Faceted Ontological Representation for a Music Domain†

Dr. Devika P. Madalli*, B. Preedip Balaji**, Amit Kumar Sarangi***

*Documentation Research and Training Centre (DRTC) Indian Statistical Institute (ISI), Bangalore 560 059, India, <devika@drtc.isibang.ac.in>
**Indian Institute for Human Settlements, Bangalore City Campus, No. 197/36 2nd Main Road, Sadashivanagar, Bangalore, 560 080, India <pbalaji@iihs.ac.in>
***ICT Doctoral School, Via Sommarive, 9 I-38123 Povo, University of Trento, Trento, <sarangiamitkumar@gmail.com>

Devika P. Madalli is an associate professor of the DRTC, Indian Statistical Institute, India, and Adjunct faculty, DISI, University of Trento, Italy. She has contributed to UNESCO's Global Open Access Portal (GOAP). She is on the advisory board of Universal Decimal Classification and a co-chair of the interest group on Agricultural Data at the Research Data Alliance. She has participated in EU funded projects including LivingKnowledge, AgInfra and worked as consultant to UNESCO and UNFAO. Dr. Madalli's interest is in the area of knowledge organization and application of facetization in information systems, information infrastructures, digital libraries, semantic web technologies, faceted ontologies, and multilingual information services.

Preedip Balaji is currently an associate librarian at the Indian Institute for Human Settlements, Bangalore, India. He holds a masters degree in library and information science and was associated with the DRTC, Indian Statistical Institute as a senior research fellow. His research interests lie in the semantic web, digital libraries and knowledge organization.

Amit Kumar Sarangi is an ICT doctoral student at the University of Trento, Italy since November 2013. He completed his MS in library and information science from DRTC (Indian Statistical Institute). As a part of his PhD thesis, he is currently working towards building data models as well as developing vocabulary for knowledge representation of bibliographic entities. His research interests also include knowledge organization for the bibliographic universe.

Madalli, Devika P., Balaji, B. Preedip, and Sarangi, Amit Kumar. **Faceted Ontological Representation for a Music Domain.** *Knowledge Organization. 42(1)*, 8-24. 45 references.

Abstract: This paper proposes an analysis of faceted theory and of various knowledge organization approaches. Building upon the faceted theory of S.R. Ranganathan (1967), the paper intends to address the faceted classification approach applied to build domain ontologies. Based on this perspective, an ontology of a music domain has been analyzed that would serve as a case study. As classificatory ontologies are employed to represent the relationships of entities and objects on the web, the faceted approach is deemed as an effective means to help organize web content. While different knowledge organization systems are being employed to address the cluttered Web in different contexts and with various degrees of effectiveness, faceted ontologies have an enormous potential for addressing this issue by performing domain analysis for knowledge modeling and ultimately facilitating semantic information retrieval.

† We are thankful to Dr. Biswanath Dutta who gave suggestions to improve this paper. We are extremely grateful to Prof. M. Cristina Pattuelli for her insightful comments and corrections, which helped to revise this paper.

Received: 20 October 2014; Revised: 7 February 2015; Accepted: 12 February 2015.

Keywords: facets, faceted, knowledge, domain, music, ontologies, classification, canons

1.0 Introduction

Classification is the process of grouping any abstract entities or objects and is an intrinsic trait of humans that developed from immemorial times. Aristotle's classical theory of categories, based on a taxonomical approach, was the first codified knowledge representation system. It lasted for millennia and it was only in the mid-1950s that it was challenged beginning with Wittgenstein and his theory of language games (Wittgenstein 1953) and new forms of knowledge organization systems (KOS) started to emerge (Taylor 2004). Our modern perception of knowledge is that the notion of knowledge is a continuum having multifaceted disparities. In the early 20th century various thinkers from philosophy and library science drew up schemes and knowledge systems, which largely influenced the mapping of the universe of subjects. Later, the birth of Internet with its scattered, unorganized resources presented the biggest challenge ever for knowledge organization (Hong 2006).

Conventionally, knowledge has been represented with classification schemes, but socio-cognitive views have widely influenced knowledge representation models. Compound and complex subjects pose several challenges to classification for differentiation and hierarchical organization. Epistemological foundations set the stage for ontological inquiry and classificatory ontologies for knowledge representation models. The need for semantic-based systems arose to influence machine-discernable, contextual and meaningful information retrieval for the web. Moreover, as the paramount task of organizing the knowledge on the web gathers momentum, the emergence of data modeling and knowledge engineering techniques are overarching. Some of the major knowledge organization approaches are described below and Figure 1 portrays major KOS in the order for a music domain—UDC Classification, an *Art & Architecture Thesaurus* entry for music, the *Ontology of Music Work* and *Library of Congress Subject Headings*.

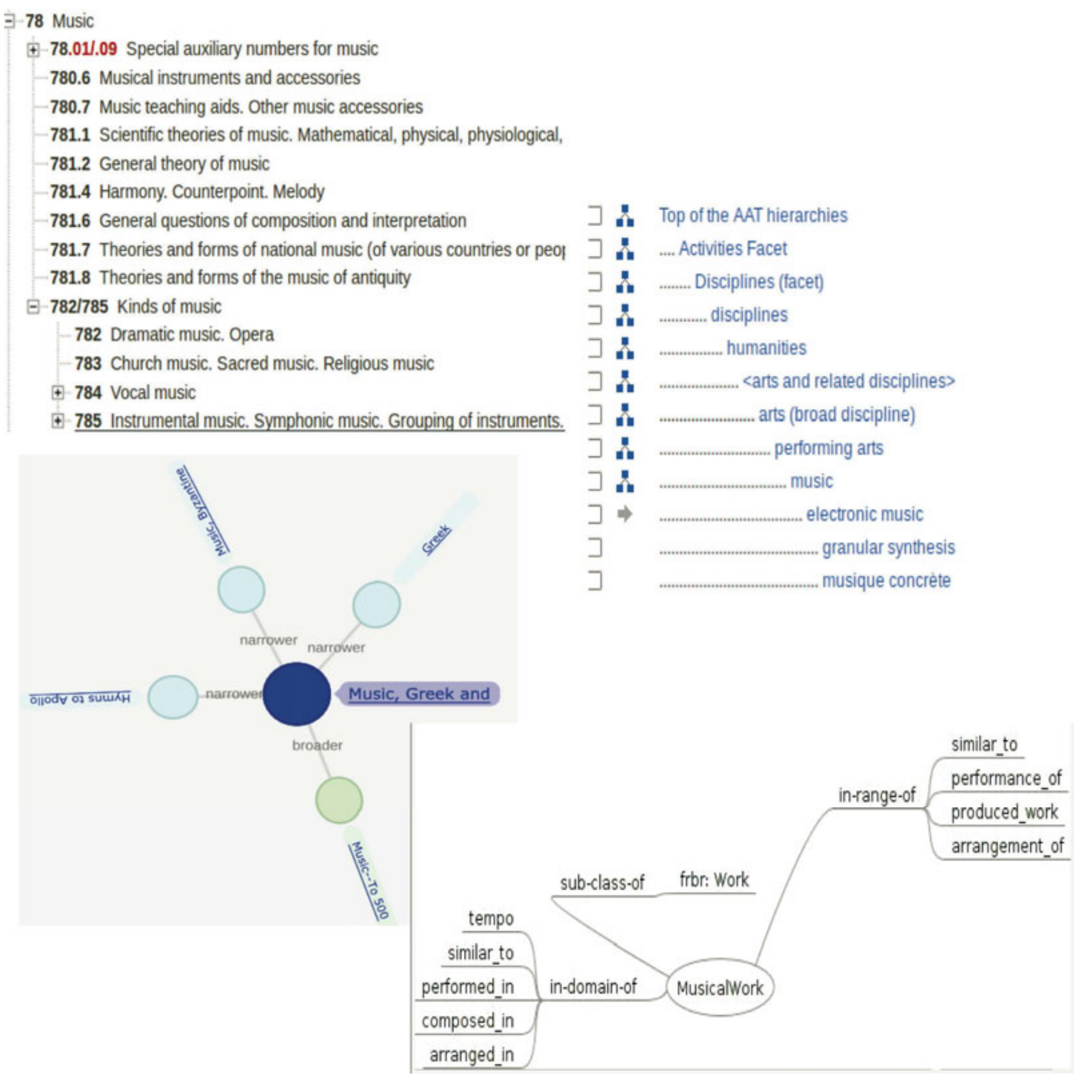

Figure 1. Example of KOS Approaches

In this paper we present a principled approach to analyzing music as a domain and explain the resulting concept scheme which is arranged by facets at a higher level of abstraction with each facet consisting of concepts with shared features with increasing intension of scope in the hierarchies. Such an exercise leads to the development of faceted ontologies, which is further discussed in section 2.0 below.

1.1 Major approaches of knowledge organization systems

From a historical perspective, the field of knowledge organization can be broadly divided into three approaches that correspond to three historical phases. Since the dawn of modern society, knowledge organization has been reflected upon these foundational approaches adapted to represent and control the universe of knowledge.

