Smith, Priscilla Jane. Toward Linked Data: The Library of Congress Medium of Performance Thesaurus. A Master's paper for the M.S. in I. S. degree. April, 2013. 64 pages. Advisor: Jane Greenberg.

The Library of Congress Medium of Performance Thesaurus is a controlled vocabulary that will be used to describe music resources and is developed by the Music Library Association in conjunction with the Library of Congress. The objective of this study is to report on the current status of the LCMPT and to outline future steps which will allow the thesaurus to operate in a linked data environment. Four research questions examined the Simple Knowledge Organization System and its use with the LCMPT: the ways in which controlled vocabularies can be transformed into the Simple Knowledge Organization System format, the technologies and methods currently in use for this process, the effectiveness of these technologies and methods, and how the LCMPT can be most effectively converted into the Simple Knowledge Organization System format. Using existing documents, this qualitative study examined official World Wide Web Consortium documentation, the Open Metadata Registry, and was informed by work on the LCMPT performed at the Library of Congress. Tasks completed during an internship at the Library of Congress included discussion and planning for further LCMPT development and progress, development of the LCMPT Authority Data Elements, consideration of existing technologies and how they might be used in the development and deployment of LCMPT, conversion of existing LCMPT documentation into a more accessible format, research and documentation of the hierarchy of LCMPT, deployment of the LCMPT to a vocabulary space and element set space in the Open Metadata Registry, and consideration for future development of the LCMPT. The results of this study help to determine whether applying Semantic Web tools such as the Simple Knowledge Organization System to the LCMPT will allow music catalogers and other users of controlled vocabularies to begin harnessing the many potential benefits of the Semantic Web.

Headings:

Simple Knowledge Organization System

Linked Data

Library of Congress

Knowledge Organization System

Controlled Vocabulary

TOWARD LINKED DATA: THE LIBRARY OF CONGRESS MEDIUM OF PERFORMANCE THESAURUS

by
Priscilla Jane Smith

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Information Science

Chapel Hill, North Carolina
April, 2013

Approved by:

_____
Dr. Jane Greenberg

Table of Contents

1. Research Objectives

This master's paper reports on an initiative about how the Simple Knowledge Organization System (SKOS) can be used with controlled vocabularies, specifically the Library of Congress Medium of Performance Thesaurus (LCMPT). The answer to this examination can be studied in many different ways, and this Master's Paper was guided by the following specific questions:

1. In what ways are controlled vocabularies converted into the Simple Knowledge Organization System?
2. What technologies and methods are currently being used to convert controlled vocabularies into Simple Knowledge Organization System?
3. Which of these technologies and methods are the most effective, for the implementers and the users?
4. What is the most effective way to convert the Library of Congress Medium of Performance Thesaurus into a Simple Knowledge Organization System format?

2. Background--Simple Knowledge Organization System

Throughout history, classification systems have been widely used in the library community (Abbas 2010). Within a Web context, several formats have been proposed for representing thesauri using XML and RDF, including the DESIRE project in 1997 and the Limber project in 2001 (Miles et al. 2005). The SKOS-Core 1.0 Guide was first introduced in 2001 by the World Wide Web Consortium (W3C) Semantic Web Deployment Working Group (SWDWG) in order to develop SKOS as a W3C standardized classification system. The title 'Simple Knowledge Organization System' was chosen to emphasize the goal of "providing a simple yet powerful framework for expressing knowledge structures in a machine-readable way," and because the scope of the system extended beyond thesauri to other types of Knowledge Organization System (KOS) such as classification schemes, subject heading lists, taxonomies, and controlled vocabularies (Miles et al. 2005).

The W3C SWDWG currently maintains several pieces of documentation on SKOS which are freely available on the Web. First, the *SKOS Reference* document, which is currently at the final W3C recommendation or standard stage, defines SKOS. Second, the *SKOS Primer* document provides a guide for users of the system. Third, the *SKOS Use Cases and Requirements* document presents a list of representative use cases and a set of requirements derived from these use cases. Both the *SKOS Primer* and the *SKOS Use Cases and Requirements* documents are at the W3C Working Group Note

stage. The SWDWG also maintains an open mailing list and a wiki through which the public may contribute to the development of SKOS.

The Simple Knowledge Organization System is a common data model for knowledge organization systems such as thesauri, classification schemes, subject heading systems and taxonomies. Using SKOS, a knowledge organization system, or controlled vocabulary, can be expressed as machine-readable data and can then be exchanged between computer application software and published in a machine-readable format on the Web. According to the *SKOS Reference*, the aims of the system are:

- "To provide a bridge between different communities of practice within the library and information sciences involved in the design and application of knowledge organization systems."
- "To provide a bridge between these communities and the Semantic Web by transferring existing models of knowledge organization to the Semantic Web technology context, and by providing a low-cost migration path for porting existing knowledge organization systems to RDF."

2.1 Official SKOS Documentation

The W3C SWDWG currently maintains three main pieces of documentation on SKOS: the *SKOS Reference* document, the *SKOS Primer*, and the *SKOS Use Cases and Requirements* document.

2.1.1. SKOS Reference

The *SKOS Reference* outlines the infrastructure required in order to create and publish a SKOS document. This infrastructure includes constructs such as concepts, labels for those concepts, and relations among them. The document also identifies

integrity conditions which a SKOS document must follow in order to be logically and semantically correct.

A SKOS concept is any 'unit of thought': an idea, an object, an event. These concepts are the building blocks of many knowledge organization systems. Because concepts are abstract ideas that exist in the mind, they are independent of the terms used to describe them. For example, the English word "Banjo" which we use to describe a stringed instrument with a piece of animal skin or plastic stretched over a circular frame is actually independent of the concept of a banjo. The idea of a concept and its descriptor (or label) being two separate entities is vital to the SKOS model. The SKOS 'concept' element allows a vocabulary builder, or architect, to describe and distinguish concepts and their descriptors (labels). A SKOS concept can be created in two steps. First, the vocabulary architect must create or reuse a URI to uniquely identify a concept. Second, the vocabulary architect must assert, in RDF notation using the property `rdf:type`, that the resource identified by this URI is of type `skos:Concept`.

In SKOS, a label is the element which is the descriptor of a concept. The three SKOS label elements, `skos:prefLabel`, `skos:altLabel`, and `skos:hiddenLabel` are sub-properties of the RDFS element `rdfs:label`. The purpose of these three elements is to link a `skos:Concept` to an RDF plain literal, or character string. Preferred Label is a SKOS element that makes it possible to assign a preferred label to a concept. For example, the preferred label for the concept accordion is the word "Accordion" in English and "Accordéon" in French. For information retrieval and organizational purposes, the *SKOS Reference* states that no two concepts in the same KOS should be given the same preferred label for any given language tag. Alternate Label makes it

possible to assign an alternative label to a concept. This label allows multiple same-language descriptors for a concept to be stored. For example, the preferred label for the concept Alto voice is the word “Alto voice” and an alternate label is the word “Alto singer.” Hidden label is a label for a resource that a KOS designer would like to be accessible to applications performing text-based indexing and search operations, but would not like to be visible otherwise. For example, the preferred label for the concept “Saxophone” is the word “Saxophone,” and two hidden labels for the concept are “Horn” and “Sax.”

The two SKOS labels skos:broader and skos:narrower assert hierarchical relationships between concepts; that one concept is broader or narrower in meaning than another. It is important to note that SKOS anticipates hierarchical problems by not defining skos:broader and skos:narrower as generally transitive properties. In other words, the semantics of the system do not support transitive inferences of this type: "if Chordophone is broader than Harp and Harp is broader than Aeolian harp, then Chordophone is assumed to be broader than Aeolian harp." It may seem logical and convenient for these types of properties to be automatically assigned to SKOS concepts. However, these types of properties may cause unexpected problems in the architecture of non-traditional or poorly designed vocabularies that are to be represented using SKOS, so the designers of the system made a conscious decision against the assumption of hierarchical relations between concepts being implied by the statement of other stated hierarchical relations. It is, however, possible to make this type of relationship by using `skos:broaderTransitive` and `skos:narrowerTransitive`.

The SKOS label `skos:related` allows a designer to assert an associative relationship between two concepts. In the SKOS data model, `skos:related` is not defined as a transitive property, and the transitive closure of `skos:broader` must be disjoint from `skos:related`. If the concepts Viola and Violin are related via `skos:related`, there must not be a chain of `skos:broaderrelationship` from Viola to Violin. In other words, concepts must not be related in both a hierarchical and an associative manner at the same time.

SKOS semantic relations are connections between SKOS concepts. This type of relation occurs when a link between two concepts is inherent in the meaning of the linked concepts. Each of the SKOS labels `skos:broader`, `skos:narrower`, `skos:broaderTransitive`, `skos:narrowerTransitive` and `skos:related` are sub-properties of `skos:semanticRelation`.

Like `skos:broader` and `skos:narrower`, the two SKOS labels `skos:broaderTransitive` and `skos:narrowerTransitive` assert hierarchical relationships between concepts, that one concept is broader or narrower in meaning than another. The transitive nature of these two labels means that statements like: "if Chordophone is broader than Harp and Harp is broader than Aeolian harp, then Chordophone is assumed to be broader than Aeolian harp" are possible in the SKOS data model.

The SKOS mapping labels are used to state mapping, or alignment, connections between SKOS concepts that exist in different concept schemes. The SKOS concept collections labels are used to describe labeled or ordered groups of SKOS concepts. For example, the Library of Congress Medium of Performance Thesaurus can be considered a

collection of SKOS concepts because it is a group of concepts that have something in common.

The SKOS label `skos:note` was created for general documentation purposes. There is a hierarchical link between `skos:note` and its different specializations which allows all the documentation associated with a concept to be retrieved in a straightforward way. The label `skos:scopeNote` supplies some (possibly partial) information about the intended meaning of a concept. It is usually used as an indication of how the use of a concept is limited in indexing practice. The label `skos:definition` supplies a complete explanation of the intended meaning of a concept. The label `skos:example` supplies an example of the use of a concept. The label `skos:historyNote` describes significant changes to the meaning or form of a concept. The label `skos:editorialNote` supplies information that is an administrative aid, for example reminders of editorial work still to be done or notifications that future editorial changes might be made. The label `skos:changeNote` documents fine-grained changes to a concept for the purposes of administration and maintenance.

