Je-LKS
PEER REVIEWED PAPERS
Vol. 12, n.4, 2016
ISSN: 1826-6223 | eISSN: 1971-8829
Journal of e-Learning and Knowledge Society
The Italian e-Learning Association Journal

LOCAL AND GLOBAL SEMANTIC NETWORKS FOR THE REPRESENTATION OF MUSIC INFORMATION

Adriano Baratè
Luca A. Ludovico

Laboratorio di Informatica Musicale (LIM)
Dipartimento di Informatica
Università degli Studi di Milano - Milan, Italy
barate@di.unimi.it, ludovico@di.unimi.it

Keywords: IEEE 1599, Multi-layer representation, Music, Semantic networks

In the field of music informatics, multilayer representation formats are becoming increasingly important, since they enable an integrated and synchronized representation of the various entities that describe a piece of music, from the digital encoding of score symbols to its typographic aspects and audio recordings. Often these formats are based on the eXtensible Markup Language (XML), that allows information embedding, hierarchical structuring and interconnection within a single document. Simultaneously, the advent of the so-called Semantic Web is leading to the transformation of the World Wide Web into an environment where documents are associated with data and metadata. XML is extensively used also in the Semantic Web, since this format supports not only human- but also machine-readable tags. On the one side the Semantic Web aims to create a set of automatically-detectable relationships among data, thus providing users with a number of non-trivial paths to navigate information in a geographically distributed

for citations:
Baratè A., Ludovico L.A. (2016), *Local and global Semantic Networks for the representation of music information,* Journal of e-Learning and Knowledge Society, v.12, n.4, 109-123.
ISSN: 1826-6223, e-ISSN:1971-8829

framework; on the other side, multilayer formats typically operate in a similar way, but at a "local" level. The goal of the present work is to discuss the possibilities emerging from a combined approach, namely by adopting multilayer formats in the Semantic Web, addressing in particular augmented-reality applications. An XML-based international standard known as IEEE 1599 will be employed to show a number of innovative applications in music.

1 Introduction

New digital media are influencing and changing social, economic and cultural lives of people. Key aspects of this phenomenon are the availability of network technologies and the increasingly pervasive presence of the Web in our lives. This work stems from the convergence of two current technological trends, the first in the field of multimedia information representation and the second in the field of advanced content experience and integration.

As it regards the former aspect, the goal is to provide rich multimedia information to the user, allowing for an integrated multi-layer and multi-modal navigation of content. Music information can provide a valuable testing ground, since a piece of music is composed by multiple and heterogeneous descriptions: catalog metadata, music symbols, graphical renderings, audio performances, structural relationships, etc. The problem of catching and describing heterogeneity in music has been traditionally addressed through a number of different file formats, each one targeting a specific facet of music information. The symbolic aspects are addressed by binary (e.g., MakeMusic Finale, MuseScore, Avid Sibelius, etc.), text-based (e.g., ABC, GUIDO, DARMS, etc.) and XML-based (e.g., IEEE 1599, MEI, MusicXML, etc.) notation formats; digital images can be saved in any of the available raster and vector formats; for audio there are many digital compressed and uncompressed formats; and so on. A problem emerging from the adoption of specific formats is the ability to relate different descriptions of the same logic entity: for instance, how to connect different writings – belonging to different score editions – of the same music symbols, how to connect different performances – belonging to different recordings – of the same aria, and finally how to synchronize the advancement of a cursor over a music score to a timed playback of a given audio track (score following). These problems can be solved through an integrated and synchronized approach to music description, as explained in Section 2. Even if not specifically designed for the purpose, the formats that currently implement this approach end up realizing a small semantic network, that will be defined as "local" in opposition to the "global" extent provided by the Semantic Web.

A semantic network is a network which represents semantic relations between concepts, often used in the field of knowledge representation. Formally it is a directed or undirected graph consisting of vertices, which represent concepts, and edges (Sowa, 2006). Early models of semantic network, intended as a way to represent structured knowledge, date back to the 1960s (Collins & Quillian, 1969; Collins, 1975; Brachman, 1977).