1.1.1 Classification Approach

Aristotle organized the universe of knowledge dividing it into ten main classes in order to:

> Unambiguously classify all the phenomena by their essential true qualities, and Francis Bacon looked at empirical scientific methods to divide primary knowledge into divisions and its subdivisions. Bacon distinguished between sacred theology revealed to man by God and human knowledge acquired by man's unaided powers wherein the human knowledge was subdivided into history, poesy and philosophy.

according to Wallace (2007).

These classical systems provided the foundations for knowledge organization that lasted until the mid-twentieth century (Taylor 2004). In the field of librarianship a highly influential system of bibliographic knowledge organization was devised by Melvil Dewey in 1876. The universe of knowledge was represented in decimal notational system, where all the syntactical and semantic relations of classes and divisions of subjects are classified to incorporate ever-evolving domains of subjects in an enumerated manner. Many other schools of thought followed this and various classification systems were developed to organize resources in libraries including the Library of Congress *Classification*, Universal Decimal Classification, Bliss *Bibliographic Classification*, and *Colon Classification*. Pioneering a freely-faceted library classification system, S. R. Ranganathan in the early 20th century focused on subject formation modes, faceted classification and the renowned faceted theory (Ranganathan 1967).

1.1.2 Subject indexing approach

As the standardized knowledge organization systems such as term lists, classification schemes, categories and relationship lists evolved, artificial subject indexing languages were also formulated (Hodge 2000). Subject indexing languages used for formulating subject terms generate topical terms in the process, therefore performing an act of classification to facilitate computerized searching in library catalogues and standalone databases (Bhattacharya 1982). As potential vocabulary aids in subject search, in order to computerize library operations, various subject indexing techniques were implemented in the computer era of the late 20th century to facilitate information retrieval. PREserved Context Indexing System (PRECIS), chain indexing, and Postulate-based Permuted Subject Indexing (POPSI) are some examples of subject indexing languages tested for automatic indexing (Austin 1974; Bhattacharya 1979; Aptagiri, Gopinath & Prasad 1995). The choice of the name of the subject of a document and the rendering of the name in the heading of the specific subject entry can be obtained by facet analysis based on postulates and principles according to Ranganathan (1967). It was later demonstrated by POPSI that "using facet analysis for subject heading does not amount to using class number" (Biswas 1998, 192). Information retrieval became of paramount importance compared to traditional library environments, because the Internet made classification free from the limitations of location, notation, and shelving, but highly dependent on the prospects of indexing for information retrieval.

1.1.3 Ontology approach

Ontology is a formal, explicit specification of a shared conceptualization (Gruber 1995). Maedche and Staab (2001) defined ontologies as "metadata schemas that provide a controlled vocabulary of concepts of explicitly defined and machine processable semantics, describing shared and common domain theories helping people and machines to communicate concisely." Ontologies help to represent the relational aspects of objects by attributes, relationships and how they are interconnected in a domain or generally. It can be said that ontology encapsulates the complete existence of an entity, its relations, and attributes within a domain by formal conceptualization and most importantly it facilitates reuse and sharing of ontologies among humans as well as machines (Fensel et al. 2007). Although knowledge representation techniques were used to explore and connect different knowledge objects, the analysis of semantics in specific domains is slowly taking off to reach the potentialities of a postulated semantic digital library. Semantic web services use models such as

Web Service Modeling Ontology (WSMO) and Web Services Description Language (WSDL), and protocols like Simple Object Access Protocol (SOAP) and data and Web ontology languages viz., RDF, OWL and SKOS etc. Basic data can be modeled with primitive features such as classes, properties and instances whereas other relationships viz., instance-of, subclass-of and subproperty-of allow structured and property hierarchies generic for ontologies. An example of a music domain ontology expressing the classes might include instrument, release, and composition, properties are collaborated_with, composed_in, remix_of , and individuals are album, single, soundtrack (see http://musicontology.com/). Hence, examining the evolutionary approaches in knowledge organization, a graphic representation of the role of ontologies in relation to the other knowledge systems is shown in Figure 2 wherein each ellipse represents the evolving nature of KOS from the core and the ontologies ellipse refers to the broader relationship it shares with other KOS.

2.0 Faceted ontologies for a semantic web

We introduce the concept of principled approach to building ontologies from theoretical classification and follow the method of facetization of domain to represent concept and conceptual relationships that lead to faceted ontologies. Faceted ontologies are deemed an effective means to create consistent and cohesive knowledge structures for data modeling and representation. They "systematically organize domain knowledge, provide the facility to envisage a given set of concepts in different contexts and relations, and so enable different subject views as required by the users" (Prasad 2008,455). As a great deal of research is focusing on semantic-based web services, faceted ontologies serve as an effective means for organizing web-based knowledge. Faceted subject analysis allows classifying an object from multiple perspectives making the faceted approach widely appreciated and adapted. Ontologies are needed to reach the desired degree of expressiveness and provide machine-readable properties which would describe web services at a sufficient level of granularity for interoperability, crosswalking and contextuality (Fensel et al. 2007).

Faceted metadata are important for the creation of Web page interfaces more intuitive for the users. Faceted metadata makes it easier to navigate a web of resources (Prasad 2007). Several usability tests have proved that users do not hover over the pages when they end up finding irrelevant results (Porter 2003). Implicit classificatory ontologies combined with faceted infrastructure would be helpful not only as knowledge organization systems but also for "rep-

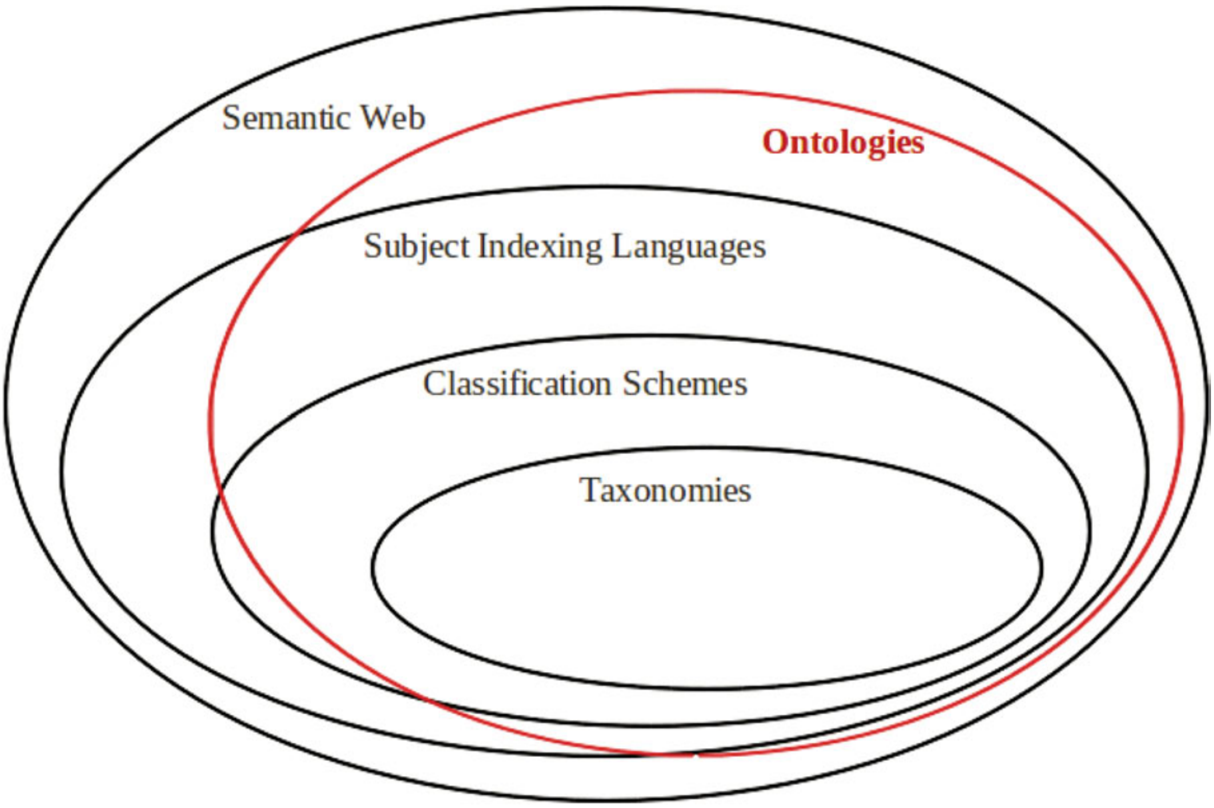

Figure 2. Relationships between KOS. Overlapping ontologies ellipse represents the broader relationship and the sizes are given only for clarity.

resenting such concept schemes in a machine-processable language that will help to realize the idea of concept-based retrieval instead of text-based searching, which is the main motto of the semantic web" (Prasad and Madalli 2008; Prasad and Madalli 2009, 225). Moreover, faceted ontologies are empirically evidenced by Prieto-Diaz (2003) who extended faceted theory for ontologies with a domain analysis approach in a semi-automated way. Giunchiglia et al. (2009, 36) also extended the theory of facet by the nomenclature faceted lightweight ontology:

> Where the labels of nodes are organized according to predefined patterns which capture different aspects of meaning i.e., facets. Here each term and corresponding concept occurring in its node labels must correspond to a term and corresponding concept in the background knowledge, modeled as a faceted classification scheme.

