State of the Art and Perspectives in Multi-Layer Formats for Music Representation

Adriano Baratè, Goffredo Haus, Luca A. Ludovico
LIM – Music Informatics Laboratory
Department of Computer Science
University of Milan
Via G. Celoria 18, 20133 Milano, Italy

{barate,haus,ludovico}@di.unimi.it

***Abstract*—This paper aims to provide an analytical comparison among the most relevant representation formats that support multi-layer descriptions of music content, namely IEEE 1599, Music Encoding Initiative, and MusicXML/MNX. After remarking the technical characteristics of such formats and highlighting their similarities and differences, we will try to shed light on their future, so as to understand the current trends in digital representation of music and multimedia.**

***Index Terms*—music, multi-layer representation, XML, IEEE 1599, Music Encoding Initiative, MusicXML, MNX**

I. Introduction

Music is one of the human activities where the concepts of multimedia and multi-modality are exploited more deeply. As a matter of fact, the goal of providing a comprehensive description of a music work is not a trivial task, since different domains and multiple types of information are involved.

The archives of a music publisher or an opera house give an idea of the quantity and articulation of music-related materials: typically, there are scores, audio/video recordings, stage photos, additional graphic materials such as posters, playbills, etc. Moreover, for each of the mentioned categories, archives can host a multiplicity of materials: multiple score editions (e.g., the manuscript, a hand-made transcription, an early print, a modern edition, etc.) and versions (e.g., full score, parts, piano reductions, etc.), audio performances heterogeneous not only in performers but also in types (e.g., video captures, multi-track recordings, etc.), and so on.

Consequently, when aiming at a comprehensive description of music content, also its digital representation is a challenging task that requires cross-domain and multi-layer approaches. This work aims to provide a review and comparative analysis of the most advanced and widespread formats addressing such a goal, also taking into consideration their future perspectives.

A point to clarify is why research, academia and industry are currently interested in this problem, as evidenced by the recent initiatives sponsored by the Institute of Electrical and Electronic Engineers (IEEE) and the World Wide Web Consortium (W3C) presented in Section III. The answers are many: first of all, a multi-layer approach can pave the way for new applications in the field of musical and multimedia enjoyment, music education, and enhancement of intangible cultural heritage; these approaches not only have a theoretical or cultural value, but correspond to new products presenting economic value. Possible examples are institutions that hold a "historicized" cultural heritage or publishers who have archives of already published materials to be enhanced through technological tools that can give new life to their contents.

The list of stakeholders potentially interested in multi-layer music representation formats includes music and multimedia publishers, radio and TV stations, manufacturers of audio equipment and gaming devices, manufacturers of mobile and wearable devices, mobile and Web service providers; in addition, we can identify cultural institutions such as opera houses, theaters, cinemas, museums, public or private music archives, institutes for music education, schools, universities, conservatories, foundations.

Currently, there are three initiatives that polarize research and development in this field, and which are characterized by the industrious efforts of the respective communities: IEEE 1599, Music Encoding Initiative (MEI), and MusicXML/MNX. These multi-layer formats for music description will be described, discussed, and compared in the following.

The remainder of this paper is structured as follows: Section II provides some basic definitions about the multi-layer representation of music, including the concepts of *instance* and *synchronization*; Section III describes the most relevant and recent initiatives supporting multi-layer descriptions; Section IV provides an analytical comparison among them, highlighting pros and cons of each approach; finally, Section V illustrates the future perspectives of the mentioned formats, whose communities are very active in the extension of their descriptive potentialities, in the development of software tools, and in the constitution of wider repositories of music.

II. Multi-Layer Representation of Music

The term *layer*, in everyday language, usually recalls a thickness of material covering a surface or body. From this point of view, a layer both hides and provides new characteristics to the underlying object. But the concept of multi-layer description implies looking at an entity from different angles, thus unveiling its heterogeneous facets and their possible

978-1-7281-1649-5/19/$31.00 ©2019 IEEE
DOI 10.1109/MMRP.2019.00014

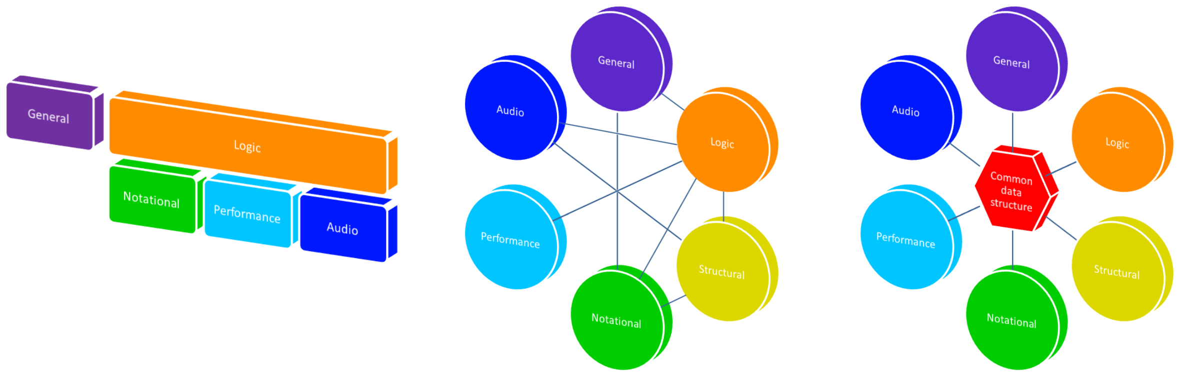

Fig. 1. A graphical comparison between different multi-layer models: from left to right, the stratified, the graph, and the star model.

connections. According to the meaning that one wants to privilege, distinct interpretations of the concept of multi-layer representation may emerge:

- the *stratified model*, where upper layers can be built on top of lower layers, using them as a base to construct specialized information;
- the *graph model*, where different layers are peers directly interconnected;
- the *star model*, where there is a common data structure acting as a reference to keep the information in layers connected and possibly synchronized.

