BMML: Braille Music Markup Language

Benoit Encelle*, Nadine Jessel, Josiane Mothe, Bachelin Ralalason and Javier Asensio

IRIT-CNRS, IUFM, Université Paul Sabatier, 118, route de Narbonne, 31062 Toulouse Cedex 4, France

Abstract: Thanks to the WAI (Web Accessibility Initiative) guidelines for producing accessible HTML documents, visually impaired people can have better access to a lot of textual information. Concerning musical score, several encoding formats are available, focusing on the representation of different aspects of this kind of content. As XML is the standard for exchanging content through the Web, several XML applications have already been specified for representing musical scores, using the traditional music notation. As a result, users can access and share a lot of different types of musical content using the Web. However, for specific notations – like the Braille one - no dedicated XML application has been developed yet. Therefore, visually impaired musicians cannot easily represent, share, and access scores using the Web. This paper presents the application we have developed to respond to this need: BMML (Braille Music Markup Language). BMML handles specificities of Braille Music notation and takes into account the core features of existing formats. The main objective of BMML is to improve the accessibility of Braille musical scores.

INTRODUCTION

There are approximately 161 million visually impaired people in the world, according to the European Blind Union [1]. For this population, it is hard to edit, access, and read musical score. For activities like word processing, e-mailing, reading electronic publications, getting information, many blind people already use computers as often as sighted persons. However, using computers in the musical domain is not yet common among the blind musician community.

Firstly, blind musicians need to produce musical scores. Such cases occur, for example, in a learning situation when a blind music teacher wants to produce graphical scores for sighted students or, when blind students want to modify specific scores or produce an exam in a graphical form for their sighted teacher.

Secondly, blind people need easy access to Braille musical scores. Even if several Braille libraries exist (e.g. AVH in France, Biblioteca Italiana per Ciechi, monza – Italy, The National Library Service for the Blind and Physically Handicapped - Library of Congress - Washington, USA), Braille musical scores are neither well referenced nor structured. As a consequence, Braille scores or parts of scores are difficult to retrieve, even on the Web.

Thirdly, the transcription of printed musical scores is time consuming and interests transcribers less and less. Even though some tools exist to automatically transcribe musical scores into Braille ones (Toccata [2], BME [3], GOODFEEL [4]), each of them uses a proprietary format. As a result, these Braille scores are not easily transformable and exchangeable. As a result, there is an urgent need to define a unique format that could be used for representing Braille musical scores, improving score exchange possibilities between blind musicians.

*Address correspondence to this author at the IRIT-CNRS, IUFM, Université Paul Sabatier, 118, route de Narbonne, 31062 Toulouse Cedex 4, France; E-mail: encelle@irit.fr

The European project Contrapunctus [5] was created to address these issues and one of its goals is to develop a unique format for encoding Braille scores, taking into account all specificities of the Braille music notation. As a result, an XML-application called BMML (Braille Music Markup language) has been designed to fulfill previously mentioned requirements.

In this paper, we firstly discuss related works in the following section and more precisely recommendations for representing musical information and relevant existing formats for encoding music. We highlight the peculiarities of Braille music notation and needed metadata in opposition to the traditional style. The guidelines we develop to handle these peculiarities are presented (BMML schema). We illustrate our model through the example of a Braille score XML document using the BMML schema we proposed. Finally, we conclude this paper and presents future works.

RELATED WORKS

Score and Music Standard

A musical score is a document that contains all musical information about a given piece of music. Musicians can write or read a musical score in order to play and perform the written music. Using scores, musicians communicate, share, learn and compose music. Various standards have been defined to help the development of language that handles musical content, their representation, and their relationship. Among musical standards, SMDL (Standard Music Description Language) [6] and SMR (Symbolic Music Representation) [7] are the most important. We give details about these standards in the following sections.

Music Standard

SMDL

The current existing musical formats follow some of the main orientations (i.e. domains) which are defined in the Standard Music Description Language, namely:

1874-1339/09

- *Logical* domain which contains the music;
- *Gestural* domain which is the current performance;
- *Visual* domain: the graphic visual display of the musical work, and
- *Analytical* domain which consists of the theoretical analyses and commentaries.

SMR

Symbolic Music Representation (SMR) is a set of recommendations to represent musical information. These recommendations suggest the representation of several aspects (named domains) for the encoding of musical information: a content-oriented domain and a presentation-oriented one. Concerning the presentation aspect of musical information, even if one recommendation considers the accessibility for blind people; it does not define how to implement this accessibility.

Music Encoding Format

There are a lot of music encoding formats, the most usual and typical are presented hereafter.

MusicXML [8] is designed to be a universal translator and an interchange format for common Western musical notation from the 17th century onwards. As it is usable by a lot of notation programs, sequencers and music performance or education programs, it has become a de facto standard.

NIFF Notation Interchange File Format [9, 10] is a file format designed to encode in a very precise way the graph used to present the musical score, so it permits the interchange of music notation data between music notation editing, publishing programs and music scanning programs.

MIDI (Musical Instrument Digital Interface) [11] is a music industry standard communications protocol that lets MIDI instruments and sequencers (or computers running sequencer software and equipped with a MIDI interface) talk to each other to play, edit and record music. MIDI is a code that is designed to produce sound.

PLAY Code [12] aims at offering the opportunity of exchanging information with sighted musicians thanks to the Braille Music Editor (BME).

The PLAY Code is a proprietary code and cannot be easily reused and extended. Musicxml, Niff and Midi don't take into account Braille music notation. These reasons justify the specification of a new format for encoding Braille musical scores.

BRAILLE MUSIC NOTATION

Peculiarities of Braille Music Notation

Braille music notation [13] is very different from ordinary music notation. A Braille score consists in a linear sequence of symbols representing notes, chords, parts, and rhythmic patterns. Due to the limited number of Braille signs available (64 in total) musical elements are produced using a combination of one or more Braille signs. So the meaning of each Braille sign is determined by its context.

Additionally, Braille is read in a linear way. When considering a musical score, this type of reading is not fully relevant. Braille writing has developed a lot of strategies and some special symbols to reduce the text length and to make it easier to read.

Hereafter is an outline of some specific Braille concepts that do not exist in printed notations and that must be taken into account when defining the code of Braille musical score.

Repeats

A major specificity of Braille music is the extensive use of repetition signs, compared to printed music. This can simplify reading, assist memorization and save space.

For example, in Braille, a special character (dots 2356) represents a full- or part-measure repeat.

