Classical Music Recording Ontology Used in a Library Catalog†

Dan Wu* and Jinsong Shi**

*Wuhan University, School of Information Management, Wuhan, Hubei, China 430072, <woodan@whu.edu.cn>
**Nankai University, Library, Tianjin, China 300071, <sksnku@nankai.edu.cn>

Dan Wu is Professor at the School of Information Management of Wuhan University. She teaches classes on information description, information retrieval, and the design of information systems. Her research is focused on information organization, information retrieval, and user seeking behavior. She holds a PhD in information science from Peking University, China. She was a visiting scholar at School of Information Sciences, University of Pittsburgh.

Jinsong Shi is a librarian at the Nankai University Library. He manages audio and video resources at the multimedia sector of the library. His research focuses on the organization of audio and video resources and multimedia services in libraries. He holds bachelor's and master's degrees from the School of Information Management, Wuhan University, China.

Wu, Dan and Jinsong Shi. 2016. "Classical Music Recording Ontology Used in a Library Catalog." *Knowledge Organization* 43: 416-430. 25 references.

Abstract: In order to improve the organization of classical music information resources, we constructed a classical music recording ontology, on top of which we then designed an online classical music catalog. Our construction of the classical music recording ontology consisted of three steps: identifying the purpose, analyzing the ontology, and encoding the ontology. We identified the main classes and properties of the domain by investigating classical music recording resources and users' information needs. We implemented the ontology in the Web Ontology Language (OWL) using five steps: transforming the properties, encoding the transformed properties, defining ranges of the properties, constructing individuals, and standardizing the ontology. In constructing the online catalog, we first designed the structure and functions of the catalog based on investigations into users' information needs and information-seeking behaviors. Then we extracted classes and properties of the ontology using the Apache Jena application programming interface (API), and constructed a catalog in the Java environment. The catalog provides a hierarchical main page (built using the Functional Requirements for Bibliographic Records (FRBR) model), a classical music information network and integrated information service; this combination of features greatly eases the task of finding classical music recordings and more information about classical music.

Received: 14 December 2015; Revised 29 February 2016; Accepted 14 March 2016

Keywords: music, musical work, classical music recording, music ontology, music recording ontology, user music information needs

† This study is supported by the National Social Science Fund of China "Research on the open sharing of government data for the national big data strategy" (NO.15ZDC025).

1.0 Introduction

According to the IFPI Digital Music Report 2013, consuming digital music has become mainstream, and musical content is no longer constrained by physical carriers; it often is published in different forms, such as a collection of several musical works, a single musical work or some parts of a complete musical work. Furthermore, classical music can take a variety of genres from a large symphony lasting an hour to a two-minute-long song, which further complicates the forms of classical music resources. Due to the long history of classical music, there may be many recordings of a classical work and a recording version may be published in different albums. All these characteristics make it difficult for classical music listeners to select the musical content that can satisfy their needs. We think that the description of classical music recording resources should go deeper into representing contents, not just at the surface level, and at the same time, the organization of the resources should focus on classical musical works and recording versions to classify or group albums in various forms.

However, according to our observations, main music websites such as Pandora, iTunes, Spotify, and Last.fm do not provide music resources organized in a way that is suitable for classical music listeners. Whether by keyword search or by browsing, users can obtain details of albums or artists, and track information within the records. But the records and track information are scattered and incomplete. For example, albums of the same recording versions, albums that contain the same works, and different recording versions of those same works are often not integrated; at the same time, the information about works and recording versions is sometimes incomplete or inaccurate. Our investigation on classical music listeners' information needs has shown that listeners have various needs, which include seeking abundant and accurate information of classical musical works, composers, performers, recording versions, and albums and the ways of finding an interesting work, a satisfying recording version of a certain work, and an album worth buying. We think that to satisfy users' needs, there should be better music content description and organization, which neither the music websites nor library OPACs at present satisfy.

Consequently, we believe that classical music recording resources need new organization schemes. Among all kinds of knowledge organization systems, we chose an ontology because it is a conceptual and structural description of the objective world or a part of the world (Dong 2008). Using an ontology for resource organization can naturally reflect characteristics of classical music recording resources.

2.0 Literature review

2.1 Research into constructing music ontologies

From research into constructing music ontologies (MOs), we have concluded that there are several approaches to constructing different types. The main approaches are:

- Constructing an ontology of music recording resources to organize music albums;
- Constructing an ontology of music elements based on knowledge of the melodies, rhythms and other elements that constitute a piece of music;
- Constructing an ontology of music listeners based on studies of their emotions, preferences and other individual characteristics when listening to music;
- Constructing an ontology of artists based on structured data; and,
- Constructing an ontology of music subjects in which music recording resources are organized by subject.

The different types of music ontologies and their relationships that we have summarized are shown in Figure 1.

An ontology of music recording resources focuses on the essence and production flow of the resources. In each stage of the production flow, there are different products

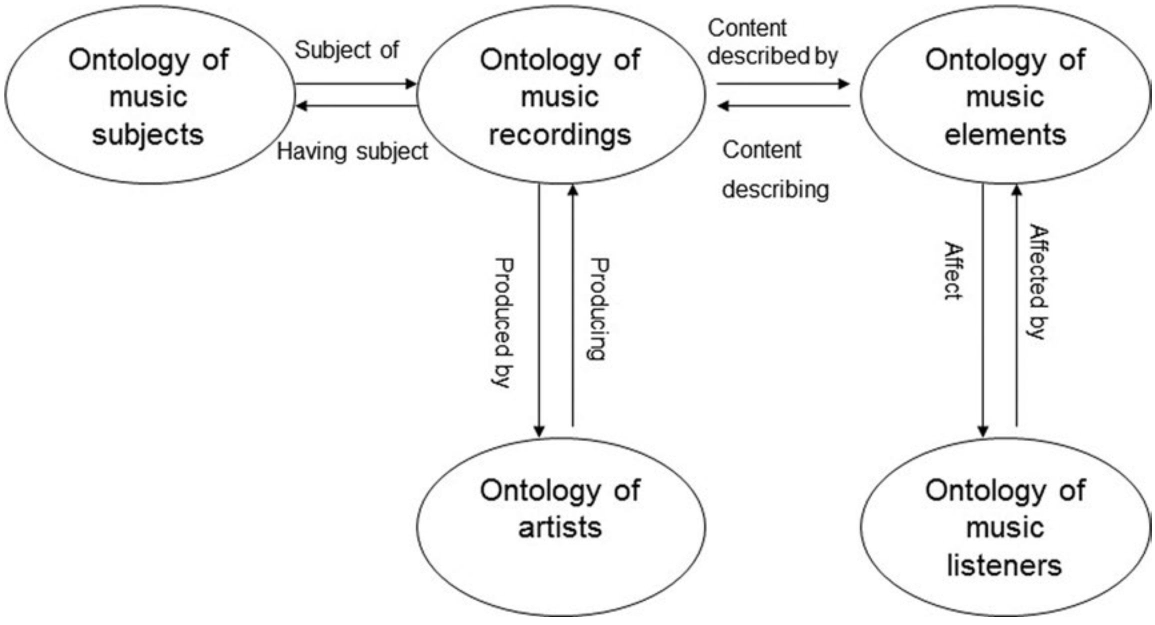

Figure 1. Types of music ontologies and their relationships.

that are closely related to each other. The Music Ontology (Raimond et al. 2007), which was created in 2006, provides a vocabulary for publishing music-related data on the web. It divides the production flow of music recording resources into three main events: composing, performing, and publishing. The products of the events are respectively musical works, recorded versions, and albums. Both the events and the products are classes of the MO.

