working Group on the Future of Bibliographic Control's 2008 action item to create a test bed for the FRBR model. The author explains that the V/FRBR grant specified the activities for completion under this project were to convert the Variations data model to a FRBR-compliant model, make sure that MARC record use conformed to this FRBR-compliant model, load FRBR Groups 1 and 2 entities (and possibly Group 3 entities) for Indiana University's score and sound recording holdings into the redesigned system, make FRBRized records available via OAI-PMH, implement an openly-accessible search interface, implement a cataloging system for FRBRized data, and perform usability testing on user and cataloger interfaces.

Riley (2011) reports on the V/FRBR project team's progress. The author states that the project team investigated FRBR Group 1, 2, and 3 entities related to music, and decided not to model relationships between aggregate works and component works in V/FRBR because this information is represented in the Manifestation entity. The author explains that V/FRBR implements some of the concepts in FRAD, but does not address FRSAD concepts because music is rarely considered to be about something—music subject headings are instead used to provide genre information. The author states that XML was selected as the schema for encoding records in the V/FRBR implementation, a second-level XML schema called efrbr was developed to support production description and discovery systems, and a third-level XML schema called vfrbr was developed for music-specific issues (this removes FRBR attributes not required for music materials and adds necessary attributes not specified under FRBR). Finally, the author shares that the V/FRBR project team also made progress in developing an algorithm that converts MARC records into FRBRized records. Riley (2011) concludes that V/FRBR's strict

FRBR implementation is useful to the broader library community, given relatively few truly FRBRized systems in place.

Zhang and Li (2008) assessed two metadata schemas for moving image collections—Organizational Directory and the Union Catalog—in a three-part study to determine how well these schemas aide users in finding moving image resources, and to determine how well the FRBR user tasks work as an evaluative tool for assessing metadata schemas. The authors evaluated the metadata schemas using an online survey and a lab experiment, and they developed evaluative questions for their three studies using the FRBR user tasks. Zhang and Li (2008) completed the first two studies—usefulness assessment of Organizational Directory metadata schema, and usefulness assessment of Union Catalog metadata schema—immediately after the schemas were designed, but not yet used. They completed the third study—usefulness assessment of Union Catalog metadata schema pertaining to users' interactions with metadata records—after approximately 1,000 Union Catalog metadata records were created. The authors divided online survey forms and experiment questionnaires into sections based on the four FRBR user tasks, and they focused on how these tasks related to participants' ideas of useful metadata fields. Zhang and Li (2008) conclude that the FRBR user tasks provide a feasible framework for evaluating metadata schema, but they state that additional studies should be performed to verify their conclusion.

Some researchers believe that library metadata have the potential to enhance the Semantic Web and Linked Data, and a broad FRBR model implementation can facilitate library metadata harvesting for the Semantic Web. Palmer (2001) defines the Semantic Web as a mesh of information linked in a way that facilitates machine processing on a global scale. Tim Berners-Lee coined the term "linked data" to describe a method for publishing structured data,

which can then be shared openly on the Web (Howarth, 2012). Gradmann (2005) and Le Boeuf (2005a) claim that the FRBR model has a lot of potential for the Semantic Web. Tillett (2005) mentions that the FRBR model can be used to deploy a "one-stop-shopping" approach to Web search through the Semantic Web or other intelligent systems. Berners-Lee (2000) stated that Web catalogs and ontologies are useful for Semantic Web technology. Howarth (2012) explains that Libraries have large amounts of structured data that can be useful to Linked Data initiatives, and triples based on library data are valuable to the Semantic Web because libraries have better-quality data than those developed by machines and nonprofessionals. Gradmann (2005) reinforces this view, stating that libraries' data are useful for building ontologies, and the Semantic Web can make good use of these data.

Library catalogs are part of the hidden Web—not searchable through general internet search engines—and, therefore, born-digital content risks residing in obscurity if cataloged by libraries (Gradmann, 2005). Howarth (2012) writes that this situation is problematic because an increasing number of digital objects are being created; however, the FRBR model will be integral in creating digital object records because of FRBR's potential ability to expose these digital objects by interfacing with the Semantic Web. Gradmann (2005) states that the Web ignores libraries and builds functional models that bypass libraries. Linked Data has the potential to expose library metadata to the open Web (Howarth, 2012). However, implementing FRBR is essential in exposing library metadata in this way. Current cataloging codes and standards are partly the cause of bibliographic overlap, which would release large amounts of redundant data to the Web if the data were opened to the Web in their current state (Gradmann, 2005). FRBR can be tied into the Linked Data concept because it provides a means to pull data

out of the central bibliographic record and allows data elements to be linked together through its entity-relationship model (Howarth, 2012).

Choi (2012) looks into mapping social tags for Web resources to FRBR work and expression entities in order to determine the effectiveness of social tagging as a means for organizing Web content. The author posits that there are two problems with current knowledge organization systems—they are based on controlled vocabulary, which lacks the scalability necessary for providing access to resources on the Web, and information is not organized by users and is therefore not intuitive of user needs. However, the author suggests that social tagging can improve Web resource search and discovery. Sample Web resources that were tagged on Delicious and indexed in two subject gateways were selected at random for inclusion in Choi's (2012) study. The subject categories and tags associated with these resources were then extracted and tags were normalized. Non-English tags, personal tags, and subjective tags were removed from the list of extracted tags. The FRBR conceptual model was used to identify the tags' bibliographic attributes at the Work and Expression levels. Manifestation and Item levels were not considered since these were all Web resources. Two people coded and mapped tags to selected FRBR attributes for Work and Expression entities, and intercoder reliability testing was performed on their coding. Intercoder reliability results were strong. The author shares that 26% of the tags examined were subject terms, 27% matched FRBR attributes, and 47% of the tags were categorized by other attributes. Tags matching the FRBR work intended audience entity had high occurrences for subjects in technology, arts, and literature; tags matching the FRBR form of expression entity had high occurrences in terms related to natural sciences and geography. Other tag attributes—those not fitting within subject or FRBR categories—were sorted into feature, utilization, and institution subtypes, with the utilization subtype leaning

toward subjective and personal tags. Choi (2012) concludes that tags have essential attributes defined by the FRBR conceptual model. Overall, taggers assign FRBR-related tags that are related to the work title and work form entities; however, tags for certain subject-specific resources showed a prevalence in the work intended audience entity, which means that taggers in these fields are considering their audience when generating tags. The author explains that there was disagreement between coders related to tags for form of work and form of expression entities, and this coding disagreement could have been avoided if the FRBR model included mechanisms for describing digital heterogeneous media resources.

A lot of work went into developing the FRBR model, and FRBR provides potential improvements to working with bibliographic data. Some researchers and practitioners argue that the FRBR model is not complete and needs additional work, while others focus on addressing factors that impact the FRBR model. Zhang and Salaba (2009) found that their Delphi study participants ranked developing cataloging rules in line with the FRBR model as the most critical issue associated with FRBR and related standards. Tillett (2005) examines resources—such as international cataloging rules, a virtual international authority file, XML data packages—that need to be put in place to successfully implement the FRBR model. Researchers agree that additional user research is critical to enhancing and implementing the FRBR model (Tillett, 2005; Zhang & Salaba, 2009). Dimec, Žumer, and Riesthuis (2005) stress that IFLA's International Standard Bibliographic Description and the FRBR model need to be harmonized to address inconsistencies between the two. FRBR application guidelines and examples, as well as FRBRization tools for converting existing data and systems to correspond to FRBR, are also important to FRBR's development (Zhang & Salaba, 2009).

Researchers, who have examined and tested the model, have outlined a variety of recommendations for improving the FRBR model. Žumer & Riesthuis (2002) state that navigation is a function of the OPAC and is feasible only when data relationships are present and functional; therefore, they recommend adding Navigate as an additional user task within the FRBR model. Albertsen and van Nuys (2005) advocate for expanding the FRBR model to include aggregate modeling that includes structural properties for aggregates (i.e., existence, dependence, recursion, ordering, referencing) based on aggregate structures encountered while harvesting digital objects. Jonsson (2005) would like to see the Expression and Manifestation definitions allow for very slight variation of Expressions and Manifestations that are "substantially the same." Jonsson (2005) also recommends collocating search results by work then editions, rather than work then expressions, because this will produce more beneficial modeling for users. Hickey and O'Neill (2005) recommend uniform titles for revised works to aide algorithms in collocating original and revised works in FRBRized catalogs, and Dimec et al. (2005) believe uniform titles, in general, need more precision.

Peponakis (2012) scrutinizes FRBR Group 1 entity conceptualization, library metadata in RDA and FRBRization projects, and cultural heritage metadata. The author looks at FRBR Group 1 entities and the abstract concept of the Work introduced by the FRBR model. Peponakis suggests that libraries are not adequately describing changes to objects that occur as a result of digitization and other issues, and that the FRBR model may not provide this level of description. The author states that library catalogs continue to describe resources with their physical aspects, but users do not use that same vocabulary to describe materials for which they are searching. Peponakis asserts that RDA is still very object-oriented and is not abstract enough to envelop the FRBR model because RDA overlooks top-down description by focusing on the Manifestation,

rather than the Work, and does not offer additional specifications for how to interpret the Work concept. The author states that Manifestations and Items are already present in existing catalog records, and the concept of Work involves clustering author and title information; therefore, uniform titles are important as they are the only data in MARC records that come close to the idea of the Work. The author explains that FRBRization relies on the quality of existing bibliographic records and cataloging practices, criticizes the FRBR model for being developed without an accompanying machine-readable schema, and articulates that a machine-readable schema should be created to fit within the FRBR framework. The author also explains that the FRBR model addresses semantics but not syntax; however, syntax will eventually need to be addressed because diverse communities will need methods to exchange data. Peponakis also explains that FRBRoo—the collaboration between libraries and museums—considers the history and context of Manifestation and Item creation, whereas the FRBR model does not. The author concludes that cataloging remains object-centric and does not deconstruct resources into logical components that describe them separately and then bring them together when users issue requests; the FRBR concepts of Work and Expression are artificial constructs that do not have a physical status—this should be treated as knowledge related to a resource, which will then be helpful to the Semantic Web; and libraries need to abandon the traditional catalog in favor of a new tool that will communicate information and knowledge about Works in a broader sense.

Social Tagging and Controlled Vocabularies

Harping (2010) defines a controlled vocabulary as "an organized arrangement of words and phrases used to index content and/or to retrieve content through browsing or searching. It typically includes preferred and variant terms and has a defined scope or describes a specific domain." Controlled Vocabularies employ authorized terms to control use of synonyms,

eliminate homonymy, and control polysemy by uniting terms into broader and narrower categories, assigning terms to each subject, and referencing variant terms within a hierarchical, enumerative system (Macgregor & McCulloch, 2006; Quintarelli, 2005; Shirky, 2005). The hierarchical relationships within these systems are intended to be stable and predictable over time (Quintarelli, 2005; Shirky, 2005). Controlled vocabularies are utilized for resource description, which requires expert catalogers, authoritative sources, and expert users to properly function (Quintarelli, 2005; Shirky, 2005). Expert catalogers use precoordinate indexing when classifying and assigning authorized terms to information resources; therefore, during the cataloging process, a cataloger must consider the terms users might employ when searching for an information resource to ensure that the resource can be found by users (Smith, 2007).

Hierarchical schemes are useful for organizing physical materials, and can be necessary for physical organization given the human memory's limitations (Shirky, 2005). These schemes work well within domains with a relatively restricted body of work, pre-coordinated categories for organization, and stable and restricted items (Quintarelli, 2005). Libraries, archives, and museums are institutions that traditionally use hierarchical schemes, such as controlled vocabularies, to organize—and facilitate access to—large physical collections. Some online services have employed traditional, hierarchical organization and classification for digital resources, but Macgregor and McCulloch (2006) assert that these do not always work for online resource discovery. Shirky (2005) argues that the digital world lacks the physicality in which hierarchical organization and classification function well; therefore, the constraints imposed by these systems are no longer necessary in the digital world.

Library categorization and classification provide efficient and effective ways to describe and organize information resources, but these systems are far from perfect. Controlled

vocabularies work well in certain applications; however, they are not without fault, and a perfect categorization system does not exist (Shirky, 2005). Categorizing information resources in advance can be problematic because professional catalogers need to anticipate how users will think about concepts when searching, and then describe resources in line with these predictions (Shirky, 2005). Bias is an issue in library classification systems such as the Library of Congress and Dewey classification systems. Majority groups have largely created the terms and naming conventions for minority groups within classification systems. This is problematic when the majority groups' identity naming for minority groups may differ from how the minority group self-describes, which generates institutionalized power relations in libraries (Bates & Rowley, 2011). The Library of Congress Subject Headings have issues with currency, exclusion, and latency, and it also imposes its own set of socio-cultural perspectives on the resources it is used to describe (Bates & Rowley, 2011; Smith, 2007). Library classification systems indicate that an information resource is about one subject, regardless of how many topics the information resource may cover. Information resources are singularly classified according to their primary subject within library classification systems. Information resources can contain content covering multiple subjects; however, they are classified as being about a single topic (Shirky, 2005).

Social tagging—also referred to as collaborative tagging and user tagging—is an approach to information organization, in which users describe resources within an information system. The descriptors, or annotations, that users apply to items are called tags, and tags are used for resource organization, sharing, and discovery (Lorince, Zorowitz, Murdock, & Todd, 2015). Tags are usually single words, but some users employ multi-word tags by using symbols in place of spaces between words (Macgregor & McCulloch, 2006). The tags used in social tagging systems are uncontrolled, natural language descriptors in social tagging systems

(Macgregor & McCulloch, 2006). Furner (2008) defines tagging as annotating resources in a collection with terms that represent a tagged item's features for search and discovery purposes, and defines user tagging as tagging performed by a search and discovery service's users, rather than professional indexers.

Social tagging systems provide users with a space to organize their own data and others' data (Kroski, 2005). The tag aggregate within a social tagging system is a user-generated categorization structure that is referred to as a folksonomy. The term *folksonomy* was coined by Thomas Vander Wal to describe the type of "informal social classification" that was emerging through social tagging on Web sites like del.icio.us and Flickr (Vander Wal, 2007). Mathes (2004) and Chen, Liu, and Qin (2008) describe folksonomies and tags as user-generated metadata. Quintarelli (2005) describes folksonomies' organic nature in referring to folksonomies as user-generated classification that emerges from users annotating content with keywords. Folksonomies allow users to share their tags with other users and use tags created by other users (Spiteri, 2007).

Social tagging systems employ uncontrolled vocabulary to describe information resources, which is a frequent criticism of these systems. Knowledge and information organization professionals designate this as a weakness in social tagging systems. However, Rorissa (2010) suggests that this weakness is also one of social tagging's strengths because these uncontrolled descriptors are the terms that users associate with the objects they are describing. Other researchers concur with this sentiment, noting that social tagging's uncontrolled vocabulary reveals how people describe information resources (Kroski, 2005; Quintarelli, 2005; Spiteri, 2007). Tags can be more robust and descriptive than controlled vocabularies due to their uncontrolled nature (O'Connor & O'Connor, 1999). Social tagging has some advantages over

controlled vocabularies in that social tagging is inclusive, current, and inexpensive to maintain (Kroski, 2005; Spiteri, 2007). Social tagging systems also allow users to describe items with multiple tags, allowing items to be categorized within multiple subjects, which is not the case with controlled vocabularies and traditional classification (Kroski, 2005). The volume and variety of tags assigned to an object have the potential to aid resource discovery by increasing the number of entry points to a resource (Macgregor & McCulloch, 2006). These added access points can also be beneficial to browsing (Kroski, 2005; Rorissa, 2010). Multiple studies examined the uncontrolled vocabulary structures—tag, folksonomy, and system—that are found within social tagging systems. Shirky (2005) looked at the structures of image tags and index terms assigned to images in a general image collection; Spiteri (2007) analyzed tag structures in relation to the National Information Standards Oganization's standards for thesaurus construction; Chi and Mytkowicz (2008) provided an approach to understand how tags evolve within a social tagging system; and Lorince, Zorowitz, et al. (2015) investigated the assumption that folksonomies are crowdsourced or created by the masses.

Quantity and quality are strongly related in social tagging systems, with increased tag quantities leading to better system quality. Aggregation and scale are essential to a social tagging system's success (Guy & Tonkin, 2006; Quintarelli, 2005; Shirky, 2005). Resource description begins to emerge as more and more people apply tags to information resources within a social tagging system (Furner, 2008). This idea is demonstrated by the application of the power law concept to social tagging (Mathes, 2004; Rorissa, 2010). The power law concept shows that a small number of identical tags are applied to a resource by a large group of users—these are also referred to as high-frequency tags. A larger number of identical tags are applied to the same resource by a small group of users, and numerous, unique tags are applied to the same resource

by individual users. A tag that is repeatedly applied to an individual resource by a large number of taggers will have a higher social value and offer greater benefit within the tagging system (Meijas, 2004). Individuals will tag their own resources if tagging is useful to those individuals, and more tags will generate robust amateur classification over time (Shirky, 2005). This idea demonstrates how the power law and scale lead to common ground within a social tagging system—the high-frequency tags emerge as a resource's main descriptors. Shirky (2005) posits that this robust amateur classification is more valuable than professional categorization schemes.

While some level of agreement and emergence of main descriptors appears in social tagging systems, there is variation between tags applied by average users and domain experts (Lu, Park, & Hu, 2010). Lorince, Zorowitz, et al. (2015) examine how average users tend to employ basic-level descriptions in tagging systems, while domain experts use subordinate terms, which affirms the presumption that expert users produce better quality tags in a social tagging system. Improved quality in this sense refers to tags that move beyond basic-level description and drill down into domain-specific terminology to describe information resources. This behavior in experts versus average users has been observed across various tagging systems. Golder and Huberman (2006) analyzed tagging systems' structures and noted such regularities in user activity, tag frequencies, and types of tags used across different tagging systems. Lorince, Zorowitz, et al. (2015) look at the impact prolific, expert taggers—referred to as supertaggers—have on the folksonomy that emerges within individual social tagging systems. This study found that small groups of supertaggers generate the bulk of the tags in social tagging systems. The authors used an expertise measure to determine that supertaggers describe resources with more terms and with greater expertise than other users.

Resource popularity plays an important role in tagging behavior. Large numbers of tags are applied to popular resources, and less popular resources see more use after users tag them. Lorince, Zorowitz, et al. (2015) found that suptertaggers are more likely to tag less popular content and average users (i.e., non-supertaggers) are more likely to tag popular content in a social tagging system. The relationship between resources' popularity and tagging activity is explored in additional studies. Rolla (2009) asserts that tagging within library systems might be more beneficial in public library catalogs and with popular material because these items are more likely to be tagged by many users. Another study of tagging on the last.fm site looked at the relationship between tag frequency and listening activity for tagged resources. This study found that tagging an artist led to small increases in listening activity; furthermore, popular tags do not lead to large increases in listening activity, but less popular tags are associated with larger increases in listening activity (Lorince, Joseph, & Todd, 2015). Therefore, popular resources within a tagging system will have more tags applied, but one can conclude that many of these tags will be basic-level descriptions. On the other hand, supertaggers and domain experts apply more descriptive terms to resources, and resources see increased use when these additional, descriptive terms are applied to them.

