Music Life Cycle Support through Ontologies

Stephan Baumann & Melanie Dulong de Rosnay

German Research Center for Artificial Intelligence
Erwin Schrödinger Str., 67663 Kaiserslautern, Germany
stephan.baumann@dfki.de

Research Center for Administrative Science
Information Technologies, Law and Linguistic Laboratory
CERSA, University Paris II-CNRS, 10 rue Thénard, 75005 Paris, France /
Medialive, 20 rue St Didier, 75116 Paris, France
melanie.dulong-de-rosnay@cersa.fr

Abstract

The focus of this paper is the exploration of ontologies as a means for knowledge sharing in music information retrieval scenarios. Our approach is intended to support the complete "life cycle of digital music" and focuses on the implications of a digital world for all the actors involved in the process of creation, consumption and modification of such entities. As a novelty we take care about legal aspects, digital rights management being heavily tangled in such scenarios. MPEG-7 and MPEG-21 vocabularies are considered as basic ingredients for our ontology which ensures in this way interoperability within the multimedia delivery and music research communities.

1.Introduction

"In the morning the alarm of Peters PDA gets off. He crawls out of the bed but he is still very tired since he had to burn the midnight oil to finish his project. The only thing that can help is the right music to get going. He grabs his PDA and says: "I need some fast Hard Rock music right now". Seconds later he listens to *Ace of Spades* from *Motorhead*. This motivating music helps him a lot to leave his house in time to drive to work. However after a couple of kilometres he finds himself in a traffic jam, which starts making him nervous. To calm down he grabs again his PDA. "I need some music to calm down" he says and listens immediately to songs from *Sade*, *Phil Collins* and *Celine Dion*. After a long days work he starts to prepare for an after work party in his favourite club. In order to prepare for the nice evening, again he grabs his PDA and says "Something funky". Only a wink later, he listens to music from *Chaka Khan* and *Kool & the Gang*. Tomorrow, Peter will try to compare his personal library with his friends' collection in order to discover new titles and build an original play-list for his girlfriend birthday party. During these days, Peter has not been bothered by copyright questions, his PDA proposed him authorized uses only and processed automated payment adapted to his usual consumption. Moreover, he was able to lend some tracks to his brother who is a music teacher and reused them easily in educational material."

Digital media access in a post-Napster world with respect of copyright and fair use may look like this [1]. Since we have to foresee a multitude of different users in such scenarios providing the pure contents and a variety of different services, knowledge sharing among all actors of the value chain becomes a major issue. In this paper we describe our joint work about the combination of an MPEG-7 related ontology - which has been used in various music information retrieval scenarios – with an MPEG-21 oriented approach for online sharing and transactions of copyrighted work. In section 2 we will introduce the digital life cycle of a piece of musical work, section 3 describes the basic ontology supporting human actors and machines, section 4 is about our prototypical application and section 5 gives a conclusion and an outlook on future work.

2. Life Cycle of Digital Music

The life cycle of digital music begins during the creation period. The author needs to exchange work-in-progress samples within his creative community (other artists, potential producers) in a way enabling to date and attribute versions, collaborative contributions and protect the unreleased work. Once the work has been edited, authors and producers expect it to be disseminated, retrieved and distributed according to legal and contractual conditions, including derivative works conditions, copyright royalties sharing processing between rights holders and their local mandates (collective societies, intermediate service providers…) and event reporting on effective uses in respect of the end-user privacy.

Music titles, albums, compilations, sound track for motion pictures and other derivative works may involve numerous contracts, making tracks use and re-use more complicated.

The task to retrieve exhaustively all different copyright owners and follow eventual rights transfer in order to ask for relevant entities authorisation, and share the royalties according to corresponding national laws and collective management, is time-consuming and burdensome.