In the era of computing, the locution Semantic Web was coined by Tim Berners-Lee (Berners-Lee *et al.*, 2001), the inventor of the World Wide Web and director of the World Wide Web Consortium (W3C). The Semantic Web is meant to extend the network of hyperlinked human-readable Web pages by inserting machine-readable metadata about contents and how they are related to each other. As a consequence, ad-hoc automated agents can access the Web more intelligently and perform specific tasks on the behalf of users. In this context, linked data are a method of publishing structured data so that they can be interlinked and retrieved through semantic queries. Web technologies and standards are employed to share information that is automatically interpretable by computers, rather than provide contents readable by humans. This enables data from different sources to be connected and queried in the framework of the Semantic Web (Bizer *et al.*, 2009).

As it regards the structure of this work, in Section 2 we will present related works concerning the two key aspects of our proposal, namely multilayer music formats and Semantic Web applications respectively. Section 3 will provide an overview of the key characteristics of the IEEE 1599 standard, chosen among other similar initiatives as particularly suitable for our purposes. Section 4 will show how a combined use of the Semantic Web on one side and a suitable multilayer format on the other may encourage innovative applications. Finally, the proposed results will be generalized and discussed in Section 5.

2 Related Works

Since the goal of this paper is to show the potential of a combined approach of local and global semantic networks, represented by multilayer formats and the Semantic Web respectively, in this section we will review the most significant initiatives in such two fields.

2.1 Multilayer Formats for Music

The word "multilayer" means relating to – or consisting of – several or many layers. Multilayer file formats are commonly in use in information technology. For instance, nowadays most image-editing applications support objects on independent overlying canvases, and they open and save multilayer files to keep information arranged in layers with appropriate offsets. Conversely, we are interested in a different meaning of the term, where layers do not represent sets of objects covering the underlying ones, but coherent aggregations of information entities that help to shed light on a specific aspect of the global description. Layers are not seen as one on top of the other, but in a certain sense they are arranged in parallel. This interpretation is useful when you have to describe complex information entities presenting a number of facets of similar importance, as in the case of multimedia in general, and music in particular.

As it regards the former field, it is worth citing the Moving Picture Experts Group (MPEG) formats. For example, MPEG-4 deals with the actual encoding of moving pictures and audio, presenting additional features such as extended VRML support for 3D rendering, object-oriented composite files, support for externally specified Digital Rights Management (DRM) and various types of interactivity (Koenen *et al.*, 1997). MPEG-7 – formally called Multimedia Content Description Interface – is a standard that allows fast and efficient searching for material that is of interest to the user. It uses XML to store metadata, and can be attached to timecode in order to tag or synchronize particular events, such as the lyrics of a song (Chang *et al.*, 2001). The combination of MPEG-4 and MPEG-7 provides a multilayer environment for the comprehensive description of multimedia.

Concerning music, the applicability of multilayer formats to music representation was postulated in a number of scientific works (Castan *et al.*, 2001; Steyn, 2002; Haus & Longari, 2005). Thanks to its intrinsic characteristics, XML is generally considered the best language to encode and exchange music data and metadata.

In this context, it is worth mentioning the Music Encoding Initiative (MEI), namely a markup language for representing the structural, renditional, and conceptual features of notated music (Roland, 2002). This format has reached full maturity as it regards notation: thanks to the MEI 2011 Schema, extension and customization can be easily applied to the core set of elements to produce custom encoding systems that extend support for new types of musical documents (Hankinson *et al.*, 2011).

Another XML-based format is MusicXML, designed from the ground up for sharing sheet music files between applications and for archiving sheet music files for use in the future. MusicXML has been explicitly defined by its creator as an "Internet-friendly format for sheet music", aiming to provide a commonly-accepted interchange format (Good & Actor, 2003). The goal of making MusicXML files readable and usable by a wide range of music notation applications has been achieved thanks to the integration of export plug-ins into leading score editors, such as MakeMusic Finale, MuseScore and Avid Sibelius.

