A collaborative platform model for digital scores annotation

Véronique Sébastien(1), Didier Sébastien(1) and Noël Conruyt(1)

(1) IREMIA - Laboratoire d'Informatique et de Mathématiques from Reunion Island University Bâtiment 2, 2 rue Joseph Wetzell 97490 Sainte-Clotilde France

Abstract

Musical scores annotation is a well-known practice from musicians, especially in educational terms. It consists in adding written comments to a score in order to make a phrase explicit and thus facilitate the learning of the piece and its performance.

However, this practice raises knowledge transmission issues: how to share and preserve the information within the annotation without altering it? Indeed, the author is often the only person who knows the meaning of his annotations. As the use of digital tools constitutes a privileged way to solve this issue, we first explore the existing solutions for sharing performances through annotations, as well as their pros and cons. Then, we propose an optimized solution to annotate scores online in order to maintain the intent of the author, i.e. make it explicit to other performers through our tools. This proposition focuses on four main aspects of score annotation: the annotation nature, its symbolic and semantic representation, its indexation on the studied piece and at last, the communication structure that supports it.

Keywords: musical score annotation, knowledge management, human-computer interfaces, multimedia, collaborative platform.

1 Introduction

Occidental music, and more precisely classical music, uses a score as the principal medium of performance sharing and transmission. Indeed, it allows any "literary" performer to grasp the composer's intent and thus to play the piece it describes more or less accurately. However, the score only holds a small part of the information necessary to play the piece correctly. That's why most musicians add annotations to it, mostly resulting from a personal thought or a teacher's advice in the frame of a lesson, in order to complete the present information in particular on technical and expressive issues.

Works such as Megan A. Winget's [Win08] prove the extent of such a practice, thus legitimating the development of collaborative tools for digital scores annotation, especially in order to share the knowledge held by this practice. At the same time, the development of scores description languages (Wedelmusic, musicXML, etc.) and associated tools (graphic editor) allowed the creation of real digital scores library on the web[1]. Thus, various applications proposing a form of assistance to score annotation emerged from these technologies. In this article, we first study the characteristics and limitations of the principal ones, then we propose a collaborative solution adapted to the practice of digital scores multimedia annotation.

2 Between Knowledge and Know-how, from notation to annotation

The following part first develops the way to create a digital score (based on traditional musical notation) which will constitute the main support to work with annotations. We then describe two projects

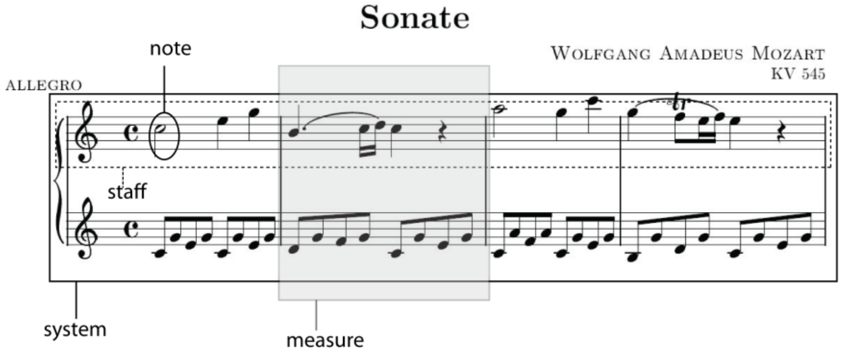
[1] [http://icking-music-archive.org, visited on the 26/10/2009].

that include annotation modules and their limitations.

2.1 Scores editor

With the development of Computer Assisted Music (CAM), various scores editors offer possibilities to annotate scores in a more or less advanced way. Most of them are based on MusicXML, an open file format for musical notation. Among them, we can count the Finale™ software's line, the Sibelius™ one, and sequencers, such as Cubase®, Logic® and ProTools®. However, the main goal of such applications is the arrangement of the score's layout as formal knowledge for composition aid (so called the notation). Helping people actually play the piece, for example in an educational context, remains secondary (annotation). Thereby, the functionalities to annotate scores are often limited to basic symbols and text. Yet, this software can be used to generate a MusicXML file describing an original score on which the annotations will be overlaid. The basic elements on this score are: the systems and the staffs, the measures, the notes and the rhythm (Figure 1). In the following paragraphs, let us examine two projects representative of our problematic.

