The NEUMA Project: Towards Cooperative On-line Music Score Libraries

L. Abrouk[1], H. Audéon[2], N. Cullot[1], C. Davy-Rigaux[2], Z. Faget[3,4], D. Gross-Amblard[1], H. Lee[1], P. Rigaux[3], A. Tacaille[5], E. Gavignet[1], V. Thion-Goasdoué[3]

[1]LE2I, Univ. Dijon, Fr. [2]IRPMF, CNRS/BnF, Paris, Fr. [3]LAMSADE, Univ. Paris-Dauphine, Fr. [4]ARMADILLO, Fr. [5]PLM, Univ. Sorbonne, Fr.

***Abstract*—The NEUMA project (*http://neuma.irpmf-cnrs.fr*) aims at designing and evaluating an open cooperative system for musician communities, enabling new search and analysis tools for symbolic musical content sharing and dissemination. The project is organized around the French CNRS laboratory of the *Bibliothèque Nationale de France* which provides sample collections, user requirements and expert validation. The paper presents the project goals, its achitecture and current state of development. We illustrate our approach with an on-line publication of monodic collections centered on XVIIe century French liturgic chants.**

I. INTRODUCTION

Current musical search engines mostly rely on audio files. Extracting significant information from such files to produce some structured description is a difficult task. Another approach is to exploit *symbolic* music representation, usually derived from musical scores. This representation contains a detailed description of all the information required to produce a precise rendering. Of course, interpreting this information is largely a matter of human expertise, and a software can hardly manipulate anything but a minor part of the intentional content conveyed by the symbolic notation. However, such manipulations can obviously go far beyond the mere on-demand printout of score copies. They include structure-based navigation, melodic fragments identification and retrieval, support for musical analysis, large-scale handling of scores databases, and community sharing and dissemination. Exploring the precise limits on such automatic or semi-automatic functionalities constitutes an exciting research area.

The goal of the NEUMA project (*http://neuma.irpmf-cnrs.fr*) is to address some of the main issues raised by the constitution of on-line digital scores libraries, and in particular:

1) create and maintain repositories of large-scale musical collections, along with appropriate browsing, searching and rendering functionalities;
2) explore automatic and semi-automatic music analysis tools applied to symbolic notation;
3) investigate problems related to copyright protection;
4) and, finally, experiment web-based sharing of musical scores archives, including cooperative development, data reconciliation and annotation.

NEUMA is a three-years project founded by the French *Agence Nationale de la Recherche* (ANR), and is built around the IRPMF laboratory (*http://www.irpmf.cnrs.fr*), a CNRS/BnF research institute devoted to the preservation and dissemination of Western music heritage. The project started in january 2009 and is currently in its early stages of development.

In the present paper, we first explain the main goals of NEUMA (Section II). We then outline the architecture of the project in Section III. Section IV is devoted to the first corpuses published by NEUMA. Finally, Section V discusses related works and Section VI provides the project's roadmap.

II. PROJECT GOALS AND REQUIREMENTS

NEUMA intends to support the development of digital collections of musical works, represented in symbolic notation, and aims at providing computer-based services on these collections. Let us consider these requirements in turn.

A. Collection of symbolic musical works

A collection is a set of homogeneous musical pieces, e.g., pieces whose contents share some stylistic, structural or notation features. The system should be flexible enough to deal with the inherent subjectivity of this concept, and allow for both intra -and inter- collection analysis and searches. For instance, the two initial collections supported by NEUMA consists of Catholic Liturgic chants (*Sequentia*, see Section IV) and French Renaissance Psalters. In terms of musical content, both consist of monodic chants, and share some historical and structural features whose description is beyond the scope of the present paper. One must therefore envisage exploratory tools that apply at the collection level, or span multiple collections for comparison and analysis purposes.