The SKOS Reference document includes several integrity conditions. These integrity conditions are statements that help to determine whether or not given data (for example, a vocabulary) are consistent with respect to the SKOS data model. The purpose of the SKOS integrity conditions are to encourage the construction of well-formed and consistent data and to promote interoperability between data represented in SKOS.

- **`skos:ConceptScheme` is disjoint with `skos:Concept`**

This condition states that SKOS concept schemes, or groups of SKOS concepts, must not be on the same hierarchical level as SKOS concepts and vice versa. For example, in the

Library of Congress Medium of Performance Thesaurus, ‘medium of performance’ is a `skos:ConceptScheme`. This means that ‘medium of performance’ must not also be a `skos:Concept` and one of the concepts, for example Violin, may not be a `skos:ConceptScheme`.

- **`skos:prefLabel`, `skos:altLabel` and `skos:hiddenLabel` are pairwise disjoint properties**

This condition states that no SKOS concept may be a member of more than one of preferred label, alternate label and hidden label.

- **A resource has no more than one value of `skos:prefLabel` per language tag**

This condition states that no SKOS concept may have more than one preferred label for each language tag. For example, the concept Bass voice has the preferred label of “Bass voice” in English and of “Basse” in French, and there may not be any other preferred labels in English or French.

- **`skos:related` is disjoint with the property `skos:broaderTransitive`**

This condition states that no two SKOS concepts may be connected by both related and broader transitive relationships.

- **skos:Collection is disjoint with each of skos:Concept and skos:ConceptScheme**

This condition states that SKOS collections, or labeled or ordered groups of SKOS concepts, must not be on the same hierarchical level as SKOS concepts and vice versa. For example, in the LCMPT, Woodwind instruments could possibly be a `skos:Collection`. This means that Woodwind instruments must not also be a `skos:Concept` and one of the concepts, like Oboe, may not be a `skos:Collection`.

- **skos:exactMatch is disjoint with each of the properties skos:broadMatch and skos:relatedMatch**

This condition states that no two SKOS concepts may related by more than one of exact match, broader match, and related match.

2.1.2 SKOS Primer

The *SKOS Primer*, like the *SKOS Reference*, is intended for users who have some previous experience with RDF. The document aims to assist in the representation and publication of SKOS concept schemes, and to provide further examples and guidance in the use of the SKOS vocabulary. The document contains four main sections: *SKOS Essentials*, *Networking Knowledge Organization Systems on the Semantic Web*, *Advanced SKOS: When KOSs are not Simple Anymore,* and *Combining SKOS with other Modeling Approaches*.

The section titled *SKOS Essentials* highlights much of the same content as the SKOS Reference document, and introduces the core concepts of the SKOS vocabulary including labels, semantic relationships, notes, and concept schemes. The section titled *Networking KOSs on the Semantic Web* discusses features of SKOS which enable the interlinking of concept schemes and how to relate conceptual resources in SKOS with other resources on the semantic web. The section titled *Advanced SKOS* discusses elements of the SKOS model which may be used to represent more complex knowledge organization systems. These include advanced documentation, coordination among concepts, relationships between labels of concepts, and accessing transitive hierarchical relationships. Finally, the section titled *Combining SKOS with Other Modeling*

Approaches discusses the re-use of SKOS labeling properties for the description of resources that might not be SKOS concepts.

2.1.3 SKOS Use Cases and Requirements

The *SKOS Use Cases and Requirements* document discusses the results of a questionnaire, issued in 2006, which was a call for representative use cases of the SKOS model. The document discusses eight separate use cases which were determined to be representative of the various uses of the SKOS model and vocabulary. The document also provides requirements which were determined after the examination of the eight use cases. These requirements include the representation of relationships between concepts, the extension of concept schemes, and the mapping of links between concepts from different concept schemes.

3. Introduction

There is an important distinction between controlled vocabularies which have been published to the Web and those which have been published in a structured way specifically for the Web. Strictly natural-language vocabularies are extremely useful for humans, but the amount of meaning that can be derived from these vocabularies by machines is very limited. The work of publishing controlled vocabularies for the Web in a way that both humans and machines can recognize meaning is being pursued by way of Linked Open Data and Linked Open Vocabularies, both Semantic Web technologies[1]. The creation of Linked Open Vocabularies allows for increased and more meaningful points of access and increased effectiveness in information retrieval.

There are several different models which can be used to structure and annotate controlled vocabularies for the web, and SKOS is a leading model. SKOS is an RDF/Semantic Web recommendation for encoding traditional controlled vocabulary data which provides a model for expressing the basic structure and content of concept schemes such as thesauri, classification schemes, subject heading lists, taxonomies, folksonomies, and other types of controlled vocabulary.

---

[1] Kaltenböck, M., & Bauer, F. (2012). *Linked Open Data: The Essentials*. Retrieved from http://www.semantic-web.at/LOD-TheEssentials.pdf

3.1 SKOS Tools

Research and development efforts on the conversion of controlled vocabularies to the SKOS format continue to grow, and a number of technologies and methods have been proposed. Although a manual conversion of a controlled vocabulary by a person or group of people is possible, it is time-consuming and likely prone to error. Tools have been developed to automatically convert vocabularies into the SKOS format. Additionally, tools to validate the quality of the formatting of a vocabulary in SKOS format have been developed[2]. Most recently, tools are being developed which can not only validate the SKOS format, but can improve the quality and validity of SKOS vocabularies[3].

[2] For example, the PoolParty online SKOS Consistency Checker. (http://poolparty.biz/)

[3] For example, Skosify is a tool that can be used to convert and improve vocabularies expressed as RDFS and OWL into SKOS format. (http://demo.seco.tkk.fi/skosify/skosify)

4. Research Methods

This is a qualitative study using existing documents. Materials used in this study include the official W3C SKOS documentation, results found by previous researchers of SKOS, and the evaluation of these results. This work has allowed me to gain a great knowledge of the SKOS framework. Additionally, knowledge taken from my work as an intern in the Library of Congress has informed my study. I participated in the production of the LCMPT, developing terms and the relationships between those terms and the greater hierarchy of the thesaurus, which will be published in the future. This work will allow me to gain a great understanding of the LCMPT thesaurus and its structure.

This descriptive study examines past implementations of the SKOS framework and uses the knowledge gained to apply the implementations to the Library of Congress Medium of Performance Thesaurus. A selection of terms and relationships from the LCMPT will be analyzed and several different approaches, informed by previous research, will be taken to convert the thesaurus into the SKOS format. Data generated from evaluating the results of my application of the previous techniques and tools to the LCMPT will be taken, and this data will be analyzed. The data, in the form of an effectiveness rating system, will be collected on the effectiveness of selected tools and techniques to convert LCMPT into the SKOS framework. The instruments used to measure the data will be informed by the official SKOS documentation and by the eleven SKOS validation criteria developed by Suominen and Hyvönen (2012). This testing

procedure will allow me to gain a great understanding of how the SKOS framework can be used for a music controlled vocabulary. The effectiveness rating system will be constructed based on two sets of SKOS vocabulary validation criteria: the eleven SKOS validation criteria developed by Suominen and Hyvönen (2012) and the six SKOS integrity conditions stated in the SKOS Reference. The use of these seventeen validation criteria to evaluate the different conversions of the LCMPT into SKOS format will allow for the comparison of the effectiveness of various conversions. This comparison will be possible by giving a rating to the final product of each conversion of the LCMPT. This rating will be calculated based on scores received in the following categories.

A selection of terms and relationships between those terms will be selected from the LCMPT, and the same selection of terms and relationships will be converted into the SKOS format using the selected conversion techniques. This will allow for accurate scoring or rating of the validation criteria.

The first six validation criteria are taken from the SKOS Reference document integrity conditions:

1. `skos:ConceptScheme` is disjoint with `skos:Concept`

Are concept schemes disjoint with concepts? For each concept scheme that is not disjoint with a concept, then the conversion will fail this integrity condition and will receive a lower score.

2. `skos:prefLabel`, `skos:altLabel` and `skos:hiddenLabel` are pairwise disjoint properties

Do concepts have the same labels in any of the pairwise-disjoint label properties `prefLabel`, `altLabel` or `hiddenLabel`? If concepts are repeated between these

pairwise-disjoint label properties, then the conversion will fail this integrity condition and will receive a lower score.

3. A resource has no more than one value of `skos:prefLabel` per language tag

Do resources have more than one preferred label values per language tag? If resources have more than one preferred label value per language tag, then the conversion will fail this integrity condition and will receive a lower score.

4. `skos:related` is disjoint with the property `skos:broaderTransitive`

Are concepts mapped together by both related and `broaderTransitive` relations? If concepts are mapped together by both related and `broaderTransitive` relations, then the conversion will fail this integrity condition and will receive a lower score.

5. `skos:Collection` is disjoint with each of `skos:Concept` and `skos:ConceptScheme`

Do resources belong to more than one of the three pairwise-disjoint classes `Concept`, `Collection` or `ConceptScheme`? If resources belong to more than one of the classes `Concept`, `Collection` or `ConceptScheme`, then the conversion will fail this integrity condition and will receive a lower score.

6. `skos:exactMatch` is disjoint with each of the properties `skos:broadMatch` and `skos:relatedMatch`

Is there a consistent use of mapping properties? If concepts are mapped with more than one of the following `exactMatch`, `broadMatch` and `relatedMatch`, then the conversion will fail this integrity condition and will receive a lower score.

The following categories are taken from the Suominen and Hyvönen validation criteria:

7. Valid URIs

Are the resource URIs valid? Are they reachable via the Web? If the resource URIs are invalid, then the conversion will fail this integrity condition and will receive a lower score.

8. Missing language tags

Do resources include appropriate language tags? If language tags are absent or contain errors, then the conversion will fail this integrity condition and will receive a lower score.

9. Missing labels

Do resources include human-readable labels? If labels are absent or contain errors, then the conversion will fail this integrity condition and will receive a lower score.

10. Loose concepts

Does the vocabulary contain loose concepts? If the vocabulary contains loose concepts, or concepts which are not top-level concepts that have no broader relationships pointing to other concepts, then the conversion will fail this integrity condition and will receive a lower score.

11. Ambiguous `prefLabel` Values

Are concepts limited to having only one `prefLabel`? If concepts contain more than one preferred label, then the conversion will fail this integrity condition and will receive a lower score.

