Music indexing and retrieval: evaluating the social production of music metadata and its use

Jean Debaecker, Laboratoire Geriico, University of Lille 3, France

Widad Mustafa El Hadi, Laboratoire Geriico, University of Lille 3, France

Abstract

This article will focus on music indexing and retrieval from different points of view. Four elements will be examined: music metadata, indexing and retrieval methods (classic indexing, collaborative indexing and social tagging), tools and users. Regarding the users, we will look at their access modes, their possible participation in indexing, and their music information-seeking behaviors. Do they look for music information in pay-for-music sites (iTunes), radio stations, blogs, or social networks (MySpace, YouTube, etc)? The question we raise is which music information systems would be suitable in a social web era. An evaluation of the existing indexing and retrieval modes was conducted. It was based on both quantitative and qualitative approaches. Our research methodology used interviews, online questionnaires and semi-directed questionnaires. We believe that the results of our evaluation will be useful for music information indexing in a Socio-Semantic Web context.

Background and rationale

This article describes the results of a project supported by the French Institute for Communication (ISCC-CNRS), which examines the problems related to the social production of music metadata. The focus of the study will be on the new devices and digital practices related to Web 2.0 and how they can be extended to the future Web 3.0. Since bridging the gap between the two Webs is an ongoing endeavour (Zacklad, 2003, 2007), we are trying to examine the possibilities for music indexing strategies and tools.

In the first place, we identified music indexing places and forms of production in order to understand and appreciate new practices and to evaluate music description and indexing modes. We began by assessing and analyzing user practices, music reception and the use of this indexing.

This first step was to understand the nature of metadata (section 2) and indexing types (section 3) in order to understand indexing mechanisms before evaluating the social indexing results (section 5). The idea is to compare metadata resulting from a hierarchical, centralized, controlled indexing to metadata which can be broadly defined as:

> 'Data that serves to provide context or additional information about other data. For example, information about the title, subject, author, typeface, enhancements, and size of the data file of a document constitute metadata about that document. It may also describe the conditions under which the data stored in a database was acquired, its accuracy, date, time, method of compilation

and processing.' (BusinessDictionary.com)

We could also identify two types of metadata as far as music indexing is concerned: metadata resulting from standard or mainstream indexing (commercial sites, recording companies) and metadata from a socially-distributed environment (for example, blogs, web radios, social networks, virtual music and many others that will be described in section 4.)

We have also been able to examine the emergent information practices related to this new shape of indexing, which provided an appreciation of its relevance to digitized music via the techniques of social indexing. The idea is to bridge the gap between the Web 2.0 and the Web 3.0 practices in order to achieve the new 'Social Semantic Web', an expression coined by Zacklad and Cahier (2003) which is defined in Wikipedia as follows:

> 'The concept of the Social Semantic Web subsumes developments in which social interactions on the Web lead to the creation of explicit and semantically rich knowledge representations. The Social Semantic Web can be seen as a Web of collective knowledge systems, which are able to provide useful information based on human contributions and which get better as more people participate. The Social Semantic Web combines technologies, strategies and methodologies from the Semantic Web and Social software and the Web 2.0'

The Social Semantic Web, focuses on the 'importance of humanly created loose semantics as a means to fulfill the vision of the semantic web. Instead of relying entirely on automated semantics with formal ontology processing and inferencing, humans are collaboratively building semantics aided by socio-semantic information systems'. This description is quite relevant for our project, which could be extended in the future to producing socio-semantic music metadata, thus contributing to the empowerment of future search engines in a Socio-Semantic Web era.

The most important feature of the Socio-Semantic Web when compared to the Semantic Web can be summarized in the following: the Semantic Web often is regarded as a system that will solve the epistemic interoperability issues we have today, thus providing ways for businesses to interoperate across domains. The Socio-Semantic Web will enable users to share knowledge through collaborative indexing.

Metadata, definition and typology

In this section we will discuss the notion of metadata in the general sense of the term, and will detail in particular the metadata approach in the field of music.

The term metadata has long been associated to the <meta> tag in HTML language, a web page indexing method. It was later extended to textual resources with the aim of supporting notions of various types - legal, technical, and semantic (Calderan, 2007). These metadata are varied: they can be content data created by humans or data targeted to only machine processing. A metadata element is defined as a 'reinterpretable representation, in a conventional form adapted to communication, interpretation or processing'. (Dalbin, 2007). Metadata can then be seen as a set of descriptive data. They can be considered within the scope of search and information access tools. The great variety of objects they describe makes their standardization impossible, but whatever their differences and variety, all metadata are used for processing, management, and dissemination of information.