The notion of symbolic notation covers the content of traditional music scores, put in digital form through a music editing software or via OCR of music sheets. However, in order to be meaningful, this content must be mapped to a semantic-oriented description, a "model" of the music content, apt at supporting a set of relevant but specific functionalities. Defining a general model of music content is probably an overambitious task, and our approach adopts a more realistic (and modest) way.

We consider that a music digital library (MDL) should propose one of several models, each identifying a specific interpretation of a musical piece content in terms of semantically meaningful elements: notes and durations, combination of horizontal (melody) of vertical (chords) combination of notes, synchronization of text and note series, etc. The MDL should

be able to map a digital score representation to an instance of one of the supported models. Well defined functions can then be applied to this instance.

The primary goal of NEUMA is to manage large collections of Western music, and as such we will define models devoted this specific area. The system design should make it possible to address a broader context by incorporating others music abstractions. Managing a collection of fork songs tablatures for instance would probably call for a specific model.

B. Computer-based functionalities for musical content

Given a model that structures the content of a musical work in terms of well-defined elements, the MDL must support content-based functionalities. We decided to focus on the following ones.

Search by content. An application relying on the MDL should be able to express complex content-based searches on music collections. Our intent is to develop each asbtraction of music pieces as a database model fully equipped with a query language. In technical terms, this means (among other features) a careful distinction between the physical and the logical level, a set of closed-form operators that can be combined to obtain expressive queries, and a user-oriented, syntax-friendly query language. We are aware that music constitutes an important challenge for such a traditional database approach, because of the unlimited flexibility of its forms and structure. However, by adopting a representation based on symbolic notation, we believe that we can find enough structure in the content to support a semantically significant search tool. The section devoted to the case study illustrates a preliminary attempt in this direction.

Cooperative work based on ontology and annotations. The MDL must allow users to share documents, and also to exchange individual knowledge concerning documents, through an ontology-based annotation system. An ontology is an explicit specification of a conceptualization. Ontological categories define, at a semantic level, the concepts that exist in the domain and relationships between these concepts (in our case music concepts, Catholic Liturgic events, etc.).

Ontology definition is a collective work bringing together ontology experts (called ontologists) and users (musicologists here). For the purpose of NEUMA a referential and global ontology is defined.

An ontology-based tagging system will be included in NEUMA. Annotation systems have been widely studied in the digital libraries community. See for example the Digital Library Annotation Service (DILAS) [1], the Collaboratory for Annotation Indexing and Retrieval of Digitized Historical Archive Material (COLLATE) [2], and the Flexible Annotation Service Tool (FAST) [3]. An annotation consists of a *tag* placed by a user on a document or a part of a document (for a better exploitability, a tag has to belong to the ontology). This enriches the MDL with personal knowledge, and shares this knowledge it with other users.

An ontology can also be used to provide a user-friendly GUI[4]. And finally, the ontology can also be exploited during the query process through inferences based on concepts relationships increasing the expressive power of the query language.

Music analysis. This generic term refers to any task that explores the symbolic musical content in order to extract some implicitly represented information. We distinguish these functionalities from query operations because they do not operate in closed-form, rely on statistical evaluations, and involve algorithms which may be specific to a certain subclass of the collections hosted by NEUMA. An interesting feature of computer-driven music analysis is its interdisciplinary nature which implies a close cooperation between musicologists and computer experts. A simple example, currently under investigation on our first collection, is a study of text/music relationships.

A broader ambition is to enable, in the realm of musical content, the analysis and classification tools already in use for text documents, where the content can be automatically analyzed, classified, and indexed by a summary of weighted keywords [5], [6]. Transposed to music content, this means identification of stylistic features, automatic extraction of rythmic or melodic patterns, genre classification, etc.

Copyright protection. A last topic of interest to our project is the protection of copyright. Since the envisioned platform will be able to store and deliver music pieces, or fragments of, we need to address the issue of preserving the ownership of each document. The currently adopted approach is to embed the owner's signature in the musical content without altering the quality of the symbolic representation ("*watermarking*"). This functionality may somehow be considered as independent from those mentioned above, as it could be useful, for instance, from web-based professional sites selling electronic scores.