An ontology of artists focuses on the people who participate in producing music recording resources, such as composers, pianists, recording engineers, and recording companies. They are the producers of the products listed above. Linked data, a kind of structured data, are often used to construct the ontology. For example, Passant (2010) extracted linked data about artists in DBpedia, which is a linked open data set, to construct an ontology for music recommendations. There are also other linked open data sets that can be used, such as files created using the FOAF ontology (http://xmlns.com/foaf/spec/).

The music subject is what a piece of music is about. Some classical musical works have subjects, such as programmatic pieces and songs. An ontology of music subjects is constructed by investigating musical works and extracting structured knowledge from them and from articles in the music research field. For example, Nisheva-Pavlova and Pavlov (2011) constructed a subject ontology of Bulgarian folk songs based on their investigation into this type of music.

An ontology of music elements describes the contents of music recording resources, including melody, rhythm, and harmony and so on. We can use some musical terms to describe general characteristics of these elements. But such an ontology is neither complete nor precise. Audio metadata standards such as MPEG-7 have been created to describe precisely the contents of music recording resources. The main approach to constructing ontologies of music elements is to convert audio metadata standards into domain ontologies. For example, Roberto Garcia and Oscar Celma (2005) converted the MPEG-7 standard into the MPEG-7 OWL Ontology using OWL syntax.

An ontology of music listeners links the characteristics of music elements with human beings' emotions. Although music is non-semantic (it has no objective meaning), it can affect listeners emotionally. And the connection is certain; specific characteristics of music elements and their combinations can arouse similar emotions among people (Zhang 2002). Ontologies of music listeners have been constructed based on the correlation, and can be used to recommend music pieces that correspond to a person's emotional needs. For example, Marques, Zuben, and Guilherme (2011) constructed the FTM Ontology to reflect how the domain of music relates to the domain of mood, personality, and physiology.

2.2 The applications of music ontologies (MOs)

MOs have a number of applications. First, MOs are used as vocabularies of music-linked data. On the semantic web, the individuals of classes in ontologies exist in the form of universal resource identifiers (URIs), which are called linked data. Constructing and sharing linked data sets is the key to the development of the semantic web. The Linking Open Data Project, which is supported by W3C, has been carried out to facilitate the process (W3C SWEO Community 2007). The Datahub website (https://datahub.io/) shows the general status of the project. At the time we wrote this article forty-two linked open datasets related to the field of music had been created. The contents of these data sets include information on music scores, recordings, musicians, and so on. Well-known data sets include MusicBrain, DBTune, BBC Music, and so on. Each linked open data set has its own vocabulary. The MOs used as vocabularies include the Music Ontology, Audio Features Ontology, the OMRAS2 Chord Ontology, and so on.

Second, MOs are used in browsing and recommendation services on music websites. The European Collected Library of Artistic Performance (ECLAP) provides users with a "Social Graph," in which entities, such as composers, musical works, and their semantic relationships, are represented by graphs (ECLAP 2013). Using the ECLAP service users can discover useful information. Ontology-based music recommendation is a widely studied application area. The main approach is to construct an ontology of music listeners, which includes the properties of emotion, preference, personality, and so on. With the ontology, the mapping relations between music listeners' individual states and relevant music recording resources can be built up, and music websites that use it can recommend music precisely. Celma and Serra (2008) used the ontology of FOAF to analyze users' music preferences and listening habits, and tried to recommend music based on their research results. Kim (2013) constructed an ontology named "UniEmotion." In the creation of the ontology, Kim classified tags in music websites, and calculated the emotional intensity of each tag. Music pieces that should satisfy users' emotional needs are recommended to them based on tags they have given. Rho et al. (2013) constructed the COMUS ontology to model users' musical preferences and context for supporting reasoning about their desired emotions and preferences. With the ontology, music websites can provide a context-based music recommendation service.

3.0 Research design

3.1 Research questions

According to the purpose of our research, there are two main questions: how to construct the classical music recording ontology and how to apply the ontology to organizing classical music recording resources.

To solve the first question, we looked for approaches based on our understanding of MOs, and the concepts of ontology and application ontology. From our research on constructing MOs we found out that the ontology of music recording resources is core among all kinds of music ontologies; the purpose of obtaining information on music subjects, artists, music elements, and music listeners is to make it easier for users to find music recording resources. Our research focuses on constructing an ontology of classical music recording resources, and we will extend our ontology in follow-up research. In the field of philosophy, ontology (Gnoli and Poli 2004) is the discipline that investigates the structure of reality. A domain ontology is the objective description of a specific domain. An application ontology (DeRidder 2007) is an ontology that has been designed to meet a specific set of practical needs using a domain ontology as a basis. If we want to organize articles on research into classical music recording resources, we need an ontology that displays concepts and their relations completely. But our purpose is to organize the resources. We need to think more about the retrieving habits and needs of users of music resource systems to improve the quality of the resource services.

For the question of applying the classical music recording ontology, we first investigated the existing applications of music ontologies, and found that these applications are only supplementary to the existing music retrieving systems. In order to improve the organization of classical music recording resources fundamentally, we decided to redesign the structure of the online music catalog completely.

3.2 Methodology

To construct the classical music recording ontology, we first investigated the existing state of classical music recording resources to define the main classes and main properties of the ontology. Then we investigated users' needs for information about classical music recording resources to obtain other properties of the ontology. With this approach, the classical music recording ontology not only remains objective but also can be used in the organization of classical music recording resources.

The construction of an ontology needs a scientific and efficient procedure. So we first formulated the whole procedure based on the skeletal methodology (Uschold and Gruninger 1996), which is often used in constructing ontologies. Figure 2 shows the whole procedure. There are three main steps. In the first step, the purpose, which is to construct an ontology of classical music recording