A graphical representation of these models is presented in Figure 1.

Whatever the meaning attributed to the term multi-layer, music is a good test bed for this kind of approach, since a comprehensive description of a music piece typically spans across multiple domains and media types. For example, music representation implies both time-based (e.g., audio performances) and space-based (e.g., scores) descriptions, and most content can be mutually synchronized.

It is worth analyzing music layers in detail. First of all, a composition can be described in terms of organized musical or sound events, directly related to the composer's idea. This level of representation is often referred to as the *logic* description. The written symbols that constitute the score, if any, are usually a transliteration of such an abstract idea into a form of music notation. Such a process can originate one or more *notational* instances. For the sake of completeness, we must also consider those pieces where the graphical representation is a fundamental part of the meaning of the composition itself (see, e.g., [1], [2], and [3]). Nevertheless, from a functional point of view we can always recognize a distinction between the logical and notational layers. Between these two levels of description we can place the *structural* layer, which organizes musical and/or sound events into aggregated objects (e.g., themes, motifs, tonal regions, macro-areas, etc.). Formats for computer-driven performances should be described in a dedicated *performance* layer. Concerning the production of sound coming from score performances, it is natural to identify a further layer devoted to *audio* and, possibly, to *video*. Finally, even metadata helps to provide a complete description of a piece of music, adding useful information about the composition's title, author, genre, etc. This kind of content, not directly related to instances of the composer's idea, can be placed in the *general* layer.

The multi-layer structure illustrated so far recalls some early research works on the subject of multi-layer representation of music, such as [4], [5] and [6], that explicitly propose the following layers: *general*, *logic*, *structural*, *notational*, *performance*, and *audio*.

Needless to say, there are examples where a clear and unambiguous association among music content and layers does not emerge. For example, is it possible, and meaningful, to identify a structural layer in a free jazz improvisation? And how to categorize the generating patch for a computer-music composition: as a non-canonical notational representation, as a structural description, or as a computer-driven performance? These questions, to cite but a few, are pushing some research groups to generalize the concept of layer or to let the user define customized layers. This issue will be discussed in more detail in Section V.

Following a multi-layer approach, the single music or sound events constituting the composition can be described multiple times, both in different layers and in different instances inside any layer. For *instance* here we intend a specialized representation made of homogeneous event descriptions. For example, the *audio* layer can host multiple audio tracks referring to different performances: according to our definition, each track is considered as an instance within the *audio* layer. Consequently, the basic idea in multi-layer formats is having a sequence of events potentially described multiple times. Considering layers and instances, relationships among event descriptions can occur:

- *internally*, when the descriptions of involved events refer to a single instance, like canvas areas in a given score or audio events in a track. In short, they are homogeneous descriptors for different events within a single instance;
- *locally*, when descriptions refer to the same event repre-

sented in different instances within a single layer, like a score symbol notated on different score editions. In short, they are homogeneous descriptors of the same event in multiple, homogeneous instances;

- *globally*, when descriptors provide one of the descriptions retrievable from different layers for the same music event, like the graphical aspect and the audio rendition of a given note. In short, they are heterogeneous descriptors of the same music event in multiple, heterogeneous instances.

An issue to solve in multi-layer formats is how to link different representations of the same event. A possible solution is to define a list of identifiers that are referred to from all layers and instances. This solution implies point-to-point connections between elements in that data structure, i.e. shared identifiers, and the occurrences of corresponding event representations in layers; considering all point-to-point links which refer to the same event, hidden relationships among different representations emerge. Please note that such hidden bindings can occur among layers, thus originating so-called *inter-layer synchronization*, and among instances belonging to the same layer, thus creating *intra-layer synchronization*. With respect to the classification reported above, inter-layer synchronization derives from global relationships, and intra-layer synchronization from local relationships, when extended to all the events encoded in the document. For the sake of clarity, an example of simplified graphical representation is shown in Figure 2, that partially illustrates the links for a single event through continuous, dashed and dotted lines.

The concept of multi-layer description – i.e. as many different types of description as possible – together with the concept of multi-instance support – i.e. as many different instances for each layer – can provide a rich and flexible way to encode music in all its aspects.