The description of metadata related to the field of music is subject to the same rules as bibliographical metadata: a tag describes a type of information (title, author, date, editor, subject, etc.), and its values are represented by keywords, descriptors and summaries. Music metadata are indexed in ID3 (IDentify

an MP3), a 128 byte character string, coded in ISO-889-1, which describes the audio file. This informal standard was developed in 1996 by Eric Kemp, and has since been widely adopted. The MP3 itself (an abbreviation of MPEG-1/2 Audio Layer 3) is a standard produced by the Moving Picture Experts Group. The compression format used is very specific. It consists of reducing the quantity of significant data required to build a sound that is a faithful reproduction.

Music metadata and Web 2.0.

A new generation of users of websites associated with the concept of Web 2.0. has witnessed the emergence of new types of collaborative applications and user participation. To meet these new demands, editors and content producers offer a wide set of new browsing and collaborative tagging tools, based on a collaborative content classification produced through free indexing tags. This practice is commonly called 'folksonomy'. Folksonomies are a collaborative classification process based on keywords freely chosen by internet users. This indexing mode is also called 'social indexing', as opposed to metadata driven from controlled lists, thesauri, ontologies, topic maps, etc. produced by a valid authority.

Music and indexing

The authoritative encyclopedic and alphabetical index RAMEAU, the French equivalent of the American Library of Congress Subject Headings, divides music indexing into nine classes (Class 000: World music, Class 099: French songs, Class 100: Jazz, Blues, Gospel, etc.). Music classes however evolve very quickly and these classifications require constant revision.

The massive use of digital technologies (the Internet, online music catalogues, personal music libraries, distributed libraries, etc.) highlights the impatience of the users (amateurs or others) who generally cannot wait until the authority catalog has been updated, and who therefore try to put forward their own indexing terms. We could give many examples, but let us start with what is generally accepted as corresponding to the definition of the term 'rock music'. We then realize that there is very little consensus on which groups fit this category, even among the groups themselves. Until recently this was problematic only for information professionals. For example, an information professional might have tagged the rock band Nirvana in class 200 (2 xxx 63), in the 'grunge' category, whereas the band itself refused such a classification and identified itself as belonging rather to the pop music category with punk influence. Technically, the problem was solved very simply: the record company launched the term 'grunge', the information professionals seized it, and the band was thus classified. The user with this information knows how to find content about this band, but if a member of the band (or one of their fans) searches a database for the music, he or she will look for 'pop music' or 'punk' and will get no results. Beyond this problem however remains the question of knowing exactly what object should be indexed: the artist, the album or the song? The artist can have very different styles and can skip easily from a pop song to electronic or even punk music. This has a direct impact on user satisfaction.

The emergence of social networks produced an evolution in this field. The user can now overstep authority registries and index his music collection with his own words. These can be new terms, or terms already used by the RAMEAU authority list, or terms which differ from current use. This is the heart of folksonomy or social bookmarking. This means that a non-specialist user becomes a central actor of content description tools processes, i.e. of metadata organization.

Music storage formats, like MP3, enrich music digitization possibilities. Metadata play a major role in social web content management. We refer to sets of data that are of a sufficiently significant size to facilitate access to the information content of a digital document and make it relevant for computer

processing. We consider the question of music information production and representation, its formalization and organization; but it is also necessary to consider how to connect media coverage and mediation, i.e. communication (for a recent review of these ideas cf. Couzinet, 2009).

Content Management of a digitized sound item is processed through tags (in ID3v1/2 format), and goes through three levels of document indexing: content – authoriality – materialization .The content refers to the description of the song itself: song title, album title, music type, etc. Authoriality refers to the intellectual property owners: the author, the copyright, composer, translator, etc.   Material characteristics could be date, format, duration, frequency, sampling, etc.

The volume of music to index has increased exponentially over the last few years, and this can be related to efficient audio compression methods, inexpensive storage space, and increased accessibility to large communication bandwidths (Laplante, 2008). As a result, tools have been developed to facilitate the exchange, management, retrieval, and purchase of music in digital format. Laplante focuses on the importance of the social network users and their role in enriching music and its discovery. His research shows how current systems are based on the intuition and personal experience of their developers.