12. Cycles in broader hierarchy

Does the vocabulary contain any cycles in broader hierarchies? If the vocabulary contains any cycles, or relationships between concepts which create loops in the logic of the vocabulary, then the conversion will fail this integrity condition and will receive a lower score.

13. Extra whitespace

Is there surrounding whitespace present in label property values? If label property values include whitespace, then the conversion will fail this integrity condition and will receive a lower score.

For each occurrence of failure of one of the thirteen validation criteria, a vocabulary conversion will accrue one point. The scores of each conversion will be compared for each validation criteria, and the total numbers of points each vocabulary conversion accrued will be compared to determine the effectiveness of each conversion technique.

5. Simple Knowledge Organization System literature review

5.1 Introduction

This literature review covers known SKOS conversion and validation techniques. Also included are numerous SKOS custom expansions and improvement techniques that have been implemented in the past. Finally, the literature review will discuss the new technique of vocabulary creation using SKOS, and the state of the field as it stands today.

5.2.1 Conversion techniques

In 2001, the Semantic Web Advanced Development for Europe (SWAD-E) published *Migrating Thesauri to the Semantic Web: Guidelines and Case Studies for Generating RDF Encodings of Existing Thesauri*, a thesaurus research prototype which presents guidelines and methods for migrating traditional thesaurus systems to RDF based thesaurus systems. The process described in this document consists of three stages. First, an RDF encoding of the thesaurus is generated. Second the encoding is taken through error checking and validation processes. Third, the encoding is published on the Web. During the first step, a traditional, or term-oriented, view of a thesaurus is converted to a concept-oriented view of a thesaurus. Thus, each ‘preferred term’ in a thesaurus becomes a ‘preferred label’ for a ‘concept.’ Each ‘preferred label’ is given a tag of `skos:prefLabel`, and each concept is given a tag of `skos:concept`. Each ‘concept’ in the thesaurus is given a unique URI which can be linked, through the Web, to the URIs for other related concepts. The technique of designating unique and persistent

URIs for all concepts in a vocabulary allows machines to understand the relationships between concepts similar to the way humans understand these relationships intrinsically.

Assem, Malaisé, and Miles (2006) expand on the technique used in step one of the SWAD-E document. This paper suggests three activities that will effectively link the term-oriented view of a thesaurus to its concept-oriented view. First, the digital format and the documentation of the thesaurus is be analyzed to determine the features of the thesaurus and how it is encoded. Second, a mapping between the thesaurus data items and the SKOS RDF is defined. Third, a transformation program, or algorithm, is created. The authors mention a sub-activity in which pre-existing URIs, if present, are identified. If no URIs exist in the term-oriented view of a thesaurus, the authors suggest the creation of randomly generated unique identifiers or the use of the name of the preferred term if it is unique in order to generate a unique URI.

Assem, Malaisé, and Miles (2006) apply their new technique to three existing thesauri: IPSV, GTAA and MeSH. These particular thesauri are chosen because of their popularity and progressive complexity. The authors find that conversion of the largest and most complex thesaurus, MeSH, is problematic but it does assist in identifying the boundaries of the applicability of their technique. The fact that MeSH contains textual notes which combine several types of knowledge, or compound concepts, lead the authors to discover that some thesauri have complex structures for which no direct SKOS counterpart exists. Additionally, IPSV and MeSH contain management information about their terms and this cannot be represented within the SKOS standard. The authors also mention that at the time of their study validation of SKOS RDF is difficult, due to the lack of validation technologies.

In 2008 Summers, Isaac, Redding, and Krech present a technique for converting Library of Congress Subject Headings (LCSH) to SKOS in their paper *LCSH, SKOS and Linked Data*. In their research, the authors use content in MARC bibliographic records for LCSH terms and mapped them to corresponding SKOS concepts. For example, the Library of Congress Control Number given to every LCSH are mapped onto skos:Concept and are used in the creation of unique URIs for each concept. Pre-coordinated LCSH terms, which have the potential to create problems in SKOS since they represent more than one term or concept, are simply flattened into one concept in this technique. In their technique, the authors write code using the Python programming language and use open-source MARCXML and RDF processing tools to create an object-oriented streaming interface to mint URIs and link the together. Additionally, the authors suggest that an extension of SKOS would allow the full meaning of these terms to be captured in SKOS form and that the integration of other Semantic Web vocabularies such as Dublin Core could allow SKOS vocabularies to be even more meaningful.

The earlier technique of SWAD-E was also adapted by Neubert, in his 2009 paper *Bringing the "Thesaurus for Economics" on to the Web of Linked Data*. Neubert found that the fact that SKOS includes built-in multilingual features were useful, given that the Thesaurus for Economics, or STW, is made up of both English and German terms. The author stated that the conversion of STW into SKOS was straightforward. However, at the time of his research, several new SKOS classes, including `skos:notation`, had been introduced by W3C. Neubert was able to take advantage of the fact that SKOS would now allow for internal management information notation, something that Assem, Malaisé, and Miles (2006) had claimed SKOS was lacking in 2006.

5.2.3 Validation techniques

In his 2009 paper, Neubert also discusses the use of SPARQL queries to check SKOS vocabularies for inconsistencies. The author describes the process of loading SKOS vocabularies into a SPARQL server and running inconsistency checks, where queries which check for illogical links between concepts would bring these inconsistencies to the attention of the implementer. Neubert also mentions that this process could potentially aid with thesaurus maintenance if inconsistency checks were to be performed routinely. The concept of automatic validation of SKOS vocabularies was new at the time of his research. However, the implementer inputting the SPARQL queries would only be able to locate inconsistencies that had previously been anticipated. Inconsistencies that are not queried would not be noticed by the SPARQL server.

In their 2012 paper *Improving the Quality of SKOS Vocabularies with Skosify* Suominen and Hyvönen propose a tool with the ability to check the quality of a SKOS vocabulary. The authors cite the lack of quality and validity of existing SKOS vocabularies as the reason for the development of this new tool, named the PoolParty online SKOS Consistency Checker, or PoolParty[4]. The researchers created a list of eleven validation criteria with which SKOS vocabularies will be checked, gathered from the W3C SKOS Reference document and through the authors' examination of fourteen freely-available SKOS vocabularies. These criteria are: the presence of valid URIs, the absence of language tags and labels, loose concepts, disjoint OWL classes, ambiguous `prefLabel` values, overlap in disjoint label properties, consistency in use of mapping

[4] http://poolparty.biz/

properties, disjoint semantic relations, cycles in the broader hierarchy, and the presence of extra whitespace. The authors found that all of the medium and large sized vocabularies failed at least one validation check, meaning that they did not meet some of the SKOS integrity constraints. Only one of the fourteen vocabularies passed all eleven validation criteria. They also found that the SKOS integrity constraint which specifies that the related relationship be disjoint with the broaderTransitive relationship was the most commonly violated.

5.2.4 Custom expansions

As seen in Assem, Malaisé, and Miles (2006) and Neubert (2009), controlled vocabularies often contain constructs which have no direct counterpart in SKOS. As stated in the *SKOS Primer,* SKOS is designed for simple extension of language constructs for the specialization of a particular vocabulary. This is made possible by expounding upon a SKOS construct using the extension `rdfs:subClassOf`. Neubert proposes two SKOS extensions for use with the STW thesaurus. First, the choice is made to split `skos:Concept` into two 'subclasses.' This change is chosen because along with the standard terms, or descriptors, STW includes a taxonomy of approximately 500 classes which are used to aid user information retrieval. Neubert chose to split `skos:Concept` into two parts: `zbwext:Descriptor rdfs:subClassOf skos:Concept` for terms and `zbwext:Thsys rdfs:subClassOf skos:Concept` for classes. This allows for well-defined semantics in terms of broader and narrower relationships between concepts and will allow users to search for both concepts and classes. Additionally, Neubert extended the newly introduced `skos:note` using the following notation: `zbwext:useInsteadNote rdfs:subClassOf skos:note` for notes which guide

users to designated preferred labels rather than other preferred labels in certain circumstances.

5.2.5 Improvement techniques

Along with their validation tool PoolParty, Suominen and Hyvönen (2012) introduce another SKOS technology. Skosify[5], a tool created to improve the quality and validity of SKOS vocabularies, is a command line tool which has the capability of reading one or more SKOS files and outputting a file in which errors and problems have not only been recognized, but corrected. Skosify is able to address nine of the eleven validation criteria mentioned earlier in their paper. After being used on fourteen vocabularies, Skosify corrects problems in all nine of these categories. The tool has the ability to correct missing language tags if a default language is provided, to detect unlabeled concept schemes and add specified labels, and to remove unnecessary whitespace surrounding property values. Skosify also has the ability to recognize and correct more sophisticated problems having to do with the relationships between concepts in vocabularies. The tool is able to identify top level concepts and add `hasTopConcept` and `topConceptOf` relationships to that concept scheme, recognize the designation of more than one `prefLabel` for one concept and correct the error, and identify and correct concepts which have been mislabeled as collections. Additionally, the tool can recognize when a concept is linked to a label using two different label properties and remove the less important property, recognize when concepts are linked together in a way that is disjoint and remove the related relationship assertion without disabling the broader

[5] http://demo.seco.tkk.fi/skosify/skosify

hierarchy, and recognize cycles, or concepts, which have a broader relationship with themselves and remove the offending relationship.

5.3 Vocabulary Creation with SKOS

In 2010, Gerbé and Kerhervé proposed a new approach to vocabulary creation with their paper *A Model-Driven Approach to SKOS Implementation*. This technique involves viewing the SKOS conceptual model as a metamodel for structured controlled vocabularies. Using a model management and model engineering approach, the authors state that flexible and extensible vocabularies can be managed and created. Model operators, or prompts used in model management, such as map, match, merge, and compose can be used to map term-oriented vocabularies onto a concept-oriented platform. Gerbé and Kerhervé introduced a metamodel and SKOS engine for use in the development of SKOS vocabularies. The metamodel is expressed in much the same way as a database schema, and is supported by the MySQL database system. The SKOS engine is built as a database with an interface for use in populating and visualizing the vocabulary content. The authors state that the tool also has the capability for importing, exporting, and merging vocabularies. This model-driven database approach to SKOS vocabulary creation takes the capabilities of tools for working with SKOS vocabularies to a new level.