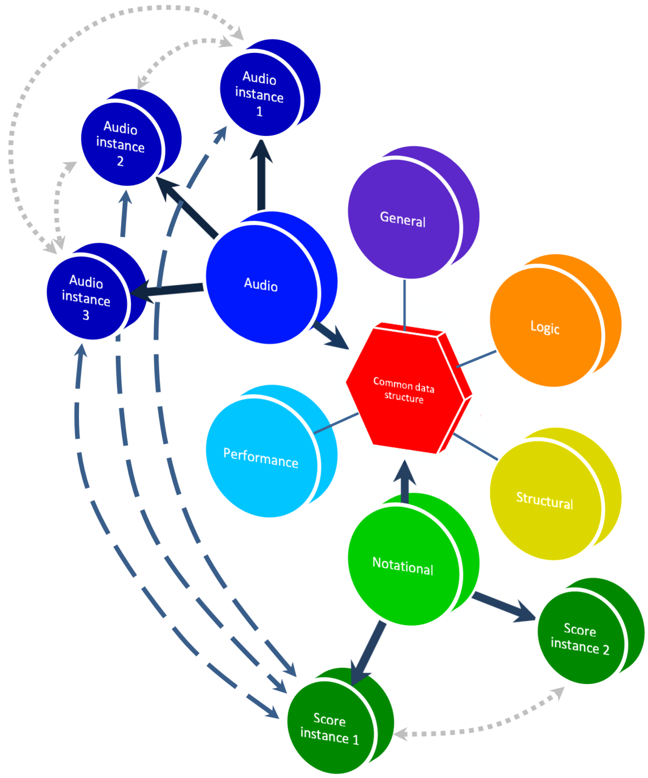

Fig. 2. A graph-modeled music representation, inspired by the layers of the IEEE 1599 format. Arrows represent links among different descriptions of the same music event. Continuous lines are the explicit links from instances to the common data structure and vice-versa, dashed lines are implicit links between instances of different layers (*inter-layer synchronization*), and dotted lines are implicit links between instances of the same layer (*intra-layer synchronization*). Implicit links do not require to be encoded, since they emerge when explicit links are parsed and resolved.

III. XML-Based Formats for Multi-Layer Music Representation

The multi-layer approach to foster an effective and comprehensive description of music has been implemented by three formats that are extensively used by the sound and music computing community: IEEE 1599, Music Encoding Initiative, and MusicXML.

They all are based on Extensible Markup Language (XML) [7]. The reasons for this choice, common to the three independent formats (and also to others), include modularity, separation of data and meta-data, reduction of learning requirements, assistance to tool development, and increase in legibility and understandability, as discussed in [5], [8], and [9]. XML was designed to represent complex, structured data in a standardized way, which is a key feature for music description; moreover, this technology solves interchange problems and allows system interoperability. Other XML-based formats have been developed in the past, but they had less commercial success and scientific impact. Some of them will be shortly presented in Section III-D.

Another point in common for the three formats is the possibility to link external digital files, typically used to encode some types of instances (e.g., for graphics and audio). On one side, this feature allows to retrieve a huge corpus of already available documents in commonly accepted file formats (e.g., JPEG or TIFF for graphics, and AIFF or MP3 for audio); on the other side, binary files are more efficient in data representation, thus solving the problem of XML verbosity.

A. IEEE 1599

IEEE 1599 is an XML-based format aiming at a comprehensive description of music information [10]. This format has been mainly developed at the Laboratorio di Informatica Musicale (LIM) of the University of Milan, in response to the *IEEE Project Approval Request 1599 (P1599)* titled "Definition of a Commonly Acceptable Musical Application using the XML Language". Since September 2008, the format is an international IEEE standard.

The purpose of the format is to achieve integration among symbolic, structural, notational, computer performance, and digital sound levels of representation, without privileging one layer with respect to others. In other words, the core of IEEE 1599 is not score information, since music symbols are considered as one of the many possible layers for music information.

A common data structure, called the *spine*, acts like a reference list of events to be referred from all layers. Consequently, this format implements the star model for content multi-layer structuring.

Concerning documentation, examples, and support tools, currently the most relevant access point is a dedicated Web portal, publicly available at http://ieee1599.lim.di.unimi.it/. Its Music Archive section embeds an on-line multi-layer player. For the preparation of materials, a Java-based multi-platform suite is available, and a number of plug-ins for digital score editors (e.g., MakeMusic Finale and MuseScore) have been released.

Finally, it is worth mentioning a book edited by the chair and the official reporter of the IEEE standardization initiative dealing with IEEE 1599 structure and applications [10].

B. Music Encoding Initiative

The *Music Encoding Initiative* is a community-driven effort to create a commonly accepted, digital, symbolic representation of music notation documents [11]. The resulting format is a set of rules for recording the intellectual and physical characteristics of music notation documents so that the information contained in them may be searched, retrieved, displayed, and exchanged in a predictable and platform-independent manner.

The MEI format focuses – though not exclusively – on the encoding of documents in musicology and libraries for research and analysis purposes. Such a format is defined through a modular, extensible XML schema, accompanied by detailed documentation. Its guidelines are published under an open-source license and periodically updated.

In recent releases, new elements, attributes, or content models may be included, allowing the addition of aspects that can address new types of documents. This customization approach allows the transition from a single, monolithic encoding schema to an extensible document-encoding framework.