For these reasons, companies are implementing advanced search algorithms such as Apple iTunes with its function Genius, or the MuMa Music Mashup project undertaken by Quaero. This project is a Franco-German venture, and its aim is the design and the implementation of music information extraction tools, digital multimedia analysis and classification tools. MuMa includes music information retrieval search engines developed and implemented by Exalead. It is integrating the IRCAM, Télécom ParisTech and Yacast technologies. These research projects and findings show how relevant our project is.

Quintarelli (2005) has extensively showed the importance of tags for individual and personal indexing. He suggested in his publications an outline of folksonomy types. He identifies a 'broad folksonomy' exemplified mainly by del.icio.us (see http://www.delicious.com/url/07f35bf8e4ea69da8fbeda4d7cb61d82)where many users tag the same digital object in relation to their personal and mental representations. For instance, an article which describes the commercialization of a new technological object such as Apple iPad 2, can possibly be tagged with 'apple-ipad-2, 'ipad', 'apple-ipad-and-ipad-2, etc.

Quintarelli also identifies a 'narrow folksonomy' as exemplified by Flickr. In this case a few users tag a digital object (photos in Flickr's case) for their own convenience. It is difficult to design a tool which enables tag associations without a framework while still enabling indexing formalization and term hierarchy. If we associate a tag for example to the word 'Apple', nothing will tell us if it is the fruit or the organisation name.

Overview of indexing methods associated with music work

Through Web 2.0, the music industry has become a privileged place to observe the practices of shifting non-professional indexing practices and consumer behaviour (Beuscart, 2007). In this article, the author identifies three different phases of relations where the music industry, the computer processing industry, and independent practices meet. He first mentions the notion of 'pure player' i.e. independent actors whose main activities take place online: they have no expertise in the field of music, but have begun to upload songs via different start-ups. The second phase started with the role of the music industry (Fnac, Universal, Virgin, etc). The third one has seen the computer processing industry participating in the trend with actors like Apple iTunes, MSN Music and many others.

We will particularly consider this recent integration of the computer processing industry that meets our professional documentary concerns. User practices are inevitably linked to the technical evolution of the

internet, particularly in the field of tag indexing. Uploading music documents requires powerful indexing tools and adapted search engines.

Three types of indexing are considered : free vocabulary indexing by the user; free vocabulary technology, with user-centred indexing, and free vocabulary indexing for the user, that is also technology centered. The way information is produced and processed is important for us, as long as it can contribute to answer some of our questions, in order to create a pattern. We will then see to what extent these indexing methods and search engines effectively collaborate. We will try to account briefly for the three methods.

Free vocabulary indexing by the user

Here the user is free to index digital content with his or her own terms which do not belong to authority lists, or to a keyword hierarchy. This is the case in particular for YouTube, MySpace or Flickr, where the internet user is free to enter his or her own keywords and content description. On MySpace, the user can even tag parts of the downloaded image. For example, on a wedding photograph, it is possible to select the bride's face and add her first name.

To illustrate, if we take YouTube as an example, we can notice the following: what some users desire is not to broadcast the music but rather to communicate their own comment. During the 2007 French presidential elections we witnessed the arrival of users who counter-attacked the opposition party through this mode of communication, by spreading such messages. What they actually valued was not content but tags. Each downloadable tune contained a text in favour or against one or the other candidate. This means that on listening to the song, our own multimedia player displayed their message. These metadata even embedded links that could automatically open a web page or a blog, etc. This means that in this case music becomes a means, a pretext, a technical medium to achieve an aim or to carry a message. How could an information professional index such a tune? The name of the composer and music type are not sufficient, especially when the tag is a pseudonym. And even when the user does not download the song, the metadata are still displayed on the page, and difficult to ignore.

Information processing tools used by internet users are increasingly varied and rich. The availability of professional tools such as metadata constitutes progress. The use made of metadata is currently going through a paradigmatic evolution, and its starting point has been a deterioration of their socially diverse context and network speed: new communication tools appear through old protocols. Initially conceived as linked to a homogenous population of objects, metadata have become entities that are finally disconnected from their original aim. The 'visual-practical' space reserved for metadata is increased by new adjuncts.