Emphasising the need for exploratory search in the post-Internet era, Tunkelang (2009, vii) described the role of faceted search pivotal for the internet as "faceted classification addresses the weakness of earlier knowledge representations—namely the rigidity of taxonomical schemes and the chaos of unstructured indexes ... faceted classification offers an approach to knowledge representation that is both faithful to its riches and practical for real-world use."

3.0 Music domain analysis

Music as a domain is vast in scope and divergent in terms of concepts and conceptual relations. Music thus poses a challenge for knowledge representation. According to Smiraglia (2001), musical works when considered as entities for information retrieval constitute varying instantiations of abstract creations. More importantly Smiraglia comments that semiotic analysis of musical works indicates a variety of cultural and social roles. This is significant in the analysis of music, as the domain is subject to cultural, temporal, geographic and sociological factors and hence can be envisaged and visualized by one or a combination of perspectives. In his analysis of the domain of music, Abrahamsen (2003, 146) argues that:

> The overall domain of music will be treated as everything that can be connected to…and as sets of related discourses and domains. Music may, of course, also be regarded as something belonging to other domains such as education (teaching music), philosophy (thinking about music in philosophical ways), business (selling music), information science (organizing and retrieving music, etc).

For the purpose of building an ontology, the music domain has to be defined in such a way as to establish its extension and intension in terms starting from generic to specific levels. According to Tennis (2003), for domain analysis to work cumulatively, transferable definitions of domains have to be made explicit. Tennis puts forth the two axes of domain analysis; one the areas of modulation and second one as the degrees of specialization. These axes correspond to the generic level indicating the scope and specific level of a domain. Hjørland (2002) describes eleven approaches to domain analysis including among others historical studies, empirical user studies and document and genre studies.

Classifying music is a challenging task as the music activities, repertories, and oeuvres transcend borders and create a conceptual framework for music faceted ontology requiring a deep analysis and domain expertise. Moreover, music genres are not always amenable to a clear-cut categorization. For instance, genres such as popular music are becoming progressively more fusional and integrated with other types of music. Terminology can be also misleading as, for example, the case of the term "glass" that can be a subject heading for the common kitchen utensil as well as a music instrument because a glass is used to produce and perform music (Library of Congress 2010a). The Library of Congress (2010b) has recognized that:

> Merely creating headings as they exist in the subject headings is not the best way, as the representation of musical works—forms, genres, types show complexity of the existing heading structure conflating genre/form terms with medium of performance and there are better means to use controlled vocabulary to promote discovery of musical works.

Prototypes of music digital libraries and catalogues have been developed that facilitate effective information retrieval of music information. Novel web services rely on music information extraction and representation underlying music-content-based, music-context-based and user-context-based approaches that inform novel types of interfaces, content-based search schemes and networked delivery mechanisms (Schedl et al. 2011). Music resources are arranged differently according to various approaches and purposes. For example, the Library of Congress uses subject, name, and title as three facets to navigate music resources under the performing arts domain (Library of Congress 2012). YouTube and Allmusic.com use genres for browsing music videos and to navigate musical content on their websites (see: http://www.youtube.com/music; http://www.allmusic.com/genres). The genres are shown in Table 1.

YouTube's Genres		AllMusic's Genres	
Alternative	K-Pop	Avant-Garde	Latin
Avant-Garde	Latin	Blues	New Age
Blues	Metal	Children's	Pop/Rock
Comedy/ Spoken	New Age	Classical	Rap
Country	Pop	Country	Reggae
Easy Listening	Rhythm & Blues	Easy Listen-ing	Religious
Electronic	Rap	Electronic	Rhythm & Blues
Folk	Reggae	Folk	Stage & Screen
Holiday	Religious	Holiday	Vocal
International	Rock	Interna-tional	
Jazz	Stage & Screen	Jazz	
	Vocal	Latin	
		New Age	

Table 1. Music genres of YouTube and AllMusic

Musipedia is an open collaborative music search engine displaying an explicit faceted interface to search music by instruments—keyboards, piano, and microphone and music elements such as contour and rhythm. It also uses melody to search for tunes and musical themes (see: http://www.musipedia.org/). The underlying ontological structure of its interface is represented in Figure 3.

4.0 Faceted ontologies for music

An attempt has been made to analyze faceted ontologies for the music domain. Faceted ontologies for music are elaborated with examples for this study as discussed in this section. A faceted system would allow for multiple ways of classifying an item, rather than the hierarchical approach of Dewey or indeed taxonomies (Coult 2010). Faceted ontologies are robust knowledge mechanisms wherein facets are described as clearly defined, mutually exclusive, and collectively exhaustive aspects, properties or characteristics of a class or specific subject. Comparing different classification schemes, Foskett (1982, 207-8) highlighted the hospitality of *Colon Classification* as hierar-

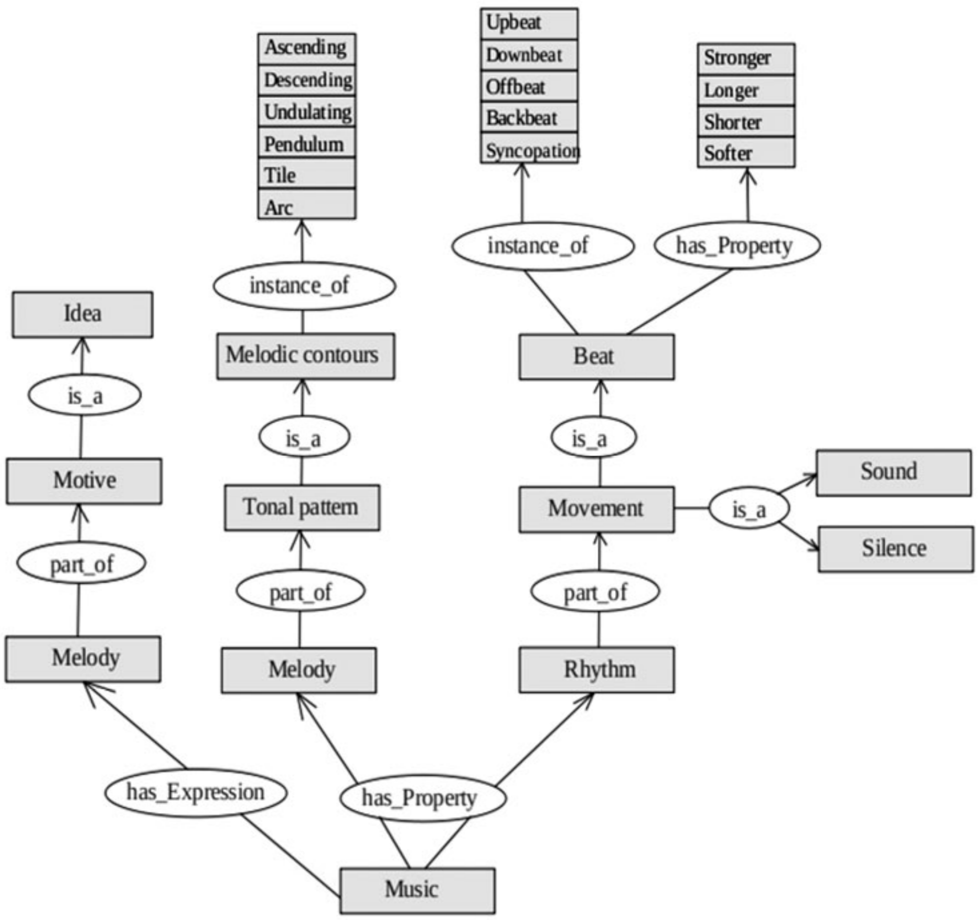

Figure 3. Ontological relationship model of Musipedia.org interface

chical and expressive, accommodating different kinds of division and displaying them both in chain and array. In faceted classification "without any influence or inhibition by the existing schedules for classification, whatever facets occur in a compound subject are all found out by the facet analysis of that subject" (Ranganathan 1967, 109). In Ranganathan's idea plane of the classification theory, he noted, "isolates are any idea or idea-complex fit to form a component of subject, but not by itself fit to be a subject." Music is the art of arranging sounds in time so as to produce a continuous, unified, and evocative composition, through melody, harmony, rhythm, and timbre (American Heritage Dictionary 2011). As a systematic domain-specific ontology, the organization of music begins with distinguishing the relations such as theory (elements: time, melody, harmony, tone, texture); themes (expressions: affection, sweetness, warmth); forms (associated with motions and notes: fugue, plainsong, canon, chant, madrigal); genres (country, classical and traditional); persons (musician, flutist, conductor) which was devised by Leach (1976) in his musical thesaurus. Figure 4 shows a basic ontology of music as a process of auditory communication and Figure 5 illustrates music as a compound subject.

Determining the relationships is essential for building faceted ontologies both implicit and explicit and even as they are embedded inherently between the entities of the main classes of the music domain. Among the many ongoing projects, Music Ontology (see http://musicontology.com/) described many relationships of music manifestations and AllMusic.com has defined its relationships attributes, as shown in the Figure 6.