Figure 1: Structural elements of a score



2.2 Score Annotations editor: The VEMUS project

VEMUS[2] is a European project which aims at developing and validating a virtual environment for musical practice [FLO07]. In particular, a score annotation module is included. This module allows annotating a score with various shapes, texts, symbols but also audio extracts and their characteristic curves (frequency spectrum, waveform, etc.) [CFLOD07]. The use of a tactile device can enrich the user experience, notably thanks to the possibility of « drawing » the annotation: encircle a passage, cross out a note or even write a phrasing. The musician experiences the same feelings as with his paper sheet music. This functionality is based on the dynamic score notion: all elements from this score can reorganize themselves according to the user's actions (enlargement, reduction of a staff, musical objects shifting, etc.). Moreover, a logic layer is laid on the graphic layer, allowing one to extract significant information about the symbols entered by the user: its context (staff, measure), its temporal register (beginning, ending), its graphic characteristics. Thus, the annotation itself is dynamic and can be shifted and modified (in the same manner as a vector path).

Nevertheless, the annotation module do not include an automatic path recognition module which would allow the association of a freehand drawing to a familiar symbol coming from a library, and thus would indicate the meaning of this annotation in its context to a third party. In addition, it seems that the application has not been tested on advanced pieces yet, where the score's complexity would permit to confirm the annotation system's sturdiness.

2.3 Score Annotations Editor: MiXa, MusicXML Annotator

MiXa[3] is a musical annotation system based on MusicXML, for the base

[2] Virtual European Music School

[3] MiXa : MusicXml Annotator

score as well as its annotations, using a typical client-server web architecture [KN04]. To do so, an XML file containing a list of indexed annotations (thanks to an XPath reference) is associated to the XML file of the original score. All the graphic part is then generated in SVG[4] in order to be displayed in a web browser. From here, any registered user on the website can view and add textual annotations classified by their content type (“Impression”, “Affection”, “Description”, “Chord” and “Structure”). The platform also allows to search present annotations, in particular by using the annotation type but also the structure of the studied piece.

However, although the system benefits from the XML standard, it also limits the annotations’ types to those proposed by MusicXML: string, numeric, Boolean and chords (group of notes).

2.4 Limitations of these projects for our purpose

The main limitation of these applications is to omit one or several dimensions of document annotation for know-how transmission. Indeed, to allow the user to explain his performance in the best conditions, the latter should be able to use any type of media to do so. This underlines the annotation’s content and type issues, which arises in the MiXa project. In parallel, the user should be able to communicate his annotations to any third-party if he requires to, so that his annotations are the base to a fruitful debate. In comparison to the VEMUS project, which is based on a peer-to-peer network of selected schools, a web based annotation module would be more opened and could be integrated to any e-learning platform, such as the Berklee College of Music[5].

As such we propose in what follows a model for a collaborative platform using multimedia annotations to constitute a real community around performance description for music e-learning.

3 Proposition of an architecture for a collaborative platform using multimedia annotations

By using the term « multimedia annotation », we signify that such an application is not limited to the traditional practice of annotating scores, which is often reduced to text or symbols, and is destined to one, even two persons within a music lesson, but also permits the addition of multimedia content according to a user centered model, be it for professors or amateurs. Indeed, music combines both intellectual and gestural exercises, and the latter may require the observation of another person performing the movements in order to reproduce them correctly. That’s why Megan Winget [Win08] proposes an Augmented Annotation Framework (AAF) that takes into account the creation context of the annotation (intellectual or gestural activity) and its nature (textual or symbolic), on which she builds a list of recommendations aimed at developers of music annotation software. Among those is the necessity to let the user dispose of a large realm of possibilities on 1) the nature of the annotations and its organization, 2) its positioning on the score, 3) its graphic representation and 4) the means of communication of the information it holds. We now present our model, which is based on those four aspects of score annotation.

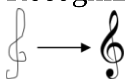
3.1 Annotations nature and content organization

The main idea of our model is that any document can constitute an annotation: a simple text, a symbol from a library, a symbol entered by the user, figures (for example for a fingering), video or audio extracts of a musical phrase which may

[4] SVG : Scalable Vector Graphics
[5] [http://www.berkleemusic.com/, visited on the 25/11/2009]