III. THE ARCHITECTURE OF NEUMA

During the early stages of the project we designed an architecture for NEUMA based on the principles outlined above and capable of fulfilling our requirements. We now describe this architecture and present a first application in the next section.

We adopted the Service Oriented Architecture (SOA) approach because of its ability to combine loosely coupled tools which can be designed and built independently as soon as they agree on a common interface. A global view of the architecture is given in Figure 1. We distinguish two families of components. The first one refers to *Applications* which can be any piece of software, located anywhere on the Internet, that manipulates musical content. The second one constitutes the NEUMA platform which consists of a set of services, each dedicated to a specific functionality. Let us briefly consider them in turn.

The *storage service* allows external applications to register musical documents in NEUMA. A specific abstraction model φ must be chosen by the application, among those proposed by NEUMA (currently only one such model has been implemented: see the next section). The *register()* service takes as input a document d, its global id (used to unambiguously

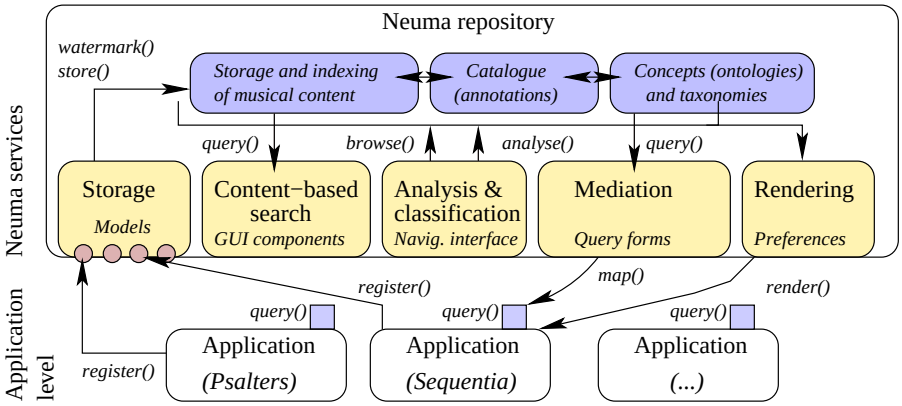

Fig. 1. Architecture of NEUMA

refer to d later on) and extracts from d an instance $i = \varphi(d)$ of the model which is stored in the repository. The set of instances $\{i_1, i_2, \cdots, i_n\}$ collected from an application constitues a *collection* which supports the other services. During this acquisition step, the owner signature can be embedded in instances for copyright protection purposes.

Content-based search is a service that takes a collection of instances and applies transforms (e.g., shifts) or predicates (e.g., musical pattern matching). As mentioned previously, our vision is to develop full-fledged musical query languages that manipulate instances of the model in closed-form and can be composed in arbitrarily complex expressions. This would allow, among other advantages, the definition of a sound query interface for this content-based service, allowing its client applications to submit query expressions to retrieve the matching instances. Our preliminary attempt in this direction is discussed in the next section. The service also provides graphical tools which can be integrated in external applications to facilitate query expression (virtual piano, virtual drums, query by humming, etc.).

The *analysis and classification* service aims at completing the symbolic content representation with annotations. This service is strongly related to the cooperative aspects that we ambition to promote in the NEUMA platform. We will integrate in the digital library tools allowing users to browse, analyse and, ultimately, tag the library content with elements from common vocabularies (or *taxonomies*). The annotation process may be manual, semi-automatic or automatic. A manual approach requires minimal support: the user must be able to browse the music collection, identify elements of interest and tag these elements with terms chosen from an appropriate taxonomy. Semi-automatic and automatic identification requires more involved tools which propose (or even decide) annotations from the result of some content-based analysis.