5.0 Faceted theory of an idea plane for music

Ranganathan (1967, 143) postulated faceted classification principles to determine the synthesis of subjects. His "idea plane" consists of canons at the very beginning of conceptualization. This can be applied while analyzing domains to build a faceted ontology. Facet is a generic term used to denote any component—be it a basic subject or an isolate—of a compound subject and also its respective ranked forms, terms and numbers. Characteristic is any attribute of any complex of attributes with reference to which the likeness or unlikeness of entities can be determined and at least two of them are unlike. For example, rhythm is a characteristic of music, but not the sound of the music, which is equally shared by all listeners. Class is a ranked group, if a set of things is divided into groups based on characteristics or attributes and those groups are ranked, then each ranked group is a class, examples are "The Arts – 7" in UDC and "Fine Arts – N" in *Colon Classification*.

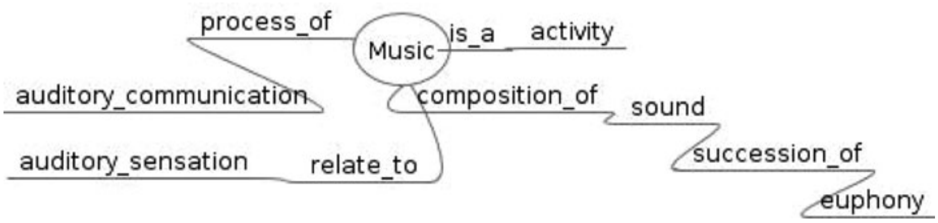

Figure 4. Basic ontology of music

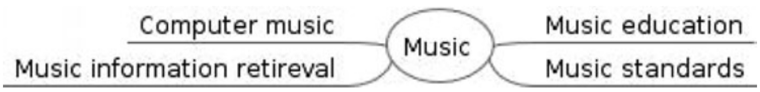

Figure 5. Music as a compound subject

	artist	label	recording	release	release_group	url	work
artist	artist-artist						
label	artist-label	label-label					
recording	artist-recording	label-recording	recording-recording				
release	artist-release	label-release	recording-release	release-release			
release_group	artist-release_group	label-release_group	recording-release_group	release-release_group	release_group-release_group		
url	artist-url	label-url	recording-url	release-url	release_group-url	url-url	
work	artist-work	label-work	recording-work	release-work	release_group-work	url-work	work-work

Figure 6. Relationship attributes of a music record at AllMusic.com

Array is the class derived from a universe on the basis of a single characteristic at any one step in the progress towards its complete assortment and arranged in the preferred sequence. Chain is a sequence of classes made up of any given class and its universe (e.g., Universe of Knowledge → The Arts → Music → Music Theory → Elements of Music → Harmony). Entities constitute any existent, concrete, or conceptual, which is a thing or an idea (e.g.: Microphone, Dance music, Mesmerizing, Music catalogue). Attribute is any property or quality or quantitative measure of an entity, action or discipline (e.g., in Music Element, Kind, Form, Musical Work, Artist, and Label are the attributes). In the universe of subjects, isolate ideas germane are manifested through verbal and notation parts by concepts and artificial numbers respectively. Just as a diamond has many facets, so does any given subject, each representing a different view, aspect or dimension, thus allowing all pertinent facets of an item to be represented in its classification (Stewart 2011). Identifying the relationships of entities is an inherent task in working out the characteristics. Ontologies are used to find the explicit relationships among the entities in many ways. In that sense, in music: available_as, track_number, instance_of, related_to, similar_to, derived_from, member_of, composed_in and so on are some examples. The idea plane comprises canons listed below and each canon has its sub-set of canons described with examples derived from music domain.

5.1 Canons for characteristics

Canons for characteristics have four sub-categories which draw attention to distinguish the properties of the entity to be classified. Each of these canons is described with examples taken from music domain analysis:

5.1.1 Canon of Differentiation

In the different kinds of music, the dance music can be differentiated, but the performance of dance music cannot be considered. Similarly, musical instruments can be categorized based on wind, string and air but not differentiated by possession, ownership, and performance.

5.1.2 Canon of relevance

Relevance of the entity is utmost importance to ensure how relatable the entity to the purpose of facetization is. When music multimedia is taken into consideration, then sound recording, computer music, motion pictures can be characterized, rather than storage, brand, duration and quality. Similarly, persons associated with music can be composers, singers, performers and researchers rather than their lifestyle, background and talent.

5.1.3 Canon of ascertainability

The characteristic in view should be definite and ascertainable. Persons associated with music can be ascertained with their date of birth, then their education, family and nationality attributes. In the same way for vocal music, sex and vocal performance are more ascertainable than the lyrics and instruments.

5.1.4 Canon of permanence

As entities in the universe of subjects are protean, permanence of the subject should be factored. For musical forms, dance music, dramatic music and religious music are permanent than by the categorization of musician and labels. Similarly, music bands can be categorized as concert bands, rock bands and jazz bands over the organization, affiliation, and producing company.

5.2 Canons for succession of characteristics

Characteristics being complex to distinguish the intricacies and fallacies, the postulates in succession of characteristics help to reduce the ambiguities of properties of item being faceted with the canons that follow.

5.2.1 Canon of concomitance

A canon of concomitance emphasizes the need for different criteria for deriving successive characteristics than dwelling upon attributes which would result in same arrays of classes, or of isolate ideas. For example if music researchers are to be classified, their area of specialization would be more relevant by ethnomusicologist, sociomusicolgist and organologist than by instruments, since a music researcher could be an instrumentalist too.

5.2.2 Canon of relevant succession

In achieving the succession of characteristics, relevance should be maintained. For music literature, language, form, kind, composer and time facets could be more relevant for getting relevant succession.

5.2.3 Canon of consistent succession

Adhering consistently with the associated scheme of characteristics is essential as long as there is no change in the purpose of ontology creation. Hence, characteristics should be propounded to maintain the permanence, and their succession in application should be consistent. For music history, space and time give rise to consistent succession, which are more general for any other domain as modifiers. For example in the query: "Use of tabla in Rajasthani classical music for thumri in 18th century," century can be analyzed as:

Basic Subject: Music
Personality: Classical music
Matter: Tabla
Energy: Thumri (it is type of rendition)
Language Isolate: Rajasthani
Space Isolate: Rajasthan
Time Isolate: 18th Century

5.3 Canons for array

5.3.1 Canon of exhaustiveness

The classes in the array of classes, and the ranked isolates in an array of ranked isolates should be totally exhaustive of their respective common immediate universe. In music, although geographical influences have been wide, it has been limited to Western music comprising of classical and popular. But given the rising of fusional music kinds, the genres are conflated with variety of elements, for accommodating international music and with geographical flavours like Afro-Jazz, and Chicago blues, etc (see Table 1 for genres).

5.3.2 Canon of exclusiveness

In order to ensure exclusivity in the classes of array, classes should be derived from its immediate universe that no two classes of the array can overlap or have an entity in common. An example of persons associated with music has been given in Figure 7. By performance, if we consider a musician as singer, pianist, accompanist and so on and the same can be further faceted (by instrument: pianist, harpist, accompanist) and (by number: soloist, troupe).