Free vocabulary technology, user-centered indexing

This kind of data indexing draws on the previous one. The software collects as much data as possible, and suggests to the user to add his or her own keywords, without limitations and without advising him or her on how to fill in the categories. In practice, this is how many online commercial sites work (Amazon, Price Minister, etc). On their home page, different 'personalized' sections are suggested to the user. On Amazon for example, we find the following:

- 'Customers who bought this item also bought…'

- 'Inspired by your navigation history…'

- 'Recommended for you…'

The web site does not explain how these sections process information that is, most of the time, relevant. Moreover the user can tag his or her own items freely. And this will make other user searches more relevant.

Delicious Library is a multimedia file processing system: song books, mainly specialized in book records management. And yet, when a new book is added into the system, it automatically suggests book recommendations in partnership with Amazon.

Evaluation methodology

As we mentioned in the introduction to this article, in order to understand music indexing strategies and tools, and to assess new practices related to Web. 2.0. and its new indexing modes, we conducted an evaluation project. Regarding the users, we looked at their access modes, their possible participation in indexing, and their music information-seeking behaviours. An evaluation of the existing indexing and retrieval modes was conducted. We have adopted a qualitative and quantitative approach. To this end we have designed an online questionnaire, together with a printed version, in order to collect extensive and representative data. We thus defined and implemented the following phases:

The objectives of the questionnaire

The objective was to find out how users practically index music information, and which strategies they use bearing in mind that they do not have prior training in professional indexing. We also wanted to discover their search strategies (who listens to music, where, which genre, etc.)

The target population taken as a sample was composed of the following:

- music amateurs

- music producers

- music libraries customers

We were interested specifically in their search strategy, thus the population was not targeted according to social and occupational criteria.

Submitting the questionnaire

Basically, the questionnaire was submitted to users directly in music stores and music distribution points, such as FNAC, Virgin, recording studios, etc. We strategically uploaded it on an on-line platform, available at http://www.icomusicquestionnaire.c.la/ which allowed us to reach a larger number of reviewers

Analyzing results

The privileged search spaces on the internet are the recognized commercial retail websites (Amazon, iTunes, etc.) and the non commercial ones, such as MysSpace, Last.fm, Blogs, etc.

We are listing these strategies in the following:

1. By *hearing* and *saying*

2. Through metadata use or by navigation and browsing

3. Through search engine queries.

Their criteria are essentially and respectively:

1. Efficiency

2. Confidence

3. Following a friend's recommendation who shares the same type of interests.

4. Curiosity

5. Speed of processing

6. The navigation organized into a hierarchy by tags allows investigating and to making new discoveries by serendipity

7. Simplicity (Google, etc.)

These results are partial and incomplete since we are still receiving answers to our questionnaire.

The user search strategy seems to them globally relevant, and simple to carry out. However in terms of evaluation, half declare themselves unsatisfied by the results. They consider that, for most of the results obtained, the first ones are useless (a high rate of noise). They however prefer giving more importance to the time they spend in searching rather than to the flexibility of the means offered by the search engines : 'we prefer typing whatever we want to rather than to be limited by search strategies imposed by the search engines aiming at the efficiency and results of the search'.

The reasons here are very simple: on one hand, those who do not practice social tagging say this activity is a waste of time, while others recognize the role of metadata. Their main concern is that people searching for a song find easily and quickly the song they are looking for.

The results of our evaluation show that, the users by tagging the various web sites can enrich music metadata and this is one of the targets. Enriched social driven metadata combined to commercial and mainstream one can be a basis for indexing music information. Other research findings in this field could be used to enhance these possibilities.

Our results, although partial, could be compared to the findings of Laplante (2008). The interviews conducted by her showed comparable results to our project, namely the following: 'analysis of the interviews shows that young adults rely heavily on their social network to discover new music. The main reason for this is twofold: 1) their friends, colleagues or relatives know their tastes and can provide personalized recommendations, and 2) as they know the tastes of their friends, colleagues or relatives, they are able to determine whether or not to trust their recommendations'.

Our results and other findings such as Laplante's confirm the importance of taking into consideration metadata driven from social network indexing and to combine it to a classical music metadata in order

to enhance search engines capacities for the future socio-semantic Web.

Perspectives : how can we empower the engines for a Socio- Semantic-Web era?