5.4 State of the art

In 2012, Manaf et al. present a survey of the current state of Simple Knowledge Organization System (SKOS) vocabularies on the Web. Candidate vocabularies are gathered through collections and web crawling, with 478 vocabularies identified which comply with the given definition of a SKOS vocabulary. Analyses conducted include

investigation of the use of SKOS constructs, the use of SKOS semantic relations and lexical labels, and the structure of vocabularies in terms of the hierarchical and associative relations, branching factors, and depth of the vocabularies.

The steps used in the survey include preparing a candidate SKOS vocabulary corpus, identifying SKOS vocabularies, collecting survey data, filtering out multiple copies of the same SKOS vocabularies, and analyzing the corpus of vocabularies. The researchers collect the following data on each SKOS vocabulary: number of SKOS concepts, depth of each SKOS concept and depth of the concept hierarchy, number of links for `skos:broader`, `skos:narrower` and `skos:related` properties, total number of concepts not connected to any other concepts, total number of concepts with `skos:narrower` relations but no `skos:broader` relations, and maximum number of `skos:broader` properties. According to the researchers, SKOS concepts and concept labeling is core to SKOS vocabularies, but not all SKOS vocabularies in the study use SKOS lexical labels for their concepts. Approximately one-third of the SKOS vocabularies studied fall into the category of term lists, with no use of any SKOS semantic relations. The researchers find that not all SKOS vocabularies published explicitly declare SKOS concepts in the vocabularies. The survey results can serve to provide a better understanding of the modeling styles of the SKOS vocabularies published on the Web, especially when considering the creation of applications that utilize these vocabularies.

6. Background--Library of Congress Medium of Performance Thesaurus

The LCMPT is a joint Library of Congress and Music Library Association project that officially began when the MLA Bibliographic Control Committee Genre/Form Task Force (MLA-BCC-GFTF) was created in 2009. The initial work of this group was to work through compiled lists of music-related terms in LCSH, dividing them into two groups of terms; those which belong in a list of genre/form terms and those which belong in a list of medium of performance terms. In 2011, work on the medium of performance terms was transferred to the MLA Bibliographic Control Committee Subject Access Subcommittee (MLA-BCC SAS). This occurred because medium of performance terms are conceptually different from genre/form terms and the members of both the MLA-BCC SAS and MLA-BCC GFTF groups agreed that the groups of terms should remain separate for development purposes.

Currently, LCSH is used to retrieve music by its medium of performance. However, because the topical nature of LCSH, a more appropriate substitute is needed, and LCMPT is being developed to fill in that gap. The MLA-BCC SAS has stated the two bibliographic reasons for the creation of LCMPT:

- "To retrieve music by its medium of performance in library catalogs, as is now done by the controlled vocabulary, Library of Congress Subject Headings (LCSH)."

- "To record the element "medium of performance" of musical works, as represented in individual music resources cataloged according to Resource Description and Access (RDA)."

The intention is to use LCMPT as a source vocabulary for RDA's "medium of performance" element. Currently, RDA contains only a small representative selection of all of the terms that could be used for medium of performance.

Medium of performance is now recognized as a separate and distinct bibliographic facet that will have its own vocabulary and vocabulary structure. Searches by medium allow users to retrieve music by a particular instrument or instrumental group, voice or vocal group, through a medium term alone or in any combination with other medium terms the searcher may specify.

LCMPT currently includes more than 850 terms, a list that is growing. The thesaurus will contain many terms familiar to users of LCSH but there will be more terms than LCSH provides. The MLA-BCC Subject Access Subcommittee members have accepted about 80 newly proposed terms, and anyone interested in the project can propose new terms through the project Wiki.[6]

The LCMPT is made up of two separate entities: the list of terms, or concepts, which are used to describe music resources, makes up a concept scheme. The list of authority data elements, which allow for description of the terms, makes up the metadata schema.

---

[6] http://musicgenrepublicforum.pbworks.com/w/page/21942009/FrontPage

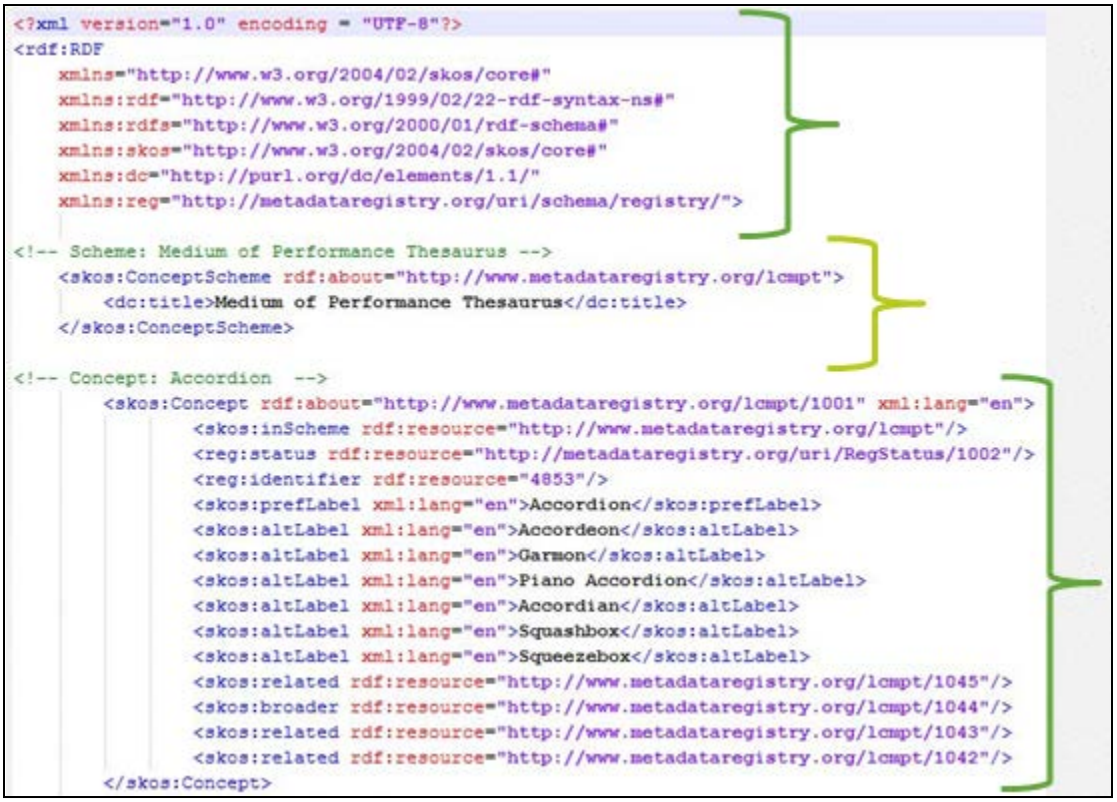
The LCMPT authority data elements are made up of a list of 30 authority data elements which will allow for the complete description of the LCMPT terms. This includes descriptive as well as administrative elements. Several of the descriptive elements, such as Broader term and Narrower term capture the hierarchical and associative relationships between terms.

#	Data element	MARC mapping	SKOS mapping
1	Source of term		skos:ConceptScheme
2	MARC record number	001	
3	Other record number	7XX (propose 762) $0	
4	Other controlled vocabulary designation	7XX (propose 762) $0	
5	Related LCCN	7XX (propose 762) $0	
6	Date record created	008/00-05	
7	Cataloger code (creator)	040 $a	
8	Date record modified	005	skos:changeNote
9	Cataloger code (modifier)	040 $d	skos:changeNote
10	Record status	Leader/05	skos:editorialNote
11	Kind of record	008/09	
12	Data type	008/07	
13	Encoding level of content	Leader/17	skos:editorialNote
14	Sponsoring agency code	040 $a	
15	Preferred term	1XX (propose 162)	skos:prefLabel
16	Variant of preferred term	4XX (propose 462)	skos:altLabel
17	Broader term	5XX (propose 562) with $w g	skos: broader
18	Related term	5XX (propose 562)	skos:related
19	Narrower term		skos:narrower
20	See also reference	15X/360	
21	Non-posted entry term	15X/260	
22	Node label	1XX (propose 162) with 008/09 coded as e	
23	Geographic code	043	
24	Associated place	370	
25	Time code	045	
26	Time period		
27	Reference source	670	
28	Scope & usage note	680	skos:scopeNote
29	Hierarchy location		
30	History note	688	skos:historyNote

6.1 Background – Open Metadata Registry

The Open Metadata Registry (OMR) provides technical infrastructure for the Semantic Web. The OMR was inspired by the Dublin Core Metadata Initiative and was originally built to support the National Science Digital Library. Now, the Registry is available publicly to all who wish to use its services. The Registry provides a means to

identify, declare and publish through registration concept schemes (controlled vocabularies, like the LCMPT terms), metadata schemas (element/property sets, like the LCMPT Authority Data Elements) and Application Profiles, which are sets of metadata elements that might be taken from multiple pre-existing element sets (like Dublin Core, Schema.org, etc.) that have been defined specifically for a particular application. When a user creates an account, they are able to designate themselves or other registered users as 'owners' of a project, allowing for collaboration.



```
<?xml version="1.0" encoding = "UTF-8"?>
<rdf:RDF
    xmlns="http://www.w3.org/2004/02/skos/core#"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
    xmlns:skos="http://www.w3.org/2004/02/skos/core#"
    xmlns:dc="http://purl.org/dc/elements/1.1/"
    xmlns:reg="http://metadataregistry.org/uri/schema/registry/">

<!-- Scheme: Medium of Performance Thesaurus -->
    <skos:ConceptScheme rdf:about="http://www.metadataregistry.org/lcmpt">
        <dc:title>Medium of Performance Thesaurus</dc:title>
    </skos:ConceptScheme>

<!-- Concept: Accordion  -->
        <skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1001" xml:lang="en">
                <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
                <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
                <reg:identifier rdf:resource="4853"/>
                <skos:prefLabel xml:lang="en">Accordion</skos:prefLabel>
                <skos:altLabel xml:lang="en">Accordeon</skos:altLabel>
                <skos:altLabel xml:lang="en">Garmon</skos:altLabel>
                <skos:altLabel xml:lang="en">Piano Accordion</skos:altLabel>
                <skos:altLabel xml:lang="en">Accordian</skos:altLabel>
                <skos:altLabel xml:lang="en">Squashbox</skos:altLabel>
                <skos:altLabel xml:lang="en">Squeezebox</skos:altLabel>
                <skos:related rdf:resource="http://www.metadataregistry.org/lcmpt/1045"/>
                <skos:broader rdf:resource="http://www.metadataregistry.org/lcmpt/1044"/>
                <skos:related rdf:resource="http://www.metadataregistry.org/lcmpt/1043"/>
                <skos:related rdf:resource="http://www.metadataregistry.org/lcmpt/1042"/>
        </skos:Concept>
```

Figure 1

The Open Metadata Registry allows for the download of an RDF markup file of an entire concept scheme. Figure 1 shows the XML declaration (in the top bracket) which describes the type of document that follows, the declaration of the Medium of Performance SKOS concept scheme (the middle bracket), and the very first term,

Accordion (the final bracket). As shown, the concept is described using RDF triples, and is associated with a permanent URI ending in 1001. The following line declares that the concept is a part of the LCMPT concept scheme, and the next following line states that the concept is at the "proposed" stage. The `reg:identifier` line identifies the URI for the term in the OMR namespace. The rest of the lines show the SKOS properties of the term which were inputted when the concept Accordion was created. All of the alternate labels are not associated with other terms, so they are simply reflected as text strings, but the related and broader terms are links to their appropriate URIs within the LCMPT namespace.