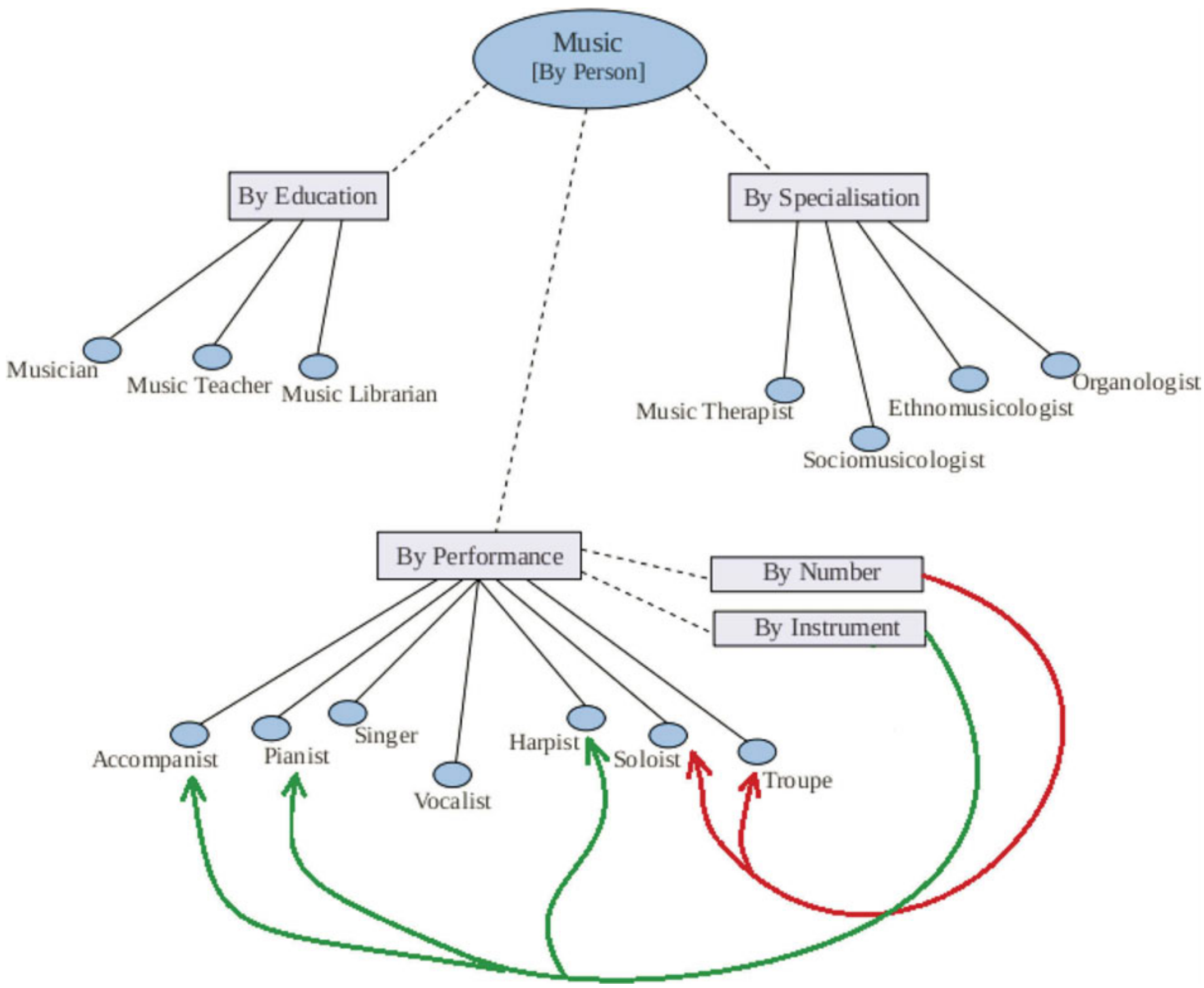

Figure 7. Facetizing exclusivity of persons associated with music

5.3.3 Canon of helpful sequence

When the subject in an array of classes evolved in a period of time, then sequencing the time facets should be rendered in a logical sequence as it progresses. For example, if music history is considered, here the principle of "later-in-time" can be applied as good prehistoric, ancient, Biblical, medieval, renaissance, Baroque, classical, romantic, 20th century, contemporary and 21st century as Wikipedia does (Wikipedia 2012).

5.3.4 Canon of consistent sequence

In order to ensure the logical conformity for sequencing the common facets, different standards are used by different classification schemas. In *Dewey Decimal Classification* (*DDC*), 22nd edition, Table 1 fulfills this with standard subdivisions. Part D of the *Colon Classification* 7th edition serves the same purpose with the help of general subdivisions and common isolates.

For example, the analysis of this query: "Pitch in Carnatic music recording in Sanskrit at Chennai December Season, 2011" gets the notation "NR15;cMB:pR.4411'P 11-964 CAR" in the *Colon Classification*. The analysis of this subject takes the aid of common isolates and general devices as follows:

NR: Music basic subject
cMB: Pitch—Common matter property isolate
pR: Energy—Recording
15: Sanskrit—Language isolate
4411: Chennai—Space isolate
P11: 2011—Time isolate
964: Winter season—Time speciator
CAR: Carnatic Music as an alphabetical device

5.4 Canons for chain

5.4.1 Canon of decreasing extension

This canon helps to obtain the broad entities in a domain into their break-down facets in the order of broad to narrow subjects as the intension of concepts increases. In the case of classical music history, the ranked isolates decreasing extension is shown in Table 2:

Music
Traditional
Folk
Eastern Folk
Indian Folk Baule, Vadu

Table 2. Canon of decreasing extension

Consider the following two chains taken from the same class "Instruments," in the main class "Music." Flute and piano belong to the ranked isolate "instruments," but are not subordinate to each other which is why they can be differentiated as monophonic instruments that can play only one note at a time and polyphonic instruments which can play several notes at a time (See Table 3).

[by note]: instruments	
Monophonic	Polyphonic
Flute	Piano

Table 3. Canon of decreasing extension

5.4.2 Canon of modulation

A chain of classes or of ranked isolates should comprise one class or one ranked isolate, as the case may be, of each and every order that lies between the orders of the first link and the last link of the chain. Here Ranganathan (1967, 176) discussed the difficulty of obscurity in subjects and the challenges posed by indeterminateness to derive class orders in any hierarchy. Table 4 shows the "sex of the voice" and the "quantity of the voice" modulating to connect the relevant characteristics and the sequence of the application of these characteristics. If any of the ranked isolates "one voice," or "more than one voice," were missing this would make the faceting defective.

Vocal music	
For one voice	
	For child's voice. For boy's voice
For woman's voice	
	Soprano
	Mezzosoprano
	Contralto
For man's voice	
	Tenor
	Baritone
	Bass
For more than one voice	
Choirs	
	Children's choirs. Boys' choirs
	Women's choirs
	Men's choirs
	Mixed choirs
	Solos with chorus

Table 4. Canon of modulation

5.5 Canons for filiatory sequence

5.5.1 Canon of subordinate classes

The subordinate classes in a class in the chain as they occur should be followed by their superordinate class immediately to ensure that they are not separated or among the classes by any other classes. Table 5 shows that by purpose music can be categorized as "background" or "foreground" but it can't be subordinate to "media" facet. This example is adapted and modified from the illustration of Spiteri (2003).

Music	
[By purpose]	
Background music	(Level 1 division)
[By media]	
Film music	(Level 2 division)
Television music	(Level 2 division)
Foreground music	(Level 1 division)
[By media]	
Music shows	(Level 2 division)
Motion pictures	(Level 2 division)
[By function]	
Entertainment	
Recreation	

Table 5. An example of subordinate classes

5.5.2 Canon of coordinate classes

In a class, coordinate classes should not be separated from each other by any classes other than their own sub-classes. From the example given in Table 5 it is clear that entertainment and recreation should not appear between the same classes as background music and foreground music because the former two classes show different characteristics by function than the application of the facet by media. Hence, only the classes formed by the media facet can be retained between background music and foreground music.

As shown above, through the application of principles and canons the domain of music is broken down into logical divisions. This approach helps in evolving logical structures that are used in constructing faceted ontologies.

6.0 Facetization of Music Ontology

In building ontologies, Preito-Diaz (2003) points out that "a classification scheme must be able to express hierarchical relationships as well as relationships created to relate two or more concepts belonging to different hierarchies. Hierarchical relationships are based on the principle of subordination or inclusion and are typical in taxonomy." Taxonomies and hierarchical schemes are basically factual, declarative and enumerative. The semantics of the relationships between their entities remain largely implicit, preventing the possibility to perform complex deductive reasoning, including logical inferences. This could represent a limitation for scientific domains where non-intuitive relationships are required to support integration of complex datasets. In this context, the explicit representation of complex relationships and the ability to perform inferences, the core principles of the Semantic Web, become particularly critical (Stewart 2011, 160). Faceted ontologies are expected to yield a major fillip to bridge these gaps by mapping relationships and characteristics with their analytico-synthetic features.

In the instance of building a faceted ontology for music concepts, domain analysis helps to comprehend how music can be categorized by different characteristics. The process of facetizing the domain of music was determined in line with facetization postulates. With the aid of knowledge organization systems the *ad hoc* scheme can be revised to obtain a validated ontology. For the purpose of this study, music ontology is developed with these top classes: theory, persons, instruments, kinds, forms and works shown in Table 6.

[theory]	[person]	[instrument]	[kind]	[form]	[work]
elements	artist	air	classical	binary	first movement
time	rapper	trumpet	Greek	ternary	
harmony	:	clarinet		da capo	allegro
tonality	:	:	Baroque	:	presto
:	album	:	:	:	vivace
:	My World 2.0	string	:	paraphrase	:
themes		guitar	popular	musical parody	:
		violin	jazz		instrument
reminiscing	Femme Fatale	fiddle	cool		symphony
in love		:	:		dance
romantic	:	:	:		ballroom
:	:		:		waltzes

Table 6. Generating facets of the top classes of music

The faceted ontology model for music gives the relationships, which are domain-specific, and some of the attributes are shown in Figure 8. The advantages offered by faceted ontologies lie in the ability to build bottom-up domain facets rather than the more common process of starting at the top and creating *a priori* categories, then slotting items, classifying the items themselves and pulling out their most essential and persistent characteristics of hierarchies into facets (Stewart 2011).

Once the background knowledge is constructed, the next step is to address the application of facets to the ontology. An example is illustrated in Figure 9 for a document titled "Using Violin Music as Auditory Practice for People Rehabilitated from Aphasia in India." In faceted