In recent times W3C has launched the Music Notation Community Group[1], an initiative that aims to unify formats syntactically and semantically different in order to establish the guidelines for a standardized approach. The original

[1] https://www.w3.org/community/music-notation/

goal was to develop and maintain format and language specifications for notated music used by Web, desktop, and mobile applications. The declared task of the Community Group was to maintain and update the MusicXML and SMuFL (Standard Music Font Layout) specifications, in order to evolve the specifications to handle new use cases and technologies, including greater use of music notation on the Web, while maximizing the existing investment in implementations of the existing MusicXML 3.0 and SMuFL specifications. However, the participation of independent experts (musicians, musicologists, computer researchers, software developers, etc.) has immediately fostered a more general discussion about the criteria a music notation standard must meet to have the quality, success, and longevity of *de-jure* and *de-facto* standards for other digital document representations.

Even if MEI and MusicXML share some common characteristics – e.g., they are both XML-based and cover music-related information other than logic symbols – their implementation of the concept of *multilayer description* is not pushed to the limit, since some kinds of information are missing.

For this reason, in the following we will adopt IEEE 1599, an XML format that supports all the layers proposed by J. Steyn (Steyn, 2002) and introduces the concept of layer as a way to keep heterogeneous information organized and the idea of *spine* as a way to interconnect different descriptions of the same information entity within the same layer or across multiple layers. The key characteristics of this format will be outlined in Section 3.

How a multilayer format for music encoding can realize a semantic network? Going back to the original definition, the semantic relationship among concepts in this case can be seen as the establishment of a complex network of interconnections among different descriptions of the same music events, as illustrated below.

2.2 Music Datasets in the Semantic Web

In Section 2.1 we have described some relevant initiatives to establish semantic relationships in music description at a local level. Let us now consider some related work at a wider level, with particular reference to global music datasets and the Semantic Web.

In the context of open data, it is worth citing the *DBpedia* project, a community effort to extract structured information from Wikipedia and to make this information available on the Web (Auer *et al.*, 2007).

The main goal of *DBpedia* is to serve as a nucleus for an emerging Web of open data, by allowing sophisticated queries against datasets derived from Wikipedia, links among Wikipedia data and other datasets available on the Web, publishing on the Web for human- and machine-consumption, etc.

This project had (and still has) an interest that goes beyond the scope of pure scientific research. As it regards the relationship between media and the Semantic Web, *DBpedia* inspired activities of data integration and document linking performed by the BBC (Kobilarov *et al.*, 2009). The idea was using Semantic Web technology – not only *DBpedia* but also *Linked Data*, *MusicBrainz*, etc. – to move across different information domains.

In an early stage of the Semantic Web, when this definition had just begun to circulate among experts, they immediately saw the possibility of applying a semantic approach to music cataloging. A pioneer project was the aforementioned *MusicBrainz*, a large database of user-contributed music metadata (Swartz, 2002).

A relevant experience regarding automatic interlinking of music datasets in the Semantic Web is described in (Raimond *et al.*, 2008). This paper presents an algorithm which takes into account both the similarities of Web resources and of their neighbors. The algorithm is tested on two different contexts: i) to link a Creative Commons music dataset to an editorial one, and ii) to link a personal music collection to corresponding Web identifiers.

3 IEEE 1599 Key Features

IEEE 1599 is an international standard sponsored by the Computer Society Standards Activity Board, designed by the Technical Committee on Computer Generated Music (IEEE CS TC on CGM), and officially recognized by the IEEE in 2008. IEEE 1599 adopts XML (eXtensible Markup Language) in order to describe a music piece in all its aspects (Baggi & Haus, 2009).

We will adopt such a language since it locally provides a semantic network of information, and this result is achieved thanks to the multilayer structure described below. In this context we use the word "local" to distinguish the information patterns available in a single XML document from the ones retrievable from the Web.