6.2 Library of Congress internship

From December 17, 2012 to January 4, 2013, I participated in an internship project in the Policy and Standards Division at the Library of Congress. Under the supervision of Senior Cataloging Policy Specialist Geraldine Ostrove, I had the opportunity to participate in the development of the Library of Congress Medium of Performance Thesaurus. The primary goal of my internship project was to establish a base file of LCMPT terms in a spreadsheet for use in future creation of SKOS, MARC or other formatted data, to examine LCSH authority records for data that would be appropriate for inclusion in LCMPT authority records and to maintain documentation of the progress made. Tasks that I completed included discussions and planning sessions, participation in the development of the LCMPT Authority Data Elements, consideration of the use of different technologies in conjunction with LCMPT, creation of a new data storage platform for LCMPT, metadata research and capture for a portion of the LCMPT terms, and a deployment of a portion of the LCMPT terms to an online metadata registry.

6.2.1 Discussion and planning

A portion of my time spent in the Policy and Standards Division was dedicated to discussion of the project and planning for further LCMPT development and progress. These discussions primarily included Geraldine Ostrove and me. Discussions included topics such as the construction of the hierarchy of the thesaurus, the benefits and drawbacks associated with linking LCSHs and LCCNs with LCMPT terms, and the process that would be required in order to research and create metadata for LCMPT terms. These conversations also allowed me to gain a greater understanding of the overall goals and scope of the thesaurus and its potential uses and users.

6.2.2 Authority data element development

A significant portion of my time was spent on the development of the LCMPT Authority Data Elements. I was able to research, in depth, the names, definitions, and descriptions of the data elements, and to participate in the revision of four draft versions of the LCMPT Authority Data Elements documentation.

6.2.3 Technologies and LCMPT

During my internship, I was able to consider many existing technologies and how they might be used in the development and deployment of the LCMPT. First, several possibilities for data storage were discussed, including spreadsheets, databases, and thesaurus software. Although the initial project proposal suggested that I use Excel spreadsheets to house LCMPT data, I made the suggestion that Google Docs spreadsheets might be an alternative, given the collaborative nature of the application.

One concern I initially had about using spreadsheets was that the nature of the data in LCMPT was not compatible with a spreadsheet format because some of the data elements are multi-valued. Multi-valued elements occur when a data element is

associated with multiple values. For example, the LCMPT term Clarinet has several narrower terms: Alto clarinet, Bass clarinet, Basset clarinet, and Basset horn. I suggested using a database to compensate for the incapability of spreadsheets. However, during a conversation with Dan Boomhower, he suggested that databases might not be the best option because the difficulties that may arise from pulling all data from a database would outweigh the perceived benefits. The decision was made to maintain the data, at least for the foreseeable future, in spreadsheets because of their ease of use and maintenance. At this point in time, we agreed that efforts should be made to seek out appropriate thesaurus software which would allow for automatic creation of RDF, XML, or HTML files that could be used to publish the thesaurus to the Web.

In addition, the use of MARC, RDA and SKOS were discussed. A complete mapping from LCMPT to MARC had already been completed previous to my time in PSD, but I made progress on creating a mapping to SKOS. The discussion of the use of MARC and RDA with LCMPT will most likely be ongoing, given the current transitory nature of the field of resource description.

Toward the end of my time in PSD, I was able to take time to meet with Kevin Ford, and to discuss the future of LCMPT and a possible deployment to the Web. We discussed several different options for thesaurus software, including TopBraid[7], Protégé[8], MultiTes[9], and TemaTres[10]. Ford suggested the use of TopBraid because of its ease of

---

[7] http://www.topquadrant.com/products/TB_Composer.html

[8] http://protege.stanford.edu/

[9] http://www.multites.com/

[10] http://www.vocabularyserver.com/

use and excellent user interface. We also discussed the possibility of migration of the LCMPT to http://id.loc.gov/ once it has been published. In addition, we discussed the use of the Extended Date/Time Format (EDTF)[11] in conjunction with LCMPT Authority Data Element 25, *time code*. The EDTF is a data format which is intended to extend the capabilities of ISO 8601(cite) by allowing for the description of approximate dates and times and other semantic qualifiers that may be used in describing temporal data.

6.2.4 Data storage platform

Another portion of the work I did while I was at LC was to convert the existing PDF document of the LCMPT into a more accessible format. I was able to transfer the existing LCMPT documentation into Google Docs format, using the spreadsheet document type. This spreadsheet, which is currently 'owned' by a newly created Google account dedicated to the LCMPT project[12], consists of two sheets, one which holds data for the LCMPT terms and associated metadata, and one which holds data for the LCMPT Data Authority Elements. The authority data elements sheet includes the names and descriptions of all elements along with a complete mapping to MARC and a partial mapping to SKOS.

6.2.5 Metadata capture

Another main task that I performed during my internship was the research and documentation of the hierarchy of LCMPT by way of authorized labels, alternate labels, broader terms, and related terms for LCMPT terms. Information was drawn from several sources: the LCSH MARC records found at the Library of Congress *Classification*

---

[11] http://www.loc.gov/standards/datetime/pre-submission.html

[12] mediumofperformancethesaurus@gmail.com

Web[13], entries in the *New Grove Dictionary of Musical Instruments* (1984) and entries in *Oxford Music Online*[14]. This work was completed by viewing the 150, 450, 550 and 670 fields of the MARC records of current LCSH terms to see what broader and narrower terms had previously been assigned to terms that were also in LCMPT.

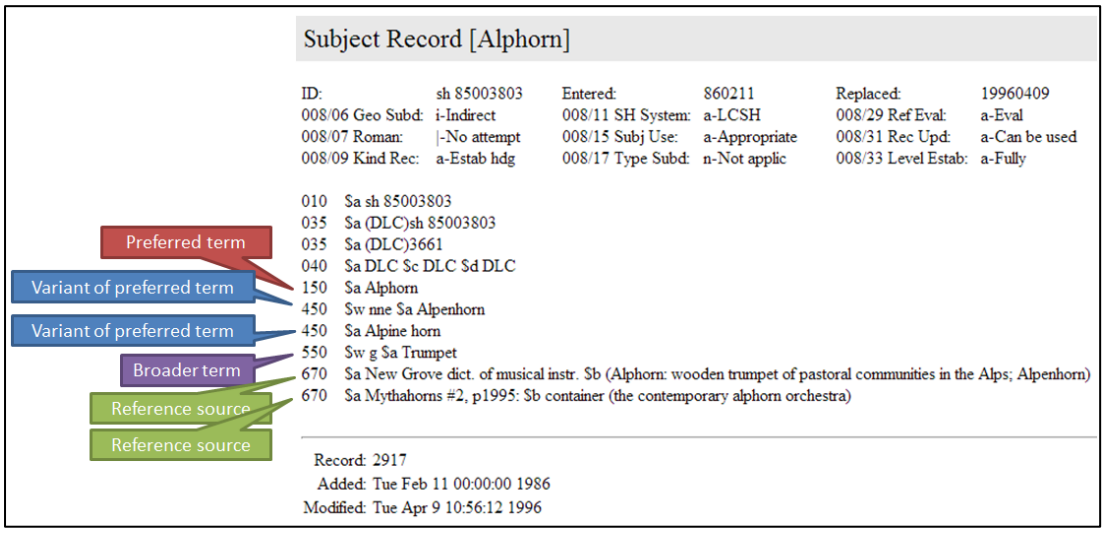


Figure 2

[13] https:/ /classificationweb.net

[14] http://www.oxfordmusiconline.com

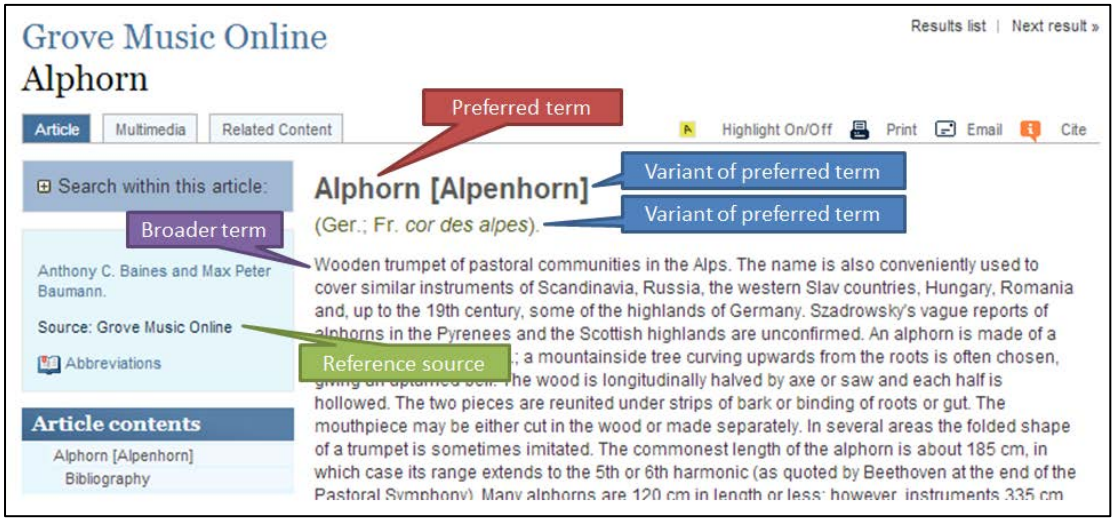


Figure 3

6.2.6 Deployment

I was able to take a portion of the data that I captured in the Google Docs spreadsheet and deploy it to a new vocabulary space[15] and element set space[16] in the Open Metadata Registry. Because the OMR only allows SKOS data elements to be recorded for its vocabularies, the OMR resources which represent LCMPT terms are incomplete, and do not have the capability of becoming complete unless the OMR allows for the use of its element sets with its vocabularies.