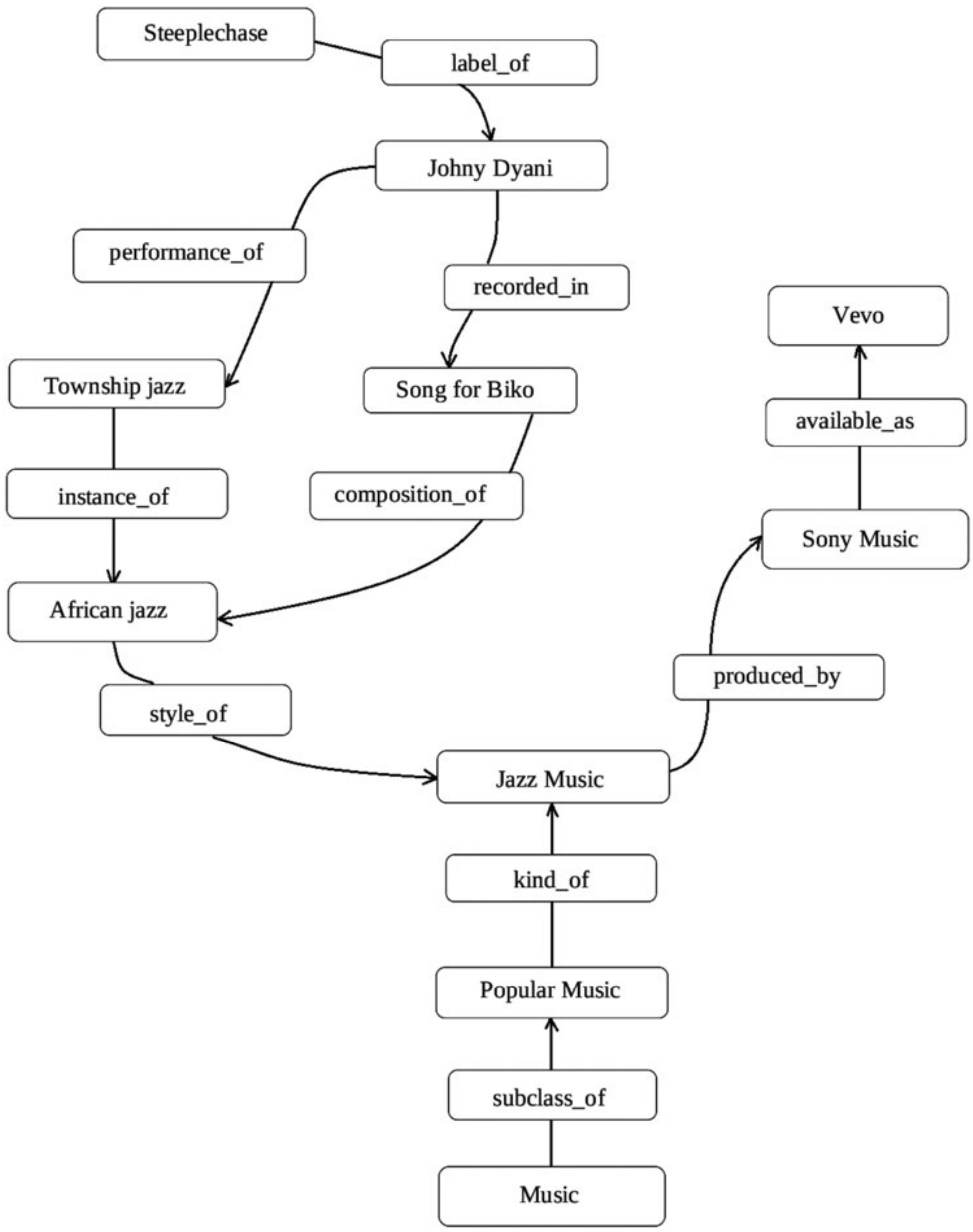

Figure 8. An example of music faceted ontology

ontology the more terms are employed, the more specific is the correspondence to the concept in each node. As a consequence, each term and corresponding concept occurring in its node labels must correspond to a term in the background knowledge (See Figure 11 adapted from Giunchiglia et al. 2009).

7.0 Faceted searching

As faceted classification is essential for organizing the Web, faceted search is the key to display the organized knowledge on the user interface. Faceted searching is helpful to sort out the different items by their groups and websites of a number of companies and institutions, from e-commerce to libraries, have increasingly adopted a faceted approach to their user interfaces (Lemieux 2009). As Broughton (2011) points out, in the past five decades research on "facet" or "faceted" has increased in the number of publications per keyword from 33 in 1960-1970 to 687 in 2000-2011. The keyword search "music" in the Stanford Library catalogue at http://searchworks.stanford.edu/ shows the faceted results navigable by facets: access, format, author, publication year, topic, location, call number, organization as author, region, era, and language. Faceted search is incorporated with enterprise search products such as Endeca (see http://www.

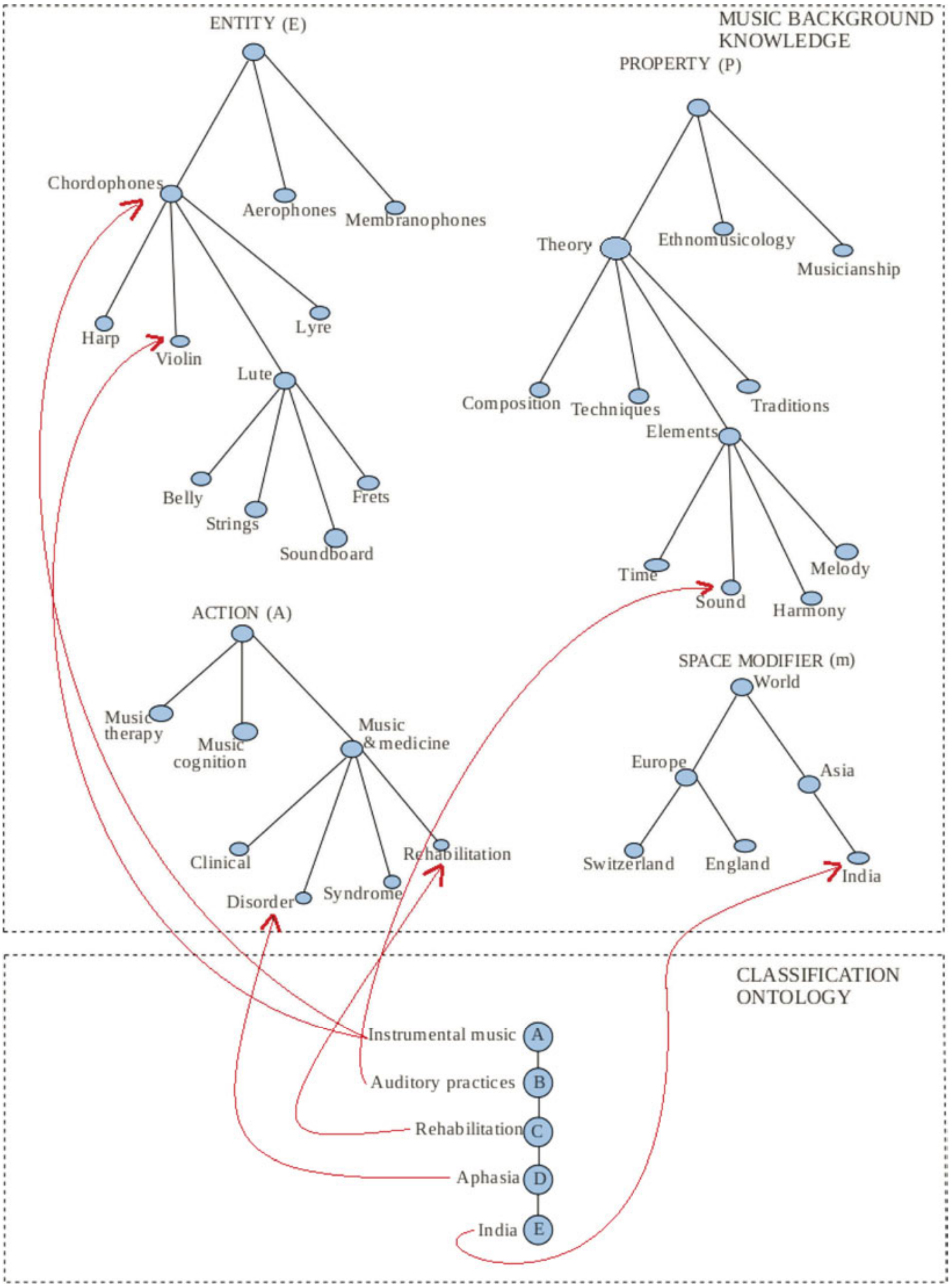

Figure 9. Example of faceted ontology analysis

oracle.com/us/corporate/acquisitions/endeca/index.html), and Solr (see http://lucene.apache.org/solr/). There are some examples of search tools deployed to increase the usability of digital resources, using faceted functionalities. User studies have shown that faceted searching provides more effective information-seeking support to users than conventional best first search. As Tunkelang (2009) highlights for faceted navigation, it is important to retain users on websites for resource discovery and value the numbers of clicks that lead to successful retrieval, and most importantly enable users to find what they are seeking.

7.1 Faceted query analysis

An example of faceted query analysis is explained here. Faceted query analysis takes into account the entity, action, action of entity etc. For instance "Using Violin Music as Auditory Practice for People Rehabilitated from Aphasia in India," is analyzed to identify the subject strings, followed by the mapping of domain knowledge wherein the properties and the relationships are explored and the subject of the query is analysed, as illustrated in Figure 10. Hence, the type of representation adapted would determine the result outcome, as the ontological representation would map different levels of relationships of documents, as illustrated in Figure 11.

When music can be modeled using the faceted metadata, schema and faceted ontologies (Tzitzikas 2002; Corthaut et al. 2008; Elhadad et al. 2010; Smith and Shadbolt 2012), their representation in standards like Dublin Core, FOAF and Simple Knowledge Organization System, can be explored as shown in Figure 12 adapted from National Library of Sweden's blog (National Library of Sweden 2008). As most of the present schematic ontologies, typically based on RDF/OWL/SKOS are subject representative, it requires extension of OWL or SKOS to incorporate analytico-synthetic features of faceted classification (Prasad and Madalli 2009; Slavic and Cordeiro 2004). However a discussion on formalization and pros and cons of existing languages is deemed as out of scope of the discussion here.