The *mediation* service supports information retrieval based on *global concepts*. NEUMA will host a set of music-oriented ontologies, general enough to span several collections managed by the digital library, and composed of commonly accepted concepts that can be used to describe musical pieces. We do not envisage to store in the NEUMA repository the instantiation of these concepts for each document. Rather, we will maintain *mappings* between these concepts and the local, application-level, representation. Each application requiring this service needs to register a mapping between its local schema and the NEUMA ontology. This mapping is then used to support description-based search by sending mediated queries (the *query()* local service in Fig. 1) to remote applications. This part of the architecture complies to standard mediation principles (See Section V for a discussion on related works).

The core components of ontological tools should consist of music-centric description: tonality, harmonic sequences, performance indications, etc. We also envisage the development of contextual ontologies that help to interpret and analyse musical pieces with respect to a historical or social environment. These ontologies should be of general value, i.e., they should apply to several collections and support inter-collection search and comparisons.

Finally, the *rendering service* is used to produce a user-friendly representation of the NEUMA services output. Typically, the set of pieces that constitute the result of a search operation must be shown as musical scores (or fragments of). Such results may be created dynamically by, for instance, transposing/merging/slicing data found in the repository.

IV. A CASE STUDY: SEQUENTIA

We implemented a first version of the envisioned architecture to test its ability to fulfil our goals. Our initial targets consist of two collections: Catholic Liturgic chants from the XVIIe century, and Psalms from the same period. Those two collections, SEQUENTIA and PSALTERS, although clearly distinct, share many features that make their association interesting. The description that follows focuses on SEQUENTIA but applies to both collections.

SEQUENTIA can be accessed on-line at *htpp://sequentia.irpmf-cnrs.fr*. In terms of functionalities, it compares to other similar sites (see for instance *http://www.globalchant.org/*) with a somehow greater search expressivity due to the ability to combine taxonomy-based criteria with similarity content-based retrieval. The

description that follows is meant as an illustration of the architectural principles outlined previously.

The data model. The data model supports the representation of polyphonics pieces (called "documents" thereafter) composed of "voices", each voice being a sequence of notes such that only one note can be played at a given instant. In other words, the model ignores polyphonic instruments. We extend this intuitive concept to any sequence of symbols taken from a finite alphabet. This allows to cover melodies (where the alphabet consists of notes) and text (where the alphabet consists of syllabes) as well as, potentially, any sequence of music-related information (e.g., fingerings).

A *musical piece* p is modeled as a set of synchronized times series $(s_1, s_2, \cdots, s_n)$ that all cover a same time frame $\mathcal{T} = [t_1, t_2]$. The time domain is discrete (the unit being the minimal note duration). Each time serie s_i is identified by a unique label λ_i which can be seen, at an abstract level, as a function $\lambda_i : \mathcal{T} \rightarrow \Sigma$ where Σ is a set of symbols (notes, text, fingerings, etc.) that constitute the "domain" of s_i.

A musical piece can then be viewed as a heterogeneous matrix where lines are instruments (or text) and columns are the sequence of instants. We insert a monophonic melody into a matrix line by converting each note to the numerical value of its pitch. In the SEQUENTIA collection, each piece is a 2-lines matrix where one line is the monophonic melody and the other is the corresponding syllable. Chamber music will have more lines, longer pieces with different movements can be split through several matrix.

The query language can then be conceptualized in terms of matrix operations. Here are a few examples:

- Get one or more instruments from a piece: get lines $\lambda_i, \cdots, \lambda_j$.
- Transpose a piece: add a constant matrix.
- Get all events at instant q: projet on column t_q.
- Get all events from a time slice: get columns $t_q, \cdots, t_{q+k}$.
- Add or remove a movement from a piece: concatenate/subtract two matrixes column-wise.
- Add or remove instruments from a piece: concatenate/subtract two matrixes line-wise.