The main goal of the format is providing a comprehensive description of music and music-related material within a unique framework. The descriptions of a music piece are multiple and heterogeneous: its symbolic content, intended here as a sequence of music symbols, all its graphical and audio reifications (e.g. scores and recordings), additional metadata (e.g. catalogue metadata, lyrics, etc.) and materials (e.g. photos, playbills, etc.), and so on.

Comprehensiveness in music description is supported by IEEE 1599 thanks to a multilayered environment based on XML hierarchical structures. In particular, music and music-related contents are placed within six layers:

- *General* – catalogue information and other metadata;
- *Logic* – the logical description of the score in terms of music symbols;

- *Structural* – identification of music objects and their mutual relationships;
- *Notational* – graphical score representations;
- *Performance* – computer-based descriptions and performances of the piece;
- *Audio* – digital or ripped recordings.

Music events are uniquely identified in the encoding, so that they can be described in different layers (e.g., the graphical aspect of a chord and its audio performance), and multiple times within a single layer (e.g., different performances of the same music event). This is the role of a common data structure known as the *spine*, which marks all music events through unique identifiers and intrinsically allows the establishment of a complex network of relationships among descriptions and occurrences of such music events.

An IEEE 1599 document contains a collection of music-event descriptions, which are related:

- *Internally*, when they together constitute one of the available descriptions. It is the case of a sequence of audio events belonging to the same track (and described within the same instance of the *Audio* layer). This kind of relationship occurs among homogeneous descriptors for different music events, all parts of a single description or digital object;
- *Locally*, when they provide one of the various descriptions – belonging to different digital instances – of the same music event within a given layer. It is the case of the opening note of an instrumental solo notated on different scores described in the *Notational* layer. This kind of relationship occurs among homogeneous descriptors of the same music event in multiple, homogeneous digital objects;
- *Globally*, when they provide one of the various descriptions – belonging to different digital instances – retrievable from different layers for the same music event. It is the case of the opening note of an instrumental solo notated on different scores, all described in the *Notational* layer. This kind of relationship occurs among heterogeneous descriptors of the same music event in multiple, heterogeneous digital objects.

By combining these aspects, it is possible to design and implement advanced frameworks for music. From a semantic point of view, the strength point of the format is the possibility to create – thanks to the spine – a network of interconnected descriptions for music events. This interconnected structure can be seen as a local semantic network. Besides, layers contain other metadata and tagging that can be used as pointers towards external data sources, as exemplified in Section 4, thus providing the possibility to integrate the local information into a global semantic network.

For further details about IEEE 1599, please refer either to the official IEEE repository or to scientific literature that discusses in detail many aspects of the standard (Baggi & Haus, 2013). In addition, an official Web site containing documentation and examples is the EMIPIU portal[2].

4 Case Studies

In the following we will present and discuss some music-related case studies where IEEE 1599 and the Semantic Web could significantly improve user experience. These applications have not been implemented yet, nevertheless they could be realized through currently-available technologies and devices. In most cases, obstacles are mainly logistical and cost-related, since semantic tagging and open data collections are not sufficiently pervasive and ad-hoc actions should be performed to make these proposals effective.

4.1 Interactive Street Posters

Posters are a common form of billboard advertising, located along roads to be viewed mainly by residents, pedestrians and commuter traffic. The goal of advertisement is to catch the attention in order to persuade an audience about a commercial offering, an idea to convey or the availability of a service.

Besides well-known street posters, digital billboards capable of displaying running text, graphics and even audio are available too. Clearly the presence of an underlying computer system may greatly enhance user experience and content customization. An example is the *Spotify Powered Interactive Music Poster* released in April 2012, an apparently-traditional poster embedded with a knock sensor to detect vibrations, which in-turn is hooked up to a micro Arduino board and connected to Spotify.