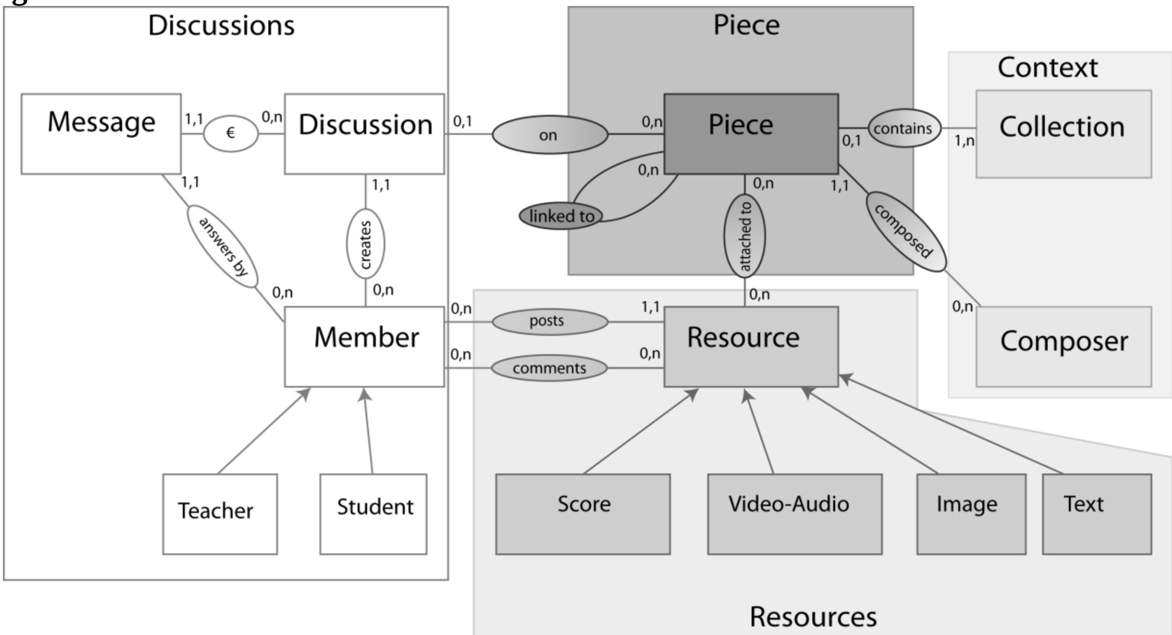
illustrate an educational advice. We thus propose the following types:

Figure 2: Annotation content types table

Type	Definition	Example	Computing object
Freehand drawing	An annotation drawn by the user with a mouse or any other capture device. The traced path can be recognized by matching it to a library including well-known symbols and thus be replaced by the right symbol (unless otherwise stated by the user)	- Unrecognized:  - Recognized: 	- Unrecognized: vector path (.svg) - Recognized: vector path (.svg) + recognized symbol (symbol object = path + meaning)
Symbol (simple or composite notation)	Common music symbols are proposed in a library, as well as symbols specific to each instrument and commonly used to annotate scores	- Path:  - Fingering: 5 1 2 3 4 2	Symbol object: Symbol = path + meaning Note: a symbol can be composed of several other symbols (for example : a fingering is represented by a series of figures : it's thus represented by several symbols which path is the shape of the figure and the meaning is the number of the corresponding finger)
Score extract (specific notation)	An extract of a score is entered by a user allowing, for example, to clarify a symbol on the original score (example: a trill symbol:)	Clarified trill 	MusicXML file
Text	A simple text	« Accentuate this note »	String data type
Multimedia	Video, audio, image		Common formats: flv, mp3, jpg

With this setting, it is imperative to dispose of a Conceptual Data Model able to support any type of annotation. Figure 3 details this structure.

Figure 3: Relational model of annotation data



The whole content is organized around a central entity: the studied piece. Three additional entities come along this central entity: the historical context of the piece (collection, composer, other metadata), the associated discussions and the resources. The *Context* entity mainly consists in the static metadata proper to the piece: the composer's name, its composition and publication date, etc. The *Resource* entity regroups all the contents associated to the piece: multimedia content, texts, scores. The *Discussion* entity contains the descriptive form of these contents, that is to say the annotations. Indeed, any annotation can be considered as a message, with a sender (a student, a professor), some recipients (the whole community) and containing one or several contents of the types enumerated previously (figure 2). Therefore, we are able to constitute several discussion/annotations threads indexed on the score, thus constituting a gloses system on its interpretation.