Several studies examine how social tagging compares to controlled vocabularies. Spiteri (2007) examined the structure and scope of folksonomies to determine how folksonomies compare to controlled vocabularies. Smith (2007) compared tags from LibraryThing—the social tagging Web catalog—to Library of Congress Subject Headings using a 25-book sample. Kipp and Campbell (2006) explored how tagging supports and enhances traditional classification, as well as how tagging diverges from traditional classification, by comparing tags from del.icio.us to terms that professional indexers would apply to information resources. Bar-Ilan, Shoham,

Idan, Miller, & Shachak (2008) compared field-based descriptions to freely-assigned tags to determine how free-text tagging compares to structured tagging. Murphy and Rafferty (2015) studied the differences between user-supplied tags and Library of Congress Subject Headings applied to select LGBTQ literature to demonstrate the Library of Congress Subject Headings' bias and highlight opposing views presented through social tagging. Rolla (2009) suggests that LibraryThing's definition of tags—an easy tool for users to categorize books based on how the user thinks of a book, as opposed to how a librarian categorizes the same book—implies that tags describe books better than the Library of Congress Subject Headings, and the authors explore this position in their study. Lu et al. (2010) explored the similarities between social tags and subject headings. The authors examined vocabulary similarities between social tags and subject headings, how social tags might enhance access to library collections, obstacles to incorporating social tags in library catalogs, and how social tags and subject headings can complement each other. The authors compared social tags to MARC 650 topical subject headings at the collection level and the individual resource level, and they compared social tags to Library of Congress Subject Heading subfields to see how Library of Congress Subject Heading subdivision concepts are used in tags. Finally, Trant (2006) found that there is more consistency among untrained users, than among professional indexers, in assigning terms to resources.

Metadata creation costs account for one of the major differences between social tagging and controlled vocabularies. Professional indexers and catalogers create high-quality metadata; however, the quality comes at a cost (Lu et al., 2010). Rolla (2009) explains that it is time-consuming for professionals to create and apply subject terms to information resources, which is why it is expensive to create quality metadata and records, and Rolla points out that librarians question the usefulness of complicated subject terms. Macgregor and McCulloch (2006) defend

the high costs associated with professional indexing because professionally-created metadata benefit many users; whereas moving resource description to the user—as is done in social tagging—makes it more difficult to search for resources that would be collocated through means used by professional indexers and catalogers, despite the economies realized by shifting resource description to users.

Controlled vocabularies are more precise than social tagging. It is generally accepted that controlled vocabularies are more effective than free-text searching for resource discovery (Bar-Ilan et al., 2008). While it is still effective to use controlled vocabularies within contained systems for search and discovery purposes, it is not feasible to use controlled vocabularies to organize information on the wide-open Web. Users increasingly lean toward search over categorization to find information resources on the Web (Shirky, 2005). The same happens in closed systems organized by controlled vocabularies. Users successfully find resources through matches on controlled vocabulary terms in search and discovery systems that rely on controlled vocabularies (e.g., library catalogs); however, this occurs because users' natural language keyword searches happen to match controlled vocabulary terms (Bates & Rowley, 2011). If users preferred to approach resource discovery through categorization, an automated approach to information organization using controlled vocabularies for Web content—which does not yet exist—would be needed; therefore, social tagging presents a useful approach to information organization on the Web (Macgregor & McCulloch, 2006). New information grows and forms at a swift pace, vocabulary and terminology associated with emergent information expands and changes quickly, too. Social tagging and controlled vocabularies respond to changes in the information landscape at different speeds and with different approaches. Controlled vocabularies are typically slower in responding to changes in information resource description, as changes are

proposed, vetted, and approved through formal structures. Social tagging is more responsive to changes in resource description because any user can introduce a new term in a social tagging system, and new terminology are not reviewed and approved by other users. Social tagging attempts to address rapidly forming information resources' organization and discovery needs, whereas information professionals question classification schema's ability to do so (Quintarelli, 2005). Other information professionals argue that social tagging's vocabulary is more current than traditional classification schema, specifically citing the Library of Congress Subject Headings (Smith, 2007). Furthermore, social tagging can describe information resources with greater subtlety than controlled vocabularies (Bates & Rowley, 2011). This is helpful in providing narrower terminology to describe and distinguish rapidly developing information resources on the Web.

Controlled vocabularies, such as subject headings in library catalogs, are used to describe what an information resource is about, but social tagging has a different approach to providing this type of description. Early social tagging studies illustrate discrepancies between terms assigned to information resources by professional indexers and users (Jörgensen, 1998; Trant, 2006). Golder and Huberman (2006) compare navigating a social tagging system to keyword searching in that users provide the descriptive terms for information retrieval in social tagging systems. Lorince, Zorowitz, et al. (2015) found that core vocabularies emerge among users in social tagging systems. Several users may apply the same tag, or similar tags, to an information resource over time; therefore, some level of consensus is generated regarding the information resource's subject. Kipp and Campbell (2006) liken this consensus in tagging to the concept of aboutness in professional indexing and cataloging. Users assigning inaccurate tags to resources is regularly cited as a problem with social tagging; however, this problem can be corrected with

additional tagging. This same problem occurs in professional indexing and cataloging, but it is not corrected as easily as in social tagging. Rolla (2009) explains that a single, incorrect tag can be corrected by the larger tag aggregate that coalesces around an information resource, but this does not happen when an incorrect subject heading is applied to an information resource in a library catalog because fewer subject headings are applied to resources.

Systems with controlled vocabularies typically use a limited number of subject terms to describe an information resource, but social tagging systems do not limit resource description in this way (Lu et al., 2010; Rolla, 2009; Rorissa, 2010). Librarians and information professionals question the limits imposed on subject terms in professional indexes considering social tagging's lack of limits. Social tagging's lack of limits highlights the philosophy that something can be about more than one topic. Shirky (2005) discusses the move away from the binary *is* or *is not* categorization traditionally applied to information resources. Rorissa (2010) writes about the differences in the number of terms assigned to information resources by professional indexers and taggers:

> …in social/collaborative tagging, tags are assigned freely without any restrictions as to type and number, whereas professional indexers adhere to guidelines that define types and minimize the number of terms assigned.

Library catalogers use the narrowest subject terms possible when describing an information resource; however, taggers apply a combination of broad and narrow terms to describe an information resource (Rolla, 2009). The freedom in social tagging allows users to express that an information resource can be about many subjects, which is something that professional indexers cannot do with controlled vocabularies (Smith, 2007). This difference is highlighted in the results from Rolla's (2009) study, which compares social tags and library subject headings applied to books. Rolla (2009) found that an average of 42.78 tags were applied to the research

sample on LibraryThing, a social tagging site for books, and an average of 3.8 subject headings were applied to the same sample in a library catalog.

Furnas (1987) surmises that the terms applied to an information resource will be used only in a small number of user searches. Therefore, one can argue that social tagging can increase potential matches on users' search terms and enhance discoverability through the unrestricted number of terms that can be applied to an information resource. Lu et al. (2010) rationalize that social tagging provides an opportunity to expand access to resources by offering alternative terms over controlled vocabularies. Social tagging systems contain a number of single-use tags—tags that are applied only once to a single resource—which some may argue inhibits resource discovery by adding noise to the discovery system. Library catalogers avoid applying single-use terms in library catalogs—in addition to limiting the number of subject terms applied to individual resources—to reduce the overall number of terms in a controlled vocabulary, facilitate resource collocation, and streamline resource discovery. Library catalogers use subject descriptors that can be used for multiple resources and avoid single-use terms; however, single-use terms do not pose a major issue to resource description in social tagging systems if users continue to apply multi-use tags to a resource because the single-use tag will either add value or become insignificant noise (Shirky, 2005). Rolla (2009) writes that library catalogers are doing a disservice to users by adhering to outdated standards and not applying more subject terms in a library catalog:

> The fact that users of LibraryThing assign tags to books representing concepts not brought out by LCSH does indicate that catalogers, by following the LC guidelines, may omit concepts that are important to users.

Increasing subject terms applied to information resources in library catalog records can improve search and discovery by potentially expanding the number of matches between user search terms and subject terms applied to resources.

Library catalogs and professional indexes appear to be at odds with social tagging because systems with professionally-developed controlled vocabularies are managed very differently from systems with user-generated social tags. Web 2.0 affords new approaches to resource description and organization, which call into question the way in which libraries provide access to collections (Rolla, 2009). Social tags can improve library catalog's usability and interactivity (Spiteri, 2007). Therefore, libraries are looking to tagging and other social applications to maintain their value to users and on the Web. Social tagging is increasingly incorporated in library catalogs, despite their antithetical approaches to information organization (Murphy & Rafferty, 2015). The Library of Congress' Working Group on the Future of Bibliographic Control issued a recommendation that libraries enable tagging for users in their library catalogs to both make library catalogs more relevant to users and improve resource discovery and access (Rolla, 2009). Studies show how social tags and subject headings intersect in library catalogs and the user benefits realized by allowing both to coexist in the same search interface. Qin (2008) discusses how social tagging can enrich and validate traditional indexing. Overlapping terms verify that professional catalogers and indexers selected subject terms that match users' mental models, which leads to successful search and discovery; tags or subject headings that do not overlap with each other are able to supplement the other system and enrich resource description. Some social tags applied to information resources overlap with Library of Congress Subject Headings applied to the same resource, which demonstrates that users and experts employ some common terms in describing resources (Lu et al., 2010; Rolla, 2009). Rolla

(2009) notes that there are some differences in how concepts are described within these overlapping terms; however, this confirms that social tags can expand subject access. Rolla (2009) also noted that some subject headings in library catalog records reference subjects that user-generated social tags did not reference, and vice versa. Overlapping, synonymous terms are valuable to expanding subject access to users by potentially increasing matches on search terms through synonymy. However, an argument can be made that non-overlapping terms offer additional value and expand subject access because each system injects concepts not found in the other.

Information professionals advocate for allowing social tagging and controlled vocabularies to coexist within information organization systems (Guy & Tonkin, 2006; Morville, 2005; Rorissa, 2010; Rosenfeld, 2005). Rolla (2009) theorizes that such a combination will produce better subject access. Wetterstrom (2008) writes, "user-assigned tags could provide additional access points, and the co-existence of tags and controlled vocabularies… could thus enhance the discovery of documents." The coexistence of these two approaches recognizes that controlled vocabularies' precision is still necessary for effective searching, but social tagging provides additional access to resources (Guy & Tonkin, 2006; Lu et al., 2010; Quintarelli, 2005; Rolla, 2009). Bates and Rowley (2011) speculate that an inclusive approach to indexing might address some of categorization's problems. Unifying social tagging and traditional indexing in a single interface allows each approach to mitigate the other's weaknesses.

Researchers also discuss the need to harness the rich, user-centered language and terms found in social tags to develop and improve controlled vocabularies. Rorissa (2010) writes, "…the problem of how to incorporate user-generated tags into the process of indexing and retrieval needs urgent attention." Macgregor and McCulloch (2006) state that terms found in

social tagging systems can provide an avenue for user input in controlled vocabularies. Information professionals can look to frequently used tags in social tagging systems to develop terms for controlled vocabularies (Mathes, 2004; Quintarelli, 2005). This will help align terms with users' language and mental models.

Scholars have discussed the merits of social tagging for organizing information on the Web. However, many of the ideas and views on social tagging have been documented outside the realm of scholarly literature (Macgregor & McCulloch, 2006). Multiple studies comment on social tagging's benefits and shortcomings, and its shortcomings are often drawn from comparisons to controlled vocabularies' benefits. Quintarelli (2005) cites lack of precision, synonym control, and hierarchy; language variability issues; and targeted searching performance as some of social tagging's shortcomings. Spiteri (2007) identifies other shortcomings including inconsistent use of singular and plural terms among tags, system impediments to multi-term tags, and ambiguous tags applied to information resources.

However, Quintarelli (2005) argues that these problems inherent to social tagging are not necessarily defects to social tagging systems because these item descriptions and searches reflect users' mental models. Shirky (2005) argues that synonyms do not exist in social tagging systems because every tag is employed for a particular reason. Users are able to match their information needs with natural language vocabulary, and social tagging vocabulary is more inclusive than that of controlled vocabularies (Quintarelli, 2005). Social tagging's lack of hierarchy and precision can also expand access points for search and discovery. Long tail concepts add ideas and avenues to a resource, which can be useful to discovery (Quintarelli, 2005).

The Web has facilitated amateur publishing in a way that has led to huge increases in information on the Web. It is economically impossible to organize the Web's content with

controlled vocabularies due to the sheer size of content on the Web; therefore, social tagging as amateur classification provides a means to organize and categorize Web content where controlled vocabularies simply cannot scale to such an application (Quintarelli, 2005). Social tagging provides a means to bridge personal and shared classification, and it allows users to create communities around classification (Quintarelli, 2005). Social tagging is not that far astray of ontological classification in scope and application of terms. Spiteri (2007) found that social tags closely correspond to the NISO guidelines for thesaurus construction in their structure, concept type, singularity, noun predominance, recognized spellings, and primary use of alphabetic characters.

Social tagging's flexible ability to organize the Web's constantly expanding information landscape is regularly cited as one of social tagging's main benefits. Ontological classification works well for physical collections—like those found in library catalogs where published books already have a place in the classification system—but does not work as well in the digital world and for Web-based resources (Shirky, 2005). Social tagging is more agile than controlled vocabularies in dealing with Web-based resources. Macgregor and McCulloch (2006) and Shirky (2005) claim that controlled vocabularies do not adapt quickly enough to new knowledge and information shared on the Web because controlled vocabularies do not match the emergent vocabulary used to describe and search for evolving knowledge and information. However, social tagging can adapt to changes in the information landscape much faster. Social tagging systems allow users to describe information objects using any terms they wish to use; whereas, controlled vocabularies contain a limited number of terms and are subject to a lengthy process to add or change terms in a controlled vocabulary system (Lu et al., 2010). Social tagging can adapt quickly to changing vocabularies in different fields, which can help avert biases present in

controlled vocabularies like the Library of Congress Subject Headings (Rolla, 2009). Golder and Huberman (2006) explain that social tagging allows both mainstream and minority opinions to coexist in describing information objects. This diversity in thought is one way in which social tagging systems can dispel bias. Macgregor and McCulloch (2006) highlight that social tagging has a significant economic advantage over professional indexing. Controlled vocabularies require professional indexers to apply descriptors to information resources, whereas anyone may apply descriptors to information objects in a social tagging system.

Researches repeatedly cite synonyms, polysemys, homographs, ambiguity, basic level variation, and lack of hierarchy as social tagging's main problems when it comes to categorizing resources (Golder & Huberman, 2006; Guy & Tonkin, 2006; Kipp & Campbell, 2006; Kroski, 2005; Lu et al., 2010; Rolla, 2009; Spiteri, 2007). Controlled vocabularies handle these issues very well. Macgregor and McCulloch (2006) and Rorissa (2010) argue that social tagging lacks controlled vocabularies' precision. Macgregor and McCulloch (2006) and Guy & Tonkin (2006) argue that users incorrectly apply tags to resources, which generates inaccurate descriptions. Kipp and Campbell (2006) address the prevalence of spelling variations among social tags. Spelling variations pose a problem for collocating resources because social tagging systems lack the ability to identify and bring together inconsistent tag spelling. In addition to spelling variations, Spiteri (2007) points out that some social tagging systems have variable approaches for handling punctuation, which impacts search and retrieval. Abbreviations, acronyms, and homographs are viewed as primary causes for tag ambiguity, and Spiteri (2007) claims that tags of this nature make up a significant percentage of tags within social tagging systems. One study demonstrated these problems by searching for social tags in two different dictionaries to discover

that over 50% of tags sampled were not found in the dictionaries, and some tags found in the dictionaries had multiple meanings (Suchanek, Vojnovic, & Gunawardena, 2008).

Social tagging proponents recognize these issues, but claim that social tagging's benefits outweigh these problems (Murphy & Rafferty, 2015). Spiteri (2007) argues that a social tagging system's user base may be able to infer ambiguous tags' contexts. While collapsing terms into a single, authorized term is considered a major strength in controlled vocabularies, Shirky (2005) argues that this is not appropriate within a social tagging system because distinct concepts will be lost if terms are collapsed. Cross references and semantic relationships among terms are essential to controlled vocabularies and ontologies, but the lack of these two key elements among tags hinders search and retrieval in social tagging systems (Chen et al., 2008; Golder & Huberman, 2006; Spiteri, 2007). In comparing social tags to the Library of Congress Subject Headings, Rolla (2009) notes that social tags lack the specificity seen in the Library of Congress Subject Headings with the Library of Congress Subject Headings' free-floating subdivisions, and tags lack the specificity that we see in the Library of Congress Subject Headings' defined time periods because there is no consensus in how users think about time periods when applying social tags to objects. Controlled vocabularies can be applied across various content domains because of the lexical control and hierarchical organization employed in creating and maintaining controlled vocabularies; however, researchers have observed variability in social tagging habits across content domains, which impedes subject interoperability in social tagging systems (Lorince, Joseph, & Todd, 2015; Macgregor & McCulloch, 2006).

What is social tagging's appeal, and what motivates so many users to contribute to resource description and organization within social tagging systems? Despite the voluminous body of social tagging research, one major question remains unanswered—why do people

participate in social tagging? Lorince, Joseph, and Todd (2015) explored the question—why do people tag? In this paper, the author states there is an assumption that someone tags a resource to help himself retrieve that resource in the future; however, this assumption is not corroborated by research or data. Lorince, Joseph, and Todd's (2015) tagging study used data on Last.fm, and the study's results suggest that people do not necessarily tag resources for personal future retrieval. The author concludes from his results that motivation for tagging, "may be socially or otherwise oriented, which may in turn result in tags that are useful for the community at large." This study, while limited in scope and not fully conclusive, demonstrates the possibility that people tag information resources in social tagging systems as a means of describing resources for other users' discovery.

A fair number of research studies have explored tagging behavior among users within social tagging systems. However, the reason why people participate in social tagging still eludes researchers (Lorince, Joseph, & Todd, 2015). Studies also show that tag quantity produced by users within a social tagging system varies greatly—with many users producing few or no tags and few users producing many tags—but the social tagging research body lacks studies on what motivates some users to be more prolific taggers than other users (Lorince, Zorowitz, et al., 2015). Some researchers have developed motivational theories for tagging, but these theories are not grounded in observational studies of user behavior (Lorince, Joseph, & Todd, 2015).

Information's proliferation and rapid growth on the Web is considered a primary reason as to why people participate in social tagging. Web publishing and online social tools facilitate information creation and dissemination at a scale previously unattainable to the average person. This has led to a shift in behavior, in which people increasingly participate in information creation, rather than solely consuming information. Rapid growth in digital information requires

users take on the task of organizing the information they create and use for themselves and others (Fox, 2006; Quintarelli, 2005; Rorissa, 2010). Yahoo attempted to organize and index Web content, but that task is too large in scope for corporations or professional indexers, given the pace at which information grows on the Web. Therefore, it has been necessary to transfer responsibility for organizing information on the Web from professional indexers to users. Svenonius (2001) describes this shift:

> The rise of the Internet is affecting the actual work of organizing information by shifting it from a relatively few professional indexers and catalogers to the populace at large…. While not consciously teleological, a self-organizing bibliographical universe nevertheless succeeds in meeting the bibliographic objectives in part, occasionally, and somewhat randomly.

Social tagging has allowed users to take on the indexer's role in the Web 2.0 environment (Rorissa, 2010). People are creating valuable organizational systems in the process of tagging information resources to keep track of, and organize, content for themselves—the aggregate of tags created is an extremely useful user-generated organizational system that comes at a fraction of the cost of professional indexing (Shirky, 2005). Users are tagging information resources as a means for dealing with information proliferation on the Web and are creating a useful byproduct in the process.

Users tag information resources to organize the prolific world of digital information resources for themselves; however, their tagging behavior has a broader social impact on other users within a social tagging system—individual tagging behavior benefits search, discovery and information organization for all users within a social tagging system. Tagging is primarily used to describe resources for one's own use or to search for resources within a system (Furner, 2008; Rorissa, 2010). Golder and Huberman (2006) outline tags' functional areas, which include aboutness, description, ownership, qualifiers, qualities, self-reference, and task organization.

