Introduction

As the amount of musical information has exploded in recent years at an insurmountable rate with the advent of the World Wide Web, resulting in the production of thousands of new musical items each year, it is not difficult to fathom the inherent difficulties in making such voluminous amounts of information easily findable and retrievable. Additionally, the mass production of music, whether as single tracks or complete musical albums, has been further spurred by major commercial entities such as iTunes and Amazon, which have done little in the way of providing more reliable means of identifying, organizing, and retrieving music. As one can imagine, these circumstances are less than ideal for addressing the needs of various user groups who are searching for music, which are often severely divided on the basis of musical knowledge and expertise (Burgoyne, Fujinaga, & Downie, 2015). Thus, with such an abundance of musical information residing both in print and digital formats, and few mature music information retrieval systems in existence to satisfy the diverse needs of searchers, it is clear that more inquiry must be directed towards the domain of music information retrieval (MIR).

Music Information Retrieval (MIR) can be defined as any means by which users search for music, either by submitting a query by example (e.g., humming or keyboard input) or searching by more traditional means such as using bibliographic information. MIR also involves the multiple ways by which music, or information about music, is retrieved. Thus, MIR consists of the entirety of the research, methodologies, search
techniques, user studies, and music information retrieval systems that allow a user to identify, locate and retrieve music, which includes recorded audio music, sheet music and scores, and information about music (i.e., metadata) (Lemström & Tzanetakis, 2009). In order to elucidate the significant issues plaguing MIR, this paper will present three overarching issues currently confronting MIR: (1) issues in fully representing all identifying aspects of music for discovery, which is severely convoluted by the multifaceted nature of music; (2) the ineffectiveness of existing content-based music information retrieval systems; and (3) the poor understanding of user needs and information seeking behaviors. Finally, discourse will be provided based on a synthesis of the literature, in order to unravel the intricacies of the above issues, and to explain the implications for libraries and information scientists.

**Discussion of MIR Issues:**

**Complete Musical Representation**

The representational completeness of a piece of music, or corpus of musical items, can be measured by the amount of descriptive facets—and their sub-facets—which are included in metadata schemas used to both describe and facilitate successful retrieval of sought music (McLane, 1996). Yet, several prevalent issues exist, which prevent full realization of representing music completely, so that numerous avenues of discovery are afforded to users. First and foremost however, is the ability to surpass the complex and multifaceted nature of music, in order to develop an appropriate metadata schema that will allow for all identifying aspects of a particular piece of music to be presented to the user (Belford, Snyder, & Randall, 2014). However, it is exceedingly difficult to develop a general schema that can effectively include all aspects of music, and still remain standardized throughout all genres and styles of music. Implications of this phenomenon
are made implicit by Belford, Snyder, & Randall (2014) in their examination of the Functional Requirements for Bibliographic Records (FRBR) and Resource Description and Access (RDA) as they apply to music (Belford et al., 2014). The authors state that music effectively evades complete representation and description on a number of fronts, primarily due to a lack of appropriate authorized access points (AAP), which are necessary to completely describe all aspects of a piece of music in any format. For instance, Burgoyne et al., (2015) identify several aspects of music that currently remain unaddressed by FRBR and RDA altogether, such aspects include publication information, pitch, timbre (i.e., tone color), rhythm, temporal considerations (i.e., duration), and harmonic summaries (Burgoyne et al., 2015). Any combination of these aspects of music would allow a user to successfully identify and locate a specific piece of music, even spanning across genre-based divisions. However, until these aspects are incorporated into RDA, issues will remain present. Taking each of these aspects of music into consideration, it is clear that existing methods are primarily tailored towards textual materials more so than music (Belford et al., 2014). As will be discussed later in this paper, coverage of these aspects may also require further developments on the part of both music retrieval systems, as well as considerations of user context in searches, which has been proven by Lee's (2010) research to constitute a major facet of user searches (Lee, 2010).

The deficiencies of existing metadata schemas and cataloging standards such as RDA, become readily apparent when the issue of identifying differing manifestations of an identical work are sought. As historical practice will portray, it is common for multiple works, and subsequent variant expressions and manifestations thereof, to frequently be
labeled with identical titles, for example. Belford et al., (2014) identified "Telemann's Overture in D Major for Strings and Continuo" as a prime example, in that six different musical works bearing this same title appear in a simple search of a library's OPAC (Belford et al., 2014). Such instances of this phenomenon are quite common in music, especially in the realm of western art music (vernacularly known as "classic music"), therefore, it is easy to observe how users could become perplexed at the lack of unique identifying information, and being unable to identify sought information accurately. Thus, alternative AAP, such as those provided in a print index of musical incipits by Barlow & Morgenstern (1948), which includes snippets of melodic information (similar to an abstract of an article) for nearly 10,000 musical works, will be necessary to increase the discriminatory power of RDA (Pietras & Robinson, 2012).