This initiative has provided not only documentation and technical specifications, but also computer-based tools widely adopted by libraries, museums, and individual scholars to encode musical scores and to render them. For example, Verovio is a portable, lightweight library, released under the GNU Lesser General Public License (LGPL) license, which transforms MEI files into Scalable Vector Graphics format.[1]

At the moment of writing, the latest version is MEI 4.0 Schema, released on Nov 1, 2018. The reference Web site – containing documentation, guidelines, a tag library, technical details, and some encoding examples – is available at http://music-encoding.org/.

From the official documentation, it clearly emerges that MEI was born as a format for musical notation. In this sense, it is an example of stratified multi-layer model, where the logic description of a music piece is the focus and other layers are optional additions.

[1]http://www.verovio.org/

C. MusicXML and MNX

MusicXML was designed for sharing sheet music files between applications and for archiving sheet music files for future use. MusicXML has been explicitly defined as an "Internet-friendly format for sheet music" [12], aiming to provide a commonly-accepted interchange format [13].

The goal is to make MusicXML files readable and usable by a wide range of music notation applications. Originally designed and developed by Recordare LLC, it was soon integrated into leading score editors, such as Finale and Sibelius. Of the three mentioned initiatives, this is probably the most commercially successful and widespread. In the official Web site, at https://www.musicxml.com/, more than 230 applications currently supporting MusicXML are reported.

With respect to documentation, the site contains tutorials, full text versions of most of the papers on MusicXML by its authors, links to selected publications authored by the user community, and conference presentations on the format. A section for developers includes MusicXML Document Type Definitions (DTDs) and W3C XML Schema Definitions (XSD), made available from the W3C MusicXML GitHub repository (see https://github.com/w3c/musicxml/).

Version 3.1 of the MusicXML format was released in December 2017. Such a version was developed by the W3C Music Notation Community Group and published as a W3C Community Group Final Report. It is licensed under the W3C Community Final Specification Agreement (FSA).

MusicXML focuses on music notation, whereas another format – currently code-named MNX – provides an overall framework for encoding works of music of many different kinds, so as to enable more than one encoding system within the same document. Also this format is under development in the framework of the W3C Music Notation Community Group. The basic idea is to describe a notation-neutral container that is concerned solely with a document's metadata, attribution and organization, with no reference to any notational system. The encoded music is not required to reference any specific visual or aural rendering, although such information can be optionally included.

Concerning the multi-layer vision, MusicXML basically adheres to the stratified model, where logic (and automatically-rendered notation) is the core description, whereas other aspects are handled by the MNX format.

D. Other Formats

Many other extensible formats for music notation have been proposed in the past. Some of them are very focused, such as ChordML for the representation of lyrics and chord symbols [14] and NeumesXML for the encoding of neumes and mensural notation [15]. Due to their very specific approach, these formats have been used sporadically by the sound and music computing community, and their impact has been limited. Conversely, other initiatives are worthy of mention, but they have not been considered in the comparison above due to their limited diffusion or disposal.

For example, WEDELMUSIC [16] is an XML-compliant format mainly focused on music notation, but equipped with multimedia capabilities as well. WEDELMUSIC objects may include: multilingual cataloguing information, music notation scores, audio files, images of music scores, multilingual lyrics, video files, documents, pictorial images, animations and sliding shows, synchronizations, etc., encoded in many different file formats. From 2001 to 2003 WEDELMUSIC organized yearly conferences to disseminate scientific results. The official Web site is http://musicnetwork.dsi.unifi.it/wedelmusic/.

The Extensible Music Format (XMF) is a family of specifications created and administered by the MIDI Manufacturer's Association. An XMF file contains one or more existing files – such as Standard MIDI Files, Downloadable Sound (DLS) instrument files, WAV files, or other digital audio files – to create a collection of all the resources needed to present a musical piece, an interactive web page soundtrack, or any other piece of media using pre-produced sound elements. The format belongs to MIDI specifications and is described at https://www.midi.org/specifications/category/extensible-music-format-xmf. It provides an example of multi-layer structure built around MIDI, namely a computer-driven performance format. A noteworthy aspect of XMF is the achievement of extensibility through a low-overhead tree data structure rather than an XML-based approach.

Finally, MPEG-7 is a multimedia content description format standardized in ISO/IEC 15938 and formally called Multimedia content description interface. It is not a standard which focuses on the actual encoding of moving pictures and audio, like its ancestors MPEG-1, MPEG-2 and MPEG-4. It uses XML to store metadata, and can be attached to timecode in order to tag particular events, or synchronize lyrics to a song. Once again, the basic concept is that description must be separate from the audiovisual content; in any case, MPEG-7 provides means to establish a relation between the content and description. In other words, the description is multiplexed with the content itself. MPEG-7 is well covered by scientific literature, but it is not extensively adopted by the sound and music computing community.

IV. Discussion

A. Supported Layers

All the formats described in Section III can be defined as multi-layered. Such an approach is explicit in IEEE 1599 documents, whose root element has 6 sub-elements named after the 6-layer structure: *general*, *logic*, *structural*, *notational*, *performance*, and *audio*. A peculiarity of IEEE 1599 is the presence of two sub-layers under the *logic* layer: the *spine* element, which is the data structure referenced by all layers in order to link heterogeneous event descriptions, and the logically-organized symbols (los) element, containing score-symbol descriptions according to different notations (common music notation, tablatures, neumes).