In order to facilitate information research among these annotations, we add to each of them a semantic type from one of the three annotations categories described in [Win08]: technical (fingering, hand position, etc.), technical-conceptual (sound intensity, acceleration, etc.) or conceptual (expression, emotion and musicality). In addition to these different granularity levels, necessary to filter the annotations attached to the score, it is interesting to take into account two pragmatic possibilities that may extend our model's functionalities. The first one is to type any commentary as being either informative or interrogative, namely which demands an answer. The second possibility concerns the grouping of several annotations depending on their common subjects (for example, the piece structure).

3.2 Indexation of annotations on the score

An annotation is only meaningful if it refers to a specific part of the piece, and thus, of the score describing its melodic, rhythmic and harmonic evolution. It can refer to a note, as well as a silence, a group of notes, of measures, or a whole musical phrase. That's why it is necessary to define an efficient way of indexing annotations.

The most simple idea consists in considering the score as an image, and thus to assign a subset of the image, in the form of a rectangle, to each annotation (Figure 5). However, this superficial approach rapidly shows its limits: lack of musical meaning, impossibility to modify the image of the score, to identify precisely the concerned elements within the rectangle, no reusability. Therefore, it is necessary to find another way of indexing the annotations while taking into account their semantic.

By observing a hand-annotated score, we notice that each annotation refers to an element (or a group of elements) of the score. Each element being accessible in a MusicXML file describing the score, we thus dispose of the same granularity level to annotate a digital score. From here, we can associate a second XML file (Figure 4) containing all the annotations to the original XML file describing the score. The indexation realized by this file uses the following hierarchy: [file a > system b > staff c > measure d > note e], the latter remaining completely consistent in a musical context. Besides, this semantic description makes the annotations independent from the original digital score: if this score is replaced by another score describing the same piece (with the same number of measures), the annotations can't lose their indexation. However, this system supposes that a logic description of the score is available. If it's not the case, the superficial method presented previously can serve as a degraded mode.

3.3 Human-Computer Interface

Concerning annotations display, Megan Winget precises the importance to distinguish the original content from the annotations themselves. That is why we propose to indicate the presence of an annotation simply by a selection rectangle surrounding the annotated elements, the latter being themselves underlined, for example by changing their color (Figure 5). The user is thus able to reveal the content of the annotation by clicking on the rectangle. Based on the semantic indexation approach described in part 3.2, the image of the score and the required links

In the same way, it might be interesting to index the annotations on audio or video extracts of the piece. Various contributions were submitted to solve this problem, like the SMR standard [BNZ05] which uses some marks common to both the score and the data stream, or by comparing the audio file with a midi performance extracted from the logic description of the score [MKR04]. Yet, those algorithms remain complex and very sensitive to the accuracy of the real performance against the score.

Figure 4: Annotation XML file example

```
<?xml version="1.0" encoding="UTF-8"?>
<annotation id="1"  fileref="./MozartPianoSonata.xml">
      <title>The trill</title>
      <references>
         <reference ref="score-partwise/part[id =
P1]/measure[4]/note[2]"></reference>
      </references>
      <content>
         <resource id="2">
            <dataType>text/html</dataType>
            <resourceContent>All the trills in that piece should be
done this way:</resourceContent>
         </resource>
         <resource id="3">
            <dataType>text/xml</dataType>
            <resourceContent>
               <title>How to perform a trill</title>
               <path>.sonata/doc/trill.mxl</path>
               <description>score extract explaining a
trill</description>
            </resourceContent>
         </resource>
      </content>
</annotation>
```

Figure 5: Annotations reading interface

are generated by an XSLT[6] processor (for example Saxon) able to create a SVG file representing the annotated score from all the XML files inputted (original score + annotation XML files).

To enter a new annotation, the interface proposes a form similar to a post entry on a message board. Once the user is connected to the platform, he simply indexes his message on the score by drawing a selection frame with his mouse. From the coordinates of the pixels thus obtained, we deduce the MusicXML elements present in the frame. To do so, it is imperative that the translation from the MusicXML file into a SVG image do not prevent the identification of the musical elements. Each path (or group of paths) described in the SVG file thus refers to the musical element it represents in the original MusicXML file. In the entry interface, we also let the user attach a multimedia file (audio or video) to his message, even a score extract, in the form of a MusicXML file, or an image. In the same way, the system should propose a simple procedure to record a video message online in the case of a multimedia annotation. The simplest way to do so remains the use of a webcam detected by the platform [SC08].