End-users expect to find a user-friendly access to the music they like. Convenient retrieval, sharing and modification of musical entities is directly related to the existence of meaningful metadata. The state-of-the-art approaches aim either at automatically extracting metadata from the audio material [2] or using manually attached metadata from expert databases (*www.allmusic.com*) or voluntary work of user communities (*www.freedb.org, www.gracenote.com*). Musical concepts and features may be described in standards such as MPEG-7 which simplifies worldwide exchange (e.g. for web services or autonomous agents). Cross-domain upper ontologies could be used to support extensions to music-related areas [3]. For the future MPEG-7 enabled creativity tools (e.g. audio or MIDI sequencing software and compositional tools) could allow the author to add the intended metadata on-the-fly during the creative process. Coming MPEG-21 applications may also enable right holders to express, update and enforce rights conditions, and all users and terminals to interact with them.

3. Ontology to support Music Information Retrieval and Automated Rights Management in the Life Cycle of Digital Music

"An ontology defines the basic terms and relations comprising the vocabulary of a topic area as well as the rules for combining terms and relations to define extensions to the vocabulary" [4]. In our application scenario we noted as terms the concepts of required know-how in the music domain. The relations consist of several types, namely is-a and part-of relations are used quite often. Is-a-relations are used to indicate specializations of concepts (e.g. *audio* is-a *media type*, *lyrics* is-a *text, remix* is-a *derivative work, author* is-a *rights owner*) while part-of-relations denote required parts (e.g. *lyrics* part-of *song*, *member* part-of *band*).

3.1 Legal requirements

The legal part of the ontology was elaborated in the context of MPEG-21 standardisation [5] with respect to Rights Expression Language [6] and associated Rights Data Dictionary [7] aiming at copyright domain knowledge representation and computed exploitation in further MIR applications associating DRM.

As it would have been possible to build several ontologies, the composition of ours was deduced from the task it is designed for [8], i.e. expression of questions and decisions about copyright for Music Information Retrieval: licensing, infringement, and exception. Derivating requirements are:

- use context description and self-categorization by all users, with a special target on non-lawyer users, and user-friendly update of Rights Expressions and use situations attached to a context and a work, allowing retrieval according to each use situation
- enforceability and validity in most national legal systems, including fair use respect [9,10], as opposed to other initiatives giving priority to business models. It is indeed technically easier to close than to open access, making a work available for specific public targets may cause security wholes for unauthorized persons. Exceptions to exclusive rights a priori modelisation is a challenging task as fair use situations are defined a posteriori by the judge, according to context and non-static fuzzy criteria.
- reusability of the knowledge database toward interoperability with foundational and core legal ontologies.

The legal database top level is a copyright law ontology. In order to be agnostic and neither preclude non-static national legislations nor favor specific business models, only a minimum core of compulsory features (the lowest common denominator of worldwide copyright legislation and contractual practices) should be obligatorily specified within standards. This minimum compulsory core is being developed in a way to enable national legal features local qualification, implementation and update, as a single term may have different national definitions and legal consequences.

Local concepts interpretation is subject to changes and uncertainties, which are unfavorable to electronic commerce development. Legal concepts are open textured concepts [11] as their understanding is related to external conditions such as the context. This diversity generates difficulties for the implementation of DRM systems. Indeed, it is important to express the conditions attached to a work's use with accuracy and flexibility.

The preliminary phase of our work was to study the state of the art of AI and law [12, 13, 14, 15] for existing core legal ontologies (FOLaw, LRI) and legal ontology hooks within foundational or upper level ontologies [16].

Other domain ontologies with the same applicative scope were also used, i.e. Rights Expression Languages (RELs) and dictionaries as indecs [17], ODRL [18], XrML [19] and IPROnto [20]. Except the latter one, all projects start from use cases rather than from legal texts. The added value of the proposed methodology is to start from an expertise of the legal knowledge domain, defined as the intersection between copyright law, law for Information Technology and media law. This expertise aims at avoiding texts partial understanding or misinterpretations,

as legal doctrine and case law knowledge enable to throw light on texts interpretation. Our reasoning process aims at modeling most music titles uses core with a special focus on re-use cases, the latter aspect being outside other projects' scope.
The ontology concepts and relations are being extracted manually in order to avoid misinterpretations, after legislative corpus automatic retrieval with syntactic parsers. Corpus semantic relations (hypernyms, hyponyms, meronyms, causality…) help to build relations between ontological concepts.
Concepts are being linked to corresponding terminological database terms with extensive definitions based on national legislations and case law. Definitions update may be delegated to some trusted third party authority in order to respect balance between stakeholders. The third expression level being the user interface, selected terms and classes have to reflect simple applicative use cases (i.e. play, send, modify).