Even if the terminology is different from the one of IEEE 1599, also MEI covers the full range of layers mentioned in Section II: a *general* layer, used for the encoding of names of persons or corporations/organizations, dates, or descriptive phrases for styles, periods or geographical indications; a *logic* layer, where music can be encoded according to a number of notation models (common music notation, mensural notation, neumes, tablatures, user-defined symbols, text such as tempo marks, directives and lyrics, etc.); a *structural* layer, where analytical information is represented as relationships between event elements (e.g., notes, chords, etc.) through pointers and references; a *notational* layer, in order to support material in graphical format; a *performance* layer, basically focusing on the Musical Instrument Digital Interface (MIDI); finally, an *audio* layer aiming to organize audio and video files of performances of a musical work. Linking and alignment functions are explicitly provided to align recorded media with elements in the musical domain. Moreover, thanks to a section devoted to critical apparatus, it is possible to encode differences between multiple exemplars of the same musical work.

MusicXML was born as a notation interchange format, thus intrinsically supporting the *general*, *logic*, and *notational* layers. The full coverage of the 6-layer structure and synchronization features come from the adoption of an additional standard called MNX, currently under development in the context of the W3C Music Notation Community Group. Such a standard will extend the integration of visual notation and aural rendering, thanks to a new type of literal encoding which links SVG graphics to audio/performance data via the time dimension. A common timeline will serve to connect notations and their audible counterparts, in a way that somehow recalls the IEEE 1599 concept of spine. In any case, MNX currently lacks a clear mapping structure and uses graphical information for different purposes, sometimes with incoming links (to map audio onto graphical symbols), sometimes with outgoing links (to map graphic elements onto semantic data).

B. Presence in the Scientific Literature

The number of quotations of each format within the scientific literature can give a broad idea of the interest raised in the research and academic community, not only in the field of sound and music computing but also in other contexts, thanks to the intrinsic trans-disciplinary nature of the subject. In order to evaluate this parameter, we chose a reliable indexing system, namely Scopus, querying it through Mendeley, a free reference manager for academic research developed by Elsevier. Queries were launched at the end of 2018.

Considering the number of works and authors mentioning the name of each format in the title and/or in the abstract, we obtained the following results:

- IEEE 1599 – 41 papers, involving a total number of 145 people in quality of authors and editors;
- Music Encoding Initiative – 25 papers, authored by 31 people;
- MusicXML – 54 papers, authored by 123 people.

It is worth noting that IEEE 1599 assumed this name in 2008, at the time of its standardization, while it was previously known with the code name MX; references to MX have not been counted.

The distribution of publications over years is shown in Figure 3, while Figure 4 illustrates the number of citations received by papers published year by year. These data remark that all formats have raised interest in the scientific community.

C. Strengths

In the authors' opinion, each format presents some distinguishing points.

IEEE 1599 is the format that best embodies the concept of multi-layer structure, where all layers have the same importance and equally contribute to music description. From this point of view, IEEE 1599 can be considered also as a general-purpose synchronization format for space- and time-localized data. Thus, such a standard is versatile enough to be applied to a number of non-musical fields, including the navigation of audio scenarios [17], the provision of advanced services on semantic Web networks [18], the augmented experience of live theatrical performances [19], and a number of educational applications [20]–[22]. Besides, the possibility of identifying and synchronizing the occurrence of musical events in graphic, audio and video files made this format particularly suitable in the field of intangible cultural heritage, as demonstrated by the adoption of IEEE 1599 in a number of international exhibitions.[2]

The Music Encoding Initiative presents a very accurate description of notated music, supporting different notational approaches. This strength has been recognized also by the W3C Music Notation Community Group, who borrows some approaches from MEI, concerning, e.g., the organization of musical texts with their attached metadata. The MEI community is very active in proposing format advancements and in encoding musical pieces. First organized in 2013, the Music Encoding Conference yearly acts as a cross-disciplinary venue for theorists, musicologists, librarians, and technologists to meet and discuss new advances in their fields. The MEI format has also some important practical implications: for instance, the Danish Royal Library is using MEI to encode the bibliographic information for the Carl Nielsen Edition[3] making it available as Linked Open Data; another very recent example is the web publication of the *Digital Mozart Edition*, a collaborative project of the Stiftung Mozarteum, Salzburg, and the Packard Humanities Institute, Los Altos (CA) aiming to make publicly accessible the works by Wolfgang Amadeus Mozart in digital formats.[4]

MusicXML is widely supported by music software, and is often considered as a *de-facto* XML standard to represent and distribute music. Among its main features, it is worth citing the exchange of sheet music between various score writing programs, the electronic distribution of musical scores, and storage and archiving of in an electronic format. Currently, the most noticeable point of MusicXML is the standardization effort in the context of the W3C Music Notation Community Group, which should turn version 3 into the commonly-accepted format for music description over the Web. The process is still far from an end, but the involvement of industry (e.g., people from MakeMusic, Steinberg, MIDI Manufacturers Association, etc.) and academia (e.g., members from MEI and IEEE 1599 working groups) suggests an interest in the ambitious project at different levels: artistic, scientific, technological, commercial, etc.