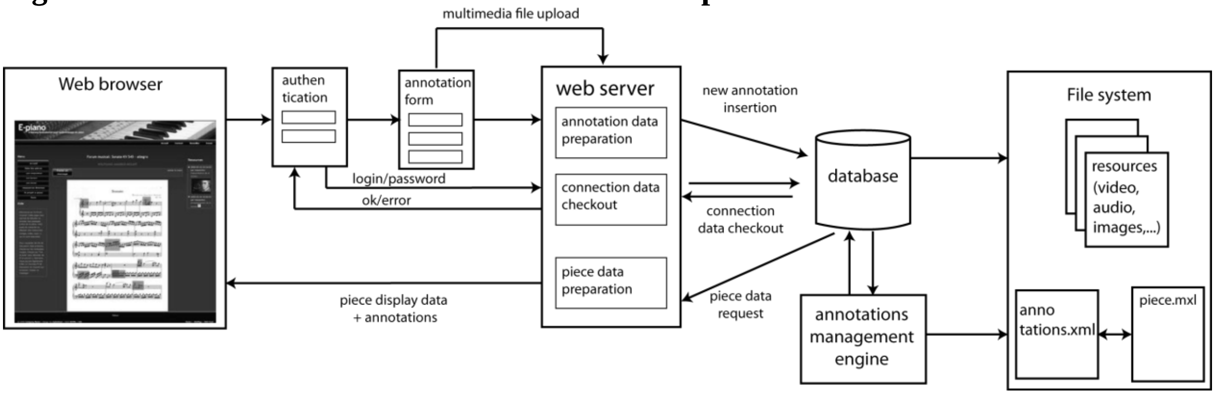
Although the interface described previously remains simple for a person who's comfortable with computers, it is still quite different than the traditional annotation practice for a musician. We thus propose to add a tactile mode to the application, associated to a module of automatic paths recognition. The interface proposes to draw the annotation directly on the electronic score. Whether it is recognized or not by the module, this type of annotation is stored in the annotation file under the symbol type or under the image type. However, the issue of the indexation of such an annotation arises. To solve this issue, we rely on the work of Y. Chapuis and D. Fober on the VEMUS project [CFLOD07]. In a worst case situation, we simply use the position in pixels of the rectangle surrounding the annotation.

3.4 Communication architecture

In order to make the annotation practice collaborative, the application takes the form of a website functioning the same way as a message board. We thus rely on the client-server architecture described on Figure 6.

[6] *eXtensible Stylesheet Language Transformations*

Figure 6: Architecture of the electronic annotation platform



This architecture should also be able to support different message types within the community. For instance, in an educational context, the platform should distinguish each user's role according to their experience in the music field. Thus, the annotations created by a sworn music professor is more credible than a student's ones. Besides, the application is user-centered and allows each user to dispose of a personal annotation space, which can be public. In the same way, we accentuate the possibility of creating working groups within the platform. Typically, a professor should be able to manage a group composed of all of his students and follow the development of each of them remotely, in parallel to the face-to-face lesson. The following part develops a few ideas on the manner to manage such a community.

4 Discussion

In order to constitute a real community, which shares its musical knowledge thanks to a dedicated online tool, our platform should be able to manage an important number of users, with various profiles and skills. Their actions on the system and their interactions during inter-user exchanges raise several issues which suggest some improvements for our model. Indeed, for each annotation, one might wonder to what extent is this proposition valid. At this level, a first binary filter can be applied: an instance of authority will validate or invalidate the proposal and, subsequently, explain its judgment in the thread associated to the resource. Naturally, this possibility raises the issue of users classification. Is it better to give more rights to users who contribute the most (web 2.0 solution, based on the user's commitment toward the system), or to those who possess the best music level (web 1.0 solution, based on the hierarchical position of a person and his recognition in the musical field)? Typically, the evaluation of a user's skills depends on an important variety of factors, which taken separately cannot determine a precise level for a given instrument:

- Experiencing the scene can be done at any level,
- The regular practice of an instrument alone doesn't mean that it is not a beginner,
- The number of years of practice does not characterize its intensity,
- The ability to perform correctly a piece may be receivable, but the performance may only result from a tireless repetition of the same movements for a very long time, attesting a lack of musical maturity,
- The sight-reading and assimilation speed of the piece may constitute an interesting parameter, but conditioned by the richness and the subtlety of the final performance.