Instead of exploring advanced evolutions of advertising, we will concentrate on standard posters to show their possible revivification in a Web 3.0 framework. In particular, we will take into account a poster advertising a jazz music event. In order to support advanced features, we will simply equip it with a machine-readable two-dimensional barcode, known as Quick Response (QR) Code. In this way, common devices such as smartphones can automatically detect and interpret that small amount of information (e.g. a Web address) sufficient to enable a number of advanced features, as detailed below.

Let us consider a jazz event dedicated to Billie Holiday. During the concerts, a number of tribute bands will perform the greatest hits of the American singer and songwriter. The poster which advertises the event, shown in Figure 1, provides only basic details about the date and place of the festival. On the one side, the attention of jazz lovers will be certainly caught by this announce,

[2] http://emipiu.di.unimi.it

but on the other a lot of people passing by will wonder what is the schedule of concerts, the name of performers or even the nature of this event. Thanks to the availability of a QR Code, they could easily focus the poster through a suitable smartphone application and get a number of metadata and services. In a context of augmented reality, this could occur even automatically thanks to wearable technology such as Google Glass.

Now let us produce and discuss a non-exhaustive list of possible applications. First, the QR Code could redirect to the Web site of the event, thus providing a lot of additional information and giving the possibility to buy tickets on line. Even if we have turned a traditional poster into an access point to Web contents, in the mentioned cases we are far from the exploitation of the Semantic Web.

Now let us link the QR code to an IEEE 1599 document containing multiple descriptions of a jazz piece, for instance Billie Holiday's 1936 performance of *Summertime*, the aria composed in 1934 by George Gershwin for the opera *Porgy and Bess*. Provided that the device in use is enabled to parse IEEE 1599 contents through a suitable application, now it is possible to exploit the huge amount of data and metadata contained in the mentioned XML document. The use cases and services that can be implemented include: letting the user enjoy the complete original recording or other free audio tracks, retrieving information about all the singers who performed *Summertime*, implementing a simple score following for the main theme, and so forth.

Fig. 1 - An example of interactive street poster.

Fig. 2 - Interaction with street-poster contents.

If the described IEEE 1599-based fruition is embedded into a network environment, or even better into the Semantic Web, supported applications may become countless (see Figure 2). For instance, it would be possible:

- to provide a preview of the concert, including audio and video contents recorder during a rehearsal session;
- to get historical photos of Billie Holiday from a publicly available repository, as well as the list of all movies containing *Summertime* in their soundtrack;
- to follow the career of one of the involved artists and to reserve a ticket for an upcoming event;
- to share information and add user-generated contents through social networks.

In this context, IEEE 1599 plays not only the role of local data storage (providing scores, audio tracks, synchronization information, etc.), but it also acts as an authoritative source to query the Semantic Web thanks to tagged metadata. These concepts will be discussed and generalized in Section 5.

In conclusion, a typically passive way to communicate with an audience can become interactive, entertaining and even addictive.

Keeping in mind the original purpose of an advertising poster, its power in terms of communication efficacy and ability to arouse interest is greatly increased.

4.2 Augmented Opera

Augmented Opera is a proposal to apply augmented reality to live opera experience. This example involves both networking – in order to support user interaction with remote contents – and a multilayer format to encode local information. Besides, tags and other semantic identifiers must be placed accordingly.

For instance, let us consider an IEEE 1599 document encoding a complete opera, e.g. *Turandot* by Giacomo Puccini. This implies the availability of:

- Catalog metadata that can be used in a Semantic Web context to link other contents. For instance, it is possible to query an open data repository to get information about Puccini's last years of life or about other works by Giuseppe Adami, marked in the document as one of the librettists;
- Symbolic information that virtually allows one to attach alternative descriptions and renditions to music events with the desired granularity, even note by note. In this way, it is possible to display libretto together with music, provided that there is a synchronization source (this process could be automated thanks to pitch tracking techniques or demanded to human control);
- A number of score versions that can be experienced in a score-following environment during the show;
- Alternative audio and video contents, which probably will not be launched during the live performance but will provide other links in terms of metadata (e.g. historical performances, great interpreters, etc.) as well as on-line purchase suggestions;
- Character identification on the stage through movable video tags;
- Additional graphic contents (such as related artwork) to improve user experience.