The aspect of musical genre has likewise been found to be an extremely vital AAP, as it can provision the additional discriminatory power needed to uniquely identify a plethora of derivative manifestations based on the same work, a common practice in popular music and Jazz (Kishimoto & Snyder, 2016). Yet in spite of the importance of genre in representing music, it has not been fully incorporated as an AAP in the capacity explained here, nor has it been further examined by either RDA or FRBR (Kishimoto & Snyder, 2016). Kishimoto and Snyder (2016) further demonstrate the great importance of genre and user context in user searchers, especially in the online environment, by examining unique aspects of music, which are only present within certain genres and styles of music. (Kishimoto & Snyder, 2016). Popular music and western art music are used to exemplify the varying descriptive characteristics that surround different genres of music. For instance, searchers seeking popular music rarely search for a composer,
instead relying solely on the performing artist and title. Moreover, popular artists do not commonly create (i.e., compose) a work, rather they only provide their own personal artistic interpretation, although their manifestation will be quite different from the original work. Conversely, musical works encompassed by western art music are chiefly identified by composer, as titles are often uniform, and manifestations of the same work are very similar, lacking any extreme variances, as dictated by western performance practice. Thus, the issues presented by the lack of additional and reformulated AAPs, such as genre, are immense (Kishimoto & Snyder, 2016).

As one can imagine, inclusion and incorporation of these aspects of music into existing metadata schemas and standards such as FRBR and RDA will be extremely difficult, as not only will AAP need to be appended, but a complete redefinition of FRBR's three working groups must also be considered. In particular, more attention must be given to group two of FRBR, which entails the responsible body for the existence of a work (Taylor, 2009). Working from a reconceptualization of RDA, Kishimoto and Snyder (2016) have developed several alternative methods for music catalogers in order for multiple creators, or responsible bodies of a work, to be included within a single bibliographic record, allowing a user to differentiate between items with identical uniform titles (Kishimoto & Snyder, 2016). Furthermore, a user may observe upon which work the exact manifestation they are attempting to locate is based upon. For example, the statement of responsibility would be depicted in a bibliographic record as follows, using Walter Murphy's 1952 remix of Beethoven's 5th Symphony: Murphy, Walter, 1952-Fifth of Beethoven→ based on (work) Beethoven, Ludwig van, 1770-1827. Symphonies, no.5, op.67, C minor (Kishimoto & Snyder, 2016).
Content-Based MIR

Content-based MIR faces several notable challenges, largely owing again to the complexities of music itself. However, there are also serious complexities in computing that must also be addressed, as MIR entails more highly convoluted computing procedures than text-based methods. Firstly, unlike textual information, which is expressed in a single format regardless of digital or print manifestations, music conversely can be represented in three profoundly different forms, each requiring its own unique considerations in terms of descriptive metadata, storage and retrieval system design, and an understanding of user needs and behaviors. However, special treatment must be provided for the type of retrieval system, and how it will adapt to the challenges of the selected format. These formats include digital audio, time-stamped events, and musical notation (Byrd & Crawford, 2002). Music in the form of digital audio is of course stored electronically, and is thus contained within a specific file format, such as an MP3 file. Time-stamped events are somewhat more complex, in that they require specialized software to be achieved, however, this format can be summarized as an analysis of wavelengths displayed upon a vector. Finally, there is musical notation, which can exist either digitally or in print, as they are both symbolic in nature (Byrd & Crawford, 2002).

As Lippincott (2001) posits in their discourse of issues in content-based MIR, there are several approaches to developing an MIR system, each with their own drawbacks and advantages; however, at present, each type of system generally focuses on only a single representation of music (Lippincott, 2001). Nonetheless, the ultimate goal of MIR is to eventually construct a retrieval system that could incorporate all major
content-based approaches, as well as careful consideration of how each aligns with user needs and preferences into a single user interface. Although, until such a system is achieved that can indeed incorporate these considerations, users will continue to struggle in grappling with the disparate methods of retrieval, which do not correlate with unequal levels of user expertise in music (i.e., the lay person vs. the professional) (Selfridge-Field, 2000). Three major content-based retrieval systems identified by Lippincott (2001) include (1) those that focus on melodic input (i.e., query by example, in which a user sings or hums into the system), which measure similarities between the input audio and incipits stored within a database to determine a match; (2) those that seek to identify unique musical signatures (by way of calculating weighted sums based on note patterns) of a certain piece of music to locate a match between query and sought music; and (3) automatic indexing methods borrowed from text-based techniques (Lippincott, 2001).