Finally, the whole factors, in addition to the educational skills and the mastery of the platform's tools, permit the identification of the profiles of the main contributors, apt to validate the system's resources. Another possibility concerns the obtaining of consequent feedback on the users' learning experience, on each proposed piece. Indeed, the former parameters characterize the users and allow the creation of profiles and categories of musicians. Thus, it is interesting to know if the way they approach, work and perform a piece is achieved in the same manner: for a given piece, is it possible to define a characteristic learning path matching a specific category of musicians? By implementing an electronic form to monitor the students' progress, step-by-step and taking into account the process and the piece appropriation's length, we could propose coherent solutions to the question of the choice and the start of a new piece for a student. An additional improvement, yet requiring heavy computation, would consist in determining the characteristic aspects of a performance by a user. Indeed, by comparing two descriptive files (MIDI or MusicXML) of a same piece, one extracted from a real performance and the second one directly translated from the piece's score (and thus being « mechanical » as performed note by note), it is possible to characterize the distance between those two performances [HNK03]. This difference corresponds (restrained by the files' respective description vocabulary) to the piece performance by the musician. This would allow to define various playing styles and to develop comparisons between different musicians.

Several of these ideas suggest an advanced use of data mining methods and the complexity of the final solution denotes the necessity of an important structuring of the proposed platform. In this vein, an evolution toward a Musical Information System (MIS), composed of specialized and inter-communicating modules, in the same manner as Biodiversity Information Systems [CSCDSR09], seems appropriate. Indeed, the use of various web applications such as a thesaurus (definition of musical ontology), a multimedia database (educational resources storage), a user directory (users management), correlated to specialized modules for annotation, experience sharing and data mining, could clearly apply to our problematic. This perspective will then be the subject of an upcoming publication.

5 Conclusion

From the study of a typical practice of the musical field: knowledge creation and sharing via scores annotation, we proposed a model of a collaborative platform for digital scores multimedia annotation. In order to facilitate its manipulation and to preserve its pertinence, we proposed to add various types to an annotation, notably according to its content (text, multimedia, symbol, etc.), and its meaning (technical, conceptual information, etc.). We then described two ways of indexing an annotation to its original score, an essential step to understand the context of the knowledge held by the annotation. Then, a human-computer interface and communication structure were described to support the proposed model. At last, several development axes were exposed, principally centered on the methods to manage a community organized around this practice, its profiles and evolution, but also on the possible evolutions of our model toward a Musical Information System.

References

[BNZ05] P. Bellini, P. Nesi, G. Zoia (2005), Symbolic Music Representation in MPEG, *IEEE Multimedia, IEEE Computer Society press, ISSN 1070-986X,* Vol.12*, N.4, pp.42-29, October-December*

[CFLOD07] Y. Chapuis, D. Fober, S. Letz, Y. Orlarey, C. Daudin (2007), Annotation de partitions musicales dynamiques, *Actes des Journées d'Infor-*

matique Musicale JIM'07 - Lyon, Grame,

[CSCDSR09] Conruyt N., Sebastien D., Courdier R., David D., Sebastien N., Ralambondrainy T. (2009), Designing an Information System for the preservation of Insular Tropical Environment in Reunion Island. Integration of Databases, Knowledge Bases and Multi-Agent Systems by using Web Services, Agent Technology Applied to Environmental Issues, *Whitestein Series in Software Agent Technologies and Autonomic Computing, Cortés, Ulises; Poch, Manel (Eds.), Birkhäuser, ISBN: 978-3-7643-8897-3*

[FLO07] D. Fober, S. Letz, Y. Orlarey (2007), VEMUS - Feedback and Groupware Technologies for Music Instrument Learning, *Proceedings of the 4th Sound and Music Computing Conference SMC'07 - Lefkada, Greece*

[HNK03] K. Hirata, K. Noike and H. Katayose (2003)**,** Proposal for a Performance Data Format, *In IJCAI2003 workshop program, APP-5, pages 65--69*

[KN04] K. Kaji and K. Nagao (2004), MiXa, A Musical Annotation System, *Proceedings of the 3rd International Semantic Web Conference, - Hiroshima, Japan*

[MKR04] M. Müller, F. Kurth, T. Röder (2004), Towards an Efficient Algorithm for Automatic Score-to-Audio Synchronization, *Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR 2004), Barcelona, Spain, 365-372, Oct.*

[SC08] Sébastien O., Conruyt N. (2008), Developing an user-centred project involving the Web 2.0 community: a case study based on video sharing websites, in *Proceedings of I-media'2008 and I-Know'2008, J. UCS, Gratz, Austria*

[Win08] Megan A. Winget (2008), Annotations on musical scores by performing musicians: Collaborative models, interactive methods, and music digital library tool development, *Journal of the American Society for Information Science and Technology*