8.0 Conclusion

This paper presented a faceted approach to domain modeling as a step in ontology building. This is an effective method to strengthen schematic knowledge representation structures and the domain of music is presented as a use case. Faceted ontologies provide a powerful tool for organizing web-based knowledge using classificatory principles that were traditionally used in knowledge organization within libraries. The levels of expressivity and hospitability make it possible to combine terms in a flexible way and thus better represent semantically complex and/or compound subjects (Hong 2006). As the application of facet theory for organizing knowledge on web is gaining prominence, faceted ontologies have high potential for evolving domain models into formalized ontologies which could be more application-oriented for the real world. Vocabularies can be engineered with faceted ontologies, which provide multiple ways of looking at the domain modeling and thus expression of ontologies would be universally acceptable.

References

Abrahamsen, Knut Tore. 2003. "Indexing of musical genres: an epistemological perspective." *Knowledge Organization* 30: 144-69.

American Heritage Dictionary. 2011. "Music." http://ahdictionary.com/word/search.html?q=music+.

Aptagiri, Devika V., Malur A. Gopinath and A. R. D. Prasad. 1995. "A Frame Based Knowledge Representation Paradigm for Automating POPSI." *Knowledge Organization* 22: 162-67.

Austin, Derek. 1974. "Progress in Documentation. The Development of PRECIS: A Theoretical and Technical History." *Journal of Documentation* 30: 47-102.

Bhattacharya, Ganesh. 1979. "POPSI: Its Fundamentals and Procedure Based on a General Theory of Subject Indexing Languages." *SRELS Journal of Information Management* 16, no. 1: 1-34.

Bhattacharya, Ganesh. 1982. "POPSI: A Source Language for Organising and Associative Classifications."