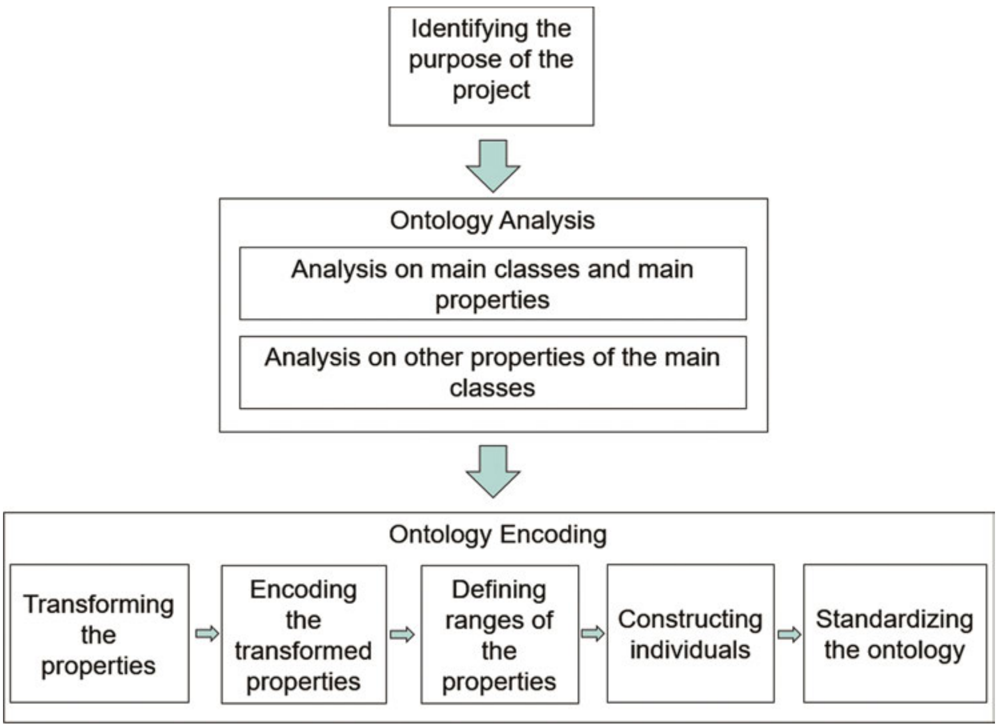

Figure 2. The procedure for constructing the classical music recording ontology.

resources, has been clarified. The second step is the core step, in which elements of the ontology were worked out. The third step focused on processing properties and representing the ontology in OWL.

To redesign the structure of the online music catalog, the main approach was to build up a multidimensional classical music catalog based on our analysis of classical music listeners' information needs, other research on their information-seeking behaviors, and the structure of the classical music recording ontology. Unlike existing music-retrieving systems, the catalog provides different types of information pages. Semantic links between the pages are built to make a music information network.

4.0 Ontology analysis

4.1 Analysis of main classes and main properties

We defined the main classes and main properties of the ontology by investigating classical music recording resources. A musical work (Smiraglia 2001) is an abstract conception of sound, which takes documentary form in a variety of instantiations. Classical music recordings are a kind of physical instantiation of classical musical works. We have to go deeper to investigate classical musical works (Smiraglia 2001), which form a key entity for music information retrieval and the relationships between classical musical works and recordings, which is critical for document-based information retrieval. Raimond, Abdallah, and Sandler (2007) described the process of producing music recordings which provided us with an approach of investigating the objects and the relationships. The main events of the process include: composing, performing and recording, and publishing. The main products of the steps are respectively musical works, recording versions, and albums. Our investigation was on the main products and relationships among them. To improve the wide applicability of our ontology in organizing classical music recordings, we chose common works, common recording versions, and common albums of them in the classical music field as the objects of our investigations.

4.1.1 Analysis of classical musical works

A classical musical work is the abstract creation of a composer, which can be recorded in notation as a music score or it can be performed by performing artists. We selected one hundred famous classical musical works that are typical representatives of different genres of classical music such as chamber music, opera, orchestral music, and so on. Using AllMusic (a well-known music metadata website, http://allmusic.com), we retrieved each work and counted the number of the albums of each work.

The result shows that the average number of albums for each of the 100 works is 566.59. The global library service provided by OCLC contains bibliographic records for all types of works including classical musical works in its WorldCat. The results show that 99% of the works in the WorldCat have fewer than 7 carriers, and only 1% of the works have more than 20 carriers. Classical musical works have many more records for carriers than most of the other kinds of works in the WorldCat. According to Smiraglia et al. (2011), at that time about 2.2% of the bibliographic records in WorldCat were musical recordings, which include classical musical recording records. Therefore, classical musical works comprise much less than 2.2% of all kinds of works in the WorldCat. The diversity and number of carriers is a prominent characteristic of classical musical works. Also it is common for an album to contain only part of a work; 40.4% of the albums contain only some parts of the 100 works.

4.1.2 Analysis of classical music recording versions

We define a classical music recording version as a recording of a performance of classical musical works by certain artists at a certain time, not a performing version of a certain work. The works performed in the same performance have the same recording version. In analyzing classical music recording versions, we first found the famous recording versions of the 100 works using the *Penguin Guide to Compact Discs and Cassettes* (Greenfield, March, and Layton 1996), which is an authoritative directory of classical music recordings. Then we selected the recording versions of Brahms' *Symphony No.4* for case analysis. This work is one of the most important works in music history (Service 2014). The number of its recording versions is at an average level among common classical musical works. We not only found its recording versions in the catalog of the Deutsche Grammophon Company (http://deutsche grammophon.com/en/cat/), which is one of the most famous classical music recording companies, but also identified the albums that include these recording versions.

Through the investigation into famous recording versions of the 100 works, we found that these works have many recording versions. There are 13.47 recording versions recorded in the *Penguin Guide to Compact Discs* on average. And at least 63.2% of all the recording versions include at least two works. The results of the case study on Brahms' *Symphony No. 4* are shown in Table 1. There are 13 recording versions in the catalog of Deutsche Grammophon. The time range of the recording versions is from 1939 to 2011. New recording versions appear in all the periods. Each recording version has albums. Seven of the 13 versions have more than one album. Here we have

Recording Versions	Albums
April, 1939: Berliner Philharmoniker; Victor de Sabata	2
December, 1953: Berliner Philharmoniker; Eugen Jochum	1
October, 1963: Berliner Philharmoniker; Herbert von Karajan	2
March, 1974: Stuttgart Radio Symphony Orchestra; Sergiu Celibidache	1
May, 1975: Wiener Philharmoniker; Karl Bohm	1
February, 1978: Berliner Philharmoniker; Herbert von Karajan	6
March, 1980: Wiener Philharmoniker; Carlos Kleiber	4
October, 1981: Wiener Philharmoniker; Leonard Bernstein	1
October, 1988: Berliner Philharmoniker; Herbert von Karajan	4
May, 1989: Wiener Philharmoniker; Carlo Maria Giulini	1
September, 1991: Berliner Philharmoniker; Claudio Abbado	2
November, 1994: Wiener Philharmoniker; James Levine	1
June, 2011: Los Angeles Philharmonic Orchestra; Gustavo Dudamel	2

Table 1. The results of the case study on Brahms' *Symphony No. 4*.

to explain that performers' various recording versions are not the same, as performances of the same work change in tempo, intensity, harmony, and so on, according to changes in the performers' understanding of the work. Twenty-four of the 28 albums contain more than one recording versions of different classical musical works.

Based on these two investigations, we identified some characteristics of the recording versions. First, common classical musical works have many recording versions. New recordings of a work are constantly appearing as time goes by, especially for famous works. In fact, there are hundreds of uncommon classical musical works that have only one recording version. The characteristic is still suitable for describing most of the classical music recordings we often see. Second, it is common that two or more works are performed and recorded in one particular performance and have the same recording version. Third, common recording versions have many albums. It is usual that famous recording versions are published many times in all kinds of forms, such as collections, popular editions, commemorative editions, and so on. Fourth, it is very common that there are two or more recording versions (usually of different classical musical works) on one album. Although digital classical music resources are more popular than CDs now, music is still published in the form of the CD album. Influenced by the capacity of a CD, it is common that there are several recording versions of several works in one album.