Sequences

In order to reduce reading time by having fewer characters, a sequence of similar elements is usually written in Braille by doubling the element at the beginning of the sequence and repeating it at the end. For instance, a sequence of two-notes chords, that all have the same 3rd interval, can be written by doubling the interval sign after the first note and putting it after the last note. As a result, for each chord (except the first and last chord) of the sequence, its second note (the 3rd interval) is not written. The same applies for a sequence of rhythmic groups.

In the case of a succession of several identical rhythmic groups, it is possible to double the grouping sign before the first group.

Octave Specification

When we use in a *Part* a modification of octave, the first note concerned has to be preceded by two octave marks. The first mark indicates the value of the octave according to the position of the note on the printed score; the second indicates the real value of the octave.

Chord, Note and Rest Durations

For a note, the dots 1, 2, 4, and 5 of the Braille character represent its pitch. The absence or presence of dots 3 and/or 6 determines its duration. Each note or rest has two duration value possibilities determined using its context. Some Braille characters can prefix a note to precisely indicate its duration.

In addition, for chords in which the notes have the same duration, only one note is written. The others are indicated by their intervals from that note.

Use of In-Accords

In Braille, musicians can only read horizontally. Therefore, vertical information must be provided as an horizontal sequence of characters. When all of the harmonic parts do not change at the same time, they are shown by dividing the measure into voices and are made by making use of *in-accords*, which are unknown in printed scores. This symbol indicates that the following notes belong to another voice belonging to the same measure.

Key Signature

Key signatures reflect the number of flats or sharps, not the actual pitches as in printed scores.

Slurs and Ties

In printed scores, all slurs and ties are represented by a similar line over or under the relevant notes. In Braille, various characters are used according to the context. Among these contexts, we can quote:

- Slur between two notes or chords,
- Phrasing slur over more than four notes or chords,
- Beginning and end of phrasing slur on one note
- Beginning and end of short slur on one note
- Slur from one in-accord part to another
- Straight line between staves for voice leading
- End of straight line
- Slur added by an editor in printed scores
- Slur that does not end on a note
- Slur for short appoggiatura;

Layout

The main difference with music in print is the concept of spatial dimensions. For music in print both the dimensions are used to convey information. In Braille both presentations are available:

- section by section layout where a group of measures of a part or of an instrument alternate with the same group of measures of another part or instrument. The number of measures in a group is defined by the transcriber and stored in the metadata of BMML.
- bar by bar layout where a measure of a part or of an instrument alternate with a measure of another part or instrument.

The information of this layout is defined in the metadata of BMML and a transformation process could propose the right presentation.

METADATA AND BRAILLE MUSIC SCORES

Visual Impaired Special Needs

In an educational or professional context, blind musicians need to easily communicate with sighted ones. Because both of them use different notations (i.e. Braille and printed notations) for the same content (a musical score), correspondence indications between these two notations should be added. For instance, the printed version of a Braille score has to be explicitly indicated in a Braille score. Indications to easily find a corresponding audio version have to be added too. As a result, specific metadata for visually impaired people have to be represented in addition to the classic ones (e.g. title, composer, etc.) that are already described in some of existing formats (e.g. MusicXML for instance).

The aim of this section is to show the most relevant information concerning the metadata that should be associated with Braille music scores. This should be the background for improving the metadata structure in the Braille Music Markup Language, and it should be also the base for the design of an online library for storing Braille music scores coded in BMML. During this section some suggestions about the use of the metadata will be highlighted, security and metadata processing are subjects that may help us to obtain a more complete metadata structure.

Traditionally regular printed scores have been stored in libraries. During decades different associations and institutions have made big efforts to define different sets of data for identifying and classifying these resources. Obviously, this work is highly useful for our needs and it has been analyzed in the first sub-section.

The next two sub-sections are focused in metadata in the electronic domains. First, several electronic formats for storing music scores are analyzed focusing in their metadata structure. After studying several sites which offer music scores (libraries and shops), a summary of the metadata is shown. In this way, we want to show the most commonly used fields for describing score characteristics.

After analyzing the domain of music scores for sighted people, an analogous study is performed for the Braille music score one. Data of several important libraries for the blind and documents from former projects are gathered to illustrate how the metadata for these resources is structured now.

There are several interesting studies about including as metadata some information extracted from the score itself. Farther than tonality, other musical features can be harvested and used for performing searches among score data bases.

Two sub-sections are focused in other topics that may be interesting for our purpose. An explanation about the semantic web and how this technology may help to enrich the whole project is given. After that, some hints on data integrity are detailed for complementing this section.

Last sub-section is the current BMML metadata proposal. The principal troubles are highlighted here and some possible solutions are proposed.

Standards and Recommendations on Metadata and Cataloging Rules

As for any other type of resource, the reason to define a metadata structure and to attach it to a music score is to provide an useful artifact to develop an effective data management. The task of designing a complete and usable metadata structure is something fundamental for achieving good performance and accurateness in processes like data classification or information retrieval.

There are several recommendations and standards on how these metadata sets should be settled, which fields should be present to identify any kind resource and give useful information about it. Other feature broadly needed in using metadata is to provide the necessary tools to relate a given resource with others, allowing to classify or locate the resource taking into account several of its characteristics.

One of the most well-known recommendation is the Dublin Core Metadata Initiative (DCMI) [14], which became an ISO standard in 2003. This initiative is settled upon fifteen metadata elements used for cross-domain resource description. A great amount of online catalogs and information repositories have adopted this standard to easily locate, share and manage their information resources.

Several universities from the United States have created The Sheet Music Consortium [15], which goal is to build an open data base of digitized music scores based on the Open Archives Initiative for metadata harvesting. Superficially, the project consists in the publishing by several universities of their music score collection metadata, so other parties can access freely to this information. The protocol to share the information relies on the DC elements for identifying every score.

There are several other standards and recommendations proposed by different associations and librarians committees. The IFLAA [16] (International Federation of Library Associations and Institutions) defined in 1968 several International Standard Bibliographic Description (ISBD) for different proposes [17], being in 1987 when a ISBD for music scores [18] was released stating a set of core elements to identify any material. This standard specifies eight groups of metadata for describing a resource:

- Title and statement of responsibility area
- Edition area
- Printed Music Specific Area
- Publication, distribution, etc., area
- Physical description area
- Series area
- Note area
- Standard number (or alternative) and terms of availability area

The library of congress of the United States, The Library and Archives of Canada and The British Library have a common format bibliographic standard known as MARC21 [19]. This general purpose format covers different kinds of resources, from books to computer files and music. As the ISBD, the entries for every record in MARC21 are grouped in several functional categories.