Kipp and Campbell (2006) look at the differences between subject access tags and tags for time-related concepts (e.g., to read, to do) and note that personal, temporal tags shift, whereas subject access tags do not shift. In other words, a *to read* tag is no longer applicable to the user who created the tag once he has read the tagged resource. Likewise, the same can be said for other personal and qualitative tags. An individual user's attitude toward, ideas about, and reactions to, an information resource can change over time; however, that information resource's aboutness and subjects are not likely to change.

Researchers divide tagging rationale into two primary categories—personal and social. Golder and Huberman (2006) speculate that most tagging is done for personal use and not to benefit the greater community. Chi and Mytkowicz (2008) conclude that tagging does help users recall information for future use, but they suggest that tagging is also used to communicate ideas about information resources to other users. Furner (2008) balances the two categories, asserting that users will tag for either personal reasons to support one's own goals or for social reasons to help other users access resources. Lorince, Joseph, and Todd (2015) cite personal retrieval, resource sharing, personal expression, performance, and activism as potential reasons people tag information resources. Personal retrieval and resource sharing fit into the personal and social categories, respectively; whereas personal expression, performance, and activism are part of a hybrid category, in a sense, with their tagging rationale crossing both the personal and social categories.

Some researchers point out that the social aspect of tagging—describing resources for others' benefit—allows people to engage with information resources and participate in online communities—usually communities of like-minded people (Furner, 2008; Macgregor & McCulloch, 2006). Furner (2008) looks at potential tagging motivations in library catalogs from

this communal lens in that tagging allows people to identify others with shared interests, help improve searches, and share knowledge.

Social tagging's critics have likened tagging to mob indexing in that users are not trained indexers and do not produce useful tags (Kipp & Campbell, 2006). Again this provides an example of how social tagging's weaknesses are also its strengths. While most users are indeed not trained indexers, their indexing behavior and choices provide researches and indexers a window into how users think about information resources. Social tagging is useful because it helps to bridge gaps between information objects and users' mental models (Furnas, Fake, von Ahn, Schachter, Golder, Fox, Davis, Marlow, & Naaman, 2006).

Indexers and catalogers attempt to describe information objects in ways that accurately reflect objects' subjects and with terms that users will employ in searches. Indexer-search consistency can be used to measure the degree to which indexers effectively anticipate which search terms users will employ when searching for a given object. Furner (2008) points out that indexer-search consistency is a method for measuring retrieval effectiveness. Researchers speculate that social tagging has a high indexer-search consistency because the people describing information objects are the same people searching for information objects, or people utilize similar vocabularies within the same information domain. Furner (2008) admits that empirical research on retrieval effectiveness for tagged resources is needed to determine if people tagging resources and people searching for resources are using the same vocabulary.

Indexers need to determine what information objects are about and how users might search for objects when describing objects as part of their indexing work (Fidel, 1994). Indexing the Web is an untenable task, and social tagging bypasses professional indexers. People creating information on the Web are also using that information on the Web (Quintarelli, 2005). Rolla

(2009) points out that catalogers rarely have time to read the books that they are describing for users; however, taggers are often familiar with the information resources they describe and are also the end-users in the social tagging systems they use. Therefore, as Smith (2007) points out, users in a community may be better than catalogers at describing certain information resources. Some catalogers and indexers are domain specialists, but they may not be as knowledgeable as those creating and using information resources within that specialization.

Many researchers state that there is a high level of indexer-search consistency within social tagging systems because the users who are tagging resources are also the users who are searching for resources. Shirky (2005) suggests that users searching for resources within a social tagging system will find the resources they are looking for if other users tag resources the same way the searcher would tag these resources. Furner (2008) and Lu et al. (2010) take this a step further, stating that it is in fact easier to achieve indexer-search consistency because the users who are tagging information objects are also the users who are searching for information objects—they are one and the same. Furner (2008) indicates that high levels of agreement between tagging terms and search terms is assured because the taggers and searchers are part of the same population.

Controlled vocabularies and classification schema are known to contain bias and do not always align with end-users' mental models. Social tagging provides a means to move past these two issues. Deodato (2010) explains that social tagging allows users to create their own knowledge and organization structures, which can be useful for marginalized users. Indexers and catalogers describe resources differently from users; therefore, the terms with which users tag objects may aid subject access, but additional research is required to see how user-generated tags are used in search and discovery (Lu et al., 2010). Terms that users apply to information

resources align with their mental models and worldviews. Marginalized populations may describe resources about their populations using different terminology than those outside their communities use to describe the same resources—this provides an inclusive way to diversify resource description and help strip away some of the bias prevalent in controlled vocabularies.

Controlled vocabularies force exclusive world views upon users. Shirky (2005) explores the difference between browse and search. He explains that browsing requires users to match their browse terms to the terms catalogers used to describe resources; therefore, users must explore resources through catalogers' world views to successfully find information resources. However, searching within a linked structure allows users to explore resources using terms within their own vocabularies, and, therefore, does not need to impose others' world views. Smith (2007) insight into the exclusivity in controlled vocabularies, as well as the dissonance between controlled vocabularies and users' mental models, also highlight how controlled vocabularies impose others' world views on users. She explains that controlled vocabularies require users to use approved terms, which exclude other potential search terms, and which may not align with users' knowledges and understandings within an information domain.

Users do not participate in creating controlled vocabularies in the way they participate in creating folksonomies, and they may not ascribe to the same world view as those responsible for creating and maintaining controlled vocabularies (Shirky, 2005). The disparity in mental models between indexers and searchers, coupled with the bias found in controlled vocabularies, highlights the inclusivity in social tagging. Spiteri (2007) describes the inclusivity in folksonomies: "they reflect the vocabulary of the users, regardless of viewpoint, background, bias, and so forth." Tagging is a form of sense-making and knowledge construction (Golder & Huberman, 2006; Shirky, 2005). Golder and Huberman (2006) explain that people incorporate

their own world views, which include personal experiences and biases, into their tagging behavior. Shirky (2005) states that it is important to accept that there are different world views without privileging any particular world view when trying to make sense of the world—an inclusive aggregate is more important and useful than an ontological goal.

Social tagging provides a window into users' mental models, including user behavior, how users think about information objects, and vocabularies used to describe information objects. Insights like these are exactly what indexers and catalogers need to improve indexing (Fidel, 1994). Fidel (1994) recognizes that inter-indexer consistency among those describing information objects is important, but argues that indexer-requestor consistency is even more important. Fidel (1994) recommends harvesting terms that users plug into databases while searching for information resources and using those harvested terms to index resources. Other researchers also promote studying how indexing terms and users' vocabularies intersect and diverge to better develop systems (Eerola & Vakkari, 2008). Social tagging provides researchers and indexers with these data, which can be used to develop user-centered indexing. Rolla (2009) advocates for examining tags to determine how users think about information objects and subjects that they are researching.

Library catalogs can take advantage of social tagging to enhance subject access to information resources. User generated tags can be used to examine the degree to which library assigned subject terms align with users' mental models, provide information professionals with a snapshot of users' mental models, and can possibly be used to assist library catalogers in selecting subject terms for library catalog records. Rolla (2009) studied how user-supplied tags shed light on how users think about subjects and whether tags can help catalogers select subject headings for resources. Lu et al. (2010) explore how user-generated tags overlap with library-

supplied subject headings and found that 85% of book records have at least one overlapping term between subject heading and tags with a significant number of terms used by both experts and users to describe the same book. It is important that libraries describe information resources in ways that make sense to their users by implementing services that draw on their users' expertise, allowing them to describe and discover resources using terms within their own cognitive models (Fox, 2006). Social tagging provides libraries with mechanisms for creating services that take advantage of their users' knowledge and developing user-centered search interfaces.

User-generated tags provide a wealth of metadata, which can be used to develop better indexing tools and better align indexing terms with users' mental models. Researchers have demonstrated that social tags can be useful in designing traditional indexing tools, such as taxonomies, thesauri, and ontologies, and social tags should inform indexing tools' future design (Rorissa & Iyer, 2008). Rorissa (2010) argues that social tagging research can be used to better align indexer-assigned terms and users' search language in indexing systems. Stock (2007) advocates for indexers to reference social tags to determine appropriate index terms, and explains that these social tags will be especially useful to indexers when the taggers are subject experts. This ties into connected knowledge concepts explored by researchers and theorists in which the broader community develops knowledge and generates user-oriented aboutness for information resources (Bates & Rowley, 2011; Hjørland, 2010; Olson, 2007). The outcome of designing indexing systems this way has the potential to reduce users' cognitive loads because users will not need to consider how informational professionals describe resources when searching for information (Sinha, 2005).

Social tagging research has led to several system design recommendations focused on improving social tagging. Some system designers are exploring mechanisms to address the

synonym issue in social tagging systems by recommending potential tags based on the search term entered (Kroski, 2005). Kipp and Campbell (2006) state that system designers can anticipate synonymous relationships between tags while designing social tagging systems, but warn that designers should not collapse synonymous tags because these different tags may hold different meanings to different users. This also affirms Shirky's (2005) view on synonyms in social tagging systems. Shirky (2005) claims that tag semantics lie with the users, rather than the systems, and therefore, a system cannot interpret overlap within a tagging system, but it can provide users with recommendations based on overlapping tags, which leaves it to the user to decide how to proceed with the system's recommendations. Compound terms are a problem in social tagging systems; therefore, these systems need to develop a better approach to dealing with compound terms (Spiteri, 2007). Chi and Mytkowicz (2008) and Guy and Tonkin (2006) believe that tagging systems should avoid promoting popular tags to avoid tagging hegemony and increase tagging diversity. Chi and Mytkowicz (2008) even argue that tagging sites should encourage users to apply tags that have not yet been used to describe objects. Social tagging systems are ripe with user-supplied metadata, of which system designers should take better advantage by pulling the most popular metadata for use in indexing design (Quintarelli, 2005; Rorissa, 2010).

Some researchers suggest that making improvements to social tagging systems can lead to improvements in folksonomies. Most recommendations related to this focus on educating, providing feedback to, and making suggestions to users applying tags to objects.

Users either create unsuitable tags or do not properly apply tags to information resources; therefore, user education is seen by some as an approach to correct some of the problems observed in tagging behavior. Guy and Tonkin (2006) advocate for educating users on spelling,

encoding, personal tags, and single use tags to address problems that these issues create in social tagging systems. Providing users with a checklist of questions to consider when tagging objects can also guide users into creating useful tags (Guy & Tonkin, 2006). Tagging behavior demonstrates that there is an inconsistency in how users employ plural and singular nouns, which can lead to search and retrieval problems in social tagging systems. Systems can try to mitigate issues caused by inconsistent plural and singular noun use by providing information about the difference between count nouns and mass nouns, and providing guidelines on how to appropriately use these different noun types when creating tags (Guy & Tonkin, 2006; Spiteri, 2007). Bar-Ilan et al. (2008) recommend suggesting elements to users when applying tags to objects to reduce ambiguous tags and improve tags' descriptive quality. The authors cite Flickr as an example of a system that provides element tagging suggestions—recommending that users add medium, genre, subject, and name tags—when adding objects to the system. Guy and Tonkin (2006) recommend a similar but more structured approach by providing users with actual tag and synonym suggestions when adding objects to a system, and then developing a feedback loop for taggers. Systems can somehow encourage users to apply multiple tags to an object instead of trying to eliminate sloppy tags and single-use tags (Guy & Tonkin, 2006). This approach does not decrease noise in tagging systems, but it can potentially increase descriptive tag quantity just by encouraging users to generate more tags for each object.

Most social tagging systems do not support compound word tags; approaches for handling compound word tags vary from system to system among those systems that do support this tag form. Creating an authorized approach to compound word tags would help address this issue (Guy & Tonkin, 2006). However, it would be difficult to enforce an authorized tag form of

any kind among users, and systems cannot determine users' intentions to correct perceived violations to an authorized form.

However, this begs the question, how would one go about educating users on tag creation and application? Would users be open to this, or would this dissuade users from tagging information resources? Constraints like these affect social tagging systems' ease of use. These approaches could potentially improve tags' descriptive quality and folksonomies' quality; however, these approaches could also decrease tagging diversity and possibly inject bias into folksonomies.

Incorporating tag bundles—tagging tags—in social tagging systems can also improve folksonomies (Guy & Tonkin, 2006). Tag bundles offer users a mechanism to group similar terms; therefore, tag bundles provide a potential solution to the synonym issue that pervades social tagging systems and could be viewed as a quasi-cross-reference mechanism within these systems. This added functionality is a system design solution and avoids user education issues.

Some studies have examined the idea incorporating social tagging in library catalogs as a way to update these systems. Library systems serve different functions because they have different types of users (Furner, 2008). The thought diversity that social tagging injects in search and discovery might be useful in meeting varying users' needs. This diversity can also help mitigate some of the bias prevalent in controlled vocabularies; therefore, helping library catalogs appear less biased in regard to non-dominant resources and more user-friendly to those interested in using these resources.

Libraries should consider harvesting tags from other systems in addition to allowing library users to tag materials in library catalogs. Rolla (2009) writes that LibraryThing could be considered the Web 2.0 of library union catalogs. Bates and Rowley (2011) explain that libraries

can import LibraryThing tags into their catalogs for a fee. The tag aggregate that comes from a site like LibraryThing or Good Reads is necessary for tagging to function effectively in a library catalog. Relying on a library's own user group to tag materials from the library's collection will not yield an adequate number, or a diverse range, of tags to reap the benefits of integrating social tagging into a library catalog.

Some librarians advocate for imposing constraints to social tagging in library catalogs to attempt to maintain high-quality metadata. Spiteri (2007) recommends that libraries explain count nouns, provide a way to make compound tags, link to an online dictionary, explain the impact of ambiguous terms, and provide an acceptable use policy when adding social tagging to the library catalog. These recommendations attempt to address issues that librarians and some researchers view as social tagging's shortcomings—particularly when placed side-by-side with controlled vocabularies. However, these problems are also what make social tagging successful for organizing information resources. Furthermore, adding constraints like these to social tagging and users' tagging creativity may discourage social tagging—this is an area that should be further explored in future research with the benefits and consequences better weighed.

Music Domain Resources

Finding music works in a library catalog or discovery interface can be an extremely difficult, and sometimes daunting, task for library users. On a general level, the principal complication lies in the concept of the musical work and its multiple expressions and manifestations, along with the elaborate bibliographic relationships that the numerous expressions and manifestations then create. More specifically, the difficulties inherent to music searching in library catalogs are created by given titles of the manifestations, music's international aspect, composers' and performers' prolific outputs, the even more prolific number

of expressions created for each work, ambiguities related to statements of responsibility and generic music titles, and the lack of suitable search types.

Developing an understanding of how users think about and search for music resources can help determine how bibliographic records can facilitate music information retrieval, whether resource description models align with users' mental models, and how other information resources can augment traditional approaches to resource description (e.g., library catalogs). Such an analysis will be helpful in designing information retrieval systems that assist users in navigating the elaborate bibliographic relationships ubiquitous to music. Doing so will benefit information search and retrieval in the long run as these elaborate bibliographic relationships are becoming more widespread among textual works. Users will need information retrieval tools capable of navigating these elaborate relationships, evaluating their search results, and selecting an appropriate instantiation of a work to satisfy their information need.

If a user encounters a piece of music for which he does not know the title, he will not be able to form a library catalog search for that piece of music until he is able to successfully fill in this missing piece of information in his musical knowledge. When the user has an expressible music information need, in other words the user knows the title of the work and composer or performer he is searching for and can represent that with a text-based search, the user can perform various catalog searches. During this process, the user may perform one search or many searches depending on search results obtained from each search and the thoroughness desired by the user. Furthermore, if the user encounters unsuccessful searches, he may consult additional resources to help refine his search terms or fill in smaller anomalies in his musical knowledge that are related to whole/part work relationships or other issues.

Authority control has long been an essential component of music cataloging practices, as it provides a partial solution to navigating the elaborate bibliographic relationships found in music works, as well as overcoming some of the difficulties inherent to searching for these materials in library catalogs. The FRBR model draws attention to the bibliographic relationships found in music materials (Vellucci, 2007) and provide a conceptual framework that is better able to depict these relationships. Much of the existing music authority control work fits nicely within the FRBR conceptual framework when it comes to illustrating the relationships among a work and its many expressions and manifestations. The increasing popularity of search and discovery platforms, with their simplified search boxes, coupled with search results that integrate both locally-owned content and materials outside of a library's collection, provides a continued need for authority control in music and an increased need for FRBRized information retrieval systems that improve users' abilities to take advantage of valuable authority control work. Music librarians have been advocating for improved search features that will help users locate music materials; however, these calls for improved functionality have not received sufficient attention, most likely because such improvements were not believed to have a significant enough impact factor on the broader, non-music-seeking user population.

Keyword searching can be somewhat effective in music searching if and when a user's search terms match terms in a bibliographic record; however, the library catalog will return results for all records containing the user's keywords. Some of these results may be relevant to the user's intended search, while other results can be related to an entirely different work—musical or non-musical. At the same time, materials relevant to the user's intended search will go undiscovered by this keyword search. Boyd (2005) describes keyword searching as a mixed blessing in that "titles and table of contents notes…are uncontrolled, i.e. they use the language

and spelling of the publisher, and are therefore liable to give incomplete or misleading search results." The effectiveness of keyword searching for music materials is limited for the reasons outlined above, but the issues of publisher's language and spelling require closer examination to understand why author and title keywords will not retrieve all relevant results.

Unfortunately, it is not possible to effectively apply the simple search techniques—author, title, and keyword—that one can employ in searching for textual works when searching for musical works because of the numerous elaborate bibliographic relationships that exist for music works. Of course it is possible for any work—textual, musical, visual, etc.—to possess multiple bibliographic relationships; however, this possibility, and the number of potential relationships, is exponentially greater in musical works than textual works because music is a performance art (Dickey, 2008; Vellucci, 1998). Musical works, like textual works, are expressed in print form (as scores), and multiple manifestations of an expression can be produced. A single musical work is performed and recorded numerous times by different performers, creating multiple expressions of the work, and these expressions can be manifested in multiple ways over time (King, 2005; Kranz, 1988; Vellucci, 1998; Vellucci, 2001). For example, Bach's *Cello Suite No. 1* was performed by the cellist Jacqueline Du Pré and recorded by the BBC. The expression's first manifestation was a radio broadcast by the BBC; additional manifestations of this expression were created at later dates, some of which include a compact disc published by EMI in 1989 titled *Jacqueline du Pré: Her Early BBC recordings 1*, a remastered version of the expression published by EMI in 1999 titled *The Early BBC recordings, 1961-1965*, a rerelease of the 1999 manifestation in 2004, and, more recently, an mp3 manifestation of the 2004 rerelease. The Beatles' *A Hard Day's Night* provides a similar example. The song was first issued on the manifestation titled *A Hard Day's Night*, and was

subsequently released on four other Beatles recordings—*The Beatles 1962-1966*, *Live At The BBC, Anthology 1*, and *1*; the song has also been covered by many other performers and appears on manifestations of other expressions.