4.1.3 Analysis of classical music albums

A classical music album is a carrier containing recording versions, which can be published in physical forms and digital forms. We investigated albums using the album data from the website of Deutsche Grammophon. There were 3,893 album records. We used the method of random sampling and obtained 100 samples. We analyzed the form and elements of albums and counted the number of musical works in each album.

Among the 100 samples in the catalog of Deutsche Grammophon, there are 7 non-recording resources. The remaining 93 records are all in the form of the CD album. The main components of an album are album cover, album name, record number, track list, recording information, publication information, performing information, and so on. As shown in Figure 3, 23 albums contain only one work, and 70 albums, which comprise 75.3% of the whole sample, contain more than one work. It is clear that most of the classical music albums are published in the form of work collections.

4.1.4 Definitions of the main classes and main properties

From the analyses above, we can see that many-to-many relationships are very often observed among classical musical works, recording versions and albums. A classical musical work has several recording versions, and several works are often performed and recorded in one recording version; a classical musical work is contained in many albums, and there are often two or more works contained in one album; a recording version is published several times in different albums, and an album often contains several recording versions. We transformed the three objects and their relationships into main classes and main properties using the RDF Schema. The classes and properties are shown in Table 2.

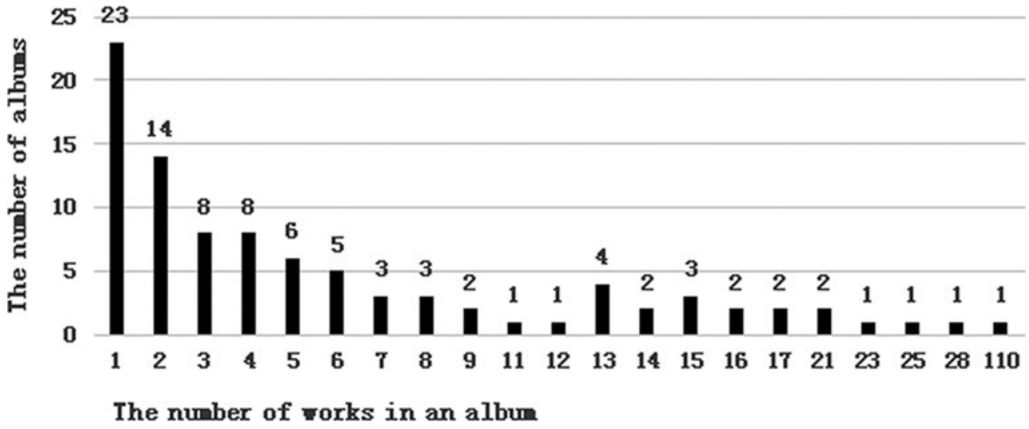

Figure 3. The number of works in each album.

Element type	Element name	Definition	
Class	Classical musical work	A classical musical work is the abstract creation of a composer, which can be recorded as a music score or be performed by performing artists.	
	Classical music recording version	A classical music recording version is a recording of a performance of classical musical works by certain artists at a certain time.	
	Classical music album	A classical music album is a carrier containing recording versions, which can be published in physical forms and digital forms.	
Element type	**Element name**	**Domain**	**Range**
Property	perform	Class: Classical music recording version	Class: Classical musical work
	performed_by	Class: Classical musical work	Class: Classical music album
	include	Class: Classical music album	Class: Classical musical work
	included_in	Class: Classical musical work	Class: Classical music album
	publish	Class: Classical music album	Class: Classical music recording version
	published_in	Class: Classical music recording version	Class: Classical music album

Table 2. Definitions of the main classes and main properties.

4.2 Analysis of other properties of the main classes

4.2.1 Survey on users' information needs

We obtained the other properties of the main classes by a questionnaire-style survey on users' information needs of the three products above. Inspired by the *FRBR* model and the *Functional Requirements for Authority Data* (*FRAD*) model, we also focused on composers and performing artists, as they have close relationships with classical musical works and recording versions. Before conducting the survey, we obtained the properties of the five objects above by investigating the structure and elements of the classical music recording resources, referring to other music ontologies, cataloging rules, the *FRBR* model, and the *FRAD* model. Then we integrated the properties obtained from different sources and classified them by the five objects. There were 58 properties in total.

The survey had three parts. The first part investigated users' needs on all the 58 properties. For each property, there were three options: "in great need," "need," and "not need[ed]." In the second part, users were asked to arrange the importance of musical work, recording version, album, composer, and performing artist. The third part asked whether the participant was a music professional or not. The participants of the survey were all classical music listeners. The music professional participants mainly came from music conservatories. Those who were not music professionals mainly came from amateur philharmonic organizations in universities and societies. We collected 150 questionnaires, of which 141 questionnaires were valid. Of the 141 valid questionnaires, 44 came from music professionals, and 97 came from non-professionals.

4.2.2 Analysis and processing of survey results

Before we analyzed and processed the results, we transformed them into numeric data. In the first part of the questionnaire, the three options of "in great need," "need," and "not need[ed]" were respectively transformed into values of 1.00, 0.50 and 0.00. In the second part, rankings from the first place to the fifth were assigned the values of 1.00, 0.75, 0.50, 0.25 and 0.00 respectively. In the third part, the participants who were music professionals were assigned the value 1.00, and those who were not were assigned the value 0.00.

Results analysis was divided into two parts. In the first part, the extent of participants' information needs for the five objects and the effect of the music professional factor on the extent were analyzed. We first used IBM-SPSS™ 21.0 statistical analysis software to calculate the average values of the two types of participants' information needs on the five objects. Then we used the bivariate correlation function provided by SPSS to analyze the effect of the music professional factor on participants' results. In the second part, the extent of participants' information needs for the 58 properties and the effect of the music professional factor on the extent were analyzed. First, average values of the two types of participants' needs for the 58 properties were calculated separately. We deleted four properties the average values of which were less than 0.50 among both types of the participants. Then we used the Pearson correlation coefficient to analyze the effect of the music professional factor on participants' needs on the remaining properties.

We processed the results for the remaining 54 properties. First, for the properties that did not correlate with the music professional factor, we calculated the average value of each property as the final value. For the properties keeping significant positive correlations with the music professional factor, we took the music professional participants' average values as the final values as we wanted to favor the music professionals' information needs to improve the applicability of the ontology. To balance the completeness and the practicability of the ontology, we set 0.60 as the threshold. The 41 properties the values of which were over the threshold were retained for ontology encoding.

We obtained several interesting results through the analysis above. First, the music professional factor has no significant correlation with users' information needs on the five objects of classical musical work, recording version, album, composer, and performing artist. The order of the five objects, which is shown in Table 3, is the same and the values are similar. Also it is obvious that what music websites provide is not consistent with what users want to know. Users need information about the classical musical work, composer, and recording version. When they retrieve these objects on the music websites, they can only get records of albums and artists with scattered, inaccurate information about the objects they are interested in. Users have to integrate the information sources or seek external information sources to get what they want. For example, if a user wants to know what different recording versions of a classical musical work exist, in order to find an interesting one to listen to, he or she has to look over each record of a related album to find information on recording versions, and often they find the same version over and over again on different albums. The process is no doubt time consuming. If the music websites could provide records of classical musical work and recording version, users could easily browse through all the versions and find the one they are most interested in.