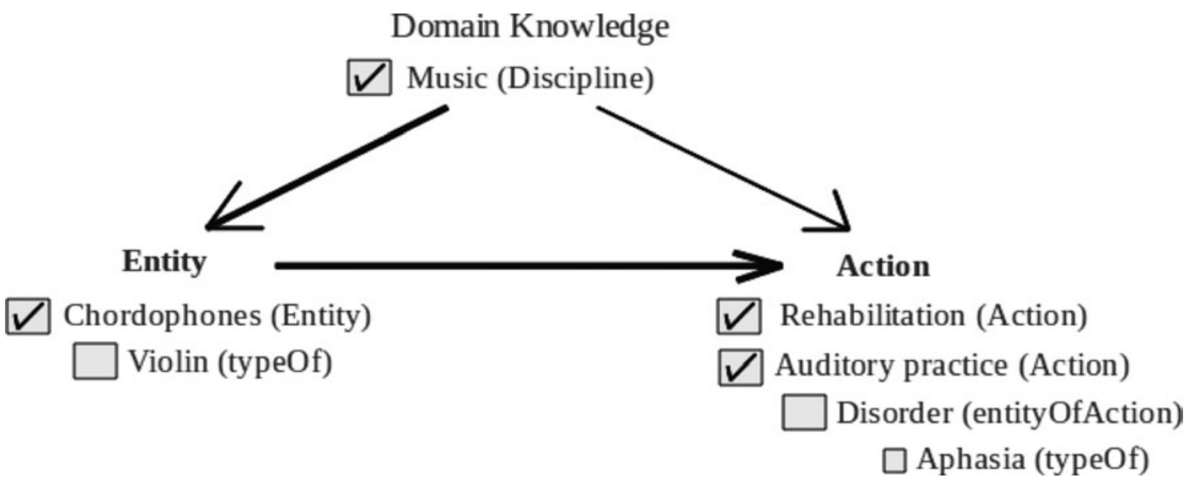

Figure 10. Query analysis of a music document

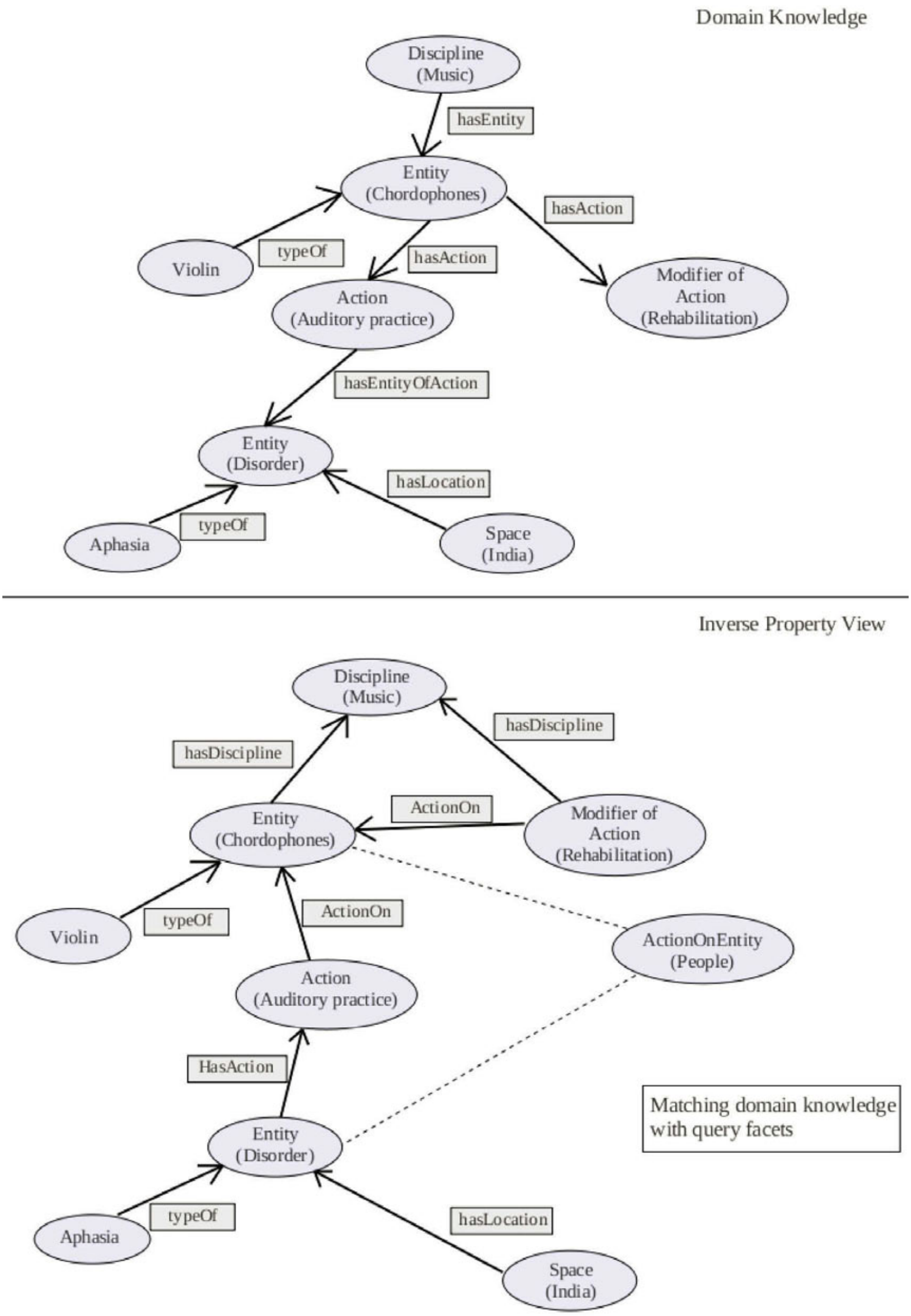

Figure 11. Faceted query analysis

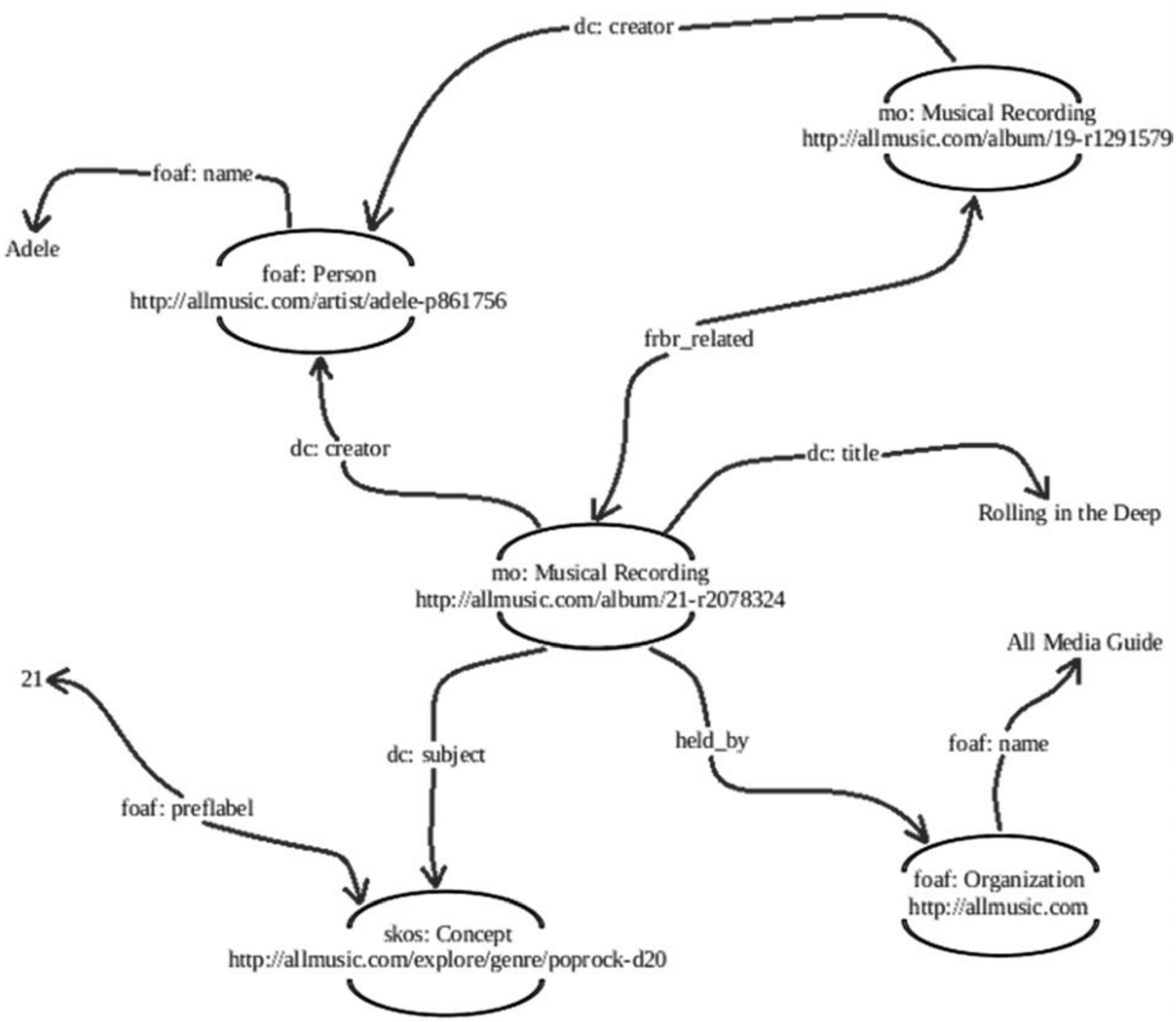

Figure 12. SKOS representation of a music recording

SRELS Journal of Information Management 19, no. 4: 240-66.

Biswas, Subal C. 1998. "Subject Heading by POPSI." In *Subject Indexing Systems: Concepts, Methods and Techniques,* edited by S. B. Ghosh and J. N. Satpathi, 188-210. Calcutta: IASLIC.

Broughton, Vanda. 2011. "Brian Vickery and the Classification Research Group: The Legacy of Faceted Classification." In *Facets of Knowledge Organization: Proceedings of the Second National ISKO UK Conference 4-5 July 2011 London, UK,* edited by A. Gilchrist. Bingley: Emerald, pp. 315-26.

Corthaut, Nik, Sten Govaerts, Katrien Verbert and Erik Duval. 2008. "Connecting the Dots: Music Metadata Generation, Schemas and Applications." http://ismir2008.ismir.net/papers/ISMIR2008_213.pdf.

Coult, Graham. 2010. "Metadata and Ontologies to Facilitate Accessibility." *Managing Information* 17, nos. 9/10: 100.

Elhadad, Michael, David Gabay and Yael Netzer. 2010. "Automatic Evaluation of Search Ontologies in the Entertainment Domain Using Text." http://www.cs.bgu.ac.il/~elhadad/papers/OntologyEvaluation.pdf.

Fensel, Dieter, Holger Lausen, Alex Polleres, Jose de Bruijn, Michael Stollberg, Dumitru Roman and John Domingue. 2007. *Enabling Semantic Web Services: The Web Service Modeling Ontology*. Berlin: Springer.

Foskett, A. C. 1982. *The Subject Approach to Information.* London: Clive Bingley Ltd.

Giunchiglia, Fausto, Biswanath Dutta and Vincenzo Maltese. 2009. "Faceted Lightweight Ontologies." In *Conceptual Modeling: Foundations and Applications,* edited by

Alex Borgida, Vinay Chaudhri, Paolo Giorgini and Eric Yu. Berlin: Springer.

Gruber, Thomas R. 1995. "Toward Principles for the Design of Ontologies Used for Knowledge Sharing." *International Journal of Human-Computer Studies* 43: 907-28.

Hjørland, Birger. 2002. "Domain Analysis in Information Science: Eleven Approaches – Traditional As Well As Innovative." *Journal of Documentation* 58: 422-62.

Hodge, Gail. 2000. "Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files." http://old.diglib.org/pubs/dlf090/dlf090.pdf.

Hong, Mei. 2006. "Potential Usage of Faceted Classification in Internet Information Retrieval." *International Information Sciences* 12, no. 1: 43-51.

Leach, Robert. 1976. *Musical Thesaurus: A Dictionary of Musical Language*. Oxford: Hannon & Co.

Lemieux, Stephanie. 2009. "Designing for Faceted Search." http://www.uie.com/articles/faceted_search/.

Library of Congress. 2010a. "Medium of Performance for Music: Working List of Terminology." http://loc.gov/catdir/cpso/medprf.html.

Library of Congress. 2010b. "Genre/form Headings for Musical Works." http://loc.gov/catdir/cpso/genremusic.html.

Library of Congress. 2012. "Performing Arts Encyclopedia." http://www.loc.gov/performingarts/.

Maedche, Alexander and Steffen Staab. 2001. "Ontology Learning for the Semantic Web." *IEEE Intelligent Systems* 16: 72-9.

National Library of Sweden. 2008. "LIBRIS available as Linked Data." http://librisbloggen.kb.se/2008/12/03/libris-available-as-linked-data/.

Porter, Joshua. 2003. "Testing the Three-Click Rule. User Interface Engineering." http://www.uie.com/articles/three_click_rule/.

Prasad, A. R. D. 2007. "Semantic Web Technologies for Meaningful Information Mapping and Retrieval." *IASLIC Bulletin* 52, no. 1: 25-34.

Prasad, A. R. D. 2008. "Guest Editorial." *Online Information Review* 32: 465-66.

Prasad, A. R. D. and Devika P. Madalli. 2008. "Faceted Infrastructure for Semantic Digital Libraries." *Library Review* 57: 225-34.

Prasad, A. R. D. and Devika P. Madalli. 2009. "Classificatory Ontologies." In *Classification at a Crossroads: Multiple Directions to Usability: Proceedings UDC Seminar 29-30 October 2009 Hague*, pp. 223-32.

Prieto-Diaz, Ruben. 2003. "A Faceted Approach to Building Ontologies." In *Proceedings of the 2003 IEEE International Conference on Information Reuse and Integration, 27-29 October 2003 Las Vegas*. New York: IEEE, 458-65.

Ranganathan, Shyali R. 1967. *Prolegomena to Library Classification. 3rd ed.* Bombay: Asia Publishing House.

Schedl, Markus, Tim Pohle, Peter Knees and Gerhard Widmer. 2011. "Exploring the Music Similarity Space on the Web." *ACM Transactions on Information Systems* 29: 14-24.

Slavic, Aida and Maria Inês Cordeiro. 2004. "Core Requirements for Automation of Analytico-Synthetic Classifications." http://cdigital.uv.mx/bitstream/123456789/6236/2/Clasificacion%20automatizacion%201.pdf.

Smiraglia, Richard P. 2001. "Musical Works as Information Retrieval Entities: Epistemological Perspective." In *Proceedings of the Second Annual International Symposium on Music Information Retrieval(ISMIR 2001), Bloomington, Indiana*, edited by Stephen Downie and David Bainbridge, pp. 85-92.

Smith, Daniel A. and Nigel R. Shadbolt. 2012. "Facetontology: Expressive Descriptions of Facets in the Semantic Web." http://eprints.soton.ac.uk/345363/1/paper.pdf.

Spiteri, Louise. 2003. "A Simplified Model for Facet Analysis." http://iainstitute.org/en/learn/research/a_simplified_model_for_facet_analysis.php.

Stewart, Darin L. 2011. *Building Enterprise Taxonomies. 2nd ed*. Lexington, KY: Mokita Press.

Taylor, Arlene G. 2004. *The Organization of Information. 2nd Ed*. Westport: Libraries Unlimited.

Tennis, Joseph T. 2003. "Two Axes of Domain for Domain Analysis." *Knowledge Organization* 30:191-95.

Tunkelang, Daniel. 2009. "Faceted Search." In *Synthesis Lectures on Information Concepts, Retrieval and Services*, edited by Gary Marchionini. San Francisco: Morgan & Claypool Publishers.

Tzitzikas, Yannis, Nicolas Spyratos, Panos Constantopoulos and Anastasia Analyti. 2002. "Extended Faceted Ontologies." In *CAiSE 2002, LNCS 2348,* edited by Banks Pidduck et al., 778-81. Berlin: Springer-Verlag Heidelberg.

Wikipedia. 2012. "History of Classical Music Traditions." http://en.wikipedia.org/wiki/History_of_classical_music_traditions.

Wallace, Danny P. 2007. *Knowledge Management: Historical and Cross-Disciplinary Themes*. Westport: Libraries Unlimited.

Wittgenstein, L. 1953. *Philosophical Investigations*. Oxford: Blackwell.