While some musical works have distinctive titles (e.g., *A Short Ride in a Fast Machine*, *Ancient Voices of Children*), many musical works use generic titles that are not distinctive (e.g., *Symphony, Piano Concerto, Sonata*) (Gentili-Tedeschi & Riva, 2004; King, 2005; Kranz, 1988), making it difficult to distinguish *Symphony No. 4* by Beethoven from *Symphony No. 4* by Brahms based solely on title. To further complicate this issue, languages, translation, and transliteration pose problems for musical works with both distinctive and generic forms. Music is created and published all over the world. For this reason, a musical work may be given a title in the composer's native language, but translated titles are often introduced with different manifestations; therefore, one work may be known by titles in multiple languages (King, 2005). Many of Igor Stravinsky's works have distinctive titles, but the work that many people know as *The Rite of Spring* in English speaking countries was titled *Le Sacre du printemps* by the composer. A user effectively searching the Library of Congress' catalog for a sound recording of Beethoven's generically titled *Symphony No. 4* will encounter many sound recordings, one of which is titled *Symphony No. 4*, and another which is titled *Sinfonie Nr. 4, B-Dur, op. 60*. Although these Beethoven sound recordings are titled *Symphony No. 4*, multiple musical works frequently appear on a single sound recording or in a score anthology, and the given title does not always reflect the manifestation's contents (Gentili-Tedeschi & Riva, 2004; King, 2005; Vellucci, 2007). Creating a standard title—known as a uniform title—for each musical work and then using the standard title to describe all manifestations of that work solves these problems related to title variation (Gentili-Tedeschi & Riva, 2004; Kranz, 1988; Poroila, 2007; Vellucci,

1990; Vellucci, 1998; Vellucci, 2001). Unfortunately, music uniform titles can be extremely difficult to construct and interpret from the user's perspective (Kranz, 1988); therefore, most users will not be able to generate a work's uniform title for a search string.

Languages, translation, transliteration, and ambiguity can also pose significant problems with musical works when it comes to composer's names. A composer's name can vary in spelling and script among different languages, and users can experience difficulty distinguishing between composers with similar or same names (Gentili-Tedeschi & Riva, 2004). The composer's name, like titles given to manifestations of a single musical work, can be spelled differently due to music's global reach. Additional ambiguity develops in instances where one composer is known by two names (e.g., C. P. E. Bach versus Carl Philip Emanuel Bach, Sean Combs versus Puff Daddy versus P. Diddy) and in instances when two composers have the same name (e.g., Johann Strauss, 1804–1849 and Johann Strauss, 1825–1899) (Gentili-Tedeschi & Riva, 2004). Catalogers have addressed these issues by using authority name records to bring order to the mayhem caused by name variations and similarities (Gentili-Tedeschi & Riva, 2004; Vellucci, 1998; Vellucci, 2001). All manifestations of a musical work by a composer are linked to the composer's authorized name file, which consists of the composer's full name and birth and death dates, to collocate all of the composer's works, regardless of how the composer's name appears on the manifestation. Despite achieving collocation through the authorized name, searching for musical works by author is extremely inefficient due to the prolific corpus of works that many composers and performers produce; furthermore, the results display in an author search lists the manifestations' given titles, which revisits all of the previously addressed problems inherent to manifestations' given titles.

Authority name records and uniform titles solve the problems that stem from variations in name and work title, respectively, but these controlled vocabulary terms are rarely useful on their own when searching for musical works for the reasons previously mentioned. Through compounding the authority name and the uniform title, catalogers create a unique work identifier—the name-uniform title—that collocates all expressions and manifestations of a work under one heading, can collocate all of a composer's works in a headings browse, provides cross-references for variant titles, and provides implicit linkages to show bibliographic relationships (King, 2005; Smiraglia, 1989; Vellucci, 1990). This, in turn, provides users with an efficient method to exhaustively search for music scores and sound recordings. Even though the name-uniform title can be complex and difficult to understand in some instances, the general principle of joining the composer's name with the work title and providing access at this broader work level, rather than the item level, is in line with how users think about musical works and what we know about music information-seeking behavior. People naturally tie the composer and the composition together (Dickey, 2008; King, 2005) to form a description of a work (e.g., Beethoven's *Symphony No. 1*) that disambiguates a generic title (i.e., Beethoven's *Symphony No. 1* is not confused with another composer's work that is titled *Symphony No. 1*). This natural name-title description of the musical work is also what users usually have in mind when searching for music—they enter the library to search for a work in hope that they will find an expression and manifestation of this work, and then leave the library with an item in hand (Dickey, 2008; Smiraglia, 2002). If a user wants a sound recording of Beethoven's *Symphony No. 4*, he will search for the symphony at the work level, rather than search by given title at the manifestation level, because searching by title for *Michael Tilson Thomas Conducts Beethoven*, for example, is not an intuitive or effective approach to music searching.

Some librarians cannot imagine providing music reference services without access to music name-uniform title headings, but other librarians feel that the profession should abandon the dated practice of using name-uniform title in favor of a new, uncomplicated bibliographic device (Poroila, 2007); however, such a device does not yet exist. Poroila (2007) describes the struggle with this heading, stating, "I believe that a shorter title is a wise decision, even if we have to forget the cataloguing rules for a moment," but concludes that, "…the world needs the finishing touch of the uniform title tool used by an ambitious and enthusiastic cataloguer. That makes information retrieval easier and more reliable." Those in favor of seeking a replacement for the music name-uniform title cite the challenge in constructing such headings (Kranz, 1988), users' lack of understanding of the heading (King, 2005; Snyder, 2010), and the struggle that users experience when trying to use the heading (Snyder, 2010).

The information need expressed by the user in the form of the composer's name, coupled with the work title, can be examined within the FRBR framework of the Group 2 name or a Group 1 statement of responsibility attribute, coupled with the Group 1 work title, which also effectively maps over to the name-uniform title heading. Although the information seeking behavior corresponds with the FRBR framework and existing cataloging practice, the name-uniform title construction still needs to be simplified in order for users to effectively use information retrieval systems to find music based on their expressed information needs. To that end, a name-uniform title headings keyword browse allows users to construct effective searches using keywords from the composer or performer name and work title, with which they usually approach a query, but they do not need to understand the name-uniform title heading's construction in as much detail.

Music compositions lend themselves well to FRBR's bibliographic relationships and are easily described using the FRBR terms *work*, *expression*, *manifestation*, and *item*. The current name-uniform title is essentially a FRBR work descriptor, a recorded performance is an expression of the work, the compact discs and mp3s that are created from these recorded performances are manifestations of the work, and the individual copies housed in a library are items. FRBRization of library catalogs and the newer search and discovery platforms will be needed to address complications related to the work-manifestations bibliographic relationship when searching for musical works, especially as reliance on local collections decreases in favor of shared resources (Boyd, 2005; Dickey, 2008; Smiraglia, 2007a; Vellucci, 1990). Perhaps FRBR will provide a more effective alternative to the name-uniform title, as may be desired by those who point out the difficulties in constructing and using uniform titles, but until the name-uniform title is replaced by a FRBR work descriptor, authority control in the form of name-uniform title is essential for an effective and exhaustive music search, as well as disambiguating the complex bibliographic relationships that exist between a musical work and all of its expressions and manifestations.

FRBR is not fully capable of handling multi-work recordings and anthologies yet. The aggregates of works that appear in sound recordings and score anthologies are classified as entity works within the FRBR framework and are viewed in two potential ways—as manifestations containing multiple works, or as standalone works (Vellucci, 2007). The incorporation of parts of works (referred to as extractions in FRBR)—such as a single song on a multi-song recording, an aria from an opera, or one movement from a multi-movement symphonic work—in aggregate works creates another complicating factor in music searching.

Uniform titles for extractions will express the extraction's whole/part relationship to the work, but they do not always provide an adequate unique identifier for the part of the work. For example, the aria *Deh vieni, non tardar* from Mozart's *Nozze di Figaro* may appear in a bibliographic record as having a specific uniform title, **Nozze di Figaro. Deh vieni, non tardar** or a non-specific uniform title, **Nozze di Figaro. Selections**. Both uniform titles show a whole/part relationship between the extraction and the work, but the latter does not describe which part of the whole is contained within the aggregate. These models are seen in bibliographic records for classical music aggregate works, but the research is devoid of uniform title constructs for popular music resources. Content notes are common among aggregates' bibliographic records; however, it remains difficult to FRBRize the data in these records in an automated manner due to data dispersion (Vellucci, 2007). If such data are present, they may appear in an uncontrolled field (e.g., MARC 500 or 505), depending on the depth and quality of the catalog record. Scripts need to be developed for information retrieval systems that will either attempt to improve inadequate uniform titles on-the-fly as part of the information retrieval process or generate reports for such inadequate headings for records maintenance projects. If a non-descript uniform title is used in a bibliographic record, such a script can scan the remainder of the record for relevant data that can add value to the uniform title heading. If relevant data are found, the system can concatenate these data with the whole work's uniform title—not the inadequate uniform title created for the extraction—to form a more descriptive uniform title, which can then be cross-checked against the authority file for accuracy. Libraries' financial and staffing constraints would have to determine whether such a script is deployed as part of the information retrieval process of each search performed or to aid a large-scale FRBRization project. Such a script will not be capable of addressing and correcting all of the poorly formed

uniform titles; however, this may be able to produce satisfactory results if sufficient quality data can be found in a record's 5XX field. Furthermore, if a library were to use such a script to generate reports for records cleanup, any records that do not contain adequate data to form a concatenated uniform title, or form uniform titles that cannot clear or cross reference the authority file, can be flagged for review, too. Further research into popular music uniform titles is needed to develop automated mechanisms for those records.

Relatively few research studies have examined the issue of multiple relationships in bibliographic entities. Those studies that have examined these relationships reinforce the notion that a significantly greater number of bibliographic relationships exists for musical works than for textual works. McNellis (1985) examined the issue of multiple manifestations in a sample of textual works found in the University of Chicago's Regenstein Library. Her findings indicated that 26% of her sample contained bibliographic relationships involving multiple manifestations of textual works; therefore, she concluded that authority control was unnecessary for titles at the work level, given this low percentage of multiple manifestations of a work (Smiraglia, 1989). Papakhian (1985) performed a similar study of bibliographic relationships, examining musical and textual works from Indiana University's Music Library. While this study showed that approximately 61% of the works he examined contained bibliographic relationships involving multiple manifestations, this study does not show the true impact of these relationships on musical works because the study examined a combination of musical and textual works. It was Smiraglia's (1989) study that demonstrated the pervasiveness of these bibliographic relationships for musical works. Smiraglia (1989) sampled musical works from works found in the book *A Basic Music Library*, and searched the Library of Congress and OCLC catalogs for all manifestations of the sampled works. The author found that close to 90%—almost the entire

sample of works—contained bibliographic relationships involving multiple manifestations, in which 95% of these relationships stemmed from variation in language (Smiraglia, 1989), effectively demonstrating that these relationships were more prevalent in musical works than in textual works at the time. Furthermore, one can infer that the prevalence of these relationships is compounded over time by comparing Smiraglia's findings to Vellucci's (1998) findings in a similar study. Vellucci (1998) examined a sample of musical works found in the Sibley Music Library's collection at the Eastman School of Music, and found that 97% of the works sampled contained bibliographic relationships involving multiple manifestations. Therefore, while some have believed authority control for titles at the work level was not essential for textual works, this manner of control has been critical for musical works. These two studies were able to infer why effective searches for textual works do not work well for musical works by outlining the types of relationships that exist in their findings (Smiraglia, 1989; Vellucci, 1998). Smiraglia (2007b) investigated the occurrence of elaborate bibliographic relationships—referred to as instantiation—in twentieth-century best-selling books. The author found that all but one of the sampled works contained elaborate bibliographic relationships similar to those frequently encountered in musical works (Smiraglia, 2007b).

Newer search and discovery interfaces do not limit users' search results to the local library collection. Instead these systems are capable of providing users with search results that combine materials from the local collection with materials found in other libraries—regionally, nationally, or even internationally—providing users with access to even more manifestations of a work and increasing the number of work-manifestation bibliographic relationships that users will encounter (Vellucci, 1998). At the same time, libraries are facing shrinking budgets and are increasing their reliance on shared collections and interlibrary loan services to accommodate

dwindling fiscal resources. This trend will also connect librarians and users with a greater number of work-manifestation bibliographic relationships to sift through as libraries lend more and more music materials to other libraries (Vellucci, 1998). Information retrieval systems and international authority control are ill-equipped for users to search for music at the work level using name-uniform title headings. Few current Integrated Library Systems offer the name-uniform title headings browse that is helpful in searching for music materials; this search method has been largely ignored by search and discovery platforms, and systems used for resource sharing—like OCLC WorldCat—do not offer the ability to search by name-uniform title. The globalization of search furthers the need for information retrieval systems to provide a name-uniform title headings browse that will allow users to search, evaluate, and locate manifestations of a musical work. International search and discovery illustrates the importance of international authority control for composer names and musical works; however, each set of cataloging rules treats authority names and uniform titles differently (Vellucci, 1998, Vellucci, 2001). International authority control has been attempted with IFLA's Universal Bibliographic Control (UBC) and OCLC's Virtual International Authority File (VIAF). VIAF contains names from a variety of national authority and bibliographic records, and is working to reduce ambiguity in authority records with a goal of 99% accuracy match rate (Hickey & Toves, 2014).

Study Methodology

Song tracks found on the last.fm Web site were used for this study's population. This music site provides streaming audio, contains resource description data for over 640 million songs, and allows users to generate their own tags for songs, recordings, and artists. While other social music sites and streaming services use tags to describe recordings, the tags found on these other sites are generated by professionals, rather than users.

This study's sample population was derived from the weekly *Billboard Hot 100 No. 1* song charts for 1958–2013. The Google Pages IMPORTHTML function was used to generate an aggregated, chronological list of these songs—with *Billboard Hot 100* issue date, song title, and artist data—from Billboard Music's (n.d.) online chart archive. The data in the aggregated list were cleaned to correct spelling errors and provide consistent song title and artist formatting for duplicate entries. Duplicates were then removed from the chronological list to generate a list of 1,034 unique *Billboard Hot 100 No. 1* songs, and a simple random sample of 50 song titles was taken from the de-duplicated list (Appendix A). This method was selected to produce a sample of popular songs that are well known—due to their playback over commercial media outlets and high sales volumes—and, therefore, more likely to be tagged on the last.fm Web site.

The song titles in the sample were searched on the last.fm Web site to collect user-generated tags for the study. Each song's tags page was accessed, and the Google Pages IMPORTHTML function was used to collect the user-generated tags applied to each song. The user-generated tags, along with their associated song titles and performers, were merged into a single spreadsheet and sorted alphabetically by tag name for analysis.

Tags were examined in relation to their associated works and artists, and subsequently analyzed in relation to the FRBR Entities and Attributes. Two coders performed content analysis on the tags. The author coded all the tags collected for the song sample, a second coder was recruited to independently code a random sample consisting of 20% of this study's tags, and intercoder reliability was calculated. Wimmer and Dominik (1994) suggest additional coders analyze a subsample of 10%–25% of coded data and then calculate intercoder reliability to enhance content anaylsis reliability. Rorissa's (2010) study uses a random sample of 20% of the data coded for secondary coding and intercoder reliability testing. This study aligns with both

Wimmer and Dominik's (1994) recommendations and Rorissa's (2010) study methodology. The second coder recruited for this study holds an M.A. in Library and Information Science from the University of Wisconsin-Madison. Intercoder reliability was calculated using percent agreement and Cohen's Kappa.

FRBR-related tags were mapped to the appropriate Attribute within the FRBR Group 1, 2, or 3 Entities; non-FRBR tags were mapped to personal, foreign language, mood, artist, and subjective tag categories. A content analysis dictionary was developed, and used, for coding song tags (Appendix B). A list of FRBR Attributes, and their corresponding definitions, that are potentially applicable to resources within the music domain was generated using the Attributes listed in Chapter 4 of IFLA's (2008) FRBR Final Report. The Item entity and some attributes—use restrictions on the expression, groove width, kind of cutting, and other designation associated with the corporate body—were not included in the content analysis dictionary because they did not apply to the song sample's tags. Supplemental definitions were created for FRBR Group 1 and Group 3 attributes, and non-FRBR attributes were defined. The following definitions were developed for mapping tags to non-FRBR categories. Artist tags provide descriptive data about singers, songwriters, or performers. Personal tags are generated by users to describe personal, abstract ideas about a song. Foreign language tags are non-English tags employed by users—these tags were not translated and coded to FRBR attributes. Subjective tags express qualitative judgments about songs, artists, or recordings. Mood tags describe or evoke a feeling or state of mind. Polysemic tags were mapped to each appropriate Entity, Attribute, or non-FRBR category. Examples for each tag attribute mapping were included in the content analysis dictionary.

Trial coding was conducted, at which time it was discovered that it was often necessary to briefly research individual songs to gain an understanding of tag-song relationships and

appropriately map tags. Coders determined each tag's context in relation to its associated song by searching each tag on the last.fm Web site to determine if last.fm users defined the tag. Coders then determined to which FRBR attributes or non-FRBR attributes defined tags corresponded, and completed mapping based on the tag definitions and coding definitions. Undefined tags, as well as tags without apparent non-FRBR attribute mappings, required further research to determine context in relation to the tagged song and appropriate mapping. Coders examined songs' Wikipedia entries for information about corresponding undefined tags to determine the tag-song relationship. If context remained unclear after consulting Wikipedia, coders performed three separate Google searches—song title and tag, recording (i.e., manifestation) title(s) and tag, and artist name(s) and tag—to determine if context could be gleaned from information on the Web and then mapped to an appropriate attribute.

Tag mapping counts were generated for the number of overall tag mappings, FRBR-related tag mappings, non-FRBR tag mappings, tag mappings to each FRBR Entity, tag mappings to each FRBR Attribute, and tag mappings to each non-FRBR tag category. These counts were used to calculate tag mapping percentages. Tag mapping percentages were calculated for each of these counts to the overall number of tag mappings. Percentages of each FRBR Entity's tag mappings to all FRBR-related tag mappings were calculated, as were FRBR Attribute's tag mappings to all FRBR-related tag mappings. Percentages of each non-FRBR category to all non-FRBR tag mappings were also calculated. Percentages of the FRBR Attributes' tag mappings within their respective Entities were calculated.

Tag mapping counts were generated for each song in the sample. These counts include total tag mappings for each song, total non-FRBR tag mappings for each song, total FRBR-related tag mappings for each song, and tag mappings to individual FRBR Entities for each song.

Mean and mode tag mappings per song were calculated for FRBR-related tag mappings, non-FRBR tag mappings, and tag mappings to each FRBR Entity (Appendix C). Frequency distributions were generated for overall FRBR tag mappings per song and for each FRBR Entity's tag mappings per song.

User-generated tags were examined by the FRBR Attribute or non-FRBR category, to which they were mapped, to determine if specific Attributes or categories enhance music resource description. Non-FRBR tag mappings were examined by category to determine if user-generated tags in particular categories have potential to enhance music resource description and search. The non-FRBR tag mappings were also examined to determine if user-generated tags in specific categories provide any reason to recommend revisions to the FRBR conceptual model.

The sample population's top-ranking on a popular consumer music list increases the likelihood that these songs are persistent works; therefore, these songs are likely familiar to music listeners and have been heavily tagged on the last.fm Web site. Other studies that examine music social tagging have also utilized the last.fm Web site's data, this allows for better comparisons between this study and other studies investigating music social tagging as a means for resource description and access. Content analysis was completed by two coders and intercoder reliability was calculated to enhance objectivity and reliability.

The last.fm Web site displays a maximum of 60 user-generated tags per resource; while this maximum provides many tags, the data collected cannot be all-inclusive. This study will examine tags related to artists and recordings that are associated with the individual songs analyzed; however, this study does not examine tags additionally assigned to the aggregate works' (i.e., recording or manifestation) pages or artists' pages on the last.fm Web site. Last.fm provides frequencies for tags applied across the site's objects, but does not provide tag

frequencies for individual songs; therefore, tag frequency by song could not be examined in this study.

Study Results and Discussion

A total of 2714 user-generated tags were collected from the last.fm Web site for 48 songs in the sample with a mean of 56.5 tags per song and a mode of 60 tags per song. One song in the sample—*The Candy Man* by Sammy Davis, Jr. with The Mike Curb Congregation—did not have an entry on the last.fm Web site, and therefore, did not have any user-generated tags associated with it. Another song in the sample—*Please Don't Go* by KC And The Sunshine Band—had an entry on the last.fm Web site but did not have any user-generated tags associated with it.