Melodic input retrieval systems operate on an approach of utilizing user-generated queries in the form of singing or humming to produce a geometrical algorithm to determine melodic difference, which is compared to existing musical fragments stored within a database, known as incipits. The system determines matches based upon factors including juxtaposition of notes, pitch difference, and note duration (ÓMaidín, 1998). Issues confronting melodic input-based systems are inherent in regards to inaccuracies that accompany any form of user-generated queries. Frequently, these inaccuracies manifest in the form of users relying on fuzzy metadata, or incomplete or inaccurate queries, where a user is relying on memory alone to supply needed information to locate music (Lee, 2010). This inevitably involves the user query to be translated by the system in order to provide an accurate match, however, computing algorithms have shown that
they are not always capable of correctly assuming what a user intended with their query, as a musical incipit could have innumerable matches (Byrd & Crawford, 2002).

Crawford, Iliopoulos, and Raman (1998) instead borrow techniques from biological and technical sciences to develop unique musical signatures for each piece of music. Assigning weights to patterns of notes, and then calculating their collective sums to identify the unique characteristics of musical structures, have allowed the authors to formulate what they hail as musical "fingerprints". Although this approach is ambitious, it has yet to be tested on heterogeneous collections of music, or a collection of music consisting of varied genres and styles (Crawford, Iliopoulos, & Raman, 1998).

Lastly, Downie (2002) has attempted to apply the text-matching technique of automatic indexing to MIR systems, in order to circumvent the defects of the previous two approaches. By presenting a monophonic melody (i.e., single voice/line) as a collection of intervals split into "n-grams"—which are simply discrete units of melodic information taken anywhere from the melody itself—Downie has sought to provide a panacea to the complexities of varying musical styles and genres (Downie, 2000). These n-grams of various lengths are used for approximate string matching against longer melodies in a database. However, this borrowed text retrieval technique does not translate well to music, as music does not retain the same meaning when distilled down to its comprising parts (i.e., notes) as words do in text. Music effectively looses any greater meaning it once possessed, making any form of a musical dictionary or index impossible (Byrd & Crawford, 2002). As can be observed from Bryd and Crawford's search of a simple musical fragment through content-based retrieval methods, countless "false" hits are retrieved, as the same fragment is found repeatedly in myriad instances.
Each of these systems involves specialized issues, which must be addressed if a system is to allow users to easily locate and retrieve sought information. However, the overarching issue enveloping all types of MIR systems that have currently been discussed, is their inability to function effectively when applied to heterogeneous collections of musical styles and genres. The fundamental component parts of music, such as individual notes and musical phrases, are constructed upon varying conceptual frameworks of both notation and human cognitive perceptions. Examination of each theoretical, underlying style of musical architecture respectively goes well beyond the scope of this paper, but suffice it to say that not all musical genres are based on traditional conventions of western music, nor are their defining characteristics inherent to all cultures. This is especially apparent in serial music of the 20th century, or music originating from Asia, both of which are constructed on antithetical systems of symbolic representation and human perception when compared to music commonly found in the occidental world of music (Byrd & Crawford, 2002). It seems that only a few commercial and web-based retrieval systems, such as Shazam and Midomi, have been successfully developed, which can accurately retrieve queries of varying musical genres. Yet even such systems as these are lacking in coverage of certain genres of music such as western art music and certain styles of folk music, resulting in false retrievals, or no positive matches whatsoever (Colvin, 2009).