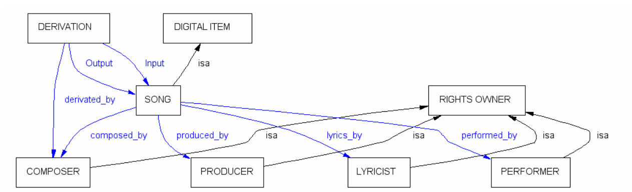

Figure 1: Excerpt of the MPEG-21 related ontology supporting music retrieval and rights management aspects

3.2 Music retrieval requirements

The starting point for the part of the ontology related to music retrieval was to model MPEG-7 ontology from scratch with an ontology editor and to add some few concepts, e.g. the lyrics. The connecting concept between the two main parts (legal, music retrieval) is the digital item. The main audio features which are used by our example application are the MPEG-7 descriptors for loudness, tempo and spectral features.

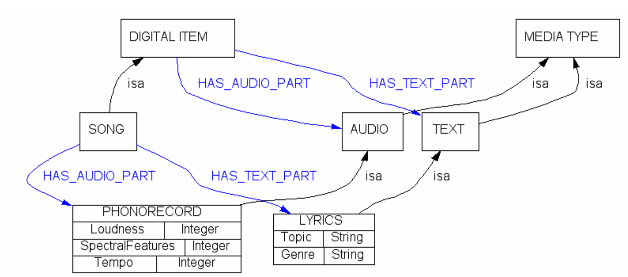

Figure 2: Excerpt of the MPEG-7 related ontology supporting music retrieval and rights management aspects

In this way we were able to set up an initial ontology for further refinement and alignment with other existing ontologies (see fig 1,2). We use the Protégé 2000 tool for convenient design of ontologies. It is mainly the flexible plug-in-architecture and the powerful import and export of standard formats such as XML and RDF(S) making it the tool of choice when it comes to ontology editors. Especially the options to import other know-how domains via importing RDF(S) files are important with respect to the ongoing developments of ontologies at different sites for different applications and in different formats. This problem has been described in the musical domain thoroughly in a recent publication of Pachet [21].
A further difficult question is how to handle the acquisition of instances which represent the actual state of the "musical world". This will be an ongoing and dynamic process since "music never stops": new artists, new releases, the latest trends in genres, etc. pop up at different corners of the world and have to be assembled. The same issue is affecting Rights Expressions at two levels. Rights conditions might be updated by their authors or cancelled by a Court if inappropriate. Besides, legal concepts definitions are not internationally harmonized but depend on states competences. Similar works will receive different legal qualification and induce different remunerations according to the national jurisdiction they depend, requiring Rights Expressions local adaptation.

4. Prototypical Application for Mobile Music Information Retrieval

Several prototypical systems combining different external web services and internal music processing agents have been realized in our previous work [22,23].

Figure 3: Mobile MIR application

The actual mobile version follows again the multi-agent framework paradigm containing the described ontology as a means for intra agent communication as well as interfacing with external web services. In addition to the previous systems we intended to offer:

- mobile access and
- interactive feedback to the recommendation agent

The mobility aspect was covered by a dense design of the website which can be displayed on a small-sized PDA screen. For the interactive feedback we decided to include a virtual joystick which can be easily accessed using the pen of the PDA. The recommendation agent uses a trimodal global similarity measure which is realized as a weigthed linear combination of three different local similarity metrics $S = w_{so} * S_{so} + w_{ly} * S_{ly} + w_{st} * S_{st}$ with

- S_{so} : sounds-alike similarity
- S_{ly} : similarity of lyrics
- S_{st} : similarity by style/cultural aspects

The position of the joystick has a direct influence on the individual weights, e.g.