Second, among the 58 properties, music professional users score higher values for their information needs than non-professional users on 55 properties, and are lower than non-professional users on the properties of "music critics," "sampling rate," and "resource access." From a macro perspective, professional users' needed information types are almost the same as those of non-professional users. But on a micro level, professional users need more specific, detailed, and in-depth information about musical works, recording versions and albums. It may be due to their professional activities, for example analyzing and comparing recordings.

Third, 23 properties have significant positive correlations with the music professional factor. These properties are: "name of the movement," "opus," "composing date," "composing place," "composing background," "musical expressional characteristics," "relationship with other works," "movement of the work," "performer," "recording date," "recording place," "recording background," "duration of the performance," "ISRC," "relationship with other recording versions," "name of the album," "packaging material," "producing technology," "barcode," "composer's gender," "composer's related place," "composer's institution," "performing artist' s important time

	Values of users' information needs for the five objects				
Type and number of users	Classical musical work	Recording versions	Album	Composer	Performing artist
Non-music professional: 97	0.8763	0.4433	0.1985	0.6005	0.3041
Music professional: 44	0.8580	0.4773	0.1364	0.6080	0.3295
Total: 141	0.8706	0.4539	0.1791	0.6028	0.3121

Table 3. Users' information needs for the five objects.

point," "performing artist's gender," and "performing artist's related place." The results also indicate that professional users are more in need of information on the details of the five objects.

5.0 Ontology encoding

5.1 Methods of ontology encoding

In order to apply the ontology to the construction of a classical music catalog, we had to present the ontology in OWL. The properties in the questionnaire, which were of high granularity, were abstract generalizations of basic properties. So, before encoding the properties, we decomposed these properties into basic properties for which we could define the domains and ranges. Also, colloquial descriptions of the properties in the questionnaire were transformed into brief, univocal terms to improve the normativity of the ontology.

In encoding the transformed properties, the main tasks were defining the properties and the ranges of the properties. To define the properties, first we divided the properties into three groups. For the key properties in the field of classical music recording resources, for example, the properties of albums, such as "album cover," we used the ontology editor tool Protégé to define them in OWL. For common properties, such as person, time, and place, we used the properties of existing ontologies or defined sub-properties of them. For example, the properties of "artist_place" and "publishing_place" were defined as sub-properties of the property "place" in the event ontology (http://motools.sourceforge.net/event/event.html). For properties in the music professional field, such as genre, instrument, mode, and rhythm, we mainly used the methods of extracting structural knowledge from authoritative music professional textbooks and referring to online music catalogs.

To define the ranges of the properties, we processed the properties in different ways. There are two kinds of property: object property and data property. The ranges of object properties are classes, and the ranges of data properties are basic data types, such as string, integer, and Boolean. For the properties we defined, we first judged whether they were object properties or data properties based on their definitions. For object properties, we created new classes as their ranges; for data properties, we chose suitable data types as their ranges. For example, "medium" is the property of album. The ranges of it are compact disk, vinyl, cassette tape, and audio files in different formats. These things are not basic data types. So we defined a new class named "Medium" and sub-classes of it as the ranges of "medium." For properties brought in from other ontologies and inherited from properties of other ontologies, we referred to these ontologies to define the ranges of the properties. We defined properties in the music professional domain as object properties, as they were all musical concepts. In defining new classes as the ranges of these properties, we controlled the hierarchies of the classes according to our purpose. For the complex concepts, we only defined the highest-level classes and referred to professional vocabularies and other kinds of knowledge organization systems as descriptions of the hierarchies of the classes. For example, the range of the property of "instrument_type" is the class of "instrument." As the hierarchies of instruments are so complex we did not define sub-classes of "instrument," and annotated the class with the URL of the "instrument taxonomy" (http://purl.org/ontology/mo/mit#). For concepts with simple structures, we transformed the conceptual structures of them in existing information organization systems to classes and sub-classes in the ontology. For example, the range of the property "genre" is the class "genre." We defined the sub-classes of it by investigating online classical music catalogs, such as Deutsche Grammophon, Decca (http://deccaclassics.com/en/cat/), AllMusic, and so on. The sub-classes include: "chamber," "choral," "keyboard," "opera/dramatic," "orchestral," and "vocal." Finally, all the defined classes and properties were encoded in OWL using Protégé.

Individuals are instances of the defined classes of the ontology. In constructing individuals, the manual method and semi-automatic methods are often used. In semi-automatic methods (Dong 2008), natural language processing technologies and text mining technologies are used to extract concepts and their relationships in natural language text, and the results are checked and fixed manually. We chose the manual method as there were problems in using semi-automatic methods. First, the information sources for constructing individuals of the classical music recording ontology were multiform and scattered. It was difficult to integrate these information sources automatically. Second, there was no proper information source for recording versions. We had to generate the descriptions of recording versions by analyzing albums manually. The procedure we followed is as follows. First, information sources were selected, including Wikipedia, online classical music catalogs, bibliographic records from libraries, and music websites. Then we extracted unstructured information from these sources and constructed individuals manually using Protégé. The procedure included: naming an individual, defining the class type of the individual, and setting the values of its properties. When naming individuals, for names of people, organizations and titles of musical works, we referred to the LC Name Authority File by using the Linked Data Service (http://id.loc.gov/) provided by the Library of Congress to make the individuals properly presented. For other individuals, we had to pay attention to the accuracy and identifiability of their names. For example, when naming a recording version, we used the form of "date/ performing artists/performed works," which was easy for users to see.

After these steps, the ontology still needed standardizing. First, we used the annotation properties to state the definitions, sources, iterative versions, and labels of the classes and properties. Then we used the tool FaCT++ provided with Protégé to find and fix the logic errors in the manual process of constructing individuals. With these operations complete, the final OWL file provided a clear, complete description of the ontology.

5.2 Overview of the classical music recording ontology

The classical music recording ontology is an ontology that describes and organizes classical music recording resources. In the process of encoding the ontology, we took records in the Chinese Machine Readable Catalog (CNMARC) as the main information source to construct individuals manually. We also referred to websites of Deutsche Grammophon, Decca, AllMusic, Wikipedia, and so on. Using Protégé, we exported the ontology in OWL and also built a term vocabulary to clearly present the ontology. The basic information of the classical music recording ontology is shown in Table 4. The definitions of the main classes and main properties are shown in Table 2. The classical music recording ontology vocabulary, which includes all the classes and properties, can be downloaded at http://1drv.ms/1LmJmA7.

The ontology has two main features: practicability and associability. The practicability is reflected in ontology analysis and ontology encoding. We processed the properties of the ontology mainly based on users' information needs. First, we balanced music professional users and non-professional users' information needs to make the ontology suitable for all the users. Second, we set a threshold value to retain those properties of high values to satisfy users' main information needs and to simply the structure of the ontology. Moreover, in encoding the ontology, we simplified the hierarchies of parts of the classes by referring to other information organization systems. The ontology of a specific field is usually very complex. It is necessary to simplify the structure of the domain ontology according to different applications, as the costs and benefits should be taken into account in developing an application.