D. Critical Issues

Each format is still under development and presents some critical issues.

Concerning IEEE 1599, the format has been critically reviewed in [23], where the main problems identified were: the necessity to adhere to a proprietary format to represent score symbols in the *logic* layer, an ambiguity in the definition of layer instances, the impossibility of describing an organized corpus of linked compositions, and very limited support to Digital Rights Management. Moreover, even if the original working group was officially composed by about 30 people and the IEEE standardization occurred in 2008 involved about 100 international experts and scholars, the staff active in research and development was basically that of the Laboratory of Music Informatics (LIM), University of Milan. Consequently, the availability of utility tools is limited and the resources required to complete a multi-layer encoding are still demanding. For this reason, the total number of IEEE 1599 pieces is very small. These considerations are leading to the establishment of a new international working group in the framework of IEEE Standard Association activities, as detailed in Section V.

As it regards MEI, the format was designed primarily to capture semantic description of music symbols, thus addressing more abstract issues than notation rendering. Being capable to support a wide variety of notation styles and symbols, it is intrinsically difficult to adopt MEI as a commonly-accepted interchange format: its completeness and its degree of detail in providing a logical description of the musical symbols is transformed into a critical aspect for the development of software tools able to support all the characteristics that can be coded in the format. As a matter of fact, rendering software support is currently a significant weakness. Finally, with respect to IEEE 1599, its interpretation of multi-layer music representation is biased towards notation-related aspects.

Conversely, MusicXML is commonly recognized as a good interchange format, also thanks to the support offered by many commercial software tools, but it is too much constrained to Common Western Notation. Its philosophy is completely different from MEI's one, and pros and cons of the two approaches are complementary. In the context of the W3C standardization initiative, where also MEI members and other independent people are present, a critical issue is the conflict among different visions of what a multi-layer music standard should be. For example, a recent proposal submitted to the community group, consisting in the evaluation of a web-

[2] http://ieee1599.lim.di.unimi.it/practice_exhibitions.php
[3] http://www.kb.dk/dcm/cnw/navigation.xq
[4] https://dme.mozarteum.at/

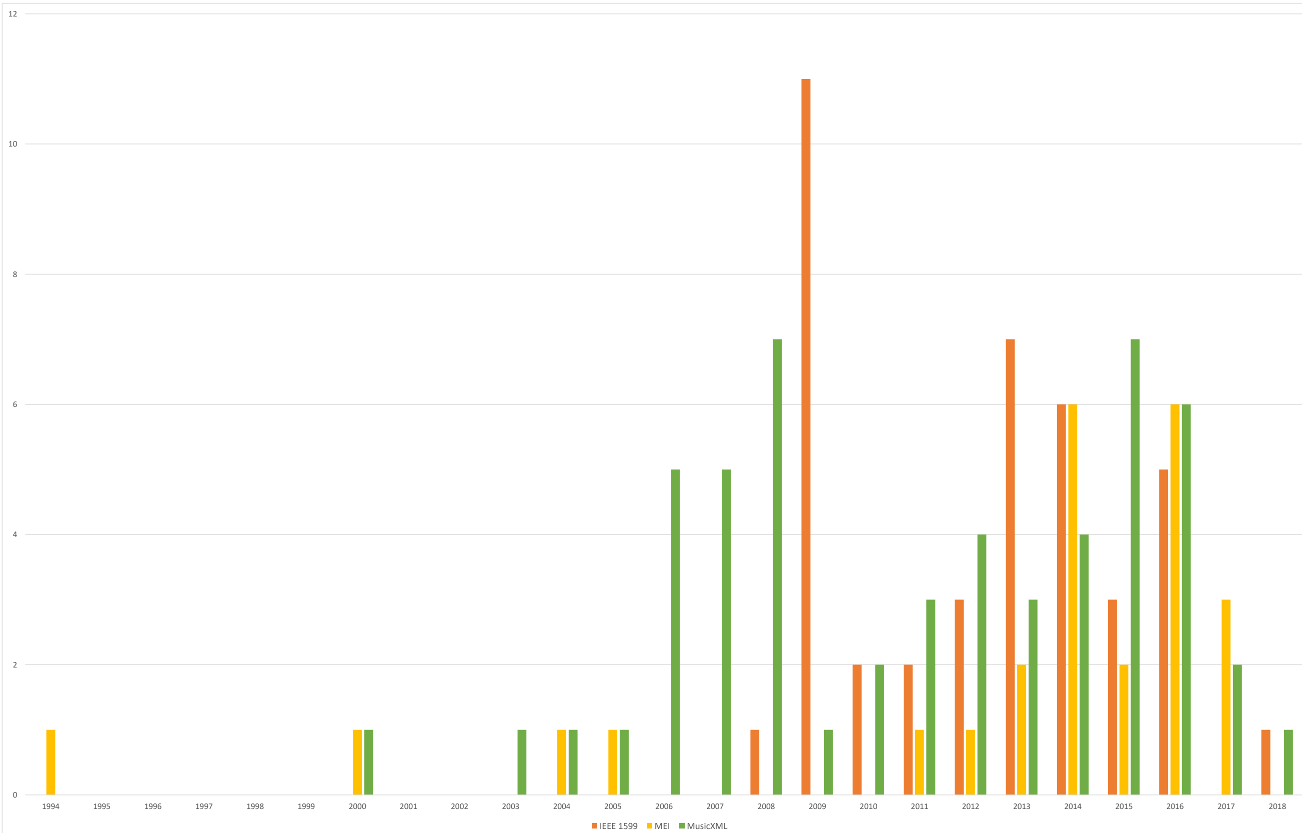

Fig. 3. Time distribution of the publications dealing with IEEE 1599 (orange), MEI (yellow) and MusicXML (green).