- center position: $w_{so}=0.33$, $w_{ly}=0.33$, $w_{st}=0.33$
- sound position: $w_{so}=1$, $w_{ly}=0$, $w_{st}=0$
- lyrics position: $w_{so}=0$, $w_{ly}=1$, $w_{st}=0$
- style position: $w_{so}=0$, $w_{ly}=0$, $w_{st}=1$

In the following three subsections some details about the computation of the local similarity measures will be given. For each of them the input are two entities representing an MP3-encoded audio file of a song with the lyrics attached as a plain ASCII text file and metadata information about the name of the song, album and artist. The output of each local similarity measure is within the range [0..1].

4.1 Sounds-alike similarity S_{so}

The automatic audio analysis recognizes song properties about loudness, tempo and timbral features. These features can be used for the determination of music similarity. For the extraction of basic features such as loudness and psychoacoustic features, we used mel frequency cepstral coefficients and a clustering approach to generate audio profiles [24, 25].
In order to perform a similarity computation based on the audio profiles we use both the Earth Moving Distance, as suggested in [26] as well as a Gaussian Mixture Model as proposed in [27]. Even working with a standard Nearest Neighbor (NN) classifier delivered interesting results for cross-genre recommendations of music "sounding" similar.

4.2 Similarity of lyrics S_{ly}

Here we use state-of-the-art document retrieval approaches. The top-level concept lyrics is broken down into a taxonomy of typical topics covered by mainstream music. In the future such handcrafted topic ontology may be supported by semi-automatic ontology learning through document clustering approaches. For the current experiments we focused first on the subsymbolic level of lyrics. In a Vector Space model, lyrics are represented as vectors. The dimension of the vectors indicate specific terms, the value of a vectors component indicates the number of times the respective term occurs in the lyrics to be represented. The similarity measure is the cosine-measure, which computes the angle between two vectors. Roughly spoken, those lyrics, which share many relevant words, will have a high similarity. Typical results deliver similar songs from concept albums of the same artist (intra-album similarity, example 1) or cross artist similarity by topic (example 2).

1.Song: Phil Collins - One More Night
Most-relevant terms: forever wait night
Similar: P. Collins – You Cant Hurry Love,
P.Collins - Inside Out,
P.Collins - This must be Love

2.Song: Cat Stevens - Father And Son
Most-relevant terms: fault decision marry son settle
Similar: P.Collins - We're Sons Of Our Fathers,
Sheryl Crow - No One Said It Would Be Easy,
George Michael - Father Figure

4.3 Cultural similarity S_{st}

In order to compute cultural similarity we perform a Google search for reviews of the musical work of the artist under investigation. The first 50 pages are downloaded and preprocessed with a filtering stage to remove HTML formatting instructions and find the main textual part of each document. On these textual parts standard techniques from the field of Information Retrieval and Computational Linguistics are subsequently applied. As a result we achieve a cultural representation of the artist which is a vector representation containing the most relevant and discriminating terms and their weightings. On top of this vector representation the cosine measure is used to compute similarity between two entities. More details can be found in [28].

5. Conclusion and Future Work

In this paper we have presented a framework supporting the life cycle of digital music using an ontology as a knowledge sharing and communication platform for human actors and agents. In our future work we will elaborate the applied automated rights management part of our proposition developing the concept of cognitive interface "taking into account a meta-knowledge implied in a situation of dialog with a view to adapting it to types of specific requests in a domain" [11], merging genre requests and legal requirements.

6. Acknowledgements

We wish to thank DFKI, CERSA, Medialive and Prof. Rousseau for supporting this work.

References

[1] Deirdre Mulligan, Aaron Burstein, *Implementing Copyright Limitations in Rights Expression Languages,* 2002 ACM Workshop on Digital Rights Management.

[2] Allamanche E., Herre J., Hellmuth O., Froeba B., Kastner T., Cremer M., *Content-based Identification of Audio Material Using MPEG-7 Low Level Description*, Proc. of the ISMIR 2001, Bloomington, IN, 2001.

[3] Hunter J., *Enhancing the Semantic Interoperability of Multimedia through a Core Ontology*, IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Conceptual and Dynamical Aspects of Multimedia Content Description, Feb 2003.

[4] R. Neches, R. Fikes, T. Finin, T. Gruber, R. Patil, T. Senator, W. R. Swartout: *Enabling technology for knowledge sharing*, in AI Magazine, 12 (3), 1991,36-56.