The LCMPT instance housed on the Open Metadata Registry currently contains the 135 LCMPT terms that I inputted. It also includes hierarchical links between these terms which were determined through my research of broader terms designated in LCSH.

The Open Metadata Registry 'home' page for the LCMPT concept scheme is freely available on the Web. By way of this home page, one can view information about

---

[15] http://metadataregistry.org/vocabulary/show/id/315.html

[16] http://metadataregistry.org/schema/show/id/70.html

the LCMPT instance on OMR: details of the vocabulary, a brows-able listing of all the concepts, history of the instance including versions, and a list of current maintainers of the instance. Under the *Detail* tab you can view general metadata about the vocabulary, for example owner of the instance, title, persistent URI, and description. Since the LCMPT is not yet a published thesaurus, the status is currently set to *proposed*. All of these attributes, besides the URI, are editable at any time. Under the Concepts tab is a list of 135 of the terms in the LCMPT concept scheme. On this page, users can click through to all of the terms and can also view the unique URIs, term status and the date last updated.

6.2.7 Issues for consideration

One of the most important considerations that must be kept in mind during the development of this thesaurus is that the balance between current practice and innovation is difficult to keep. The mappings to both MARC and SKOS that are present in the LCMPT Google Doc spreadsheet are a perfect example of this. The MLA-BCC Subject Access Subcommittee is concerned with LCMPT being interoperable in the immediate future in current MARC-based systems, and also once RDF/RDA becomes common practice and SKOS can be used.

Several of the LCMPT authority data elements, like *Preferred term*, *Variant of preferred term*, *Broader term* and *Narrower term* map directly to SKOS elements, making SKOS an excellent model for reflecting relationships among LCMPT terms. Many of the authority data elements, however, do not fit within the SKOS vocabulary, which means that only a handful of the 30 authority data elements for LCMPT can be stored in the SKOS data model.

LCMPT's default language is English. However, there is much valuable information about these concepts/terms in other languages. For retrieval purposes, I think it would be very valuable to include, at the very least, common foreign language names for LCMPT terms. The language of a property in the Open Metadata Registry may be chosen from an extensive list of languages, and SKOS handles this issue by using 2-letter language tags to designate the language of the text string of a property. One related issue that should be mentioned is the fact that many of the preferred terms in the LCMPT are foreign language terms that are used commonly in English as well. Should the preferred term automatically be designated in English, with an identical variant form in the original foreign language?

Because of the nature of the thesaurus, many LCMPT terms have broader terms that are at differing hierarchical levels. For example, the term "Bandora" has two broader terms, taken from LCSH and Oxford Music Online, which are not on the same hierarchical level: "Plucked instrument" and "Lute." For the sake of hierarchical consistency and term organization, it seems that the best solution would be that only broader terms that are one level above in the hierarchy should be designated.

There are many terms/concepts in LCMPT which do not currently have a broader term. Because one of the functional requirements of the thesaurus states that every term must have at least one broader term, these currently 'orphan' terms have been given a broader term of "Music BT needed." This new term is proposed, and it also speaks to the fact that the hierarchy of the thesaurus is still in development.

Currently, there is no concrete hierarchy to the LCMPT. Broader terms for 135 of the terms have been taken from LCSH or have been researched in Oxford Music Online.

The construction of this hierarchy can continue to be built from the ground up in this fashion, or top hierarchical terms could be designated and all narrower terms could be fit into the hierarchy below. The construction of the hierarchy of the LCMPT is a work in progress, and all options for hierarchy construction should be explored.

Various data storage platforms have been discussed for LCMPT. While the Open Metadata Registry allows terms to be linked to other terms through hierarchical and associative relationships, much of the other authority data element metadata is lost due to limited description properties. A system which allows for the linking of terms to other terms and the storage of complete data authority elements for all terms is the ultimate goal for LCMPT. Many thesaurus software exist, some which are proprietary, some which are free and open to the public.

Another one of the tasks I completed while at LC was to deploy the LCMPT Authority Data Elements to a metadata schema on the Open Metadata Registry. However, I was unable to discover a way to use the newly created metadata schema with the LCMPT concept scheme also on OMR.

During my time at LC, I contacted Diane Hillmann and Jon Phipps, the current maintainers of the Open Metadata Registry, regarding the linking of the LCMPT concept scheme and the Authority Data Elements metadata schema on OMR. Diane responded, stating that at that time, the linking of a concept scheme and a metadata schema is not yet possible on OMR.

With its complete metadata schema, the LCMPT cannot be stored in the Open Metadata Registry since the OMR application only allows for description of concepts in a

limited fashion, using SKOS properties. Another data storage solution needs to be agreed upon by those working on this project.

7. Conclusions and future project development

The objective of this study was to assess the current status of the Library of Congress Medium of Performance Thesaurus and to consider steps that would allow the thesaurus to operate in a linked data environment. The study informed the process of conversion of LCMPT to the Simple Knowledge Organization System by examining past and present cases of SKOS conversion. Four research questions examined SKOS and its use with the LCMPT, and worked to determine whether applying linked data tools such as SKOS to the LCMPT would allow music catalogers and other users of the LCMPT to make use of the benefits of the Semantic Web.

Using existing documents, this qualitative study examined official W3C documentation, the Open Metadata Registry, and was informed by work on the LCMPT performed at the Library of Congress. Tasks completed during my internship at the Library of Congress included the following:

1. Discussion and planning for further LCMPT development and progress.
2. Development of the LCMPT Authority Data Elements.
3. Consideration of existing technologies and how they might be used in the development and deployment of LCMPT.
4. Conversion of existing LCMPT documentation into a more accessible format.
5. Research and documentation of the hierarchy of LCMPT by way of authorized labels, alternate labels, broader terms, and related terms.

6. Deployment of the LCMPT to a vocabulary space and element set space in the Open Metadata Registry.
7. Consideration for future development of the LCMPT.

A proposed next step for this project might be to develop an application profile unique to this thesaurus. Developing an application profile, or a set of metadata elements that might be taken from multiple pre-existing element sets that have been defined specifically for LCMPT will allow for full description of the metadata associated with LCMPT terms, and will allow those metadata terms to be meaningful in a linked data environment. Standardization of the LCMPT Authority Data Elements metadata schema will allow LCMPT to be interoperable and linked with other applications and thesauri, and will allow catalogers and other users of the thesaurus to make full use of its content and structure.

References

Abbas, J. (2010). *Structures for organizing knowledge: exploring taxonomies, ontologies, and other schemas*. New York: Neal-Schuman Publishers, Inc.

Assem, M., Malaisé, V., Miles, A., & Schreiber, G. (2006). A method to convert thesauri to SKOS. *Lecture notes in Computer Science, 4011*, 95 – 109. doi: 10.1007/11762256_10

Baker, T., Schreiber, G., Swick, R., & Herman, I. (2012). W3C semantic web deployment working group development & participation. Retrieved from http://www.w3.org/2004/02/skos/development

Gerbé, O. & Kerhervé, B. (2010). A model-driven approach to SKOS implementation. Paper contributed to the International Conference on Internet and Web Applications and Services (5th) St. Maarten, AN. doi:10.1109/ICIW.2010.79

Iseminger, B. (2012). The music genre/form project: history, accomplishments, and future directions. In P. H. Lisius & R. Griscom (Eds.), Directions in music cataloging (pp. 63-77). Middleton: A-R Editions, Inc.

Isaac, A., Phipps, J., & Rubin, D. (2009). SKOS use cases and requirements. Retrieved from http://www.w3.org/TR/skos-ucr

Isaac, A. & Summers, E. (2009). SKOS simple knowledge organization system primer. Retrieved from http://www.w3.org/TR/skos-primer/

Kaltenböck, M., & Bauer, F. (2012). *Linked Open Data: The Essentials*. Retrieved from http://www.semantic-web.at/LOD-TheEssentials.pdf

Library of Congress Music Genre/Form Project Group. (2012). Medium of performance terms: draft authorized vocabulary for medium of performance statements in bibliographic records as agreed on by the library of congress and the music library association [press release]. Retrieved from http://www.loc.gov/catdir/cpso/medprf_lcmla.pdf

Manaf, N., Bechhofer, S., & Stevens, R. (2012). The current state of SKOS vocabularies on the web. Lecture notes in Computer Science, 7295, 270-284. doi:10.1007/978-3-642-30284- 8_25

Miles, A., Rogers, N., & Beckett, D. (2001). Migrating thesauri to the semantic web: guidelines and case studies for generating RDF encodings of existing thesauri. Retrieved from http://www.w3.org/2001/sw/Europe/reports/thes/8.8/

Miles, A., Rogers, N., & Beckett, D. (2001). SKOS-Core 1.0 guide: an RDF schema for thesauri and related knowledge organisation systems. Retrieved from http://www.w3.org/2001/sw/Europe/reports/thes/1.0/guide/20040504/

Miles, A., Matthews, B., Beckett, D., Brickley, D., Wilson, M., & Rogers, N. (2005). SKOS: a language to describe simple knowledge structures for the web. Paper

contributed to the XTech 2005: XML, the Web and Beyond conference (5th) Amsterdam, NL. Retrieved from http://epubs.cclrc.ac.uk/work-details?w=33893

Miles, A., Bechhofer, S. (2009). SKOS simple knowledge organization system reference. Retrieved from http://www.w3.org/TR/skos-reference

Neubert, J. (2009). Bringing the "Thesaurus for Economics" on to the web of linked data. Paper contributed to the WWW Workshop on Linked Data on the Web (2nd) Madrid, ES. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/similar;jsessionid=36C9F3DE0244DA85D5CFAD2 A744E771D?doi=10.1.1.184.2819&type=ab

(G. Ostrove, personal communication, June 5, 2012-February 2, 2013)

public-swd-wg@w3.org Mail Archives. (n. d.). Retrieved November 5, 2012, from Semantic Web Deployment Working Group Mailing List Archives: http://lists.w3.org/Archives/Public/public-swd-wg/

Simple Knowledge Organization System (SKOS). (n. d.). Retrieved November 4, 2012, from Semantic Web Deployment Working Group Wiki: http://www.w3.org/2001/sw/wiki/SKOS

Summers, E., Isaac, A., Redding, C., & Krech, D. (2008). *LCSH, SKOS, and linked data*. Paper contributed to the International Conference on Dublin Core and Metadata Applications (8th) Berlin, DE. Retrieved from http://arxiv.org.libproxy.lib.unc.edu/abs/0805.2855

Suominen, O. & Hyvönen, E. (2012) Improving the quality of SKOS vocabularies with Skosify. *Lecture notes in Computer Science*, 7603, 383-397. doi:10.1007/978-3-642-33876-2_34

Washington, M., Notess, M., & Dunn, J. (2010) *Taking music metadata from MARC to FRBR to RDF*. Paper contributed to the International Conference on Dublin Core and Metadata Applications (11th) The Hague, NL. doi:10.1109/ICIW.2010.79

Appendix A

Glossary of terms:

Content negotiation

Content negotiation refers to the practice of making available multiple representations via the same URI. Negotiation between the requesting agent and the server determines which representation is served (usually with the goal of serving the "best" representation a receiving agent can process). HTTP is an example of a protocol that enables representation providers to use content negotiation.