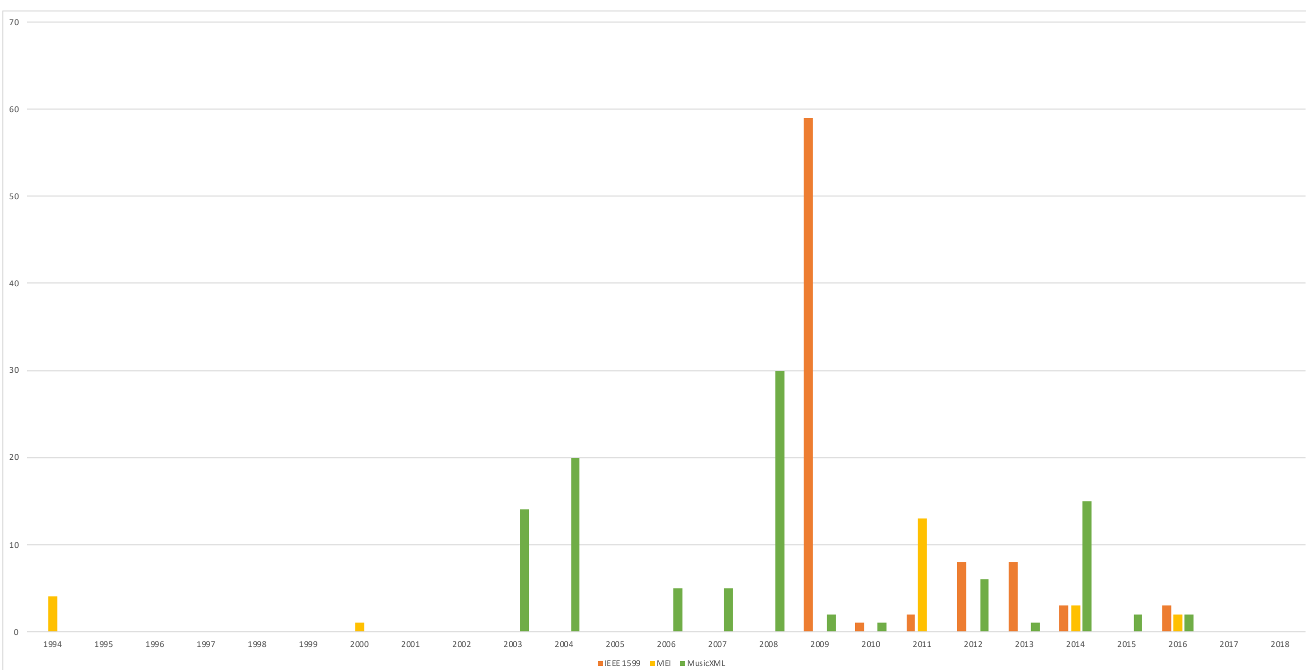

Fig. 4. Number of citations for the publications dealing with IEEE 1599 (orange), MEI (yellow) and MusicXML (green).

centered approach (i.e., building music notation on top of existing Web standards such as SVG, Web MIDI, Web Audio, etc.), has been rejected since, according to the group's co-chairs, efforts should rather focus on the development of MNX. The real risk is that the W3C initiative, born to encourage peer discussion on the best Web standard to be adopted, suffers too much the influence of MusicXML supporters and turns into a standardization of a revised version of MusicXML.

V. The Future of Multi-Layer Music Representation

From the analysis of the state of the art on XML-based multi-layer formats, different directions for future work emerge. The main goal is to involve a wider group of experts from industry, research, and academia (including musicians, musicologists, computer scientists, software developers, marketing experts, etc.) in the processes of standardization, so as to be inclusive with respect to other music encoding strategies and communities. It is difficult to assume that, in the future, the three formats will disappear or converge into a single standard, since they embody different visions of music description, even if some efforts are currently on the way.

First, the already mentioned activities of the W3C Music Notation Community Group are aiming to identify a commonly-accepted standard for the Web representation of music. Currently, MusicXML 3.1 is emerging as the best candidate to represent the logic and notational layers within such an initiative, whereas 2019 will be devoted to a significant improvement of the MNX format and, presumably, to the development of related software tools. In this framework, MEI's (and other groups') contributions are fundamental to provide views different from MusicXML's one, mainly focused on visual rendering of notation.