Finally, the aspect of polyphony (i.e., multiple voices or lines that occur simultaneously; the antithesis of monophony) must also be addressed in respect to content-based music retrieval systems, because serious inaccuracies often result due to this complex aspect of music (Lemström & Tzanetakis, 2009). As "full-text" audio is as of yet too computationally complex and expensive to be fully realized, and music itself is inherently complex due to its composition of multiple independent lines or voices
occurring simultaneously, retrieval systems are often hard-pressed to match a monophonic query to polyphonic items (Lippincott, 2001). Dovey and Crawford (1999) have proclaimed polyphony as the source of all issues, with which music retrieval systems must struggle, owing to the complexity and interwoven nature of the voices and lines comprising a single piece of music. Perhaps the best way to illustrate this concept is by imagining a textual source, in which multiple lines of text overlap one another throughout, obscuring meaning with an abundance of extraneous noise (Dovey & Crawford, 1999). Dovey and Crawford (1999) exemplify the obstacles presented by polyphony by searching for specific melodic fragments within a Beethoven symphony, only to find that of the 98 instances the sought fragment appears throughout the piece, the system was only able to successfully identify a single occurrence (Dovey & Crawford, 1999). Thus, greater matching algorithms, in conjunction with more advanced systems—able to effectively separate polyphonic voices to perform cross-voice matching—must be instilled in future retrieval systems, if these barriers to retrieval are to be remedied (Lippincott, 2001).

**Poor Understanding of User Needs and Behaviors**

It is a deeply held assumption that MIR research is severely lacking in both empirical research that seeks to understand the real-life needs and searching behaviors of users searching for music, as well as an analysis of real-life queries submitted by users (Lee, 2010). Moreover, this inadequate understanding of user behaviors and preferences has caused profound detractions in users' abilities to successfully interact with existing retrieval systems to identify and locate music (Belford et al., 2014). Thus, implications resulting from the culmination of misguided assumptions on the part of MIR researchers
and professionals, has led to the creation of retrieval systems that are ill-suited in satisfying user needs, and adoption of general metadata schemas that do not take into account specific aspects of music (McNab, Smith, Witten, Henderson, & Cunningham 1996). Retrieval systems typically only employ those metadata that are considered important from the standpoint of domain experts, rather than considering those aspects of music that are most important to users themselves. However, according to what little empirical research that actually has been conducted, there are serious gaps between which metadata users actually prefer and require to carry out searches effectively, and what researchers consider as important. These gaps in understandings will only continue to detract from the effectiveness and efficiency of MIR, as users are increasingly forced to aberrate from more reliable means of finding the information they need (Futrelle & Downie, 2003). To counteract these deficiencies on the part of MIR researchers and the information profession as a whole, more empirical research is strongly recommended, offering a more exhaustive examination into aspects of user needs and behaviors (Lee, 2010).

Instances of poor understandings of the needs and behaviors of users searching for music are made explicit in Bainbridge, Cunningham, and Downie's (2003) empirical research, which involved the analysis of over 502 music queries posted to the Google Answers Website. The authors found that overall, users experienced adverse difficulties in formulating clear descriptions of musical items they were searching for, often relying on fuzzy metadata values and query-by-example methods to search for music (Bainbridge, Cunningham, Downie 2003). Further, users by and large forwent the selection of metadata completely (e.g., artist, title, publication date) due to difficulties in describing
music, instead relying on contextual information, such as information pertaining to other cultural objects that were related somehow to the sought music. For example, analyzed user queries revealed frequent use of descriptions of TV commercials, artists and songs that "sounded like" the sought music, and movies that users had related to sought music, in order to resolve uncertainties of descriptive metadata (Lee, 2010).

Therein lies the predominate issue plaguing MIR at present, as such search techniques involving user context as a searchable parameter are typically not supported by retrieval systems, or even the Web, as can be referenced in earlier passages of this paper concerning issues of complete representation due to inadequate descriptive coverage of music (Lee, 2010). These issues are further complicated due to the absence of any specialized taxonomy of user needs, which should thoroughly define which aspects of music should be represented in retrieval systems to ensure optimal success in locating sought information. Thus, with no clear inclination on the part of MIR researchers as to which needs are most critical, thereby receiving greater attention, continued evaluations of user studies will continue to yield no meaningful results, which could otherwise be utilized to further improve MIR (Ingerwesen, 1992).