- Control information, identification and classification numbers, etc.
- Main entries
- Titles and title paragraph (title, edition, imprint)
- Physical description, etc.
- Series statements
- Notes
- Subject access fields
- Added entries other than subject or series; linking fields
- Series added entries, holdings, etc.
- Reserved for local implementation

All these cataloging recommendations are totally generic and some features may apply to electronic music scores for blind people and some not, but it is a good idea to abstract the goals of these designs and use them in the design of any kind of metadata focused in similar targets.

As a last recommendation review, a reference must be done to the International Association of Music Libraries [20] (IAML) and its Core Bibliographic Record for Printed Music [21] (CBRPM). The IAML is a member of the already mentioned IFLA, and its goal is to promote international cooperation and to support the interests of the music librarians. The last review for the CBRPM was released in 1996. Its set of bibliographic elements is based on the MARC standard, but some features for the score music domain are added. The main structure of this record is set as follow:

- Control information, identification and classification
- Main entries
- Title section
- Publication, distribution and edition statements.
- Notes and Added entries (Focused in the music score features).

All these cataloging techniques are widely used in physical music libraries all over the world. A deeper study has been already made in the adoption of this standards in the HARMONICA [22] European project where in the deliverable D.1.2.1 a relation between libraries and cataloging rules is settled. Most of them use either ISBD or the AARC2 [23] (an analogous cataloging system used in the US and UK), with some adaptations to the country rules.

Electronic Formats for Sighted People Scores

There have been a lot of attempts to design a complete and universal music score sheet format for printing, sharing, exchanging and editing music scores using a computer. Although there are dozens of programs which allow to perform these actions, no agreement is currently achieved among musical software companies about a common file format. In these last days it seems that Musicxml [8] is arousing as the new standard and the support for this format is spreading through musical software.

Now let us analyze several formats to store music scores electronically, some of them are completely outdated and others are the future in music encoding. This study is always driven by the metadata features included in these formats.

During the last years the most extended format was (and, perhaps still is) the Musical Instrument Digital Interface also known as MIDI [11]. This format does not allow to store in MIDI files any kind of metadata related with the score contained.

For more than ten years different researchers and institutions have tried to build new formats to store music scores. Most of them have disappeared but others are still in use, and some of them have made real efforts to introduce metadata in their score descriptions.

Two plain text music score formats are MuseData [24] and **kern [25] representation. MuseData (1993), apart from the musical representation, allows to store several useful information as the date on which the file was encoded and the name of the person that carried out the process, work number and movement, title and source. Although it is not based on any metadata recommendation, it is a first approach to add information to the score in its electronic representation.

The **kern representation is part of the wide-proposed project Humdrum [26]. As part of the project, a formal grammar is defined for representing sequential symbolic information known as the *humdrum syntax*. **Kern representation scheme is a concrete use if this abstract grammar. Although the main target of this language is to represent the syntactic information inside a music score, it allows to store several complementary information about the music work. This information or metadata is divided in different groups [27]:

- Common bibliographic codes as Composer, date of composition and Title.
- Analysis related codes:
- Composer related codes.
- Edition related codes.
- Lyricist related codes.
- Performance related codes.
- Work (opus) related codes.
- Publication related codes.
- User related codes.
- Scholarly related codes.

As in many other fields the tendency to represent music notation has been gradually moving towards mark-up languages formats. An outdated music representation which began the effort of using a mark-up language was the Standard Music Description Language (SDML) [6] based on the XML's parent SGML. The first draft of this meta-language targeted to music representation appeared in 1988 but it was not until 1995 that the final draft was stated. SMDL is a meta-language designed for the musical needs, so there is no standard definition in which metadata is attached to the scores. SMDL architecture was derived from HyTime (a technology for metadata and accessing hyperlinks similar to XPath) and it is easy to obtain useful information performing queries against the SDML documents.

The project Web Delivering of Music (WEDELMUSIC) [28] started in the year 2000 had as main goal the establishment of a framework for distributing and sharing music scores through internet, being deeply respectful with the publisher's rights and protecting them from the copyright violation. All the information of the scores is stored in a XML document which specification is open and available in their web site. Inside the WELDELMUSIC file a variety of metadata can be specified:

- Identification: ISBN, ISMN, Catalogue identification, publisher, date, distribution information.
- Classification: Several fields for classifying the work as tonality, author, title and parallel title, act, movement, ensemble, genre, style, propose, original language, duration, composition date, epoch, start year, end year, composition location, nationality, dedicated to, commitment, first execution, review description, Performers and even critical reviews.
- Protection: Result of hash functions for watermarking the score.
- Media attached to the score: Images, audio and video files attached to the score.

The most spread electronic music score format, said to be the standard "de facto" format, is the MusicXML. From a metadata point of view, MusicXML has a simple metadata model that allows to store information about the creator of the work, rights, encoding information, source of the score, relations with other works and miscellaneous information. The author of MusicXML highlights that some of these elements are taken from the Dublin Core, although no identification slot is included and there is a lack of fixed structure for other interesting information.

Another XML approach for storing music scores proposed from the North American University of Virginia in 2002 is known as Music Encoding Initiative (MEI) [29]. This format is analogous to the Text Encoding Initiative: a standard for representing texts in a digital form but music-scored oriented. It is pretty clear that the author made an effort to include a complete metadata section inside the file schema. A great variety of data can be detailed inside this section and the structure has been thought to follow several international standards on cataloguing and resource description as the ISBD, AARC2 and Dublin Core.

The metadata header includes a complete bibliographic description of the score including title, responsible of the work, edition issues, publication information (data, localization, identifier, copyright, etc.), information about the series this work belongs (if any), extent information as number of pages, a finger print and complementary notes. The MEI format allows to append also information about the propose with the score was encoded, editorial description, non-bibliographic issues (e.g. languages, historic aspects, music classification) and a revision history of the MEI file. Also it is possible to specify the complete bibliographic reference of the source work from this electronic file has been encoded.

Existing Libraries of Music Scores for Sighted People

On the internet a big amount of public libraries, music score shops, and other music storage projects can be found. The next step consists in analyzing some of them to extract the metadata involved in the querying process and the metadata stored in the records.

Most on-line music stores have a completely lack of structured searches. The site *sheetmusicarchive.net* and. *music-scores.com* do not have a search box: the only way to find a score is to browse through every composer and select the desired score. Other sites like *musicsheetplus.com* or *stagepass.com* offer a Google-like search box: this kind of interfaces has been demonstrated to be very user-friendly but it is not possible to make tuned queries based on the features of a closed domain as the music scores one.