Controlled vocabulary

An established list of preferred terms from which a cataloger or indexer must select when assigning subject headings or descriptors in a bibliographic record to indicate the content of a work in the library catalog or in an index or bibliographic database.

EXtensible Markup Language (XML)

A subset of the SGML markup language in which the tags define the kind of information contained in a data element rather than how it is displayed. “Extensible” means that XML tags are not limited and predefined as they are in HTML–they must be created and defined through document analysis by the person producing the electronic document. Designed to meet the needs of large-scale electronic publishing, XML is a flexible text format that can be used with HTML in the same Web page. Document structure can be defined in a Document Type Definition (DTD) or XML Schema capable of handling document hierarchies. The most elaborate XML vocabularies have been developed to support business-to-business transactions.

Functional Requirements of Bibliographic Records (FRBR)

The principles espoused in the 1998 report of the IFLA Study Group on Function Requirements for Bibliographic Records titled *Function Requirements for Bibliographic Records*. Although the report covers the user-oriented functions that

bibliographic records should fulfill, and the data elements necessary to fulfill these functions, the term *FRBR* is usually used to refer to the *entity-relationship model* described in the report, which defines the characteristics of *works*, *expressions*, *manifestations*, and *items*.

Glossary

An alphabetical list of the specialized terms related to a specific subject or field of study, with brief definitions, often appearing at the end of a book or at the beginning of a long entry in a technical reference work.

Information Retrieval

The process, methods, and procedures used to selectively recall recorded information from a file of data. In libraries and archives, searches are typically for a known item of for information on a specific subject, and the file is usually a human-readable catalog or index, or a computer-based information storage and retrieval system, such as an online catalog or bibliographic database. In the design of such systems, a balance must be attained between speed, accuracy, cost, and effectiveness.

Library of Congress Control Number (LCCN)

When the Library of Congress began printing catalog cards in 1898 and distributing them in 1901, a unique Library of Congress Card Number was assigned to each item for identification and control. With the development of machine-readable cataloging in the late 1960s, LCCN became the Library of Congress Control Number. It is used in bibliographic records and also in authority and classification records. The LCCN is assigned to a publication after the deposit copy is received by the U.S. Copyright Office or in advance of the publication date if the publisher requests cataloging-in-publication.

Library of Congress Subject Heading (LCSH)

A descriptive word or phrase selected by a subject specialist at the Library of Congress from the list of Library of Congress Subject Headings and assigned to a

book or other item when first published to indicate its subject. Multiple subject headings are assigned when necessary or desirable. The complete list of subject headings is published annually in a multivolume set colloquially known as “the big red books,” usually available in the reference section of most large public and academic libraries and in the cataloging department of smaller libraries.

Linked data

The term Linked data refers to a set of best practices for publishing and connecting structured data on the Web. Linked data refers to data published on the Web in such a way that it is machine-readable, its meaning is explicitly defined, it is linked to other external data sets, and can in turn be linked to from external data sets. Linked data relies on documents containing data in RDF, and uses RDF to make typed statements that link arbitrary things in the world.

Linked Open Vocabularies (LOV)

Controlled vocabularies created using Semantic Web technologies that make up a part of the growing linked data Web. The best known of these vocabulary technologies (SKOS, Dublin Core, FRBR) now form a core of Semantic Web standards. These vocabularies are linked in that they rely more and more on each other through reusing, refining or extending, stating equivalences, declaring metadata. The objective of Linked Open Vocabularies is to provide easy access methods to this ‘ecosystem’ of vocabularies by defining the ways they link to each other and providing metrics on how they are used in the linked data cloud, to help to improve understanding, visibility and usability, and overall quality of LOVs.

MAchine-Readable Cataloging (MARC)

An international standard digital format for the description of bibliographic items developed by the Library of Congress during the 1960s to facilitate the creation and dissemination of computerized cataloging from library to library within the same country and between countries. By 1971 the MARC format had become the national

standard for dissemination of bibliographic data and by 1973 an international standard.

The MARC record had three components:

Record structure – an implementation of national and international standards

Content designation – codes and conventions that explicitly identify and characterize the data elements within a record to facilitate the manipulation of data.

Data content – defined by external standards such as AACR2, Library of Congress Subject Headings, etc.

Metadata

Structured information used to describe information resources/objects for a variety of purposes. Metadata can be categorized as descriptive, structural, and administrative. *Descriptive metadata* facilitates discovery, identification, and selection. *Structural metadata* describes the internal structure of complex objects. Administrative metadata aids in the management of resources and may include rights management metadata, preservation metadata, and technical metadata describing the physical characteristics of a resource.

Natural Language

A human language in which the structure and rules have evolved from usage, usually over an extended period of time, as opposed to an artificial language based on rules prescribed prior to its development and use, as in a computer language.

Resource Description and Access (RDA)

A standard for cataloging that provides instructions and guidelines on formulating data for resource description and discovery. RDA provides a set of guidelines and instructions on formulating data to support resource discovery, and provides a

comprehensive set of guidelines and instructions covering all types of content and media.

Resource Description Framework (RDF)

A language for representing information about resources on the World Wide Web. It is particularly intended for representing metadata about Web resources, such as the title, author, and modification date of a Web page, copyright and licensing information about a Web document, or the availability schedule for some shared resource. However, by generalizing the concept of a "Web resource", RDF can also be used to represent information about things that can be identified on the Web, even when they cannot be directly retrieved on the Web.

Semantic Web

The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries. It is a collaborative effort led by W3C with participation from a large number of researchers and industrial partners. It is based on the Resource Description Framework (RDF).

SPARQL Protocol and RDF Query Language (SPARQL)

An RDF query language able to retrieve and manipulate data stored in Resource Description Framework format. SPARQL can be used to express queries across diverse data sources, whether the data is stored natively as RDF or viewed as RDF via middleware. SPARQL contains capabilities for querying required and optional graph patterns along with their conjunctions and disjunctions. SPARQL also supports aggregation, sub-queries, negation, creating values by expressions, extensible value testing, and constraining queries by source RDF graph. The results of SPARQL queries can be result sets or RDF graphs.

Structured data

Data that resides in fixed fields within a record or file. Relational databases and spreadsheets are examples of structured data. Although data in XML files are not

fixed in location like traditional database records, they are nevertheless structured, because the data are tagged and can be accurately identified.

Taxonomy

The science of classification, including the general principles by which objects and phenomena are divided into classes, then into sub-classes, and so on. Taxonomies have traditionally been used in the life sciences to classify living organisms, but the term has been applied more recently within the information sector to the classification of resources available via the World Wide Web.

Thesaurus

A book of synonyms and near-synonyms in a written language, usually arranged conceptually, although dictionary arrangement is not uncommon. Also refers to an alphabetically arranged lexicon of terms comprising the specialized vocabulary of an academic discipline or field of study, showing the logical and semantic relations among terms, particularly a list of subject headings or descriptors used as preferred terms in indexing the literature of the field. In information retrieval, a thesaurus can be used to locate broader terms and related terms if the user wishes to expand retrieval or narrower terms to make a search statement more specific. A well-designed thesaurus also enables the indexer to maintain consistency in the assignment of indexing terms to documents.

Web Ontology Language (OWL)

The OWL Web Ontology Language is an ontology language for the Semantic Web. OWL ontologies provide classes, properties, individuals, and data values and are stored as Semantic Web documents. OWL ontologies can be used along with information written in RDF, and OWL ontologies themselves are primarily exchanged as RDF documents.

World Wide Web Consortium (W3C)

A nonprofit organization whose mission is to lead the Web to its full potential by developing technologies (standards, specifications, guidelines, software, tools) that will create a forum for information, commerce, inspiration, independent thought, and collective understanding. Its members include corporations, research institutions, government agencies, universities, libraries, and nonprofit organizations.

Sources:

Bizer, C., Heath, T., Berners-Lee, T. (2009). Linked Data - The Story So Far. Retrieved from http://tomheath.com/papers/bizer-heath-berners-lee-ijswis-linked-data.pdf

Jacobs, I., Walsh, N. (2004). Architecture of the World Wide Web, Volume One: Content negotiation. Retrieved from http://www.w3.org/TR/webarch/#def-coneg

Joint Steering Committee for Development of RDA. (2012). RDA: Resource Description and Access background. Retrieved from http://www.rda-jsc.org/rda.html#background

Manola, F., Miller, E. (Eds.). (2004). RDF Primer. Retrieved from http://www.w3.org/TR/rdf-primer/

RDF Working Group. (2004). Resource Description Framework (RDF) Overview. Retrieved from http://www.w3.org/RDF/

Reitz, Joan M. (2004). Dictionary for Library and Information Science. Westport: Libraries Unlimited.