Each tag was analyzed in relation to FRBR Group 1, 2, and 3 Entities and Attributes. FRBR related tags were mapped to the appropriate FRBR Attribute, and tags that could not be related to the FRBR model were mapped to non-FRBR categories. 1384 of the collected tags mapped to at least one FRBR Entity or Attribute, while 1330 tags did not map to a FRBR Entity or Attribute. A total of 3161 mappings were applied to the 2714 collected tags with a mean of 66 tag mappings per song and a mode of 67 tag mappings per song. The number of tag mappings is greater than the total number of tags because polysemic tags were mapped to each appropriate FRBR Entity or Attribute or non-FRBR category. There were 387 polysemic tags among the 2714 tags that were analyzed; 327 polysemic tags mapped to two tag categories, and 60 polysemic tags mapped to three tag categories.

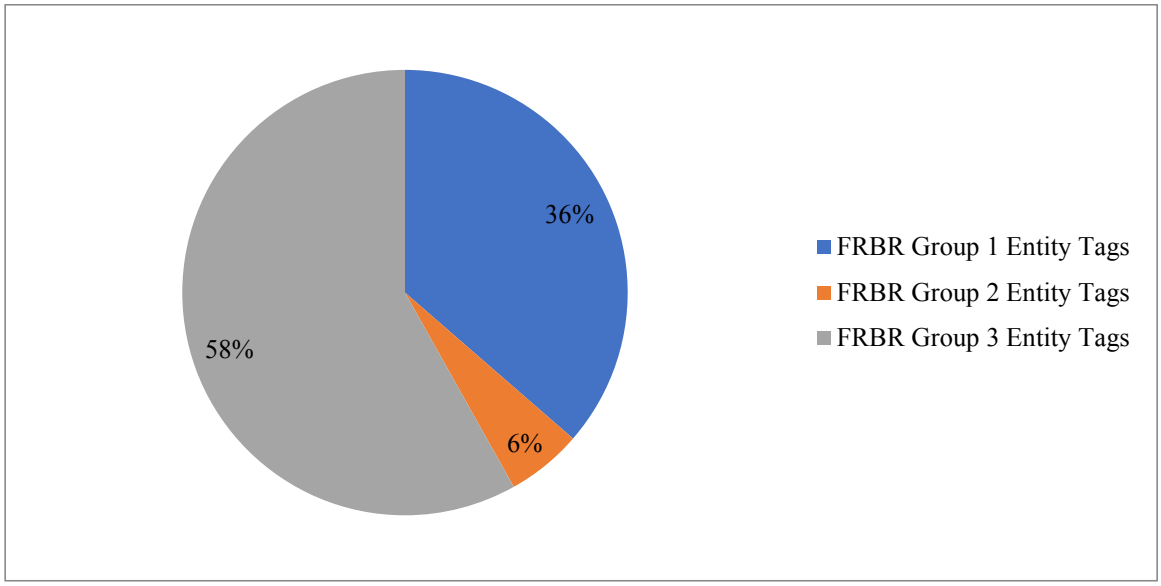
Polysemy among the tags is mainly explained by inheritance and overlap in the FRBR model. Tags that mapped to Statement of Responsibility Attributes in the Group 1 Entities also mapped to Group 2 Entities, and there was some occasional overlap with Statement of Responsibility Attributes within the Group 1 Work and Manifestation Entities. Title of the

Expression was inherited from Title of the Work, and occasionally matched the Title of the Manifestation; therefore, title tags mapped to at least two and sometimes three FRBR attributes. There was some overlap with date attributes in FRBR Group 1 Entities, particularly Date of Expression and Date of Publication/Distribution for songs that were recorded and released within the same year. A portion of the overall tags also contained combinations of ideas, such as genre and time period concatenations.

1384 collected tags (51% of the tags) were analyzed as FRBR-related tags with a total of 1767 tag mappings (56% of the tag mappings) made to FRBR Entities and Attributes. 1330 collected tags (49% of the tags) were analyzed as non-FRBR tags with a total of 1394 tag mappings (44% of the tag mappings) made to non-FRBR categories—these were coded as artist tags, foreign language tags, mood tags, personal tags, or subjective tags. The user-generated tags associated with the songs in the sample had a mean of 37 FRBR-related tag mappings per song and a mean of 29 non-FRBR tag mappings per song; the mode for FRBR-related tag mappings was 38, and the mode for non-FRBR tag mappings was 28. These results demonstrate that a majority of the user-generated tags map to the FRBR conceptual model and that users' mental models are aligned with the FRBR conceptual model in some ways.

The majority of FRBR-related tag mappings—1027 tag mappings (33% of the overall tag mappings, and 58% of the FRBR-related tag mappings)—were coded to Attributes within the FRBR Group 3 Entities. This was followed by FRBR Group 1 Entity tag mappings, with 643 tag mappings (20% of the overall tag mappings, and 36% of the FRBR-related tag mappings). 97 tag mappings (3% of the overall tag mappings, and 6% of the FRBR-related tag mappings) were made to Attributes within the FRBR Group 2 Entities (Figure 2).

Figure 2: Percentages of FRBR-Related Tag Mappings by Entity

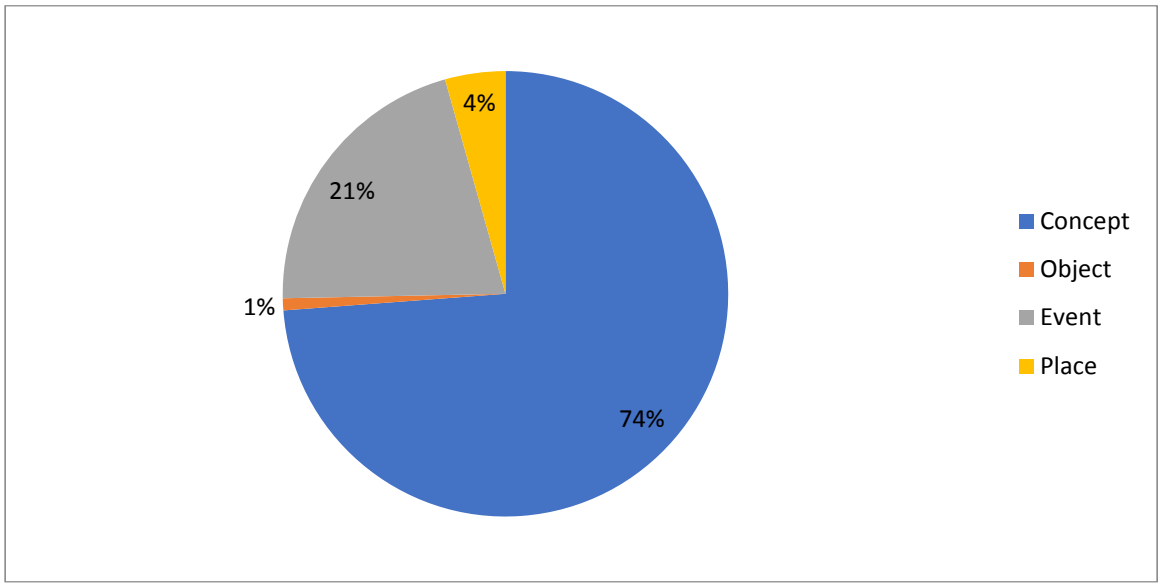


The FRBR Group 3 Concept Entity was not only the most prevalent FRBR Entity among the FRBR-related tag mappings, but also the most heavily mapped category overall (Figure 3). A total of 758 tag mappings were made to the FRBR Group 3 Concept Entity, accounting for 24% of the overall tag mappings, and 43% of the FRBR-related tag mappings.

There was a high rate of tagging, with 494 tag mappings, associated with music genres, subgenres, and styles among those tags mapped to the FRBR Group 3 Concept Entity. A notable occurrence here is that users employed combinations of broader, narrower, different tags to describe genre, subgenre, and style for the same song, indicating that people think differently about genre within the music domain. Library catalogs provide descriptions of, and access to, music resources at the manifestation level with limited musical form-genre headings encompassing all of the content in a given manifestation. User-generated genre tags applied at the song level can therefore enhance access to music resources by providing a larger variety of genre tags to a single resource; furthermore, the resource description provided by user-generated

tags at the song level provides more granularity than library catalogs' descriptions of the Manifestation on which the song appears. This can be seen in the following two examples.

Figure 3: Percentages of FRBR Group 3 Entity Tag Mappings



The song *A Hard Day's Night* by The Beatles has thirteen user-generated genre tags—*alternative*, *britpop*, *classic rock*, *dance*, *funk*, *indie*, *merseybeat*, *pop*, *pop – classic*, *rock*, *rock and roll*, *rock n roll*, *pop-rock*, *pop rock*, and *rockpop*—associated with it on the last.fm Web site (last.fm, 2014a). Whereas OCLC's WorldCat has 2 genre subject headings—*Rock music* and *Rock music--1961-1970*—associated with the Manifestation on which this song appears (OCLC, 2014); in this case, the song title is also the Title of the Manifestation. The song *Good Times* by Chic has seventeen user-generated genre tags—*70s disco*, *70s soul*, *acid jazz*, *alternative*, *classic rock*, *dance*, *disco*, *funk*, *house*, *jazz*, *pop*, *r&b*, *rb*, *rhythm and blues*, *rnb*, *rock*, and *soul*—associated with it on the last.fm Web site (last.fm, 2014). Whereas OCLC's WorldCat has 2 genre subject headings— *Popular music* and *Popular music--1971-1980*—associated with the Manifestation on which this song appears, which is titled *Risqué* (OCLC, 2014a).

The remaining 264 FRBR Group 3 Concept Entity tag mappings are for user-generated tags that describe what the individual songs are about—with tags such as *death*, *love*, and *breakup*—and topics associated with the individual songs—with tags such as *british invasion*, and *call and answer vocal harmony*. Current practice in applying subject headings to music resources does not permit catalogers to apply topical subject headings to describe a music resource's *aboutness*. Even if cataloging rules allowed for this type of subject heading application, it would be difficult to implement this type of resource description in library catalogs because resource description occurs at the sound recording level, rather than the individual song level. This reveals another way in which user-generated tags at the song level enhance access to music resources.

215 user-generated tags mapped to the FRBR Group 3 Event Entity, with 99% of these tags pertaining to time periods. The tags mapped to this Entity (e.g., 1970s, eighties) are primarily used to describe the decades within which the tagged songs were recorded and released. The remaining 1% of user-generated tags mapped to the FRBR Group 3 Event Entity are related to a specific event—the Country Music Association Festival. There was some concatenation of time periods and genres within single tags (e.g., 60s rock n roll), which is a phenomenon also seen in Library of Congress Subject Headings for music resources.

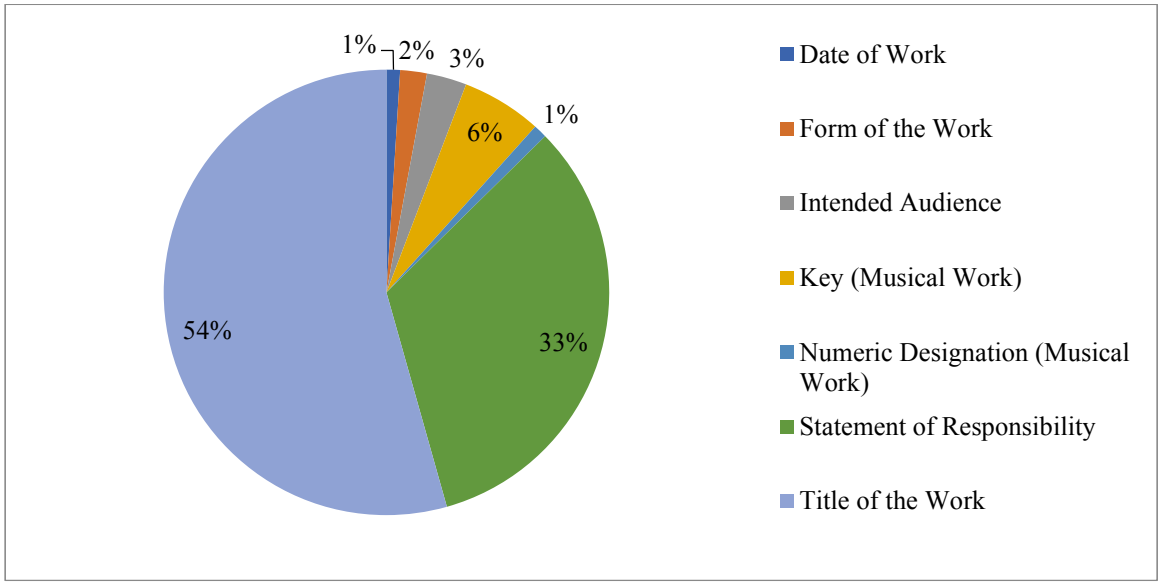
A small number of tags mapped to FRBR Group 3 Place and Object Entities. 45 tags, primarily for cities, states, countries, and geographic areas (e.g., *san francisco*, *new jersey*, *usa*) mapped to the Place Entity. Only 9 tags mapped to the Object Entity.

Within the FRBR Group 1 Entities, user-generated tags mapped to Attributes in the Work Entity, the Expression Entity, and the Manifestation Entity. None of the user-generated tags

mapped to Attributes in the Item Entity. A total of 643 tag mappings were made to the FRBR Group 1 Entities.

103 user-generated tags mapped to Attributes in the FRBR Group 1 Work Entity (Figure 4). The Title of the Work and Statement of Responsibility Attributes had the highest tag concentrations within this Entity with 56 and 34 tags mapped to the Attributes, respectively. Date of the Work, Form of the Work, Intended Audience, Key (Musical Work), and Numeric Designation (Musical Work) Attributes also had tags mapped to them; however, tagging for these attributes was minimal.

Figure 4: Percentages of FRBR Work Entity Tag Mappings by Attribute



353 user-generated tags mapped to Attributes in the FRBR Group 1 Expression Entity, making it the most prevalently tagged FRBR Group 1 Entity (Figure 5). User-generated tags that were mapped to the Title of the Work Attribute were also mapped to the Title of the Expression Attribute because the Work and Expression titles are the same in the case of the sample data, with the Expression inheriting its title from the Work.

The Critical Response to the Expression Attribute had the highest concentration of tag mappings within the FRBR Group 1 Expression Entity with 133 tags mapped to this Attribute. Most of these tags mapped to this Attribute were related to the songs' rankings on the Billboard Hot 100 list—the song sample's source. Even if the 78 Billboard-related tags mapped to the Critical Response to the Expression Attribute are removed from the mappings, the FRBR Group 1 Expression Entity still has the most tag mappings of the FRBR Group 1 Entities with 243 tags mapped to this Entity.

The Medium of Performance Attribute had the second highest concentration of tags mapped to it within the FRBR Group 1 Expression Entity with 93 tags mapped to this Attribute. This Attribute provides information about instrumentation and voice type for the Expression. These data are particularly useful to users interested in finding music resources with specific instrumentation or voice types. Instrumentation and voices can vary from song to song on a given Manifestation; therefore, this information is not always presented in library catalog records for musical sound recordings, whereas it is usually found in library catalog records for musical scores. Instrumentation and voice types can also vary among Expressions of the same song. These data can be useful to users searching for song recordings with particular instrumentation or voice types.

31 user-generated tags mapped to the Other Distinguishing Characteristic Attribute, which has the purpose of differentiating Expressions of the same Work. 18 of the tags mapped to this attribute were a concatenated author-title description of the song (e.g., *robert palmer addicted to love*). The remaining tags mapped to this Attribute provide alternate titles by which the Expression is known—for example, a user applied the tag *wimoweh* to the song *The Lion*

Sleeps Tonight by the Tokens—or are used to indicate that an Expression is not the original performed Expression of the Work (e.g., the tag *cover*).

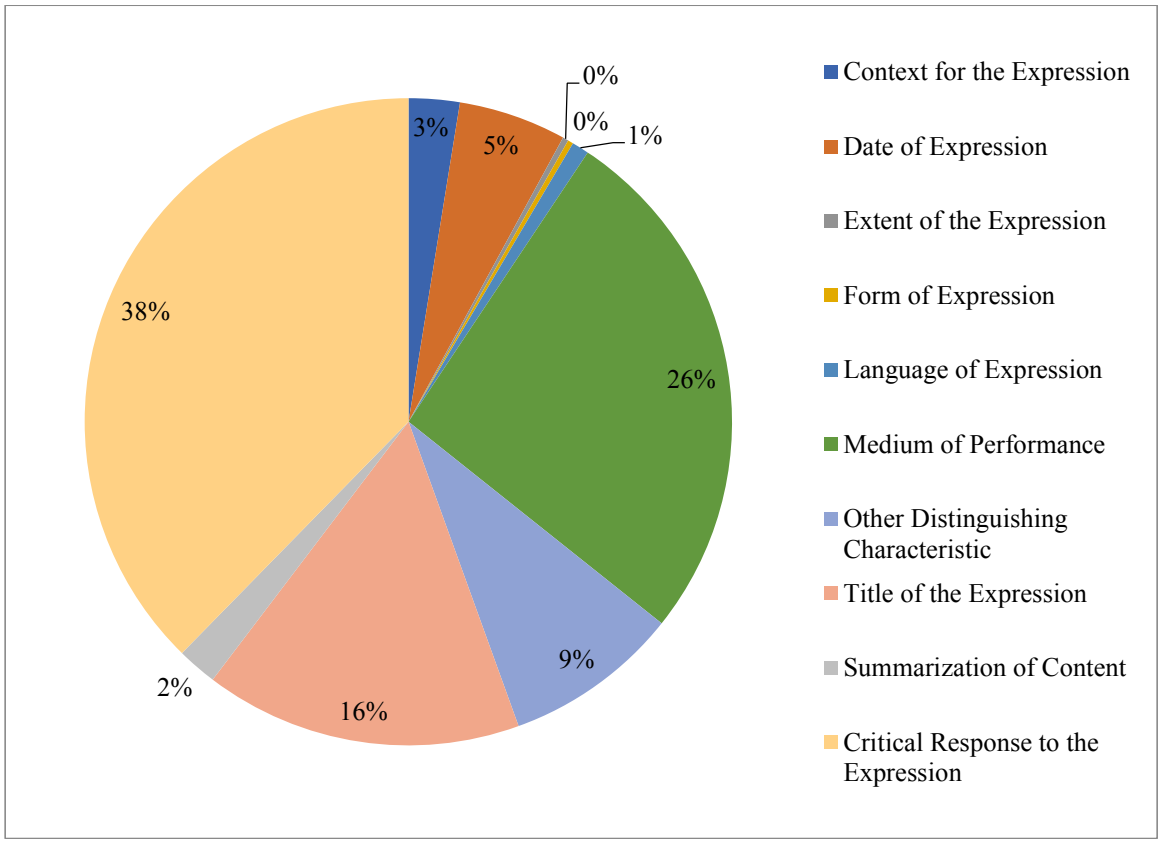
Of the remaining user-generated tags mapped to the FRBR Group 1 Expression Entity, 19 tags were mapped to the Date of Expression Attribute, and tags were mapped at minimal levels to Context for the Expression, Language of Expression, Extent of the Expression, Summarization of Content, and Form of the Expression Attributes. Tags mapped to the Summarization of Content Attribute while not frequent, can provide users, who are searching for a specific type of content in a recorded song, with information. Examples of user-generated tags mapped to this Attribute in the sample include, but are not limited to, *an electric guitar solo*, *drum break*, and *repetitive melodic phrasing*.

187 user-generated tags mapped to Attributes in the FRBR Group 1 Manifestation Entity. The Statement of Responsibility Attribute had the highest concentration of tag mappings within the FRBR Group 1 Manifestation Entity with 90 tags mapped to this Attribute. Higher tag mappings for the Manifestation Statement of Responsibility Attribute than the Work Statement of Responsibility Attribute can be explained by the fact that the person or people who perform a song are not always the same person or people who write and compose the song. Furthermore, performers generate a larger output than do composers and writers (Papakhian, 1985). Tags with performer and producer names were mapped to the Manifestation Statement of Responsibility Attribute.