On the other hand, the on-line shop *sheetmusicdb.net* has a very interesting search interface. In spite of offering a quick search box, a more detailed query can be performed. The allowed fields are: title, publisher, year of publication, series title, composer, arranger and performer and other additional information. If the scores are purchased for educational purposes, a combo box for selecting the difficulty level of the score and possibly a selector of the evaluation levels of many countries are provided.

A site which has been up for a long time is *musicanet.com*. This site is focused in choral music and offers more than 150,000 electronic score sheets. The search options allow the user to choose among more than 25 search criteria. Most of the parameters are targeted only for choir music like *difficulty for the singer*, *type of choir* and *voicing*. The complexity of a sub-genre of the domain is showed when an exhaustive query is wanted to be performed. As a conclusion, some of the fields used here may be taken into account in order to design a general and complete metadata set. Other feature highlighted from this web site is their customizable search engine. The user can add any criteria from the wide set available, allowing a flexible and accurate search.

Apart from commercial online store, we have to focus on the music libraries all around the world. An interesting initiative is The International Music Score Library Project (imslp) [30] that aims at collecting every public domain score and creating an open database. The main trouble with this site is that they follow a wiki-like structure. This kind of sites was not designed for performing detailed searches: you have to know the name of the score you are looking for or its author's name. They attach the following metadata for every record stored:

- Work Title
- Composer
- Opus/Catalogue Number
- Number of Movements/Sections
- Dedication
- Year/Date of Composition
- Year of First Publication
- Genre
- Librettist
- Language
- Piece Style
- Instrumentation

Several databases take advantages of some of the cataloging standards mentioned before as the MARC21 one used in the California Sheet Music Project [31] (university of California) and in the university of Indiana's project Variations2 [32].

Blind People Libraries and Formats

Several libraries for the blind have been analyzed to detect the metadata stored and attached to Braille scores. As a result, most of the libraries do not have any specific data structure for scores: in most cases they are treated as regular Braille books and only information about the author, title, publication, description and media type are detailed.

A very interesting work in analyzing cataloging needs of Braille music scores has already be done during the MIRACLE project [33]. MIRACLE project results from the collaboration of seven important partners: FNB, SBS, ONCE, RNIB, DBB and Stamperia Braille. All partners agreed in a consensus to set the necessary fields for cataloging Braille music scores. Two groups of field were defined:

- Ink Print related fields:
  - Name/Author
  - Uniform Title
  - Main title
  - Other titles
  - Form
  - Key
  - Instrument set
  - Place of publication
  - Date
  - Place of performance
  - ISBN
  - ISMN
  - ISSN
- Braille information
  - Availability
  - Braille Conventions
  - Medium type
  - Braille music pres (braille format)
  - Languages
  - Grade
  - Lines per page
  - Characters per line
  - Braille Character set
  - Producer
  - City and Year of Production.
  - Production Number
  - Status
  - Completeness
  - Currency
  - Price
  - Archive file
  - File type

These fields arose as an agreement among the most important librarians for the blind in Europe and have to be taken into account when designing the metadata set for the BMML format.

Metadata Harvested from the Music Data

Other projects focus their efforts in extracting metadata directly from the score itself: the most complete example is the Stanford University's *themefinder.org*. This web page allows to search in their scores database, setting query parameters. Query parameter can be for example an absolute

note sequences (like "ABBC#G"). Many other querying possibilities exist. The interval search for instance needs a query where an interval sequence is detailed through three elements: the interval direction (rising or falling), the interval quality (as minor, minor or diminished) and the interval size. As a result, it is not necessary to know exactly the notes of the tune but only the skeleton of the melody. It is also possible to describe the melody in terms of scale degrees indicating the position of the note in the scale through numbers from 1 to 7. But the most interesting of all these special queries are probably the *contour searches*: the *themefinder* engine find matches among the scores stored with a description of the melody made in terms of ascending and descending intervals.

Many other researchers have made similar proposals to *themefinder*: for instance, the "Context- based indexing of music scores" [34] use a relative interval search similar to the *contour search* but starting from a recorded hummed melody from the user. The idea is based in making an index of melody shapes which represent how the music line evolves during time. When a user hums the melody to be searched, a shape is taken out from the wave recorded and it is matched against the index stored.

This kind of metadata may be hard to extract and require a lot of time and resources. As a consequence, an in depth study has to be performed concerning the feasibility and future possible usages given by this kind of service.

Other Issues Concerning Music Metadata

Two concepts that may be introduced together are the semantic web and the music ontology specification. The semantic web (SW) [35, 36] is an attempt to enrich the current world wide web adding metadata relating to many concepts as possible. In this way, many of the tasks that are performed nowadays manually by users might be performed automatically. Several technologies are the foundations of the SW: XML provides the syntax, the Resource Description Framework (RDF) provides the data model for describing the relations between concepts based in a subject-predicate-object structure and the Ontology Web Language (OWL) can be used for describing attributes and classes in order to add vocabulary to the whole. Using these tools, anyone can publish data as usual but can also add semantics that may be used by automatic agents.

The SW was conceived in 1999 and, although its use grows constantly and several efforts have been made by the World Wide Web Consortium, it does not seem to settle as a widely used technology. There are anyway several examples of SW used in the musical domain: the most famous is the web site *musicbrainz.org*. This site is an open music metadata base where the data is maintained by users. Information about bands, songwriters, composers is stored and widely related among each other. The interesting feature is that *musicbrainz.org* publishes this data in the RDF format and can be freely used by external tools or sites.

Concerning the SW field, an effort has been made to specify in the Music Ontology Specification (MOS) [37] all the vocabulary needed to define any music work or performance. By having a common vocabulary, different systems should be able to communicate with each other, to share data and harvest information upon this data. The big trouble with the MOS is that it is not a standard. As a consequence, only some systems currently use this technology.

Security and Metadata

Security is commonly split into confidentiality, integrity and availability. Preserving the confidentiality and the availability are subjects out of the scope of this section but a study must be performed concerning integrity. A great interest of preserving copyrights and authorship of the work has been shown so far, and some techniques can be used for this issue, which have a deep relation with metadata.

The application of checksum functions and digital signatures is widely used for validating the source of a document or its authorship, checking that no one has altered the content of the data. The World Wide Web Consortium have developed a recommendation for including checksums and signatures into an XML document called "XML Signature Syntax and Processing" [38], even namespaces for including signatures in a standard way are open and available in the recommendation. A common procedure for signing BMML documents could be the following: the transcriber finishes a Mozart Braille score and exports it to BMML, he also introduces a signature using the private key of his institution and finally he uploads the score to the Contrapunctus library. Any user with access to the work may use the public key to validate the authorship of work.