The associability is reflected by the interlinking degree of the ontology. The interlinking degree of an ontology is the number of ontologies to which the ontology links. Improving the interlinking degrees of related data sets and ontologies is an important factor in realizing a unified, open semantic web. In the Linking Open Data Cloud 2014 Report, the interlinking degree is a key indicator for evaluating linked data sets or vocabularies (including ontologies). In order to improve the interlinking degree of the classical music recording ontology, we linked to existing ontologies and vocabularies by using properties and classes of them in defining the common properties, such as people, time, and places. With the development of the Linking Open Data Project, there will be more and more ontologies and vocabularies in the field of music. The interlinking degree of the classical music recording ontology needs to be improved to meet the needs of the development of the semantic web.

Element type	**Number**
Classes	79
Object properties	72
Data properties	34
Annotation properties	3
Individuals	254（still increasing）
Type of the information organization systems referred to in the ontology	**Information organization systems**
Ontologies	Event Ontology(event), FOAF vocabulary(foaf), Geo Ontology (geo, http://w3.org/2003/01/geo/wgs84_pos), Music Ontology(mo), Timeline Ontology (tl, http://motools.sourceforge.net/timeline/timeline.html), Music(music, http://www.kanzaki.com/ns/music).
Thesaurus in Simple Knowledge Organization System (SKOS)	Instrument taxonomy
Authoritative textbook in music major	Tutorial of Music Analysis(Qian, 2003)
Classification table in online encyclopedia	The classification table of audio formats in Wikipedia(http://en.wikipedia.org/wiki/Audio_file_format)
Online catalogs of record companies and music service websites	Catalogs of Deutsche Grammophon, Decca and AllMusic

Table 4. Basic information of the classical music recording ontology.

6.0 Constructing the classical music catalog using the classical music recording ontology

6.1 Methods of constructing the classical music catalog

We constructed the catalog in two phases: logical layer design and physical layer design. For the logical layer, we designed the structure and functions of the catalog based on our analysis of classical music listeners' information needs, and other research on their information-seeking behaviors. The catalog consists of different types of record pages. The pages correspond to classes in the ontology. The information and links in the pages correspond to data properties and object properties of the ontology respectively. Our design ideas are discussed thoroughly in the section Discussion.

In the physical layer, we realized the structure and functions designed in the logical layer. First, we used the API provided by Apache Jena to read the OWL file of the ontology and constructed a model of the ontology in the Java environment. Second, we adapted the inquiry functions of Apache Jena and also defined new functions for querying individuals of different classes. Third, we constructed the record pages based on inquiry functions, a Java Server page (JSP) for retrieval and a full text retrieval system on the OWL file based on Lucene. The retrieval system receives users' queries from the JSP. Then, the system selects suitable retrieving strategies and returns URIs of individuals to the JSP. Then the JSP calls inquiry functions to process the URIs and generate a list of individuals, which links to the record pages of each individual. We also constructed JSPs for browsing to provide users with an integrated browsing service on different kinds of individuals.

6.2 The general structure of the classical music catalog

The classical music catalog, the structure of which was mainly built on the classical music recording ontology, consists of a main record page, other record pages, browsing lists, and a retrieving page. The pages correspond to the classes of the ontology. The information in the pages corresponds to the data properties of the classes, and the links among the pages correspond to the object properties.

The main record page provides the structural information about classical musical works, recording versions and albums. The upper half of a main record page shows the information about a classical musical work. The tabs of recording versions are in the bottom half of the main record page. Users can browse the tabs to see the general information about each version and albums of each version. By clicking on the versions and albums, users can open the pages of recording versions and albums to see more information. The other record pages are recording version page, album page, composer page, performing artist page, and recording company page. Individual information about these classes is shown in these pages. Links built based on object properties in the pages bring the pages together, and form a classical music information network, which is shown in Figure 4.

The classical music catalog provides the information browsing service mainly by providing browsing lists, which include classical musical works, genres, performers, albums, and recording companies. The information browsing service also exists in record pages, in which integrated information is provided. For example, from a main record page, users can browse different recording versions of a work and different albums of a recording version; from an album page, users can browse all the works in the album; from a performing artist page, all of the artist's recording versions can be browsed. The retrieving page provides unified search and field limiting searches on works, composers, performing artists and albums, and links to each browsing list.

7.0 Discussion

The classical music recording ontology is constructed to describe and organize classical music recording resources. But this is not enough to provide classical music listeners with satisfactory online music services. The user interface, retrieval system, and functions of a music website should also be redesigned. So we explored the approach of constructing an online classical music catalog based on the structure and concepts of the classical music recording ontology. We also designed the catalog based on the *FRBR* model, the results of our survey on classical music listeners' information needs, and other research on music listeners' information-seeking behaviors. The following discussion covers the user-based design features of the classical music catalog.

7.1 The hierarchical structure of the main page

Inspired by the *FRBR* model, the main page was designed with a hierarchical structure. The complex relationships of classical musical works, recording versions, and albums are realized on the main page. It is much easier for users to find different versions of a work and different albums of a recording version on the main page than in OPACs of libraries and on music websites, which might provide only scattered records of albums. What's more, from our investigation into users' information needs, we know that in-

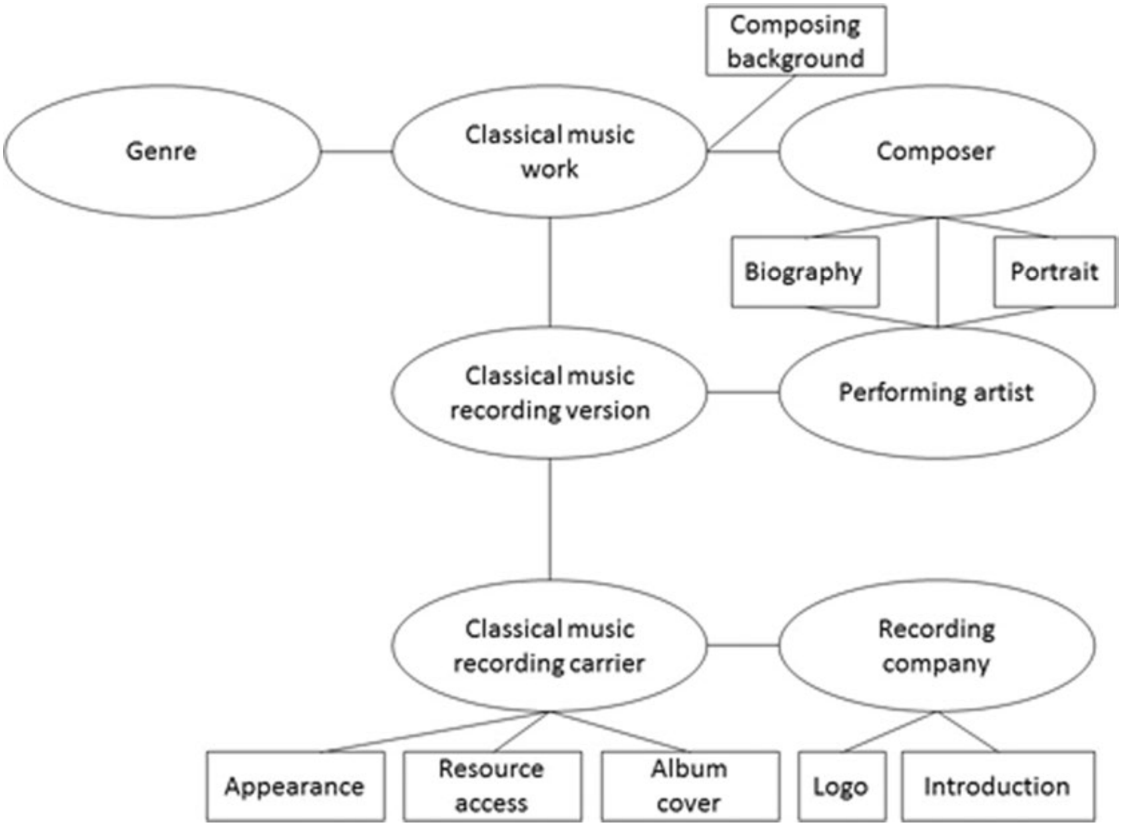

Figure 4. The classical music information network.