All these operations operate in closed form and can therefore be combined to form complex expression. Note that operations on the lines of a matrix can generate information not explicitly featured in the music score. Computing an harmonic progression involves a difference operator between that takes several lines as input and produces a new one. Such an algebra constitutes the backbone of a user-friendly query language proposing more meaningful expressions.

Query primitives. The data model is implemented in C++, with serialization functions that allow to put/retrieve a piece to/from a persistent repository. The matrix operations (currently under implementation) are based on a few primitives: subsequence matching, slicing, shifting, etc. The main primitive implemented so far is an approximate pattern matching based on the *Dynamic Time Warping* (DTW) algorithm. It is used for content-based search, either for ranking query results, or during data retrieval (in that case a pre-defined threshold is applied) when a music pattern is provided by the user using a virtual piano (see Fig 2).

Music analysis. We did not investigate automatic music analysis algorithms so far. An ongoing work on our two collections addresses the issue of text-music relationship. The goal is to extract musical fragments associated to text sentences, and to study how the melodic shape matches the meaning of the text.

Taxonomies. The platform (in its current state of implementation) supports a simple mechanism for content annotations based on controlled vocabularies, or *taxonomies*. A taxonomy consists of a set of terms together with a subsumption relation between terms. If $s \preceq t$ then we say that s is *subsumed* by t, or that t *subsumes* s. A taxonomy is usually represented as a graph, where the nodes are the terms and there is an arrow from term s to term t iff s subsumes t. Figure 3 shows an excerpt of one of the taxonomies used in SEQUENTIA (description of Liturgic offices). Two terms s and t are *synonyms* iff $s \preceq t$ and $t \preceq s$. Given a taxonomy $(T, \preceq)$, an *annotation* in T is any set of terms from T. Users can attach annotations to the documents, and annotations can be used as a support to find and retrieve the document.

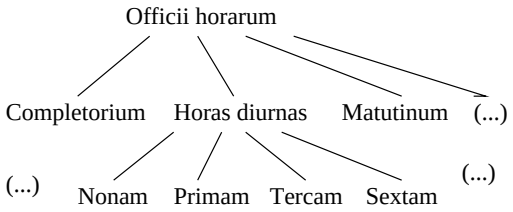

Fig. 3. A taxonomy (excerpt)

The annotation of a d is stored in a repository which can be thought of as a set of pairs (t, d) for each term t appearing in the annotation of d. The set of all such pairs (t, d), for all documents is the *catalogue* (see Fig. 1).

A *query* over the catalogue is any string derived by the following grammar, where t is a term:

$$q ::= t | q \wedge q' | q \vee q' | q \wedge \neg q' | (q)$$

If the query is a single term, then the answer is the set of all documents related either to t or to a term subsumed by t. If the query is not a single term then we proceed as follows. First, for each term appearing in the query, replace the term by the set of all documents computed as explained above; then replace each boolean combinator appearing in the query by the corresponding set-theoretic operator; finally, perform the set-theoretic operations to find the answer.

SEQUENTIA uses on three taxonomies: *Offices*, *Liturgic calendar* and *Solemnity level*. Term-based queries can be submitted via a web form that allows to select the terms of interest to retrieve a set of documents.

Current state of implementation and discussion.