The former method has an important weakness, some malicious user may get rid of the signatures tags of the file and change a couple of tags without modifying the effective content of the score, and no one could probe the real authorship of the document. For this reason it would be better to use a watermarking method. Watermarks are slight changes in a digital file data which insert some hidden information as a digital signature or the result of a checksum function. If the method is good enough, this information will remain even if someone modifies the file. In this way the authorship of a score could be demonstrated even if someone has introduced alterations.

Watermarking is widely used on audio, video and image files. Applying this method on text files is rarer but there are several examples [39]. Over XML files the process is quite more complex, the closed structure of these files makes that even little modifications will corrupt the entire file. Several general purposed methods for watermarking XML have been developed [40, 41] but none of them seems to be applicable for all domains.

Current BMML Metadata Structure

The current BMML metadata structure is based on the DC and is complemented with several fields of the MusicXML schema. However, there are several points that should be improved to have a complete and useful metadata set. This sub-section aims to be an analysis of the current metadata set definition (cf. Table **1**).

Outside the already present elements, the ISDB recommendation for printed music defines an interesting section for music related issues. Some elements of this section are the piece tonality, instruments, number and kinds of singers, etc. Obviously this set must be flexible enough for adding

Table 1. Current Proposal for BMML Metadata Elements

BMML Element	BMML Labels	Comments
Identifier	Identifier	No comments
Title	Title	Many musical works have different or complementary names, or maybe the name has been translated to any language and the original one must be pointed out. So a field for parallel titles and other title information may be included as the ISBD recommends and MIRACLE project includes as part of its meta data.
Subject	Subject	No comments
Creator	Creator.Author Creator.Composer Creator.Arranger	Three types of responsible can be stated: composer, author and arranger. However an unlimited number of authorships may be attached to a music work as lyricist, librettist, etc. The BMML metadata must give support to a wider responsible statement, and although some fields may fixed, the possibility to add non predefined author types has to be available.
Format	Display.Symbol-number-in-a-line Display.Line-Number-In-a-Page Display-Page-Number	The format element is well defined as it gives support to Braille score features not present in other domains as it was proposed in the MIRACLE project. But a field for indicating Braille conventions may be added. In different countries the may use Braille signs for slightly different proposes, and a transcriber may want to add this information to his work.
Source	Print-Link Audio-link Braille-link Album.Title	This has to be definitely improved is the *source* label. Most of the BMML scores will be transcriptions from another resource to braille, so detailed information about the source work is extremely useful. MEI treats the source as an bibliographic element itself, and a complete ISDB record is attached to this element. It would be useless to add the information regarding to score already contained in the file as Title or authors if the do not change, but editorial and publication statements must be included for every score. The elements outside the source element must refer to the file itself, publication info, copyright, format, etc must be information about the braille score, but the same information regarding the source will be set inside the element.
Publisher	Album.Publisher	Publisher statements should be more complete, adding information about dates, places and responsible agents of the publication as the ISDB recommends.
Rights	Rights.copyright Rights.license	Some kind of method for assuring the data integrity must be included in order to preserve the work copyrights.
Encoding	Encoding.Date Encoding.Software Encoding.Encoder Encoding.Description Encoding.Contributor Encoding.Supports	The encoding element is taken directly from the MusicXML metadata definition, it adds some useful information but it should be redefined for several reasons. First a revision history of the file may maintained inside the XML file itself with all the relevant information, so changes in the life file could be easily traced. The information in every revision may be the current encoding element but some data about the transcription may be added as well.
Miscellaneous	Miscellaneous-field	No comments
Description	Description	No comments
Language	Language	No comments

yet unknown music features that certainly would enrich the BMML metadata.

And as general recommendation, the different metadata sets may be regrouped as defined in the ISDB or MARC21 recommendations, so they will be more easily accessible and they will have a clearer presentation.

A first approach to a more complete metadata set is described as follows:

- **Identification area**: In case a signature is included inside this file it should be located inside this element.
  - **Identification number**: Unique number to address this file.
- **Title and statement of responsibility area:**
  - **Title**: The chief title of this item.
  - **Parallel titles**: When on the prescribed source of information there are titles in more than one language and/or script.
  - **Composer:** Composer of the work.
  - **Other responsible agents of the work**: Other agents may be included as arranger, lyricist, etc.

- **Subject**: Topic of the resource.

- **Format Area**: In a braille context this information about the format must be included.
  - **Number of braille symbols in a line.**
  - **Number of lines in page**
  - **Total amount of pages.**
  - **Braille Conventions:** Here the transcriber must indicate the possible Braille conventions depending of the country, grades, etc.
- **Music Specific Area:**
  - **Music information**: Open music information elements, type and value are required for each one. As for instace type="tonality" value="A" or type="instrument" value="piano".
- **Publication area:**
  - **Copyright**: Description of the copyright of the BMML file.
  - **File history**: History of changes of the file.
  - **Encoding**: Every time someone modifies the file a encoding entry is added with the following data.
  - Date: date of the modification.
  - Software: Software involved in the process (may be several)
  - Encoder: Responsible of the encoding process (may be several)
  - Description: Description of the changes involved in this revision.
  - Contributor: Contributor in the modification of the file (may be several)
- **Source Area**: Information referring the resource transcribed to Braille or the Braille score encoded in BMML format. The following elements must not appear if their content does not change from the data of the BMML file itself.
  - **Identification area** : ISBN,ISMN or other unique identifier for the source.
  - **Title and statement of responsibility area**: Containing the title, parallel titles, responsible agents and subject.
  - **Edition Area:**
    - **Edition statement**: Statement referring to copies of an item formally identified as constituting a named edition .
    - **Statements of responsibility relating to the edition**: Responsible agents of the edition
    - **Additional edition statement**: An additional edition statement is given when the item carries a formal statement identifying it as belonging to an edition within an edition.
  - **Publication Area**: These elements may be optional depending the resource itself.
    - **Place of publication**: The place of publication is the name of the town or other locality.
    - **Name of publisher**: May be several publishers.
    - **Date of publication**: The date of publication or production of the item described is given.
    - **Place of manufacture**: The place of manufacture is given.
    - **Name of manufacturer**: The name of the manufacturer is given.
    - **Date of manufacture**: The date of manufacture is given.
  - **Notes Area**: notes about this resource
    - **Language.**
    - **Miscellaneous.**
    - **Description.**
- **Series area**: In case the work belongs to a series the following data must be included.
  - **Title of series**: The proper title of the series.
  - **Parallel title of series**: When the title appears in the prescribed source of information in more than one language and/or script.
  - **Other title information of series:** Other title information relating to the series.
  - **Statements of responsibility relating to the series:** A statement of responsibility can be given with respect to any entity responsible for or contributing to the creation of the series
  - **International Standard Serial Number of series**: The International Standard Serial Number (ISSN) relating to the series.
  - **Numbering within series:** The numbering of the item within a series.
- **Notes area**
  - Description
  - Miscellaneous
  - Language

A Markup Language for Braille Music

There is currently no unified format for encoding the Braille representation of musical score. Indeed, existing formats mostly concern the representation of printed notation and dedicated ones are proprietary, not extensible and not readable. With the help of XML technologies, the drawbacks of existing codes are missed. As a result, the BMML code is a solution to store and exchange Braille musical score.