40 user-generated tags mapped to the Title of the Manifestation Attribute (Figure 6). The Work and Expression Title Attributes match the Title of the Manifestation Attribute when a song is the title track for the recording—this was the case with 20 of the tags mapped to this attribute. Conversely, the Title of the Manifestation Attribute diverges from the Work and Expression Title

Attributes when a song is not the title track for the recording; there were 20 tags that mapped to the Title of the Manifestation Attribute but did not map to the Title of the Work or Title of the Expression Attributes.

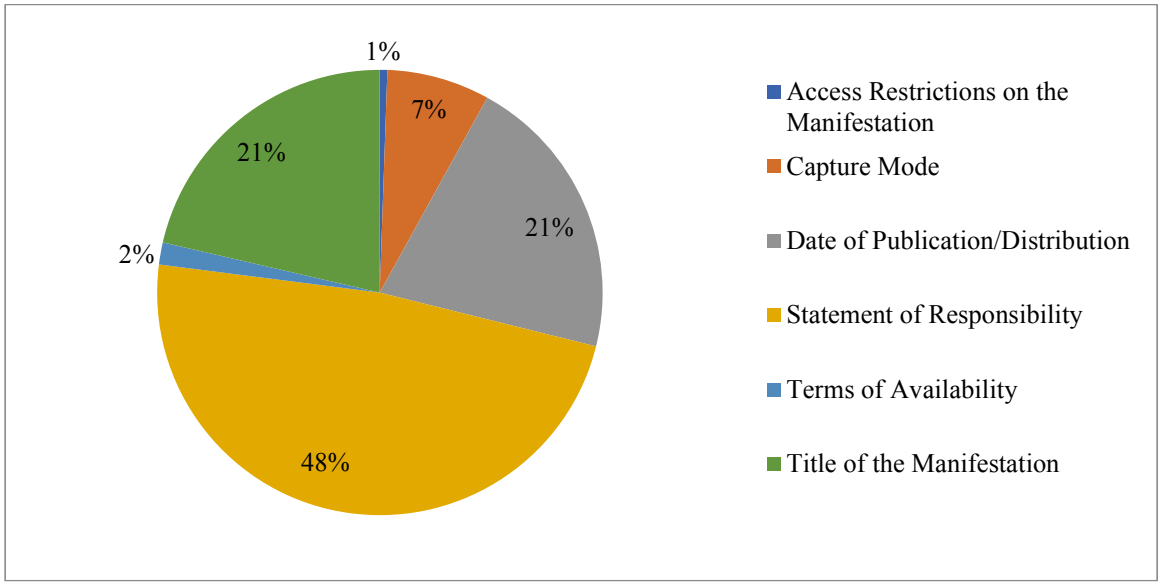
Figure 5: Percentages of FRBR Expression Entity Tag Mappings by Attribute



39 user-generated tags mapped to the Date of Publication/Distribution Attribute, which is a higher number of date-associated tags than mapped to the Date of Work or Date of Expression Attributes. In the case of recorded music, dates will vary for when the song as a Work is written or composed, when that song is realized through an Expression, and when it is embodied in a Manifestation. It is the Manifestation's release date, or publication date, that is most readily available to the user on the physical recording and in current the library catalog records, which can explain the higher number of tag mappings for the Date of Publication/Distribution Attribute than for the Date of Work or Date of Expression Attributes.

14 user-generated tags mapped to the Capture Mode Attribute, which describes how the song was recorded (e.g., *acoustic*). A small number of user-generated tags mapped to the Terms of Availability and Access Restrictions on the Manifestation Attributes. Three tags described songs' availability for the *Rock Band* video game, and one tag described a song's streamability with an online music provider.

Figure 6: Percentages of FRBR Manifestation Entity Tag Mappings by Attribute

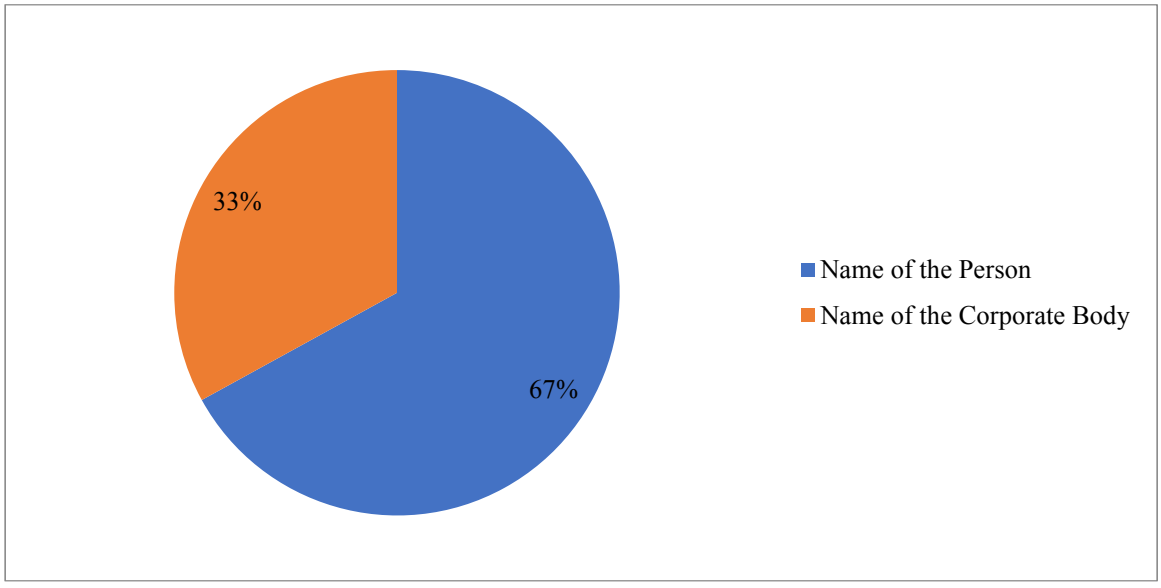


97 user-generated tags mapped to two Attributes within the FRBR Group 2 Entities (Figure 7). 65 tags were mapped to the Name of Person Attribute, and 32 tags were mapped to the Name of the Corporate Body Attribute. All of the tags mapped to these two Attributes were also mapped to Statement of Responsibility Attributes within the FRBR Group 1 Entities.

OCLC's approach to FRBRizing WorldCat utilizes an algorithm to generate an author-title key from MARC21 data; the key is then used as a Work identifier (Hickey & O'Neill, 2005). This combination of author and title with classical music resources—the composer's name followed by the Work title—aligns with users' mental models in how users approach searching library catalogs for classical music resources (Dickey, 2008; King, 2005). Data

collected in this study show that some users think about popular music resources in this way, too; however, these name-title concatenations are primarily Expression-level performer-title combinations, rather than the Work-level creator-title combinations. This is best demonstrated by the tags mapped to the Expression Entity's Other Distinguishing Characteristic Attribute. This is another example of how user-generated song tags can enhance access to music resources, and this shows how the FRBR conceptual model can improve access to popular music sound recordings. These user-generated tags, in combination with the tags mapped to either Statement of Responsibility or Title Attributes within the FRBR Group 1 Entities, could potentially be used by information professionals to generate additional references to authorized headings in the event that resource description can occur at the song level. Users have difficulty constructing Uniform Titles at the Work-level (Kranz, 1988); therefore references derived from user-generated tags might assist users with known-item searching for songs.

Figure 7: Percentages of FRBR Group 2 Entity Tag Mappings by Attribute



Given that the user-generated tags in this sample show that people create tags associated with the performers who realize a Work's Expression, it would be useful to expand the FRBR

Group 1 Expression Entity to include a Statement of Responsibility Attribute to enhance description of, and access to, songs and similar performance-based resources. These tags were mapped to the Manifestation Entity's Statement of Responsibility Attribute in this study; however, a sound recording as a Manifestation of songs is an aggregate work with each song as an individual work in its own right. The song performers are not necessarily entitled to a Statement of Responsibility entry at the Work-level; therefore, a Statement of Responsibility Attribute within the Expression Entity can improve the song descriptions.

The idea of time also plays an important role in how users think about resources within the music domain. This is evidenced by the 272 tag mappings (11.24% of overall tag mappings, and 26.48% of FRBR-related tag mappings) made to Date Attributes within the FRBR Group 1 Entities and FRBR Group 3 Event Entity related to time periods. Time-related tag mappings are more heavily concentrated in the FRBR Group 3 Event Entity with 213 Event Entity tag mappings related to time periods. In the FRBR Group 1 Entities, a total of 59 tag mappings were made to Date of Work, Date of Expression, and Date of Publication/Distribution Attributes. These data can also serve as a level of validation for the FRBR conceptual model's time-related Entities and Attributes.

The majority of user-generated tags that were mapped to non-FRBR tag categories were coded as personal tags and subjective tags. 719 user-generated tags were mapped to the personal tag category; examples of personal tags include *amayzes loved*, *bianca smiles*, and *cali baby*. 362 user-generated tags were mapped to the subjective tag category; examples of subjective tags include *best*, *great songs*, and *ridiculous*. While tags mapped to these two non-FRBR categories may help individual users in recalling songs, they do not aide the broader user population in searching for music resources nor do they enhance resource description from a knowledge

organization perspective. However, a portion of the non-FRBR tag mappings—tags mapped to artist tag and mood tag categories—demonstrate potential for enhancing access to music resources and expanding upon the FRBR entity-relationship model. 135 user-generated tags mapped to the non-FRBR artist tag category, and 133 tags mapped to the non-FRBR mood tag category. These combined tag mappings accounted for 11.07% of the overall mappings and 19.23% of the non-FRBR tag mappings.

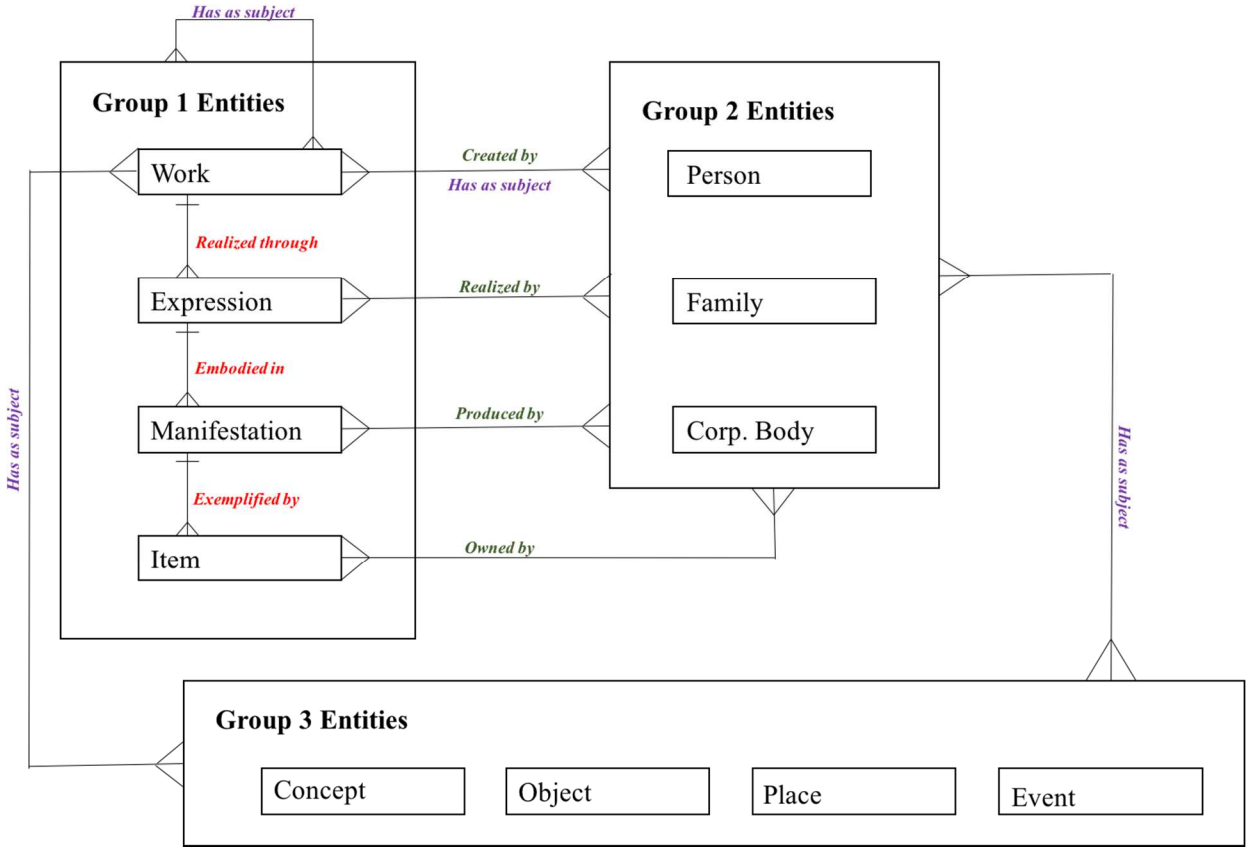
Although mood can be considered subjective, mood tags like those in this sample—*cheerful*, *happy*, *mellow*, *romantic*, *sad*, *upbeat*—can assist users in finding music resources by mood. Current library catalogs do not offer users the option to search for music resources by mood. Searching by mood is not explicitly prohibited in the FRBR model, rather it is the individual cataloging codes that do not employ this type of resource description. Mood could potentially be positioned within the FRBR Group 3 Concept Entity. Such tags are better suited to resource description at the song-level and would be difficult to apply at the Manifestation-level because a Manifestation can contain multiple songs, and each song could potentially have different mood tags with which it is associated. Therefore, while this is a potentially valuable search facet for users, it would be difficult to apply this level of resource description to current library catalogs, which primarily catalog musical sound recordings at the Manifestation-level. However, a FRBRized interface capable of searching at the song- or track-level—displaying the hierarchical connection of a song to its Manifestation and then linking to Items in a library catalog—would allow users to search for songs by mood and locate a manifestation with the song or track. This approach also adheres to the FRBR user tasks—find, identify, select, and obtain.

User-generated tags that were mapped to the non-FRBR artist tag category consist of tags that describe the performing artist or musical group responsible for an Expression's performance. Given that the artists and bands fall within the FRBR Group 2 Person or Corporate Body Entities, the artist tags from this sample are essentially subject tags for FRBR Group 2 Entities in that they describe the Group 2 Entities' *aboutness*, *ofness*, and *isness*. Examples of these tags within the sample include—but are not limited to—*60s girl group*, *british*, *dead rockstars*, *duos*, *singer-songwriter*, and *tv star*. The ideas expressed in these tags best fit within the FRBR Group 3 Entities; however, the FRBR model does not directly relate Group 3 Entities with Group 2 Entities (Figure 1). The *aboutness* of a specific Group 2 Entity must be conveyed via a separate Work, which then has that Group 2 Entity as its subject, within the FRBR model. This means that a user interested in finding a song performed by a 1960s girl group must first identify 1960s girls groups before searching for Expressions of Works performed by 1960s girl groups. In other words, the user tasks would be sequenced as find, identify, and select a Group 2 Entity meeting the criteria of being a 1960s girl group, followed by find, identify, select, and obtain a sound recording with a song performed by the 1960s girl group identified in the first half of this sequence.

In light of the above example of finding an Expression of a song performed by a 1960s girl group, I recommend a revision to the FRBR conceptual model to include a relationship of Group 3 Entities to Group 2 Entities. The ability to describe Group 2 Entities in such a way would enable users to quickly find Works, Expressions, or Manifestations created by people or corporate bodies defined as Group 2 Entities by describing subject Attributes associated with these Group 2 Entities. The resulting model would resemble (Figure 8). This approach to searching music resources would also require enhanced linking of name authorities with their

appropriate subject authorities in library catalogs. There has been some discussion about the FRBR conceptual model's potential applications in the Semantic Web and Linked Open Data (Gradmann, 2005; Howarth, 2012). Incorporating this recommended revision into the FRBR conceptual model would better position the model as a tool for the Semantic Web and Linked Open Data. Such an expansion of the FRBR model also better adheres to the four user tasks—find, identify, select, and obtain—when searching for songs performed by people who meet a certain subject criteria. The sequence—find, identify, select, find, identify, select, and obtain—outlined in the 1960s girl group example would be reduced to simply find, identify, select, and obtain—thus meeting one of the FRBR conceptual model's principle goals.

Figure 8: Recommended Revision to the FRBR Entity-Relationship Model



Conclusions

General Conclusions

Users describe music resources in a way that aligns with the FRBR model as evidenced by this study's tag mapping results. The majority of tags applied to sampled songs mapped to FRBR-based attributes within all three FRBR group entities. Users describe resources primarily by subject, title, and creator (i.e., statement of responsibility) within these FRBR-based mappings. Users predominantly apply subject tags to popular songs to describe a song's aboutness, genre, and related subjects. This study's tag mapping helps validate most of the FRBR attributes outlined in the study's content analysis dictionary. Tags were mapped to all but four FRBR attributes used in the content analysis dictionary. The Manifestation entity's attributes for Playing Speed, Tape Configuration, Kind of Sound, and Special Reproduction Characteristics did not correspond to any of the tags examined in this study.

Social tagging enhances access to music resources because it provides more granular indexing data for songs than do library catalogs. The songs in this study's sample are essentially *bona fide* works—to use Le Boeuf's (2005b) and Smith and Varzi's (1997) terminology—within aggregate works. Library catalogs contain records for the aggregate works—or complete sound recordings—and may index the individual songs in a free-text note field, but library catalog records do not provide the level of description for individual songs that tags provide. Social tagging can describe resources with greater descriptive inclusivity through the combination of the broader and narrower subject and genre terms users apply to resources.

Cataloging codes and practices are oriented toward manifestation-level music description with subject headings providing genre information. Popular music and songs are frequently about a topic or subject, and the different songs on a single recording can vary in subject;

however, library catalog records cannot capture and express this information under current codes and practices. Future cataloging code revisions should investigate how to provide song-level indexing and subject description to make music search and discovery more user-oriented, as this study demonstrates that song-level subject description is integral to how people think about and describe music resources. This change is also important due to shifts in music consumer behavior that places greater emphasis on individual songs over albums, given consumers' abilities to purchase individual songs through digital music platforms (Elberse, 2010)

Study strengths

It is important that cataloging practices meet the needs of both information professionals and users. This study aids the cataloging and metadata communities in validating portions of the FRBR conceptual model from the user perspective. This study also demonstrates the importance of non-required bibliographic entities and elements to user search and discovery, as well as how social tagging enriches resource description. Music materials can be particularly difficult to find in library catalogs, and this study provides insight into improvements for music cataloging to aid discovery.

Study limitations

The sample in this study consists of only popular music. The genres are therefore limited in scope, and may not be indicative of all music tagging behavior. This study does not examine how social tags for non-music objects (e.g., textual, visual art, dance, video) align with the FRBR conceptual model.

Future research possibilities

A similar content analysis study using laypeople as coders to map social tags to FRBR attributes can further aide in determining how well user-generated social tags describe the

resources to which they are applied and whether there are commonalities among users' mental models. Social tags for other music genres and non-textual resources should be examined and analyzed in relation to the FRBR conceptual model to further this research. It will also be useful to examine Last.fm tags on artists' (i.e., performers, song writers, composers) pages in relation to FRBR Group 2 entity attributes to determine how user description maps to, and potentially validates, this portion of the FRBR model.

References

Albertsen, K., & van Nuys, C. (2005). Paradigma: FRBR and digital documents. *Cataloging & Classification Quarterly*, *39*(3-4), 125–149.