What is the role of search engines? Can they carry on their indexing tasks without considering social metadata processing? The stake is in the evaluation of metadata relevance in relation with the indexed content. Multimedia contents (videos, images, songs) are particularly difficult to analyze and be processed by search algorithms. If we take the example of the presidential election mentioned above, what relevant information should the search engine index: the video itself? the song? the metadata (and the new role they play as seen previously)? the whole set of these data? Search engines progress in the way they process metadata. We are forward to their integrating socio-semantic metadata in the near future.

References

Beuscart, J.S, (2007) Les transformation de l'intermédiation musicale, la construction de l'offre commerciale de musique en ligne en France. In *Réseaux,* 25, 141-142/2007, 143-176.

Beuscart, J.S., (2008), Sociabilité en ligne, notoriété virtuelle et caractère artistique: les usages de MySpace par les musiciens contemporains autoproduits. In Cardon, D. (ed) (2008) *Réseaux* 26, 152/2008, 139-168.

Business Dictionary. [Metadata definition.] Retrieved from http://www.businessdictionary.com/definition/metadata.html

Cahiers du LCPE, 4, à paraître.

Calderan, L., Hidoine, B., Millet, J. (eds), (2008), Métadonnées: mutations et perspectives, Séminaire INRA, 29 septembre – 3 octobre 2008, Dijon, ouvrage coordonné.

Calderan, L., Métadonnées: mutations et perspectives, Introduction. In Calderan, L., Hidoine, B., Millet, J. (eds), (2008), *Métadonnées: mutations et perspectives*, Séminaire INRA, 29 septembre – 3 octobre 2008, Dijon, p. 5-8.

Cardon, D. (dir) (2008) *Réseaux sociaux de l'Internet, Réseaux*; vol 26, 152/2008, dossier coordonné par Dominique Cardon, Hermès-Lavoisier.

Chartron, G., & Broudoux E., (eds) (2006). Document numérique et société, *Actes de la conférence organisée dans le cadre de la Semaine du document numérique à Fribourg (Suisse) les 20 et 21 septembre 2006*. ADBS Éditions, 2006. 344 p. Collection Sciences et techniques de l'information, Fribourg.

Couzinet, V., (ed), (2009), *Dispositifs info-communicationnels, questions de médiations documentaires*, Introduction, Hermès-Lavoisier, Paris.

Crepel, M. (2008), Les folksonomies comme support émergent de navigation sociale et de structuration de l'information sur le We. In Cardon, D. (ed) (2008) *Réseaux*, 26, 152/2008, p. 169-2004.

Dalbin, S., (2008) Représentation et accès à l'information: transformation à l'œuvre, In Calderan, L., Hidoine, B., Millet, J. (ed), (2008), *Métadonnées: mutations et perspectives*, p. 10-58.

David, S. (1997). Représentations sensorielles et marques de la personne: contrastes entre olfaction et audition. In *Catégorisation et cognition. De la perception au discours*, ed. D Dubois, pp. 211-242. Paris: Kimé.

David S, Dubois D, Rouby C, Schaal B (1997). L'expression des odeurs en français: analyse lexicale et représentation cognitive. *Intellectica*, 24 (1), 51-83.

Dubois D (ed.) (1991). *Sémantique et cognition*. Paris: Editions du CNRS.

Dubois D (ed.) (1997). *Catégorisation et cognition. De la perception au discours*. Paris: Kimé.

Ertzscheid, O., & Gallezot, G. (2006), Etude exploratoire des pratiques d'indexation sociale comme une

renégociation des espaces documentaires. Vers un nouveau big bang documentaire ? In *Document numérique et société sous la dir. de Ghislaine Chartron et Evelyne Broudoux*, p. 291-306.

Frith, S., & Le Guern, Ph. (eds), 2007, Sociologie des musiques populaires, *Réseaux* 25, 141-142/2007

Laplante, A. (2008). Everyday life music information-seeking behaviour of young adults: an exploratory study. Thèse de doctorat, McGill University.

Quintarelli, E. Folksonomies: power to the people. In *ISKO Italy-UniMIB Meeting, Milan, Jun. 2005* Retrieved from http://www.iskoi.org/doc/folksonomies.htm

The Social Semantic Web. *Wikipedia*. Retrieved from http://en.wikipedia.org/wiki/Social_Semantic_Web.

Vander Wal (2005). *Folksonomy explanations*. Retrieved from http://www.vanderwal.net/random/entrysel.php?blog=1622

Corresponding author :

Widad Mustafa El Hadi can be contacted at widad.mustafa@free.fr