BMML Description Logic

According to the specificities of Braille music, we treat a Braille musical score as a sequence of notes followed and preceded by describers corresponding to octave, tie, slurs, fingering, nuances, etc... as illustrated in the Fig. (**1**).

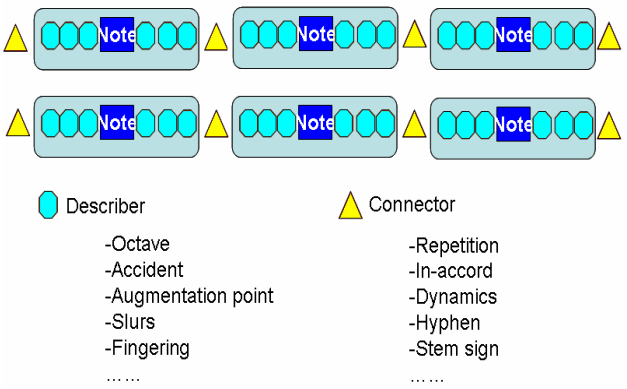


Fig. (1). Structure of a Braille score.

In the same way, a Braille score is considered as a sequence of measures which are interconnected by the means of measure connectors like repetition, In-accord, Dynamics, etc. symbols (see Fig. **1**).

Equally important, BMML takes into account the abbreviation used in Braille music in order not only to reduce the score size but also to facilitate its reading and memorizing. As an illustration, here are some examples of Braille abbreviation shapes.

Shape 1

The first type of shape expresses a sequence of actions which will be repeated a number of times (Fig. **2**).

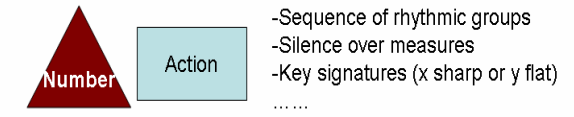
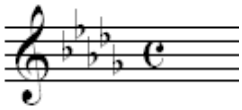


Fig. (2). A Braille abbreviation structure.

Example (Fig. 3)

If there are four or more accidentals in a key signature, the number sign is used instead of writing the flat sign many times.



#E<
Number 5
flats

Fig. (3). Key signature information, using Braille abbreviation structure.

Shape 2

Another type of shape, as illustrated in Fig. (**4**), permits the maintaining of a property on a sequence of elements by doubling the symbol of the property before the sequence and putting it only once after the sequence.

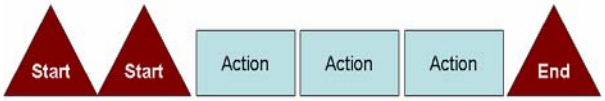


Fig. (4). Structure of a sequence abbreviation.

Example

The example below (i.e. Fig. **5**) shows how each note in the sequence is followed by the corresponding 3rd interval. Thus, the first note of the sequence is preceded by the couple of third interval symbols. Subsequently, the last note is followed by the third interval symbol.

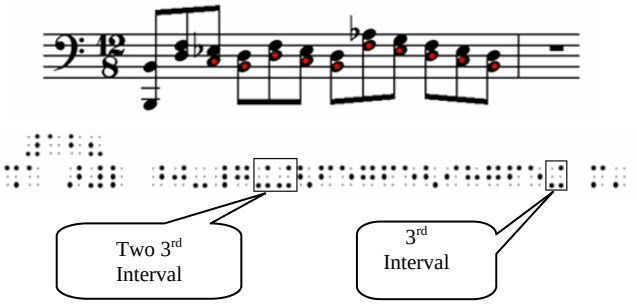


Fig. (5). Sequence of Third interval using sequence abbreviation structure.

It is worth reconsidering the example illustrated in Fig. (**5**) to show an important aspect of the coding process. This example can be Braille-coded in two ways. The first solution consists in using the abbreviation structure: only the information that is on the Braille musical score will be encoded. On the contrary, the second solution will encode all the notes. We have chosen the first option that encodes the abbreviated form, leaving to the reading program the generation of the printed content.

BMML Schema Definition

We developed the initial definition of the code using the W3C XML Schema specification [42].

According to our model, a Braille score is composed of metadata and one or several *Parts*. In addition to metadata that is already used in the previously mentioned formats, elements like: Print link, Audio link, ISBDPM (International Standard Bibliographic Description for Printed Music) [43] have been added to improve, on the one hand, communication between visually impaired and sighted musicians and, on the other hand, the understanding of the score. Finally, specific information regarding Braille score such as the line number on a page and the symbol number in a line have been added too. Such metadata is illustrated in Fig. (**6**).

We stress that in Braille music, a *Part* consists of one or several *Sections* which are composed of one or several *Measures*. Each section is characterized by its *Key*, *Key signature* and *Time signature*. Each measure can be connected to others using *Connectors*.

Notes element is a set of *Notes* which is, according to the writing rules of Braille music, preceded by *Pre-describer*

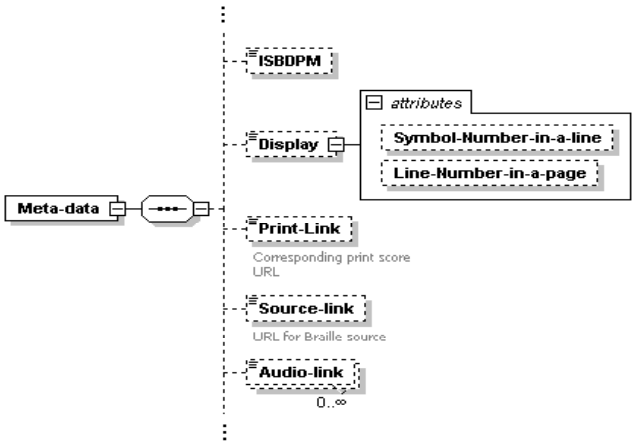


Fig. (6). Excerpt of the metadata structure.

and followed by *Post-describer* (cf. Fig. **7**). Moreover, in Braille score, *Note* and *Rest* are represented in the same way.