After building a semantic network of information, all these data can be suitably presented thanks to an augmented-reality device such as Google Glass. A use case is shown in Figure 3. The interface contains a number of advanced features, such as dynamic identification of characters, automatic score and libretto following with user-defined translation, links to external contents, and so on. Of course the interface must be designed to enhance user experience rather than produce information overload, a typical risk of augmented-reality initiatives.

Fig. 3 - An example of augmented reality application for interactive opera experience.

As mentioned above, some additional work could be required to configure the environment accordingly. For instance, the automatic identification of the artists on the stage, as it regards not only their presence but also their current position, would enable a number of advanced features: video captions could be placed near the corresponding singer, a feature particularly relevant when independent voices are simultaneous in operatic ensemble pieces. Unfortunately, available technology does not permit to embed a reliable face-recognition algorithm in a wearable device, due to both limited computational resources and environmental characteristics (darkness, distance, etc.). Nevertheless, alternative technological solutions are available, such as Wi-Fi based positioning systems (WPS) that provide a way to attach *tracking tags* to moving objects.

A similar situation occurs for automatic score (and libretto) following. In fact, even if relationships among music symbols, graphic scores and libretto are made explicit by the IEEE 1599 document (i.e. they are locally synchronized), these "packets" should be synchronized with the current live performance, whose timing cannot be determined *a priori*. Also in this context an automatic approach – e.g. pitch/beat tracking algorithms – would be helpful, but the required level of reliability discourages its adoption. For instance, some opera houses (e.g. Teatro alla Scala of Milan) delegate the automatic advancing of lyrics to a manual process performed by a dedicated expert during the live show.

A positive side effect of this approach is the possibility to create a usertailored environment that does not interfere with the rest of the audience. For example, the debate on installing stage libretto-lyrics displays in concert halls and opera houses mainly depends on the possible inconvenience caused to the public not interested in the scrolling text. On the contrary, this kind of solution is personal and can be customized with regard to many aspects: language, font size, information position on the screen, etc.

5 Generalization and Discussion

A matter of debate could be the implementability of what described in Section 4 by the adoption of either a suitable multilayer format or the Semantic Web alone. In fact, a number of advanced applications for multimedia fruition, music education, cultural heritage revivification, etc. have been designed without building complex and distributed semantic structures. In response to this concern, we recall that both approaches – corresponding to local vs. global semantic relationships – have their own characterization and typical use. An in-depth analysis of the applications described above clearly shows that only a subset of their features could be designed and implemented without this integration, but the most advanced services and use cases depend on the availability of strong interconnections among local and global information.

Such case studies can be generalized and extended to non-music fields. What emerges from their analysis can be summarized as follows:

- A multilayer format – here intended as a way to describe the multiple facets of a given information entity – provides a network of interrelated data having its own semantics. An application capable of managing such a format should provide an interface between local semantics and users in order to make relationships emerge and allow semantic queries;
- If the mentioned multilayer format contains authoritative tagging, as in most XML-based languages, this kind of local information can be used to link and explore external semantic data sets;
- The Semantic Web and open-data initiatives can benefit from the integration with multilayer formats from a number of perspectives, both as content and service providers in response to locally-driven semantic queries, and as potential receivers of multilayer contents contained in local resources to be embedded into a wider semantic network.

Conclusions

In this work we illustrated some advanced music-oriented applications based on the integration of local semantic networks, realized through an XML-based multilayer format, within a global semantic network, represented by the Semantic Web. This integration can foster advanced music experience and allows the implementation of innovative services. Music with its multiple facets represents a suitable domain to test the adoption of multilayer formats in network-based and open-data scenarios.