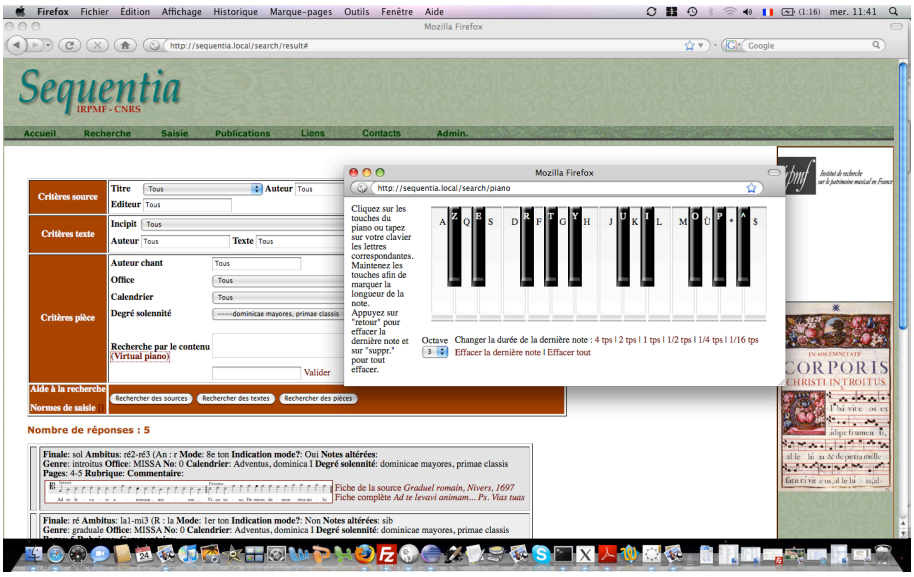

Fig. 2. The search form of SEQUENTIA

Figure 2 shows the main query form in SEQUENTIA, along with a few results. The form allows to combine content-based search (melodic fragments can be entered with a virtual piano) and term-based search. The list of music pieces matching the query criteria is displayed as a ranked list, each item being a link to navigate in the collection.

The *register()* function consists in sending a MusicXML containing the description of a new musical document. A mapping function extracts from the XML representation an instance of our data model which is then stored in the repository. This document can be annotated with terms from the Catalogue. MusicXML (*http://www.recordare.com*) is now a widely used exchange format for musical symbolic content, supported by almost all the music editor softwares. Other symbolic formats could be envisaged as well, providing that the appropriate extraction function is supplied.

The *render()* function relies on Lilypond (*http://lilypong.org*). Currently the score is produced during the registration of a new piece, and canot be changed later on. This prevent the dynamic generation of scores from query results. We plan to devote a part of the project resources to implement a convenient API upon Lilypond to obtain a more flexible and powerful tool.

As mentioned above, our primary goal was to validate the main lines of our approach. The services that compose our envisioned architecture have been implemented with basic functionalities, and interact to support this simple application. In terms of query language and query evaluation techniques, the current implementation lacks from efficient mechanisms such as, for instance, indexing of musical fragments. However, the combination of approximate content-based retrieval on one side, with the term-based query language built on taxonomies on the other side, turns out to provide a useful search tool and constitutes an encouraging result for further development of the NEUMA components. In particular, we are now working on a stronger querying service based on a query algebra, powerful primitive implementations, and a user query language that isolates the application from the internal mechanisms. Access paths to efficiently support common searches on large music collections is also underway [7], [8].

V. RELATED WORK

The past decade has witnessed a growing interest in techniques for representing, indexing and searching (by content) music documents. The domain is commonly termed "Music Information Retrieval" (MIR) although it covers many aspects beyond the mere process of retrieving documents. We refer the reader to [9] for an introduction. Roughly speaking, systems can manipulate music either as audio files or in symbolic form. NEUMA belongs to the latter category. The symbolic representation offers a structured representation which is well suited for content-based accesses and sophisticated manipulations and analysis [10].

Several projects have been or are currently devoted to MIR and digital libraries. Close projects to NEUMA are the OMRAS (Online Music Recognition and Searching)[11] and OMRAS2 (Ontology-driven Music Retrieval and Annotation Sharing Service) [12] projects. OMRAS's ambition was to build a system for content-based searching of online musical databases via an intuitive interface that uses music in a visual or aural form. OMRAS paid a particular attention to music transcription that might be seen on a score [13], [14]. OMRAS2 (Ontology-driven Music Retrieval and Annotation Sharing Service) is a framework for annotating and searching collections of both recorded music and digital score representations (like MIDI). Of particular interest to our goals are the logic-based representation and management [15], [16] and social networks analysis [17], [18].