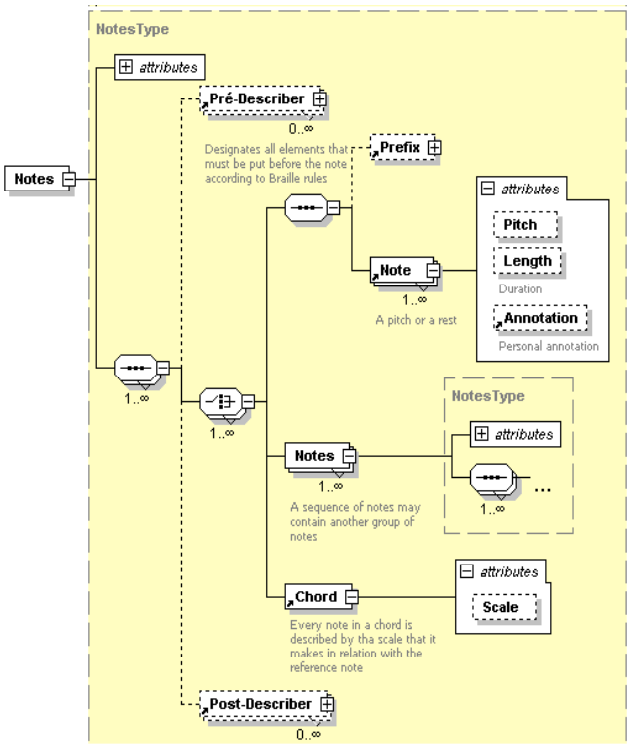


Fig. (7). Notes element Structure.

When the context of some Braille signs (e.g. note/rest durations) is not sufficiently clear to precisely determine their meanings, prefix elements are used (cf. Fig. **8**).

A Sample of a BBML Braille Score

An example of a score containing 6 flats in the key (which means that the musical score is in the tonality of G flat major) and a sequence of thirds is given in Fig. (**9**). This score is then transcribed into Braille using its BMML representation. In this example, we highlight the use of Braille abbreviation in both cases (Key signature and chords sequence) according to the Braille notation rules mentioned.

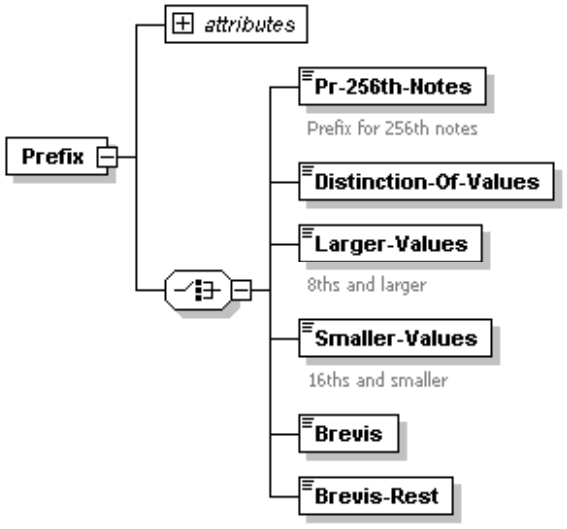


Fig. (8). Prefix element structure.

Both in Fig. (**9**) and in the XML document, the chord sequence and its corresponding Braille abbreviation are highlighted in green while the key signature and its corresponding abbreviation are highlighted in cyan.

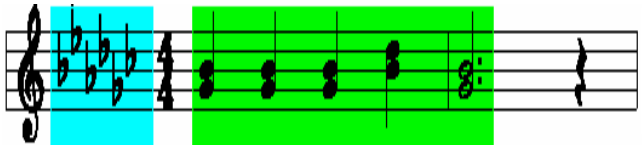


Fig. (9). A print score example. Example.

Here is the corresponding valid BMML score representation.

```
<?xml version="1.0" encoding="UTF-8"?>
<Score                    xmlns="http://www.punctus.org/bsml"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.punctus.org/bsml/bsml.xsd">

<Part Instrument="Piano" Number="1">
<Section Number="1" Movement="Moderato">
...
<Key-Signatures>
<Key-Signature Number="6" Accidental-Sign="Flat" Concise-
Presentation="true"/>
</Key-Signatures>
...
<Measure Number="1">
        <Notes>
            <Chord Scale="Third"/>
            <Chord Scale="Third"/>
            <Pré-Describer>
            <Octave Number="4" Display="true"/>
            </Pré-Describer>
            <Note Pitch="G" Length="Quarters-64ths"/>
            <Note Pitch="G" Length="Quarters-64ths"/>
            <Note Pitch="G" Length="Quarters-64ths"/>
            <Note Pitch="B" Length="Quarters-64ths"/>
            </Notes>
</Measure>
<Measure Number="2">
        <Notes>
```

```
            <Note Pitch="G" Length="Halves-32nds" />
            <Post-Describer>
                <Symbol-Point3/>
            </Post-Describer>
            <Chord Scale="Third"/>
            <Note Pitch="Silence" Length="Quarters-64ths"/>
            </Notes>
        </Measure>
    </Section>
  </Part>
</Score>
```

CONCLUSION

Since BMML complies with the International Braille Music Manual [13] and is based on the XML format which is the standard format for structuring and exchanging data, we think that the impact of this work is significant and will enhance access to Braille music for professionals and visually impaired people.

Indeed, by using BMML language, we could perform the following tasks:

- Information searches, because it contains metadata allowing the identification of the corresponding printed score, the tools and software used.
- Teaching and learning of Braille scores and/or music.
- Score navigation and queries about content.
- Music analysis.

Thus, the Braille format we have defined will allow the representation of musical Braille scores in different libraries to be unified; it will thus facilitate the exchange of scores from various origins.

In order to convert existing Braille scores into this new BMML format, a recognition and conversion tool has been developed. This will permit the conservation of musical heritage as many more scores will become readable.

On the whole, BMML is the first stage in the improvement of the access to Braille music. We intend now to develop tools for publishing and querying BMML scores.

We also will develop a BMML Easy-reader which permits a blind musician to read a Braille score with the help of a Braille display and/or a vocal synthesis.

REFERENCES

[1] "European Blind Union Guide," [Online]. Available: http://www.euroblind.org/fichiersGB/visincfr.html [Accessed: March 20, 2009].

[2] "Toccata", [Online]. Available: http://members.optusnet.com.au/~terryk/toccata.htm [Accessed: March 20, 2009].

[3] JESSEL Nadine, TORNIL Bertrand, ENCELLE Benoît. "BME Braille Music Editor : une solution pour l'accessibilité à l'information musicale par les non voyants ". Available: http://www.snv.jussieu.fr/inova/villette2002/act12.htm [Accessed: March 20, 2009].