Arastoopoor, S., Fattahi, R., & Parirokh, M. (2011). Developing user-centered displays for literary works in digital libraries: Integrating bibliographic families, FRBR and users. *Proceedings of the Second International Conference of Asian Special Libraries, February 10-12, 2011, Tokyo*. Retrieved from http://asia.sla.org/Sholeh%20Arastoopoor.doc

Bar-Ilan, J., Shoham, S., Idan, A., Miller, Y., & Shachak, A. (2008). Structured versus unstructured tagging: a case study. *Online Information Review*, *32(5),* 635–647.

Bates, J., & Rowley, J. (2011). Social reproduction and exclusion in subject indexing. *Journal of Documentation*, *67*(3), 431–448.

Bennett, R., Lavoie, B. F., & O'Neill, E. T. (2003). The concept of a work in WorldCat: An application of FRBR. *Library Collections, Acquisitions, and Technical Services*, *27*(1), 45–59.

Berners-Lee, T. (2000, December 6). *Semantic Web on XML [Conference presentation]*. Retrieved from http://www.w3.org/2000/Talks/1206-xml2k-tbl

Billboard Music. (n.d.). *The hot 100*. Retrieved from http://www.billboard.com/charts/hot-100

Borgman, C. L. (1996). Why are online catalogs still hard to use? *Journal of the American Society for Information Science, 47(7),* 493–503.

Bowen, J. (2005). FRBR. *Library Resources & Technical Services, 49(3),* 175–188.

Boyd, A. (2005). The worst of both worlds: How old rules and new interfaces hinder access to music. *CAML Review*, *33(3).* Retrieved from https://pi.library.yorku.ca/ojs/index.php/caml/article/viewArticle/1389

Carlyle, A. (1999). User categorisation of works: Toward improved organisation of online catalogue displays. *Journal of Documentation*, *55*(2), 184–208.

Carlyle, A. (2001). Developing organized information displays for voluminous works: A study of user clustering behavior. *Information Processing and Management, 37(5),* 677–699.

Chen, M., Liu, X., & Qin, J. (2008). Semantic relation extraction from socially-generated tags: A methodology for metadata generation. *Proceedings of the International Conference on Dublin Core and Metadata Applications*, *DC-2008, Berlin, September 22-26, 2008.* (pp. 117–127). Retrieved from http://dcpapers.dublincore.org/pubs/article/view/924/920

Chen, Y., & Chen, S. (2004). A metadata practice of the IFLA FRBR model: A case study for the National Palace Museum in Taipei. *Journal of Documentation, 60(2),* 128–143.

Chi, E. H., & Mytkowicz, T. (2008). Understanding the efficiency of social tagging systems using information theory. In *Proceedings of the Nineteenth ACM Conference on Hypertext and Hypermedia* (pp. 81–88). New York, NY: ACM.

Choi, Y. (2012). A practical application of FRBR for organizing information in digital environments. *Knowledge Organization, 39(4),* 233–254.

Deodato, J. (2010). Deconstructing the library with Jacques Derrida: Creating space for the "other" in bibliographic description and classification. In G. J. Leckie, L. M. Given, & J. E. Buschman (Eds.), *Critical theory for library and information science: Exploring the social from across the disciplines* (pp. 75–88). Santa Barbara: Libraries Unlimited.

Dickey, T. J. (2008). FRBRization of a library catalog: Better collocation of records, leading to enhanced search, retrieval, and display. *Information Technology & Libraries, 27(1),* 23-32.

Dimec, Z., Žumer, M., & Riesthuis, G. J. A. (2005). Slovenian cataloguing practice and functional requirements for bibliographic records: A comparative analysis. *Cataloging & Classification Quarterly*, 39(3-4), 207–227.

Elberse, A. (2010). Bye-bye bundles: The unbundling of music in digital channels. *Journal of Marketing, 74(3),* 107–123.

Ercegovac, Z. (2006). Multiple-version resources in digital libraries: Towards user-centered displays. *Journal of the American Society for Information Science and Technology, 57(8)*, 1023–1032.

Eerola, J. & Vakkari, P. (2008). How a general and a specific thesaurus cover expressions in patients' questions and physicians' answers. *Journal of Documentation*, *64(1),* 131-142.

Fidel, R. (1994). User-centered indexing. *Journal of the American Society for Information Science*, *45*(8), 572–576.

Fox, R. (2006). Cataloging for the masses. *OCLC Systems & Services: International Digital Library Perspectives*, *22*(3), 166–172.

Furnas, G. W., Fake, C., von Ahn, L., Schachter, J., Golder, S., Fox, K., Davis, M., Marlow, C., and Naaman, M. (2006). Why do tagging systems work? In *Proceedings CHI '06 Extended Abstracts on Human Factors in Computing Systems* (pp. 36-39). New York, NY: ACM Press.

Furner, J. (2008). User Tagging of Library Resources: Toward a Framework for System Evaluation. *International Cataloging & Bibliographic Control*, *37*(3), 47–51.

Gentili-Tedeschi, M., & Riva, F. (2004). Authority control in the field of music: Names and titles. *Cataloging & Classification Quarterly*, *39(1/2),* 399-412.

Golder, S., & Huberman, B. (2006). Usage patterns of collaborative tagging systems. *Journal of Information Science, 32(2),* 198-208.

Gradmann, S. (2005). rdfs:frbr--Towards an implementation model for library catalogs using semantic web technology. *Cataloging & Classification Quarterly*, *39*(3/4), 63–75.

Guy, M., & Tonkin, E. (2006). Folksonomies: Tidying up tags? *D-Lib Magazine, 12(1)*, 7–20. http://doi.org/10.1045/january2006-guy

Hickey, T. B., & O'Neill, E. T. (2005). FRBRizing OCLC's WorldCat. *Cataloging & Classification Quarterly*, *39*(3-4), 239–251.

Hickey, T. B., & Toves, J. A. (2014). Managing ambiguity in VIAF, *D-Lib Magazine, 20(7/8)*. Retrieved from http://mirror.dlib.org/dlib/july14/hickey/07hickey.print.html

Hider, P., & Liu, Y. (2013). The use of RDA elements in support of FRBR user tasks. *Cataloging & Classification Quarterly, 51(8)*, 857–872.

Hoffman, G. L. (2010). Negotiating normative institutional pressures and maintaining legitimacy in a complex work environment: A multiple case study of three academic cataloging units. In D. E. Williams & J. Golden (Eds.), *Advances in Library Administration and Organization* (v.29) (pp. 243-292). Bingley, UK: Emerald Group Publishing Limited.

Hjørland, B. (2010). The foundation of the concept of relevance. *Journal of the American Society for Information Science and Technology, 61(2)*, 217-37.

Howarth, L. C. (2012). FRBR and linked data: connecting FRBR and linked data. *Cataloging & Classification Quarterly*, *50*(5-7), 763–776.

Hufford, J. R. (1991). The pragmatic basis of catalog codes: Has the user been ignored? *Cataloging and Classification Quarterly*, *14*(1), 27–38.

IFLA Study Group on the Functional Requirements for Bibliographic Records. (2008). *Functional Requirements for Bibliographic Records: Final Report*. Retrieved from https://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf

Jonsson, G. (2005). Cataloguing of hand press materials and the concept of expression in FRBR. *Cataloging & Classification Quarterly*, *39*(3-4), 77–86.

Jörgensen, C. (1998). Attributes of images in describing tasks. *Information Processing & Management, 34(2/3),* 161–174.

Kilner, K. (2005). The AustLit Gateway and scholarly bibliography: A specialist implementation of the FRBR. *Cataloging & Classification Quarterly*, *39*(3-4), 87–102.

King, D. M. (2005). Catalog user search strategies in finding music materials. *Music Reference Services Quarterly*, *9*(4), 1-24.

Kipp, M. E. I., & Campbell, D. G. (2006). Patterns and inconsistencies in collaborative tagging systems: An examination of tagging practices. In *Proceedings of the ASIST Annual Meeting*, *43* (pp. 1–18). doi:10.1002/meet.14504301178

Kranz, J. (1988). The music uniform title: Sources for the novice cataloger. *Cataloging and Classification Quarterly, 9(2),* 73-80.

Kroski, E. (2005, December 7). The Hive Mind: Folksonomies and User-Based Tagging [Web log post]. Retrieved from http://infotangle.blogsome.com/2005/12/07/the-hive-mind-folksonomies-and-user-based-tagging/

Last.fm. (2014). Chic—Good times. *last.fm*. Retrieved from http://www.last.fm/music/Chic/_/Good+Times/+tags.

Last.fm. (2014a). The Beatles—A hard day's night. *last.fm*. Retrieved from http://www.last.fm/music/The+Beatles/_/A+Hard+Day%27s+Night/+tags.

Le Boeuf, P. (2005a). FRBR: hype or cure-all? Introduction. *Cataloging & Classification Quarterly*, *39*(3-4), 1–13.

Le Boeuf, P. (2005b). Musical works in the FRBR model or "Quasi la Stessa Cosa": Variations on a theme by Umberto Eco. *Cataloging & Classification Quarterly*, *39*(3-4), 103–124.

Lorince, J., Joseph, K., & Todd, P. M. (2015). Analysis of Music Tagging and Listening Patterns: Do Tags Really Function as Retrieval Aids? In N. Agarwal, K. Xu, & N. Osgood (Eds.), *Social Computing, Behavioral-Cultural Modeling, and Prediction* (pp. 141–152). Springer International Publishing. Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-16268-3_15

Lorince, J., Zorowitz, S., Murdock, J., & Todd, P. M. (2015). Wisdom of the Few? "Supertaggers" in Collaborative Tagging Systems. *The Journal of Web Science 1(1)*. Retrieved from http://www.webscience-journal.net/webscience/index

Lu, C., Park, J., & Hu, X. (2010). User tags versus expert-assigned subject terms: a comparison of LibraryThing tags and Library of Congress Subject Headings. *Journal of Information Science*, *36(6),* 763-779.

Madison, O. M. A. (2000). The IFLA functional requirements for bibliographic records. *Library Resources & Technical Services, 44(3),* 153–159.

Macgregor, G., & McCulloch, E. (2006). Collaborative tagging as a knowledge organisation and resource discovery tool. *Library Review*, *55*(5), 291–300.

Mathes, A. (2004). *Folksonomies - Cooperative Classification and Communication Through Shared Metadata*. Unpublished manuscript, University of Illinois, Urbana-Champaign, IL.

McNellis, C. H. (1985). Describing reproductions: Multiple physical manifestations in the bibliographical universe. *Cataloging & Classification Quarterly, 5(3),* 35–48.

Meijas, U. A. (2004, December 27). A del.icio.us study [Web log post]. Retrieved from http://blog.ulisesmejias.com/2004/12/27/a-delicious-study/

Miller, D., & Le Boeuf, P. (2005). "Such stuff as dreams are made on": How does FRBR fit performing arts? *Cataloging & Classification Quarterly*, *39*(3-4), 151–178.

Morville, P. (2005). *Ambient Findability*. Sebastopol, CA: O'Reilly.

Murphy, H., & Rafferty, P. (2015). Is there nothing outside the tags? *Journal of Documentation*, *71*(3), 477–502.

Neuendorf, K. A. (2002). *The content analysis guidebook*. Thousand Oaks, CA: Sage Publications.

Nicolas, Y. (2005). Folklore requirements for bibliographic records: Oral traditions and FRBR. *Cataloging & Classification Quarterly*, *39*(3-4), 179–195.

O'Connor, B., & O'Connor, M. (1999). Categories, photographs & predicaments: Exploratory research on representing pictures for access. *Bulletin of the American Society for Information Science, 25(6),* 17–20.

OCLC Online Computer Library Center. (2014). A hard day's night (Musical CD, 1987). *WorldCat*. Retrieved from http://www.worldcat.org/oclc/15247456.

OCLC Online Computer Library Center. (2014a). Risqué (Musical CD, 1979). *WorldCat*. Retrieved from http://www.worldcat.org/oclc/32175669.

Olson, H.A. (2007). How we construct subjects: a feminist analysis. *Library Trends, 56(2)*, 509-41.

Palmer, S. B. (2001). The semantic Web: An introduction. *infomesh.net*. Retrieved from http://infomesh.net/2001/swintro/.

Papakhian, A. R. (1985). The frequency of personal name headings in the indiana university music library card catalogs. *Library Resources & Technical Services*, *29,* 273-285.

Patton, G. E. (2005). Extending FRBR to authorities. *Cataloging & Classification Quarterly*, *39*(3-4), 39–48.

Peponakis, M. (2012). Conceptualizations of the cataloging object: A critique on current perceptions of FRBR group 1 entities. *Cataloging & Classification Quarterly, 50(5–7)*, 587–602.

Pisanski, J., & Žumer, M. (2007). Functional requirements for bibliographic records: An investigation of two prototypes. *Program: Electronic Library & Information Systems, 41(4)*, 400–417.

_____. (2010a). Mental models of the bibliographic universe. Part 1: Mental models of descriptions. *Journal of Documentation, 66(5)*, 643–667.

_____. (2010b). Mental models of the bibliographic universe. Part 2: comparison task and conclusions. *Journal of Documentation,* 66(5), 668–680.

_____. (2012). User verification of the FRBR conceptual model. *Journal of Documentation, 68(4)*, 582–592.

Poroila, H. (2007). Waste of time or time saver?: Catalogues of uniform titles in finnish music libraries. *Fontes Artis Musicae*, *54(1),* 127-136.

Qin, J. (2008). Controlled semantics vs. social semantics: An epistemological analysis. In C. Arsenault, & J. T. Tennis (Eds.). (2008). *Culture and identity in knowledge organization: Proceedings of the Tenth International ISKO Conference* (pp. 229-234). Würzburg: Ergon Verlag.

Quintarelli, E. (2005, June). Folksonomies: power to the people. *Proceedings of the ISKO Italy-UniMIB meeting, June 24, 2005, Milan, Italy*. Retrieved from http://www.iskoi.org/doc/folksonomies.htm

Riley, J. (2011). Leveraging the FRBR model for music discovery and data sharing. *OCLC Systems & Services*, *27(3),* 175–189.

Riva, P, & Oliver, C. (2012). Evaluation of RDA as an implementation of FRBR and FRAD. *Cataloging & Classification Quarterly, 50(5–7)*, 564–586.

Rolla, P. J. (2009). User tags versus subject headings: Can user-supplied data improve subject access to library collections? *Library Resources and Technical Services*, *53*(3), 174–184.

Rorissa, A. (2010). A comparative study of Flickr tags and index terms in a general image collection. *Journal of the American Society for Information Science & Technology, 61(11),* 2230–2242.

Rorissa, A., & Iyer, H. (2008). Theories of cognition and image categorization: What category labels reveal about basic level theory. *Journal of the American Society for Information Science and Technology, 59(9),* 1383–1392.

Rosenfeld, L. (2005, January 6). Folksonomies? How about Metadata Ecologies? [Web log post]. Retrieved from http://louisrosenfeld.com/home/bloug_archive/000330.html

Shirky, C. (2005), "Ontology is Overrated: Categories, Links and Tags", Shirky.com, New York, USA, available: http://shirky.com/writings/ontology_overrated.html (accessed 27 July 2015)

Sinha, R. (2005), "A cognitive analysis of tagging", Rashmi Sinha's weblog, available: https://rashmisinha.com/2005/09/27/a-cognitive-analysis-of-tagging/ (accessed 17 May 2013)

Smiraglia, R. P. (1989). Uniform titles for music: An exercise in collocating works. *Cataloging and Classification Quarterly, 9(3),* 97-114.

_____ (2002). Musical works and information retrieval. *Notes*, *58(4),* 747.

_____ (2007a). Bibliographic families and super works. In A.G. Taylor (Ed.), *Understanding FRBR: What it is and how it will affect our retrieval tools*, pp. 73-86. Westport: Libraries Unlimited.

_____ (2007b). The 'works' phenomenon and best selling books. *Cataloging & Classification Quarterly, 44(3/4)*, 179–195.

Smith, B., & Varzi, A. C. (1997). Fiat and bona fide boundaries: Towards an ontology of spatially extended objects. In S. C. Hirtle, & A. U. Frank (Eds.), Proceedings from *Conference on Spatial Information Theory-COSIT 1997 International Conference*. (pp. 103–119). Berlin: Springer.

Smith, T. (2007). Cataloging and you: Measuring the efficacy of a folksonomy for subject analysis. In Lussky, Joan, Eds. *Proceedings 18th Workshop of the American Society for Information Science and Technology Special Interest Group in Classification Research*, *Milwaukee, Wisconsin*. Retrieved from http://arizona.openrepository.com/arizona/handle/10150/106434

Snyder, T. (2010) Music materials in a faceted catalog: Interviews with faculty and graduate students. *Music Reference Services Quarterly*, *13(3-4),* 66-95.

Spiteri, L. (2007). The structure and form of folksonomy tags: The road to the public library catalog. *Information Technology and Libraries, 26(3)*, 13–25.

Stock, W.G. (2007). Folksonomies and science communication: A mashup of professional science databases and Web 2.0 services. *Information Services & Use, 27,* 97–103.

Suchanek, F. M., Vojnovic, M., & Gunawardena, D. (2008). Social tags: meaning and suggestions. *Proceedings of the 17th ACM International Conference on Information and Knowledge Management October 26-30, 2008, Napa Valley, CA*. 223–232. Retrieved from http://dl.acm.org/citation.cfm?id=1458082

Svenonius, E. (2000). *The intellectual foundation of information organization.* Cambridge, MA: MIT Press.

Taylor, A.G. (2007). An introduction to functional requirements for bibliographic records (FRBR). In A.G. Taylor (Ed.), *Understanding FRBR: What it is and how it will affect our retrieval tools*, pp. 1–19. Westport: Libraries Unlimited.

Tillett, B. (2005). FRBR and cataloging for the future. *Cataloging & Classification Quarterly*, *39*(3-4), 197–205.

Trant, J. (2006). Exploring the potential for social tagging and folksonomy in art museums: Proof of concept. *New Review of Hypermedia and Multimedia, 12(1),* 83–105.

Vander Wal, T. (2007, February 2). Folksonomy [Web log post]. Retrieved from http://vanderwal.net/folksonomy.html

Vellucci, S. L. (1990). Uniform titles as linking devices. *Cataloging & Classification Quarterly*, *12(1),* 35-62.

_____ (1998). Bibliographic relationships and the future of music catalogues. *Fontes Artis Musicae*, *45(3/4),* 213-226.

_____ (2001). Music metadata and authority control in an international context. *Notes*, *57(3),* 541-554.

_____ (2007). FRBR and music. In A.G. Taylor (Ed.), *Understanding FRBR: What it is and how it will affect our retrieval tools*. (pp. 131-152). Westport, CT: Libraries Unlimited.

Wetterstrom, M. (2008). The complementarity of tags and LCSH: A tagging experiment and investigation into added value in a New Zealand library context. *The New Zealand Library & Information Management Journal, 50(4)*, 296–310.

Wimmer, R. D., & Dominick, J. R. (1994). *Mass media research: An introduction, (4th ed.)*. Belmont, CA: Wadsworth Publishing Company.

Zavalina, O. (2007). Collection-level user searches in federated digital resource environment. *Proceedings of the American Society for Information Science and Technology, 44(1),* 1–16.

Zavalina, O. (2012). Subject access: Conceptual models, functional requirements, and empirical data. *Journal of Library Metadata, 12(2/3),* 140–163.

Zeng, M., & Žumer, M. (2010). Introducing FRSAD and mapping it with SKOS and other models. *International Cataloging & Bibliographic Control, 39(3),* 53–56.

Zhang, Y., & Li, Y. (2008). A user-centered functional metadata evaluation of moving image collections. *Journal of the American Society for Information Science and Technology, 59(8),* 1331–1346.