Architecture. The architecture of NEUMA follows the current trend of building structured views over loosely structured documents. A representative example is the COUCHDB Apache system [19] which supports the definition and manipulation of views built using the Map/Reduce paradigm. A NEUMA model can be seen as a view implementing a particular interpretation of some musical content.

The mediation approach has been widely studied in databases and AI domains (see [20] for an overview). A mediator provides a uniform query interface for querying collections of pre-existing data sources that were created independently. Two approaches of mediation can be considered, depending on the way mappings are defined: the GAV (Global As Views) and the LAV (Local As Views) approach. A comparison of this approaches can be found in [21]. Number of mediator systems exist (e.g. PICSEL [22], Xyleme [23], TSIMMIS[24], etc).

Ontologies. Ontologies are means to describe agreed and shared knowledge for communities of people. Different levels of knowledge representation can be considered in the specification of an ontology, ranging from *taxonomies* to define an agreed set of terms used in a specific domain, to more complex conceptual models that can interact as mediation models in a cooperation of information retrieval systems. The Music Ontology [25] addresses a large panel of knowledge including editorial, cultural and acoustic information. The MPEG-7 ontologies [26] deal to model standard metadata to interlink different multimedia resources. The MX-Onto ontology [27] is more related to music description and classification. Ontologies in NEUMA ambition to provide a support model to interlink several musical collections.

VI. CONCLUSION

NEUMA aims at exploring the range of possibilities offered by the application of well established digital libraries services to large archives of symbolic musical content. The project roadmap is based on the step-by-step development of functionalities supporting several collections, ranging from simple monodic pieces to large and complex musical works, e.g., piano concerti.

We are aware that music is inherently a quite various and unpredictible material, and the design of our platform must be flexible enough to cope with it. Our ambition is to produce results of interest to the community of users as well as to contribute to advances in the realm of mutimedia digital libraries.

ACKNOWLEDGMENTS

This work is supported by the French ANR project NEUMA (*http://neuma.irpmf-cnrs.fr)*).

REFERENCES

[1] M. Agosti, H. Albrechtsen, N. Ferro, I. Frommholz, P. Hansen, N. Orio, E. Panizzi, A. M. Pejtersen, and U. Thiel, "Dilas: a digital library annotation service," in *IWAC*, J.-F. Boujut, Ed. CNRS - Programme société de l'information, 2005, pp. 91–101.

[2] J. Keiper, H. Brocks, A. Dirsch-Weigand, A. Stein, and U. Thiel, "Collate - a web-based collaboratory for content-based access to and work with digitized cultural material," in *ICHIM (1)*, 2001, pp. 495–511.

[3] M. Agosti and N. Ferro, "A system architecture as a support to a flexible annotation service," in *DELOS Workshop: Digital Library Architectures - LNCS Volume*, ser. Lecture Notes in Computer Science, C. Türker, M. Agosti, and H.-J. Schek, Eds., vol. 3664. Springer, 2004, pp. 147–166.

[4] L. Cinque, A. Malizia, and R. Navigli, "Ontodoc: An ontology-based query system for digital libraries," in *ICPR (2)*, 2004, pp. 671–674.

[5] K. S. Jones and D. M. Jackson, "The use of automatically-obtained keyword classifications for information retrieval," *Information Storage and Retrieval*, vol. 5, no. 4, pp. 175–201, 1970.

[6] M. J. Dovey, "Adding content-based searching to a traditional music library catalogue server," in *Proc. Joint Conf. on Digital Libraries (JCDL)*, 2001, pp. 249–250.

[7] Y. Zhu and D. Shasha, "Warping Indexes with Envelope Transforms for Query by Humming," in *Proc. ACM SIGMOD Symp. on the Management of Data*, 2003, pp. 181–192.