The Computer Language Company Inc. (2012). Definition of structured data. Retrieved from http://www.pcmag.com/encyclopedia_term/0,2542,t=structured+data&i=52162,00.asp

Vatant, B., Vandenbussche, P. (2012). Linked Open Vocabularies - Big Picture. Retrieved from http://lov.okfn.org/dataset/lov/about/

W3C OWL Working Group. (2009). OWL 2 Web Ontology Language Document Overview. Retrieved from http://www.w3.org/TR/owl2-overview/

W3C Semantic Web group. (2012). What is the Semantic Web? Retrieved from http://www.w3.org/2001/sw/

Appendix B

XML Schema excerpt:

```
<rdf:RDF>
   <!-- Scheme: Medium of Performance Thesaurus -->
   <skos:ConceptScheme rdf:about="http://www.metadataregistry.org/lcmpt">
      <dc:title>Medium of Performance Thesaurus</dc:title>
   </skos:ConceptScheme>
   <!-- Concept: Accordion  -->
   <skos:Concept rdf:about=”http://www.metadataregistry.org/lcmpt/1001” xml:lang="en">
      <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
      <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
      <reg:identifier rdf:resource="4853"/>
      <skos:prefLabel xml:lang="en">Accordion</skos:prefLabel>
      <skos:altLabel xml:lang="en">Accordeon</skos:altLabel>
      <skos:altLabel xml:lang="en">Garmon</skos:altLabel>
      <skos:altLabel xml:lang="en">Piano Accordion</skos:altLabel>
      <skos:altLabel xml:lang="en">Accordian</skos:altLabel>
      <skos:altLabel xml:lang="en">Squashbox</skos:altLabel>
      <skos:altLabel xml:lang="en">Squeezebox</skos:altLabel>
      <skos:related rdf:resource="http://www.metadataregistry.org/lcmpt/1045"/>
      <skos:broader rdf:resource="http://www.metadataregistry.org/lcmpt/1044"/>
      <skos:related rdf:resource="http://www.metadataregistry.org/lcmpt/1043"/>
      <skos:related rdf:resource="http://www.metadataregistry.org/lcmpt/1042"/>
   </skos:Concept>
   <!-- Concept: Accordion band  -->
   <skos:Concept rdf:about=”http://www.metadataregistry.org/lcmpt/1002” xml:lang="en">
      <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
      <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
      <reg:identifier rdf:resource="4854"/>
      <skos:prefLabel xml:lang="en">Accordion band</skos:prefLabel>
      <skos:altLabel xml:lang="en">Accordion orchestra</skos:altLabel>
      <skos:broader rdf:resource="http://www.metadataregistry.org/lcmpt/1046"/>
   </skos:Concept>
   <!-- Concept: Acoustic bass guitar  -->
   <skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1003" xml:lang="en">
      <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
      <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
      <reg:identifier rdf:resource="4855"/>
      <skos:prefLabel xml:lang="en">Acoustic bass guitar</skos:prefLabel>
      <skos:broader rdf:resource="http://www.metadataregistry.org/lcmpt/1047"/>
      <skos:related rdf:resource="Bassoguitar"/>
   </skos:Concept>
   <!-- Concept: Actor  -->
   <skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1004" xml:lang="en">
```

```
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4856"/>
   <skos:prefLabel xml:lang="en">Actor</skos:prefLabel>
   <skos:related rdf:resource="Acting"/><skos:related rdf:resource="Theatre"/>
   <skos:broader rdf:resource="http://www.metadataregistry.org/lcmpt/1048"/>
</skos:Concept>
<!-- Concept: Aeolian harp  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1005" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4857"/>
   <skos:prefLabel xml:lang="en">Aeolian harp</skos:prefLabel>
   <skos:altLabel xml:lang="en">Eolian Harp</skos:altLabel>
   <skos:broader rdf:resource="http://www.metadataregistry.org/lcmpt/1049"/>
   <skos:broader rdf:resource="http://www.metadataregistry.org/lcmpt/1050"/>
</skos:Concept>
<!-- Concept: Aerophone  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1006" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4858"/>
   <skos:prefLabel xml:lang="en">Aerophone</skos:prefLabel>
   <skos:broader rdf:resource="http://www.metadataregistry.org/lcmpt/1117"/>
</skos:Concept>
<!-- Concept: Ajaeng  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1007" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4859"/>
   <skos:prefLabel xml:lang="en">Ajaeng</skos:prefLabel>
   <skos:altLabel xml:lang="en">A-jaeng</skos:altLabel>
   <skos:broader rdf:resource="http://www.metadataregistry.org/lcmpt/1051"/>
   <skos:related rdf:resource="Sanjo ajaeng"/>
</skos:Concept>
<!-- Concept: Alboka  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1008" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4860"/>
   <skos:prefLabel xml:lang="en">Alboka</skos:prefLabel>
   <skos:broader rdf:resource="http://www.metadataregistry.org/lcmpt/1052"/>
</skos:Concept>
<!-- Concept: Alesis HR-16 drum machine  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1009" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
```

```
   <reg:identifier rdf:resource="4861"/>
   <skos:prefLabel xml:lang="en">Alesis HR-16 drum machine</skos:prefLabel>
   <skos:broader rdf:resource="http://www.metadataregistry.org/lcmpt/1053"/>
</skos:Concept>
<!-- Concept: Alesis MMT-8 drum machine  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1010" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4862"/>
   <skos:prefLabel xml:lang="en">Alesis MMT-8 drum machine</skos:prefLabel>
   <skos:broader rdf:resource="http://www.metadataregistry.org/lcmpt/1053"/>
</skos:Concept>
<!-- Concept: Babadok  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1082" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4928"/>
   <skos:prefLabel xml:lang="en">Babadok</skos:prefLabel>
   <skos:altLabel xml:lang="en">Babadok drum</skos:altLabel>
   <skos:broader rdf:resource="http://www.metadataregistry.org/lcmpt/1066"/>
</skos:Concept>
<!-- Concept: Bagpipe  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1084" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4930"/>
   <skos:prefLabel xml:lang="en">Bagpipe</skos:prefLabel>
   <skos:altLabel xml:lang="en">Great pipe</skos:altLabel>
   <skos:broader rdf:resource="http://www.metadataregistry.org/lcmpt/1052"/>
   <skos:related rdf:resource="http://www.metadataregistry.org/lcmpt/1112"/>
</skos:Concept>
<!-- Concept: Bambuso sonoro   -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1091" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4937"/>
   <skos:prefLabel xml:lang="en">Bambuso sonoro </skos:prefLabel>
   <skos:altLabel xml:lang="en">Bamboo Organ</skos:altLabel>
   <skos:broader rdf:resource="http://www.metadataregistry.org/lcmpt/1068"/>
</skos:Concept>
<!-- Concept: Button-key accordion  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1151" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4987"/>
   <skos:prefLabel xml:lang="en">Button-key accordion</skos:prefLabel>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1136"/>
```

```
</skos:Concept>
<!-- Concept: Cello  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1071" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4922"/>
   <skos:prefLabel xml:lang="en">Cello</skos:prefLabel>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1031"/>
</skos:Concept>
<!-- Concept: Chordophone  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1050" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4901"/>
   <skos:prefLabel xml:lang="en">Chordophone</skos:prefLabel>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1005"/>
</skos:Concept>
<!-- Concept: Drum  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1066" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4917"/>
   <skos:prefLabel xml:lang="en">Drum</skos:prefLabel>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1026"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1033"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1082"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1124"/>
</skos:Concept>
<!-- Concept: Drum machine  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1053" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4904"/>
   <skos:prefLabel xml:lang="en">Drum machine</skos:prefLabel>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1010"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1009"/>
</skos:Concept>
<!-- Concept: Dulcimer  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1075" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4925"/>
   <skos:prefLabel xml:lang="en">Dulcimer</skos:prefLabel>
   <skos:related rdf:resource="http://www.metadataregistry.org/lcmpt/1025"/>
</skos:Concept>
<!-- Concept: Electronic organ  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1055" xml:lang="en">
```

```
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4906"/>
   <skos:prefLabel xml:lang="en">Electronic organ</skos:prefLabel>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1012"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1035"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1088"/>
</skos:Concept>
<!-- Concept: Euphonium  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1115" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4958"/>
   <skos:prefLabel xml:lang="en">Euphonium</skos:prefLabel>
   <skos:related rdf:resource="http://www.metadataregistry.org/lcmpt/1101"/>
   <skos:altLabel xml:lang="en">Tenor tuba</skos:altLabel>
</skos:Concept>
<!-- Concept: Flute  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1054" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4905"/>
   <skos:prefLabel xml:lang="en">Flute</skos:prefLabel>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1011"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1015"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1036"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1125"/>
</skos:Concept>
<!-- Concept: Guitar  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1047" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4898"/>
   <skos:prefLabel xml:lang="en">Guitar</skos:prefLabel>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1003"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1085"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1126"/>
</skos:Concept>
<!-- Concept: Harp  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1049" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4900"/>
   <skos:prefLabel xml:lang="en">Harp</skos:prefLabel>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1005"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1037"/>
</skos:Concept>
```

```
<!-- Concept: Music box  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1076" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4926"/>
   <skos:prefLabel xml:lang="en">Music box</skos:prefLabel>
</skos:Concept>
<!-- Concept: Music BT needed  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1117" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4959"/>
   <skos:prefLabel xml:lang="en">Music BT needed</skos:prefLabel>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1006"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1089"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1093"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1137"/>
</skos:Concept>
<!-- Concept: Woodwind instrument  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1052" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4903"/>
   <skos:prefLabel xml:lang="en">Woodwind instrument</skos:prefLabel>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1008"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1039"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1084"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1134"/>
</skos:Concept>
<!-- Concept: Xylophone  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1063" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4914"/>
   <skos:prefLabel xml:lang="en">Xylophone</skos:prefLabel>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1021"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1023"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1090"/>
</skos:Concept>
<!-- Concept: Zither  -->
<skos:Concept rdf:about="http://www.metadataregistry.org/lcmpt/1051" xml:lang="en">
   <skos:inScheme rdf:resource="http://www.metadataregistry.org/lcmpt"/>
   <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1002"/>
   <reg:identifier rdf:resource="4902"/>
   <skos:prefLabel xml:lang="en">Zither</skos:prefLabel>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1007"/>
   <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1025"/>
```

```
      <skos:narrower rdf:resource="http://www.metadataregistry.org/lcmpt/1040"/>
   </skos:Concept>
   <!-- Status properties used in this document  -->
   <skos:Concept rdf:about="http://metadataregistry.org/uri/RegStatus/1002">
      <skos:prefLabel xml:lang="en">New-Proposed</skos:prefLabel>
   </skos:Concept>
</rdf:RDF>
```