Zhang, Y., & Salaba, A. (2012). What do users tell us about FRBR-based catalogs? *Cataloging & Classification Quarterly, 50(5–7),* 705–723.

_____. (2009). What is next for functional requirements for bibliographic records? A Delphi study. *The Library Quarterly, 79(2),* 233–255.

Žumer, M., & O'Neill, E. T. (2012). Modeling aggregates in FRBR. *Cataloging & Classification Quarterly*, *50*(5-7), 456-472.

Žumer, M., & Riesthuis, G. (2002). Consequences of implementing FRBR: Are we ready to open Pandora's box? *Knowledge Organization*, *29*(2), 78-86.

Appendices

Appendix A: Song Sample

Billboard Issue Date	Song	Artist
8/2/86	Glory Of Love (Theme From "The Karate Kid Part II")	Peter Cetera
12/25/71	Brand New Key	Melanie
6/10/72	The Candy Man	Sammy Davis, Jr. with The Mike Curb Congregation
8/18/79	Good Times	Chic
7/7/73	Will It Go Round In Circles	Billy Preston
10/21/72	My Ding-A-Ling	Chuck Berry
4/21/73	Tie A Yellow Ribbon Round The Ole Oak Tree	Dawn Featuring Tony Orlando
10/15/77	You Light Up My Life	Debby Boone
2/19/66	Lightnin' Strikes	Lou Christie
9/10/66	You Can't Hurry Love	The Supremes
12/3/66	Winchester Cathedral	The New Vaudeville Band
9/26/59	Sleep Walk	Santo & Johnny
10/28/78	Hot Child In The City	Nick Gilder
1/21/84	Owner Of A Lonely Heart	Yes
7/13/85	A View To A Kill	Duran Duran
5/28/83	Flashdance...What A Feeling	Irene Cara
6/3/89	Rock On (From "Dream A Little Dream")	Michael Damian
11/15/58	It's Only Make Believe	Conway Twitty
5/3/86	Addicted To Love	Robert Palmer
1/12/74	The Joker	The Steve Miller Band
12/20/86	Walk Like An Egyptian	Bangles
2/8/75	Fire	Ohio Players
6/27/87	I Wanna Dance With Somebody (Who Loves Me)	Whitney Houston
2/7/76	50 Ways To Leave Your Lover	Paul Simon
1/19/80	Rock With You	Michael Jackson

Billboard Issue Date	Song	Artist
10/22/88	Groovy Kind Of Love	Phil Collins
4/7/90	Love Will Lead You Back	Taylor Dayne
3/4/89	Lost In Your Eyes	Debbie Gibson
5/1/76	Let Your Love Flow	Bellamy Brothers
12/11/04	Drop It Like It's Hot	Snoop Dogg Featuring Pharrell
7/14/62	Roses Are Red (My Love)	Bobby Vinton
5/18/96	Tha Crossroads	Bone Thugs-N-Harmony
11/21/87	Mony Mony	Billy Idol
11/23/91	When A Man Loves A Woman	Michael Bolton
12/23/61	The Lion Sleeps Tonight	The Tokens
12/24/77	How Deep Is Your Love	Bee Gees
9/10/11	Moves Like Jagger	Maroon 5 Featuring Christina Aguilera
11/13/10	We R Who We R	Ke$ha
1/8/77	You Don't Have To Be A Star (To Be In My Show)	Marilyn McCoo & Billy Davis, Jr.
9/15/07	Crank That (Soulja Boy)	Soulja Boy
6/29/74	Sundown	Gordon Lightfoot
10/9/76	A Fifth Of Beethoven	Walter Murphy & The Big Apple Band
7/2/05	Inside Your Heaven	Carrie Underwood
1/5/80	Please Don't Go	KC And The Sunshine Band
5/13/78	If I Can't Have You	Yvonne Elliman
8/1/64	A Hard Day's Night	The Beatles
1/21/06	Grillz	Nelly Featuring Paul Wall, Ali & Gipp
7/24/99	Wild Wild West	Will Smith Featuring Dru Hill
4/3/76	Disco Lady	Johnnie Taylor
4/14/90	I'll Be Your Everything	Tommy Page

Appendix B: Content Analysis Dictionary for Tag Coding

Entity	Attribute	FRBR Definition	Definition	Example
Work	Title of the Work	"word, phrase, or group of characters naming the work"	Song title	someone like you
Work	Form of the Work	"class to which the work belongs (e.g., novel, play, poem, essay, biography, symphony, concerto, sonata, map, drawing, painting, photograph, etc.)"	Musical form (usually applicable in classical music)	sonata
Work	Date of the Work	"date (normally the year) the work was originally created "	Year or more specific date song was written/composed	1993
Work	Other Distinguishing Characteristic	"any characteristic that serves to differentiate the work from another work with the same title"	Song writer/composer name combined with song title	bruno mars locked out of heaven
Work	Intended Audience	"class of user for which the work is intended, as defined by age group (e.g., children,	Group of people expected to listen to the song	club

Entity	Attribute	FRBR Definition	Definition	Example
		young adults, adults, etc.), educational level (e.g., primary, secondary, etc.), or other categorizatio n"		
Work	Medium of Performance (Musical Work)	"instrumental , vocal, and/or other medium of performance for which a musical work was originally intended (e.g., piano, violin, orchestra, men's voices, etc.)"	Voice type(s) or instrumentation originally intended to perform song	guitar; male vocals
Work	Numeric Designation (Musical Work)	"serial number, opus number, or thematic index number assigned to a musical work by the composer, publisher, or a musicologist"	Number assigned to work (by composer or musicologist) to distinguish it from other works (more typical with classical music)	opus 10
Work	Key (Musical Work)	"set of pitch relationships that establishes a single pitch class as a tonal centre	Key in which the song was composed	a minor

Entity	Attribute	FRBR Definition	Definition	Example
		(e.g., D major)"		
Expression	Title of the Expression	"word, phrase, or group of characters naming the expression"	Song title	someone like you
Expression	Form of Expression	"the means by which the work is realized (e.g., through alpha-numeric notation, musical notation, spoken word, musical sound, cartographic image, photographic image, sculpture, dance, mime, etc.)"	How song is realized (music notation, lead sheet, musical sound)	sheet music
Expression	Date of Expression	"the date the expression was created (e.g., the date the particular text of a work was written or revised, the date a song was performed, etc.). The date may be a single date or a range of dates. In the	Date associated with song's recording or release	1978

Entity	Attribute	FRBR Definition	Definition	Example
		absence of an ascertainable date of expression, the date of the expression may be associated with the date of its publication or release."		
Expression	Language of Expression	"the language in which the work is expressed"	Language in which song is sung/recorded/printed	english
Expression	Other Distinguishing Characteristic	"any characteristic of the expression that serves to differentiate the expression from another expression of the same work"	1. Descriptor or distinctive element used to delineate between multiple versions (i.e., expressions) of a song; 2. performer name/song title combination	cover vs original; glory of love peter cetera
Expression	Extent of the Expression	"a quantification of the intellectual content of the expression… For works expressed as sound and/or motion the extent may be a measure of duration (e.g., playing time)."	a song's playing time	over 5 minutes long

Entity	Attribute	FRBR Definition	Definition	Example
Expression	Summarization of Content	"an abstract, summary, synopsis, etc., or a list of chapter headings, songs, parts, etc. included in the expression"	summary explanation of song; list of songs on recording; description of smaller parts of song	prominent bass riffs
Expression	Context for the Expression	"the historical, social, intellectual, artistic, or other context within which the expression was realized (e.g., the Art Deco period, etc.)"	Context/circumstances under which song can be understood	mod; new romantic
Expression	Critical Response to the Expression	"the reception given to the expression by reviewers, critics, etc., as encapsulated in an annotation"	awards received, record chart rankings, text from critics' reviews	billboard number ones
Expression	Medium of Performance (Musical Notation or Recorded Sound)	"instrumental and/or vocal medium of performance represented in the expression of a musical work (e.g., two pianos, soprano and alto, etc.)"	instrumentation or voice type used in song recording	acoustic rhythm guitars

Entity	Attribute	FRBR Definition	Definition	Example
Manifestation	Title of the Manifestation	"the word, phrase, or group of characters naming the manifestation. There may be one or more titles associated with a manifestation"	1. Song title (if song released as a single, or if song title and recording title are the same); 2. Recording/CD/LP title if song title is not the recording title; 3. television show/movie title if song is part of soundtrack	animal; 1; gilmore girls
Manifestation	Statement of Responsibility	statement appearing in the manifestation (normally in conjunction with the title) that names one or more individuals or groups responsible for the creation or realization of the intellectual or artistic content embodied in the manifestation.	Person or band performing song on recording; sound engineer for recording; producer for recording	billy idol; dr luke
Manifestation	Place of Publication/Distribution	"the city, town, or other locality associated in the manifestation with the name of the	city where recording was produced/pressed	london

Entity	Attribute	FRBR Definition	Definition	Example
		publisher/distributor"		
Manifestation	Publisher/Distributor	"the individual, group, or organization named in the manifestation as being responsible for the publication, distribution, issuing, or release of the manifestation "	record label name	sony
Manifestation	Date of Publication/Distribution	"the date (normally a year) of public release of the manifestation "	date recording was released	1964
Manifestation	Fabricator/Manufacturer	"the individual, group, or organization named in the manifestation as being responsible for the fabrication or manufacture of the manifestation "	record label name	sony
Manifestation	Form of Carrier	"the specific class of material to which the physical carrier of the	Media type for physical recording; file type for electronic recording	lp; mp3

Entity	Attribute	FRBR Definition	Definition	Example
		manifestation belongs (e.g., sound cassette, videodisc, microfilm cartridge, transparency, etc.)"		
Manifestation	Capture mode	"the means used to record notation, sound, or images in the production of a manifestation (e.g., analogue, acoustic, electric, digital, optical etc.)"	Technique used to record song	acoustic; live
Manifestation	Manifestation Identifier	"a number or code uniquely associated with the manifestation that serves to differentiate that manifestation from any other manifestation."	Record catalog number assigned to recording; appears on LP or CD spine	troy 779
Manifestation	Terms of Availability	"Terms of availability are the terms indicated in the manifestation	how a song is made available	rock band dlc

Entity	Attribute	FRBR Definition	Definition	Example
		under which the supplier (i.e., the source for acquisition/access authorization) will normally make the manifestation available, or the price for which the manifestation sells."		
Manifestation	Access Restrictions on the Manifestation	"restrictions on access to and use of a manifestation "	description of any restrictions placed on song's recording	no longer streamable
Manifestation	Playing Speed (Sound Recording)	"the speed at which the carrier must be operated to produce the sound intended (e.g., 33 1/3 rpm, 19 cm/s, etc.)"	LP record playing speed	33 1/3
Manifestation	Tape Configuration (Sound Recording)	"the number of tracks on a sound tape (e.g., eight track, twelve track)"	number of tracks used to make tape	8 track
Manifestation	Kind of Sound (Sound Recording)	"reflects the number of sound channels used to make the recording (monaural, stereophonic,	number of channels used for sound recording	mono; stereo

Entity	Attribute	FRBR Definition	Definition	Example
		quadraphonic, etc.)"		
Manifestation	Special Reproduction Characteristic (Sound Recording)	"the equalization system, noise reduction system, etc. used in making the recording (e.g., NAB, DBX, Dolby, etc.)"	audio compression technology used for the recording	dolby
Group 2	Name of Person	"the word, character, or group of words and/or characters by which the person is known"		george harrison; dr luke
Group 2	Dates of Person	"may include the precise or approximate date of the person's birth and/or death, or dates indicating the period in which the person was known to be active in a given field of endeavour"		1943-2001
Group 2	Title of Person	"a word or phrase indicative of rank, office, nobility, honour, etc. (e.g., Major, Premier,		Sir

Entity	Attribute	FRBR Definition	Definition	Example
		Duke, etc.), or a term of address (Sir, Mrs., etc.) associated with the person."		
Group 2	Other Designation Associated with the Person	"a numeral, word, or abbreviation indicating succession within a family or dynasty (e.g., III, Jr., etc.), or an epithet or other word or phrase associated with the person (e.g., the Brave, Professional Engineer, etc.)"		sound engineer; jr
Group 2	Name of the Corporate Body	the word, phrase, character, or group of words and/or characters by which the body is known		beatles; supremes
Group 2	Number Associated with the Corporate Body	"the numerical designation sequencing a meeting, conference, exhibition, fair, etc. that constitutes		13th

Entity	Attribute	FRBR Definition	Definition	Example
		one of a series of related meetings, conferences, exhibitions, fairs, etc., or any other numerical designation associated with a corporate body."		
Group 2	Place Associated with the Corporate Body	"the city, town, or other designation of location in which a meeting, conference, exhibition, fair, etc. was held, or the location with which the corporate body is otherwise associated"		newport
Group 2	Date Associated with the Corporate Body	"the date or range of dates on which a meeting, conference, exhibition, fair, etc. was held, or a date with which the corporate body is otherwise		2008

Entity	Attribute	FRBR Definition	Definition	Example
		associated (e.g., the date of its incorporation)."		
Group 2	Other Designation Associated with the Corporate Body	"a word, phrase, or abbreviation indicating incorporation or legal status of the body (e.g., Inc., Ltd., etc.), or any term serving to differentiate the body from other corporate bodies, persons, etc. (e.g., firm, musical group, etc.)."		
Group 3	Term for the Concept	"the word, phrase, or group of characters used to name or designate the concept (e.g., economics, existentialism , radioactivity, etc.). A concept may be designated by more than one term, or by more than	a concept that serves as a subject; musical genre	british invasion; dance music

Entity	Attribute	FRBR Definition	Definition	Example
		one form of the term"		
Group 3	Term for the Object	"word, phrase, or group of characters used to name or designate the object (e.g., a building, a ship, etc.). An object may be designated by more than one term, or by more than one form of the term."	an object that serves as a subject	roller skates; animals
Group 3	Term for the Event	"the word, phrase, or group of characters used to name or designate the event (e.g., Battle of Hastings, Tour de France, etc.). An event may be designated by more than one term, or by more than one form of the term" Also a time period.	an event that serves as a subject	2008 cma fest; 1972

Entity	Attribute	FRBR Definition	Definition	Example
Group 3	Time period	a span of dates; a word or phrase describing a period of time (e.g., renaissance)	a date range or term describing a date range/time period; not a date associated with a Date of the Work, Date of Expression, Date of the Manifestation, or date associated with a Critical Response to the Expression	60s; seventies
Group 3	Term for the Place	"the word, phrase, or group of characters used to name or designate the place (e.g., London, St. Lawrence River, etc.)"	a place that serves as a subject	cleveland; uk
Non-FRBR	Artist tag		provide descriptive data about singers, songwriters, or performers	british; deceased
Non-FRBR	Personal tag		describe personal, abstract ideas about a song	bianca smiles; workout
Non-FRBR	Foreign language tag		non-English tags	gute laune; puedo
Non-FRBR	Subjective tag		express qualitative judgments about songs, artists, or recordings	10 of 10 stars; catchy
Non-FRBR	Mood tag		describe or evoke a feeling or state of mind	fun; uplifting

Appendix C: Tag Mapping Counts by Song Title

Song	General Tags	Work Tags	Place Subject Tag	Expression Tags	Manifestation Tags	Concept Subject Tag	Concept Subject Tag 2	Object Subject Tag	Event Subject Tag	Group 2 Entity	Tag Mappings for song	FRBR tag mappings/song
50 Ways To Leave Your Lover	37	4	1	5	4	5	8		2	2	68	31
A Fifth Of Beethoven	41	1		4	4	8	5		2	2	67	26
A Hard Day's Night	22	2	1	4	8	9	13	1	6	5	71	49
A View To A Kill	24	2	1	7	5	17	8		6	1	71	47
Addicted To Love	31	2		11	3	9	6		6	1	69	38
Brand New Key	31	2	2	3	2	5	9	1	6	2	63	32
Crank That (Soulja Boy)	37	4	5	10	2	4	7		6	1	76	39
Disco Lady	28	1	2	8	4	8	13		3	3	70	42
Drop It Like It's Hot	31	6	1	7	8	5	11	1	1	6	77	46
Fire	29	4		9	5	6	12	1	3	1	70	41
Flashdance...What A Feeling	35	4	1	13	2	6	3		2	1	67	32
Glory Of Love (Theme From "The Karate Kid Part II")	17	2	2	16	3	15	7		3	2	67	50
Good Times	27	1		6	4	5	17		7	1	68	41
Grillz	33	2		10	3	11	7		2	2	70	37
Groovy Kind Of Love	29	4	2	11	2	11	6		4	2	71	42
Hot Child In The City	38	3	1	9	1	4	4		8	2	70	32
How Deep Is Your Love	28	1		3	4	14	9		5	2	66	38

Song	General Tags	Work Tags	Place Subject Tag	Expression Tags	Manifestation Tags	Concept Subject Tag	Concept Subject Tag 2	Object Subject Tag	Event Subject Tag	Group 2 Entity	Tag Mappings for song	FRBR tag mappings/song
I Wanna Dance With Somebody (Who Loves Me)	28	1	6	8	3	6	10		5	2	69	41
I'll Be Your Everything	1			3		2	1		2		9	8
If I Can't Have You	22	2		6	7	13	14		8	3	75	53
Inside Your Heaven	31	1	3	10	4	7	8		3	2	69	38
It's Only Make Believe	28	5		10	7	5	11		5	3	74	46
Let Your Love Flow	34			6	4	6	9		7	3	69	35
Lightnin' Strikes	36	2	1	7	2	7	6	1	6	1	69	33
Lost In Your Eyes	32	2	2	8	3	7	6		6	1	67	35
Love Will Lead You Back	28	5		5	2	11	10		7	3	71	43
Mony Mony	26	2	2	13	7	6	10		5	1	72	46
Moves Like Jagger	19	1	1	5	3	6	5		2	2	44	25
My Ding-A-Ling	30	1	2	6		7	11	1	8		66	36
Owner Of A Lonely Heart	14	3	1	10	7	14	15		8	3	75	61
Rock On (From "Dream A Little Dream")	13		1	6		6	4		3		33	20
Rock With You	35	1		4	3	3	13		6	2	67	32
Roses Are Red (My Love)	33	3		8	7	11	5	1	6	1	75	42
Sleep Walk	42			6	1	3	7		5		64	22
Sundown	34	1		7	3	3	14		4	1	67	33
Tha Crossroads	32	3	3	7	7	5	7	1	1	5	71	39
The Joker	30	3		10	7	9	7		3	2	71	41

Song	General Tags	Work Tags	Place Subject Tag	Expression Tags	Manifestation Tags	Concept Subject Tag	Concept Subject Tag 2	Object Subject Tag	Event Subject Tag	Group 2 Entity	Tag Mappings for song	FRBR tag mappings/song
The Lion Sleeps Tonight	30	2		8	5	10	8	1	4	4	72	42
Tie A Yellow Ribbon Round The Ole Oak Tree	20	1	2	6	3	4	5		8	1	50	30
Walk Like An Egyptian	34	1		7	5	12	5		6	2	72	38
We R Who We R	34	7	1	5	8	3	9		3	4	74	40
When A Man Loves A Woman	33	2		6	3	8	8		3	2	65	32
Wild Wild West	34	3	1	7	4	11	9		1	2	72	38
Will It Go Round In Circles	40	3		7	2	5	6		4	1	68	28
Winchester Cathedral	36			7	3	9	4		4	2	65	29
You Can't Hurry Love	31	2		3	5	9	14		3	3	70	39
You Don't Have To Be A Star (To Be In My Show)	8			5	2	3	3		3	2	26	18
You Light Up My Life	28	1		11	6	13	3		4	3	69	41
Totals	**1394**	**103**	**45**	**353**	**187**	**366**	**392**	**9**	**215**	**97**	**3161**	**1767**