[8] E. J. Keogh and C. A. Ratanamahatana, "Exact Indexing of Dynamic Time Warping," *Knowl. Inf. Syst.*, vol. 7, no. 3, pp. 358–386, 2005.

[9] M. Muller, *Information Retrieval for Music and Motion*. Springer, 2004.

[10] G. Haus, M. Longari, and E. Pollstri, "A Score-Driven Approach to Music Information Retrieval," *Journal of American Society for Information Science and Technology*, vol. 55, pp. 1045–1052, 2004.

[11] "Online Music Recognition and Searching (OMRAS) project," http://www.omras.org.

[12] "Ontology-driven Music Retrieval and Annotation Sharing Service (OMRAS2) project," http://www.omras2.org.

[13] J. P. Bello, G. Monti, and M. B. Sandler, "Techniques for automatic music transcription," in *Proc. Intl. Conf. on Music Information Retrieval (ISMIR)*, 2000.

[14] J. Bello, L. Daudet, and M. Sandler, "Time-domain polyphonic transcription using self-generating databases," in *Proc. of the 112th Convention of the Audio Engineering Society*, 2002.

[15] A. Anglade and S. Dixon, "Characterisation of harmony with inductive logic programming," in *Proc. Intl. Conf. on Music Information Retrieval (ISMIR)*, 2008.

[16] ——, "Towards logic-based representations of musical harmony for classification, retrieval and knowledge discovery," in *MML*, 2008.

[17] A. Anglade, M. Tiemann, and F. Vignoli, "Virtual communities for creating shared music channels," in *Proc. Intl. Conf. on Music Information Retrieval (ISMIR)*, 2007.

[18] K. Jacobson, B. Fields, and M. Sandler, "Using audio analysis and network structure to identify communities in on-line social networks of artists," in *Proc. Intl. Conf. on Music Information Retrieval (ISMIR)*, 2008.

[19] J. C. Anderson, N. Slater, and J. Lehnardt, *CouchDB: Rough Cuts Version*. O'Reilly Editions, 2008.

[20] G. Wiederhold, "Mediators in the architecture of future information systems," *IEEE Computer*, vol. 25, no. 3, pp. 38–49, 1992.

[21] E. Rahm and P. A. Bernstein, "A survey of approaches to automatic schema matching," *VLDB J.*, vol. 10, no. 4, pp. 334–350, 2001.

[22] F. Goasdoué, V. Lattès, and M.-C. Rousset, "The use of carin language and algorithms for information integration: The picsel system," *Int. J. Cooperative Inf. Syst.*, vol. 9, no. 4, pp. 383–401, 2000.

[23] A. Y. Levy, A. Rajaraman, and J. J. Ordille, "Query-answering algorithms for information agents," in *AAAI/IAAI, Vol. 1*, 1996, pp. 40–47.

[24] S. S. Chawathe, H. Garcia-Molina, J. Hammer, K. Ireland, Y. Papakonstantinou, J. D. Ullman, and J. Widom, "The tsimmis project: Integration of heterogeneous information sources," in *IPSJ*, 1994, pp. 7–18.

[25] Y. Raimond, S. Abdallah, M. Sandler, and F. Giasson, ""the music ontology"," in *Proc. Intl. Conf. on Music Information Retrieval (ISMIR)*, 2007, pp. 417–422.

[26] R. Troncy, S. Celma, O.and Little, R. Gracia, and C. Tsinaraki, "Mpeg-7 based multimedia ontologies: Interoperabity support or interoperability issue?" in *Proc. of the 1st Worshop on Multimedia Annotation and Retrieval enabled by shared Ontologies (MAReSO'07) in conjunction with SAMT'07*, 2007, pp. 2–14.

[27] A. Ferrara, L. A. Ludovico, S. Montanelli, S. Castano, and G. Haus, "A semantic web ontology for context-based classification and retrieval of music resources," *TOMCCAP*, vol. 2, no. 3, pp. 177–198, 2006.