[5] MPEG-21 *Overview v5*, ISO/IEC JTC 1/SC 29/WG 11 N5231, http://chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm

[6] ISO/IEC FDIS 21000-5, *MPEG-21 Rights Expression Language (REL)*, ISO/IEC JTC 1/SC 29/WG 11/N5839, July 2003.

[7] ISO/IEC FDIS 21000-6, *MPEG-21 Rights Data Dictionary (RDD)*, ISO/IEC JTC 1/SC 29/WG 11/N5842, July 2003.

[8] Trevor Bench-Capon, *Task Neutral Ontologies, Common Sense Ontologies and Legal Information Systems*. Second International Workshop on Legal Ontologies, JURIX 2001, Amsterdam.

[9] Séverine Dussolier, *Fair use by design in the European Directive of 2001: an empty promise*, 12th Conference on Computers, Freedom and Privacy, 2002. http://www.cfp2002.org/fairuse/dusollier.pdf

[10] Melanie Dulong de Rosnay, *Digital Right Management systems toward European law*, In Proceedings of the Second International Conference on Web Delivering of Music, Darmstadt, Germany, December 2002.

[11] Filipe Borges, Danièle Bourcier, Evelyne Andreewsky, Raoul Borges, *Conception of cognitive Interfaces for legal knowledge – Evolution of the JURISQUE project on the risks of avalanches,* ICAIL 2001 The Eighth International Conference on Artificial Intelligence and Law, May 21-25, 2001 St. Louis.

[12] Valente, A., *Legal Knowledge Engineering*. Amsterdam, IOS Press, 1995.

[13] Breuker, J., Valente, A., Winkels R., *Legal Ontologies: A Functional View*, in First International Workshop on Legal Ontologies, University of Melbourne, Australia, 1997.

[14] Sartor, G., Branting, K., *Judicial Applications of Artificial Intelligence*, Kluwer Academic Publishers, 1998.

[15] Visser, P.R.S., Bench-Capon, T.J.M., *A Comparison of Four Ontologies for the Design of Legal Knowledge Systems*, Artificial Intelligence and Law Computer/Law Series, No. 16, Kluwer Law International, 1998.

[16] DOLCE: *A Descriptive Ontology for Linguistic and Cognitive Engineering*, http://www.loa-cnr.it/DOLCE.html

[17] Indecs, http://www.indecs.org/pdf/framework.pdf

[18] ODRL, http://odrl.net/

[19] XrML, http://www.xrml.org/

[20] IPROnto, http://dmag.upf.es/ontologies/ipronto/index.html

[21] Pachet F., Cazaly D., *A Classification of Musical Genre*, Proc. of Content-Based Multimedia Information Access (RIAO) Conference, Paris, France, 2000.

[22] Baumann S., *Using Natural Language Input and Audio Analysis for a Human-oriented MIR System,* In Proceedings of the Second International Conference on Web Delivering of Music, Darmstadt, Germany, December 2002.

[23] S. Baumann and O. Hummel. *Increasing consumer value in virtual music environments*. In International Workshop for Technology, Economy, Social and Legal Aspects of Virtual Goods,, Ilmenau, Germany, May 2003.

[24] S. Baumann, *Music Similarity Analysis in a P2P Environment*, Proceedings of the WIAMIS 2003, London, UK, April 2003.

[25] S. Baumann and Tim Pohle. *A comparison of music similarity measures for a p2p application*. In Proceedings of the Sixth International Conference on Digital Audio Effects(DAFx-03), London, UK, September 2003.

[26] B. Logan*, Mel Frequency Cepstral Coefficients for Music Modeling ,* Proceedings of the Int. Symposium on Music Information Retrieval (ISMIR) 2000, Plymouth, USA, October 2000.

[27] J. Aucouturier, F. Pachet, *Music Similarity Measures: What's the Use*?, In Proceedings of the ISMIR 2002, Paris, France, Oct 13-17, 2002.

[28] S. Baumann and O. Hummel. *Using Cultural Metadata for Artist Recommendation*. In Proceedings of the Third International Conference on Web Delivering of Music, Leeds, UK, September 2003.