Concerning the IEEE initiative, 10 years have passed since the standardization of IEEE 1599 occurred in 2008. In the meanwhile, technology has evolved and other multi-layer encoding initiatives have appeared. During this period, both requests for improvement in the definition of the standard and new needs emerged. Consequently, the aim of the reborn working group is to analyze and adjust a number of issues so as to significantly extend the representation possibilities offered by the standard. A clarifying example is the proposal of generalizing the organization of layers, letting the user customize them. By overcoming the original 6-layer structure, IEEE 1599 should solve some logical problems (e.g., how to scaffold a Max/MSP patch or Csound code) and provide support to representation domains that are currently unpredictable. Another key goal is to allow multiple formats to describe music symbols in the logic layer, so as to potentially embed MusicXML and MEI scores. This would mark a clear difference with respect to these formats, extending the field of applicability also to domains other than music.

References

[1] J. Cage, *Notations*. Something Else Press, 1969.

[2] K. Stone, *Music notation in the twentieth century: a practical guidebook*. WW Norton New York; London, 1980.

[3] E. Ulman, "The music of Sylvano Bussotti," *Perspectives of New Music*, pp. 186–201, 1996.

[4] A. Lindsay and W. Kriechbaum, "There's more than one way to hear it: multiple representations of music in MPEG-7," *J. of New Music Research*, vol. 28, no. 4, pp. 364–372, 1999.

[5] J. Steyn, "Framework for a music markup language," in *Proc. of the 1st Int. Conf. on Musical Application using XML (MAX2002)*. IEEE, 2002, pp. 22–29.

[6] G. Haus and M. Longari, "A multi-layered, time-based music description approach based on XML," *Computer Music J.*, vol. 29, no. 1, pp. 70–85, 2005.

[7] T. Bray, J. Paoli, C. M. Sperberg-McQueen, E. Maler, and F. Yergeau, "Extensible markup language (XML)." *World Wide Web J.*, vol. 2, no. 4, pp. 27–66, 1997.

[8] G. Castan, M. Good, and P. Roland, "Extensible markup language (XML) for music applications: An introduction," *The Virtual Score*, vol. 12, pp. 95–102, 2001.

[9] P. Roland, *Design patterns in XML music representation*. Johns Hopkins University, 2003.

[10] D. L. Baggi and G. M. Haus, *Music navigation with symbols and layers: Toward content browsing with IEEE 1599 XML encoding*. Hoboken: John Wiley & Sons, 2013.

[11] P. Roland, "The music encoding initiative (MEI)," in *Proc. of the 1st Int. Conf. on Musical Applications Using XML (MAX2002)*. IEEE, 2002, pp. 55–59.

[12] M. Good *et al.*, "MusicXML: An internet-friendly format for sheet music," in *XML Conf. and Expo*, 2001, pp. 3–4.

[13] M. Good and G. Actor, "Using MusicXML for file interchange," in *Web Delivering of Music, 2003. 2003 WEDELMUSIC. Proc. 3rd Int. Conf. on*. IEEE, 2003, p. 153.

[14] G. Haus and M. Longari, "Towards a symbolic/time-based music language based on XML," in *Proceedings of the 1st International Conference on Musical Application using XML (MAX2002)*. IEEE, 2002.

[15] L. W. Barton, "The NEUMES project: Digital transcription of medieval chant manuscripts," in *Web Delivering of Music, 2002. WEDELMUSIC 2002. Proceedings. Second International Conference on*. IEEE, 2002, pp. 211–218.

[16] P. Bellini and P. Nesi, "WEDELMUSIC format: An XML music notation format for emerging applications," in *WEDELMUSIC*. IEEE, 2001, p. 0079.

[17] L. A. Ludovico and D. A. Mauro, "Sound and the city: Multi-layer representation and navigation of audio scenarios," in *Proc. of the SMC 2009 - 6th Sound and Music Computing Conf., 23-25 July 2009, Porto - Portugal*, A. Barbosa, F. Gouyon, and X. Serra, Eds. Porto: SMC, 2009, pp. 19–24.

[18] A. Baratè and L. A. Ludovico, "Local and global semantic networks for the representation of music information," *J. of e-Learning and Knowledge Society*, vol. 12, no. 4, pp. 109–123, 2016.

[19] A. Baratè, G. Haus, L. A. Ludovico, and D. A. Mauro, "IEEE 1599 for live musical and theatrical performances," *J. of Multimedia*, vol. 7, no. 2, pp. 170–178, 2012.

[20] A. Baratè, M. G. Bergomi, and L. A. Ludovico, "Development of serious games for music education," *J. of e-Learning and Knowledge Society*, vol. 9, no. 2, pp. 93–108, 2013.

[21] L. A. Ludovico and G. R. Mangione, "An active e-book to foster self-regulation in music education," *Interactive Technology and Smart Education*, vol. 11, no. 4, pp. 254–269, 2014.

[22] L. A. Ludovico and C. Zambelli, "Web-based frameworks for CLIL in primary school: Design, implementation, pilot experimentation and results," in *Computers Supported Education - 8th Int. Conf., CSEDU 2016, Rome, Italy, April 21-23, 2016, Revised Selected Papers*, ser. Communications in Computer and Information Science, G. Costagliola, B. M. McLaren, J. Uhomoibhi, and S. Zvacek, Eds. Springer, 2017, vol. 739, pp. 139–158.

[23] A. Baratè, G. Haus, and L. A. Ludovico, "A critical review of the IEEE 1599 standard," *Computer Standards & Interfaces*, vol. 46, pp. 46–51, 2016.