Aside from the apparent lack of search techniques that could be derived from more intense real-life user studies, there too exists negligence to the fact that users as a majority seek information about music (i.e., metadata) rather than the music itself (e.g., CD, individual tracks, sheet music, etc.). Also, it is no surprise that the bulk of musical queries can be classified as known-item searches (Lee, 2010). The complex role of musical queries in this issue is evident, because each derivate manifestation of a piece of music is inherently different from the original work, upon which it is based (Belford
et al., 2014). Bearing this in mind, emphasis must be placed on the notion that when users are searching for music, they are actually searching for a very specific version or manifestation of a piece of music—usually one that is special to them—rather than for a general work. This is a deviation from characteristics of text retrieval; although textual objects may be duplicated into several manifestations bearing different book jackets and re-prints, they are still embodiments of the *same* work, sharing the equivalent characteristics. However, derivative manifestations of a musical work can vary immensely from the original, especially as they are realized in other genres and styles, which may deviate altogether from the deeply held musical traditions of western music (Kishimoto & Snyder, 2016). One need only reference the previous example of Walter Murphy's 1952 remix of Beethoven's 5th Symphony, which is in the genre of funk/disco. Whereas the original work is a classical symphony by Beethoven from the 19th century, the differences of these two works are obviously striking.

**Implications for Libraries and the Field of Information Science**

Implications of the discussed issues in MIR are abundant in the realm of libraries and information science, especially in terms of pursuing further empirical research, which would allow closer examination of user needs and behaviors as suggested by Lee (2010). By taking into consideration the results of past explorative studies, and the trends which they have uncovered, MIR researchers and practitioners in the field alike could begin to develop retrieval systems, search techniques, and finding aides to incorporate improved means of searching. This may perhaps be best realized by undertaking a synthesis of all common metadata schemas used to describe music, in order to identify those which are commonly chosen by users; these could be further subdivided by user groups, most
prominently amongst the lay person and professional musician (Lee, 2010). Moreover, these commonly used and consolidated metadata could then be cross-referenced with those search terms gleaned from user case studies as being held most conducive to user search success, thereby facilitating the creation of an all-inclusive metadata schema, which could be flexible enough to describe all major aspects of music (Kelly, 2015).

A more ambitious solution is posited by Kelly (2015), in that the discoverability and retrieval of music could be significantly enhanced through incorporation of linked open data (LOD) into the library environment. By consolidating resources within the library with those located externally (e.g., iTunes, MIREX, the Library of Congress) a rich set of discovery tools could be developed, effectively circumventing the issues currently plaguing MIR, (Kelly, 2015). Additionally, retrieval systems could be redesigned, incorporating multiple avenues of query input, which are conducive to both the highly trained professional, as well as the lay person. These may include improved query-by-example input techniques, which mutually incorporate the aspects of pitch, rhythm, and melodic contour, rather than relying on each exclusively. Additional search facets specific to the domain of music, as well as novel means of identifying multiple extant manifestations of identical works would also allow users to select search parameters, such as genre (Kelly, 2010). It may also be of the utmost interest of researchers in the field of MIR to borrow methodologies extraneous to the domain of music, which are commonplace in other domains, such as the sciences and other areas of the humanities. Expanding the reach of MIR research to include developments in the realm of the semantic web—and the major accomplishments therein that have been achieved in recent years—could assimilate the greatly improved techniques of assigning
permanent URIs to each and every distinct manifestation of a musical work, and subsequently linking these URIs through use of HTTP, allowing greater discoverability and discriminatory powers on the part of the user (Rodriguez, 2009).

Finally, if special consideration is directed at providing methods of including the aspect of user context, and the overwhelming trend of users to search for information about music, rather than music itself, significant headway could be achieved in expanding existing search retrieval systems and processes. These implications are just as imperative to reference music librarians as they are to MIR researchers and retrieval system developers, as previous user queries, contextual information regarding other cultural objects (e.g., TV shows, movies, similar artists and music), and user demographics could have the potential to improve music retrieval exponentially (Gracy, Zeng, & Skirvin, 2013).

**Conclusion**

The issues posited within this paper are by no means simplistic in nature, nor are they effortlessly combated with sporadic and shallow methodologies. Thus, a holistic approach, engaging all major aspects of music, and those ubiquitously exhibited user behaviors, is recommended as a panacea to these issues. For if these issues remain unaddressed, they could potentially transform into insurmountable barriers to access and user frustration, further eroding the effectiveness of MIR altogether. Nonetheless, it is essential that greater examinations are undertaken in an effort to unravel the complex aspects of music, while simultaneously addressing issues of computing inefficiencies, if MIR issues are to be stemmed and solutions appended to existing methods. However, ongoing research proves promising in shedding light on these issues, and perhaps with the combined efforts of linked open data, in conjunction with greater attention to user
needs, MIR can begin to establish a major presence at the forefront of information retrieval research, consolidating the shared issues and uncertainties of all sub-disciplines of this ever expanding field.
References