[4] Bill McCann, "GOODFEEL", Available: http://www.dancingdots.com/ [Accessed: March 20, 2009].

[5] "Contrapunctus ", [Online]. Available: http://www.punctus.org/ [Accessed: March 20, 2009].

[6] Charles F. Goldfarb, " SMDL (Standard Music Description Language)", Available: http://xml.coverpages.org/smdl10743-1995.pdf. [Accessed: March 20, 2009].

[7] P. Bellini, P. Nesi, and G. Zoia, "Symbolic music representation in MPEG," *IEEE Multimedia*, Vol. 12, Issue 4, pp. 42-49, Oct.-Dec. 2005.

[8] Michael Good (Recordare), "MusicXML Definition," Available: http://www.musicxml.org/xml.html [Accessed: March 20, 2009].

[9] "NIFF", [Online]. Available: http://www.music-notation.info/en/niffml/niffml.html [Accessed: March 20, 2009].

[10] D. Langolff, N. Jessel, and D. Levy, "NIFF Transcription and Generation of Braille Musical Scores,".*The virtual Score, Representation, Retrieval, Restoration idem*, Available: http://www.ccarh.org/publications/books/cm/vol/12/07 [Accessed: March 20, 2009].

[11] "MIDI Specification," [Online]. Available: http://www.midi.org/about-midi/specshome.shtml

[12] "Play Code idem," [Online]. Available: http://www.dodiesis.com.

[13] "Braille format: New International Manual of Braille Music Notation," [Online]. Available: http://www.opustec.com/products/newintl/ [Accessed: March 20, 2009].

[14] "Dublin Core Metadata Imitative," [Online]. Available: http://dublincore.org/ [Accessed: March 20, 2009].

[15] "The Sheet Music Consortium," [Online]. Available: http://digital.library.ucla.edu/sheetmusic/project.jsp [Accessed: March 20, 2009].

[16] "The International Federation of Library Associations and Institutions," [Online]. Available: http://www.ifla.org [Accessed: March 20, 2009].

[17] "International Standard Bibliographic Description," [Online]. Available: http://www.ifla.org/VII/s13/pubs/isbd3.htm#14a [Accessed: March 20, 2009].

[18] "International Standard Bibliographic Description for Printed Music," [Online]. Available: http://www.ifla.org/VII/s13/pubs/ISBDPM_Nov10_2004.pdf [Accessed: March 20, 2009].

[19] "MARC 21," [Online]. Available: http://www.loc.gov/marc/ bibliographic/ [Accessed: March 20, 2009].

[20] "The International Association of Music Libraries," [Online]. Available: http://www.iaml.info/ [Accessed: March 20, 2009].

[21] "Core Bibliographic Record for Printed Music," [Online]. Available: http://www.iaml.info/activities/projects/core_record_for_ music/printed_music

[22] "Harmonica concerted action," [Online]. Available: http://projects.dedicon.nl/harmonica/Harmonica%20Deliverables/introduction.htm [Accessed: March 20, 2009].

[23] "Anglo-American Cataloguing Rules," [Online]. Available: http://www.aacr2.org/ [Accessed: March 20, 2009].

[24] "MuseData," [Online]. Available: http://www.musedata.org/ [Accessed: March 20, 2009].

[25] "**Kern," [Online]. Available: http://www.musiccog.ohio-state.edu/Humdrum/representations/kern.html [Accessed: March 20, 2009].

[26] "Hundrum," [Online]. Available: http://dactyl.som.ohio-state.edu/Humdrum/ [Accessed: March 20, 2009].

[27] "Humdrum Bliographic Codes," [Online]. Available: http://www.ccarh.org/courses/254/files/HumdrumBibliographicCodes.pdf [Accessed: March 20, 2009].

[28] "WELDELMUSIC," [Online]. Available: http://www.wedelmusic.org/ [Accessed: March 20, 2009].

[29] "Music Encoding Initiative," [Online]. Available: http://lib. virginia.edu/digital/resndev/mei/ [Accessed: March 20, 2009].

[30] "International Music Score Library Project," [Online]. Available: http://imslp.org/ [Accessed: March 20, 2009].

[31] "19th Century Califirnia Sheet Music," [Online]. Available: http://people.ischool.berkeley.edu/~mkduggan/history.html [Accessed: March 20, 2009].

[32] "Variations2," [Online]. Available: http://variations2.indiana.edu [Accessed: March 20, 2009].

[33] "MIRACLE project," [Online]. Available: http://projects.dedicon.nl/miracle/default.htm [Accessed: March 20, 2009].

[34] R.A. Medina, L.A. Smith, and D.R. Wagner, "Content-based indexing of musical scores," *Joint Conference on Digital Libraries*. pp. 18-26. May 2003.

[35] "Wikipedia entry for Semantic Web," [Online]. Available: http://en.wikipedia.org/wiki/Semantic_web [Accessed: March 20, 2009].

[36] "Semantic Web," [Online]. Available: http://semanticweb.org/wiki/Main_Page [Accessed: March 20, 2009].

[37] "Music Ontology Specification," [Online]. Available: http://musicontology.com/ [Accessed: March 20, 2009].

[38] "XML Signature Syntax and Processing," [Online]. Available: http://www.w3.org/TR/xmldsig-core/ [Accessed: March 20, 2009].

[39] Mikhail J. Atallah, Victor Raskin, Michael Crogan, Christian Hempelmann, Florian Kerschbaum, Dina Mohamed and Sanket Naik. Natural Language Watermarking: Design, Analysis, and a Proof-of-Concept Implementation. Lecture Notes in Computer Science. Vol 2137. pp. 185-200. 2001.

[40] R. Yao, Q. Zhao, and H. Lu, "A novel watermark algorithm for integrity protection of XML documents," *International Journal of Computer Science and Network Security*, Vol. 6, No. 2B, pp. 202-207. February 2006.

[41] W. Ng and H.-L. Lau. "Effective Approaches for Watermarking XML Data", *Lecture Notes in Computer Science*. Vol. 3453 pp. 68-80.

[42] "W3C Schema definition," [Online]. Available: http://www.w3.org/XML/Schema [Accessed: March 20, 2009].

[43] ISBD(PM) "*International Standard Bibliographic Description for Printed Music*," [Online]. Available: http://www.ifla.org/VII/s13/pubs/ISBDPM_Nov10_2004.pdf [Accessed: March 20, 2009].

Received: February 15, 2009 Revised: May 25, 2009 Accepted: June 03, 2009