formation on classical musical works is what users mostly want to obtain. The main pages provide users with abundant information about classical musical works and links to other main objects. The design of the main page is consistent with users' information needs and can naturally guide a user's music-exploring journey from a work to an album that can be enjoyed in real life. Figure 5 shows the example of Beethoven's *Symphony No. 7*. The upper half shows the information about the work. In the bottom half, there are five recording versions listed. The albums of the recording versions are listed under each recording version. Users can obtain information about the work first to decide whether to listen to it or not. If users are interested in the work, they can select recording versions and albums by clicking on the tabs. If users want know more information about the composer and works of the same genre, they can click on the links "Beethoven" and "symphony."

7.2 The classical music information network

Besides the hierarchical main page, we also designed pages for composers, performing artists, and other objects to form a classical music information network based on the structure of the ontology. These pages and links between them can provide users with much more information about music than OPACs and music websites, which only provide limited information about albums and artists. According to our investigations, there are only 3 properties of which the values are less than 0.5 among the 58 properties we identified. The results are consistent with the conclusion of other research (Laplante and Downie 2011) on music-information-seeking behaviors that obtaining not only music recording resources but also information about music can satisfy users during their music-information-seeking activities. Classical music lovers do need abundant information about music. What's more, the network structure can enhance users' satisfaction. Hedonic outcomes (Laplante and Downie 2011), which are mainly manifested as pleasure and engagement in users' music-information-retrieving activities, are important in evaluating user satisfaction with music-retrieving systems. The engagement (O'Brien and Toms 2008, 949) is "a quality of user experiences with technology that is characterized by challenge, esthetic and sensory appeal, feedback, novelty interactivity, perceived control and time, awareness, motivation, interest, and affect." In the specific environment of using music-retrieving systems (Laplante and Downie, 2011), the engagement mainly comes from users' aesthetic experience with the interfaces of systems and the process of discovering new musical works and information on music consis-

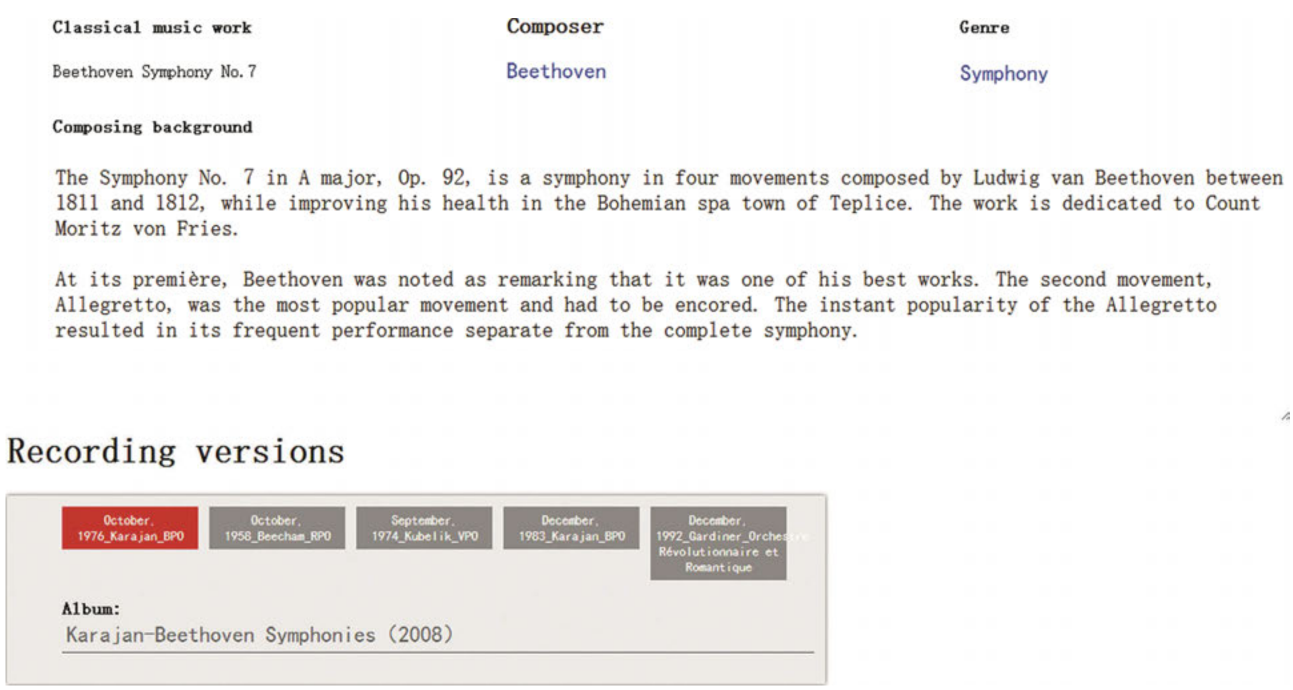

Figure 5. The hierarchical structure of the main page.

tently. The network structure of the catalog can provide abundant links to stimulate users to discover more.

7.3 The ubiquitous browsing services

Integrated information and browsing services help tremendously when users want to find information on a specific kind of object. For example, often users want to find all the albums of a recording version to decide which one to buy. However, the music websites and OPACs can only provide simple lists, such as lists of a singer's repertoire, musical style lists, and so on. The lists cannot satisfy classical music listeners' needs. The classical music catalog has realized information integration on all kinds of objects in multiple dimensions by providing abundant links between different pages. This feature can help users find information more efficiently. For example, when a user opens a page of a recording version, he or she can browse the integrated information about performers, albums of the version, and musical works performed in the version.

The ontology-based structure and the user-based design of the catalog will make classical music listeners feel more satisfied during their online music time. But, it is still not enough. We will conduct more research into classical music listeners' behaviors, psychological states, and other characteristics when they seek for classical music information and resources. These factors are of significance in designing our online catalog. Although we have now discovered some of the behavior patterns of classical music listeners (for example, listeners often spend a long time seeking for information about recording versions when they want to listen to a classical musical work they have never listened to), these patterns need follow-up research to verify and explain. Moreover, our understanding of music listeners' information-seeking behaviors are from studies on general music listeners. Classical music listeners' specific characteristics still need to be discovered. We still have many things to do to refine the catalog, and to expand the classical music recording ontology by encompassing more concepts from the field of classical music listeners' information-seeking behaviors.