The research results and the examples presented here can be easily extended to other domains of information structuring and encoding. In particular, the multilayer representation, the identification of relationships among different descriptions and the possibilities offered by information interchange are key aspects for the design and implementation of innovative applications.

When wearable technologies will be spread among users, augmented reality will finally leave the stage of theoretical research to enter the exploitation phase, and semantic tagging will be extensively performed, then the approach of user communities towards information interconnection and data exchange will be revolutionized, and proposals like those mentioned above will dramatically improve our lifestyle.

REFERENCES

Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., & Ives, Z. (2007), *Dbpedia: A nucleus for a web of open data* (pp. 722-735). Springer Berlin Heidelberg.

Baggi, D. L., & Haus, G. (2009), *IEEE 1599: Music Encoding and Interaction*. IEEE Computer, 42(3), 84-87.

Baggi, D. L., & Haus, G. M. (2013), *Music Navigation with Symbols and Layers: Toward Content Browsing with IEEE 1599 XML Encoding*. John Wiley & Sons.

Berners-Lee, T., Hendler, J., & Lassila, O. (2001), *The semantic web*. Scientific american, 284(5), 28-37.

Bizer, C., Heath, T., & Berners-Lee, T. (2009), *Linked data-the story so far*. Semantic Services, Interoperability and Web Applications: Emerging Concepts, 205-227.

Brachman, R. J. (1977), *What's in a concept: structural foundations for semantic networks*. International journal of man-machine studies, 9(2), 127-152.

Castan, G., Good, M., & Roland, P. (2001), *Extensible markup language (XML) for music applications: An introduction*. Computing in Musicology, 12, 95-102.

Chang, S. F., Sikora, T., & Purl, A. (2001), *Overview of the MPEG-7 standard*. Circuits and Systems for Video Technology, IEEE Transactions on, 11(6), 688-695.

Collins, A. M., & Loftus, E. F. (1975), *A spreading-activation theory of semantic processing*. Psychological review, 82(6), 407.

Collins, A. M., & Quillian, M. R. (1969), *Retrieval time from semantic memory*. Journal of Memory and Language, 8(2), 240.

Good, M., & Actor, G. (2003), *Using MusicXML for file interchange*. In Proceedings of the Third International Conference WEB Delivering of Music (WEDELMUSIC'03)

(p. 153). IEEE Computer Society.
Hankinson, A., Roland, P., & Fujinaga, I. (2011), *The Music Encoding Initiative as a Document-Encoding Framework*. In ISMIR (pp. 293-298).
Haus, G., & Longari, M. (2005), *A multi-layered, time-based music description approach based on XML*. Computer Music Journal, 29(1), 70-85.
Kobilarov, G., Scott, T., Raimond, Y., Oliver, S., Sizemore, C., Smethurst, M., Bizer, C., & Lee, R. (2009), *Media meets semantic web–how the BBC uses DBpedia and linked data to make connections*. In The semantic web: research and applications (pp. 723-737). Springer Berlin Heidelberg.
Koenen, R., Pereira, F., & Chiariglione, L. (1997), *MPEG-4: Context and objectives*. Signal Processing: Image Communication, 9(4), 295-304.
Raimond, Y., Sutton, C., & Sandler, M. B. (2008), *Automatic Interlinking of Music Datasets on the Semantic Web*. In Proceedings of the Linked Data on the Web Workshop, Beijing, China, April 22, 2008, 369.
Roland, P. (2002), *The Music Encoding Initiative (MEI)*. In Proceedings of the First International Conference on Musical Applications Using XML (pp. 55-59).
Sowa, J. F. (2006), *Semantic networks*. Encyclopedia of Cognitive Science.
Steyn, J. (2002), *Framework for a music markup language*. In Proceeding of the First International IEEE Conference on Musical Application using XML (MAX2002) (pp. 22-29).
Swartz, A. (2002), *Musicbrainz: A semantic web service*. Intelligent Systems, IEEE, 17(1), 76-77.