8.0 Conclusion

In this paper, we described how we constructed the classical music recording ontology to improve the organization of classical music recording resources and constructed the classical music catalog based on the ontology. Compared with the music ontologies mentioned in the literature review, the classical music recording ontology is closer to the concept of application ontology. We developed new methods in the steps of ontology analysis and encoding, which can not only ensure objectivity, but also improve the practicability of the ontology. The methods of ontology analysis still need to be improved, especially the method of investigating classical music recording versions. We will find new ways to increase the objects of the investigation to make the results more convincing. In addition, providing a good degree of interlinking in the ontology was always a main principle during the step of ontology encoding. The classical music recording ontology has a better interlinking degree than the music ontologies above, as we brought in classes and properties of the existing ontologies and vocabularies, as many as possible. The semi-automatic method of constructing individuals was not implemented, which has been utilized in constructing different kinds of music ontologies; this needs subsequent research.

In constructing the classical music catalog, we took users' information needs and information-seeking characteristics into consideration and reflected them in the structure of the catalog. The hierarchical main page, many kinds of record pages with abundant links among them, and extensive browsing functions should satisfy users better than existing music retrieving systems during their music retrieving activities. But we are still lacking in a complete understanding of the classical music listeners. Inskip (2011) stressed the importance of the music users as the ultimate participant of musical communication process. Inskip, MacFarlane, and Rafferty (2012) have also found that subjective facets such as mood and genre are very important in users' query formation, which were not included in our investigation on classical music listeners' information needs. The functions of the catalog cannot suit users' music-retrieving habits perfectly. Further research on classical music listeners should be carried out to improve the catalog.

References

Celma, Oscar and Xavier Serra. 2008. "FOAFing the Music: Bridging the Semantic Gap in Music Recommendation." *Web Semantics: Science, Services and Agents on the World Wide Web* 6: 250-56.

DeRidder, Jody L. 2007. "The Immediate Prospects for the Application of Ontologies in Digital Libraries." *Knowledge Organization* 34: 227-46.

Dong, Hui. 2008. *Ontology and Digital Library*. Wuhan: Wuhan University Press.

ECLAP. 2013. "Social Graph." www.eclap.eu/portal/?q=en-US/node/116088

Garcia, Roberto and Oscar Celma. 2005. "Semantic Integration and Retrieval of Multimedia Metadata." In *5th International Workshop on Knowledge Markup and Semantic Annotation, Galway, Ireland, November 7, 2005*, ed. Sylvain Tenier, Amedeo Napoli, Xavier Polanco and Yannick Toussaint. http://ceur-ws.org/Vol-185/semAnnot05-07.pdf

Gnoli, Claudio and Roberto Poli. 2004. "Levels of Reality and Levels of Representation." *Knowledge Organization* 31: 151-60.

Greenfield, Edward, Ivan March and Robert Layton. 1996. *The Penguin Guide to Compact Discs and Cassettes*. London: Penguin Books.

Inskip, Charles. 2011. "Music Information Retrieval Research." In *Innovations in Information Retrieval: Perspectives for theory and practice,* ed. Allen Foster and Pauline Rafferty. London: Facet publishing, 69-84.

Inskip, Charles, Andy Macfarlane and Pauline Rafferty. 2012. "Towards the Disintermediation of Creative Music Search: Analyzing Queries to Determine Important Facets." *International Journal on Digital Libraries* 12: 137-47.

IFPI. 2013. "IFPI Digital Music Report 2013." http://ifpi.org/downloads/dmr2013-full-report_english.pdf

Kim, Hyon Hee. 2013. "A Semantically Enhanced Tag-Based Music Recommendation Using Emotion Ontology." In *5th Asian Conference, ACIIDS 2013, Kuala Lumpur, Malaysia, March 18-20, 2013, Proceedings, Part II,* ed. Ngoc Thanh Nguyen Ali Selamat and Habibollah Haron. Intelligent Information and Database Systems. Berlin: Springer, 119-28.

Laplante, Audrey and J. S. Downie. 2011. "The Utilitarian and Hedonic Outcomes of Music Information-Seeking in Everyday Life." *Library & Information Science Research* 33: 202-10.

Marques, Caio Miguel, João Von Zuben and Ivan Rizzo Guilherme. 2011. "FTMOntology: An Ontology to Fill the Semantic Gap Between Music, Mood, Personality, and Human Physiology." In *On the Move to Meaningful Internet Systems: OTM 2011 Workshop, Hersonissos, Greece, October 17-21, 2011,* ed. Robert Meersman, Tharam S. Dillon and Pilar Herrero. http://dblp.uni-trier.de/db/conf/otm/otm2011-w.html

Nisheva-Pavlova, Maria and Pavel Pavlov. 2011. "Ontology-based Search and Document Retrieval in a Digital Library with Folk Songs." *Information Services and Use* 31: 157-66.

O'Brien, Heather L. and Elaine G. Toms. 2008. "What is User Engagement? A Conceptual Framework for Defining User Engagement With Technology." *Journal of the American Society for Information Science and Technology* 59: 938-55.

Passant, Alexandre. 2010. "Dbrec--Music Recommendations Using DBpedia." In *The Semantic Web-ISWC 2010: 9th International Semantic Web Conference, Shanghai, China, November 7-11, 2010,* ed. Peter F. Patel-Schneider, Pan Yue, Hitzler Pascal, Mika Peter, Zhang Lei, Jeff Pan and Ian Horrocks. http://iswc2010.semanticweb.org/pdf/ISWC2010_PD_proceedings.pdf

Qian, Renkang. 2003. *Tutorial of Music Analysis*. Shanghai: Shanghai Music Publishing House.

Raimond, Yves, Samer Abdallah, Mark Sandler and Frederick Giasson. 2007. "The Music Ontology." In *ISMIR 2007: Proceedings of the 8th International Conference on Music Information Retrieval, Vienna, Austria, September 23-*

30, 2007, ed. Simon Dixon, David Bainbridge and Rainer Typke. http://dblp.uni-trier.de/db/conf/ismir/ismir2007.html

Rho, Seungmin, Seheon Song, Yunyoung Nam, Eenjun Hwang and Minkoo Kim. 2013. "Implementing Situation-Aware and User-Adaptive Music Recommendation Service in Semantic Web and Real-Time Multimedia Computing Environment." *Multimedia Tools and Applications* 65: 259-82.

Service, Tom. 2014. "Tom Service's Survey of the 50 Symphonies that Changed Classical Music." http://www.theguardian.com/music/series/50-greatest-symphonies

Smiraglia, Richard. 2001. "Music Works as Information Retrieval Entities: Epistemological Perspectives." In *ISMIR 2001: Proceedings of the 2nd Annual International Symposium on Music Information Retrieval, Bloomington, Indiana, USA, October 15-17, 2001,* ed. J. Stephen Downie and David Bainbridge. http://ismir2001.ismir.net/proceedings.pdf

Smiraglia, Richard P. et al. 2011. "Music in the OCLC WorldCat: A Replication." In *Directions in Music Cataloging*, ed. Peter H. Lisius and Richard Griscom. Music Library Association Technical Reports Series, v. 32. Middleton, Wisc.: A-R Editions, 3-14.

Uschold, Mike and Michael Gruninger. 1996. "Ontologies Principles, Methods and Applications." *Knowledge Engineering Review* 11: 56-67.

W3C SWEO Community. 2007. "Linking Open Data Project." http://w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData

Zhang, Qian. 2002. *Tutorial of Music Aesthetic.* Shanghai: Shanghai Music Publishing House.