Collaborative and Semantic Enrichment of Musical Libraries in the VARIAZIONI project

Carlos A. Iglesias	Francesco Spadoni	Jens Grivolla	Joachim Neumann
Germinus XXI (Grupo Gesfor)	*Rigel Engineering*	*Music Technology Group Univ. Pompeu Fabra*	*Music Technology Group Univ.t Pompeu Fabra*
cif@germinus.com	*spadoni@rigel.it*	*jgrivolla@iua.upf.edu*	*jneumann@iua.upf.edu*

Abstract

This article introduces the content enrichment techniques that have been explored and combined in the eContentPlus VARIAZIONI project[1]. The project has proposed a content metadata model for musical assets based on FRBR, which has been integrated with a standard content management system. Wizards for manual metadata feeding have been developed within the content management system. In addition, it supports social annotation, and includes automatic enrichment based on the sound properties of the contents and on smart clipping from available Internet resources, such as Wikipedia or Yahoo.

1. Introduction

The web 2.0 phenomenon and its social approach is reaching new application domains, such as the enterprise systems (so called *Enterprise 2.0*) [La06,Cr06] and, more recently, is approaching the scope of Digital Libraries, so called *Library 2.0* [Ca06]. The main peculiarity of this approach is its user orientation. The Library 2.0 approach proposes that the user is not longer a pure passive entity that consults a catalogue. Instead, the user acquires greater protagonism, and contributes with his opinion about the items and an active role in their cataloguing.

One interesting notion of the Web 2.0 is "the Long Tail". In the context of the Library 2.0, while most bookshops or even libraries can only host the most demanded books, due to space and budget restrictions, the Internet has shown the value of the specialisation. Niche content providers and business can find a profitable business thanks to the attraction of minorities, which in Internet terms, constitute an attractive commercial target. In the same way, these specialised communities create social networks and can be even organised into movements such as the Open Source Movement.

This article presents the content enrichment techniques of the project Variazioni in order to apply the emerging Library 2.0 approach to the Musical Digital Libraries.

The rest of the article is organised as follows. Section 2 introduces the context of this research, the eContentPlus project Variazioni. Section 3 describes the content enrichment techniques that are being explored in the project. Section 4 describes the basics of the Variazioni Content Model and its user orientation, as well as the workflows of content enrichment that have been identified within the project. Section 5 introduces the role of social annotation in content enrichment. Section 6 and 7 describes briefly automatic content enrichment through sound analysis and smart Internet clipping, respectively. Finally, section 8 draws conclusions and presents future works.

2. The Variazioni project

The Variazioni Project is an eContentPlus Project funding as Content Enrichment Project with a duration of 30 months, starting on September 2007. The project is being coordinated by the musical private institution Fundación Albéniz and counts with several additional musical institutions (Lithuanian Academy of Music and Theatre, Koninklijk Conservatorium Brussels, Escolal Superior de Música e Artes do Espectáculo do

[1] This research has been co-funded by the European Community under the programme eContentPlus. The authors are solely responsible for this article and it does not represent the opinion of the European Community. The European Community is not responsible for any use that might be made of information contain within.

co-funded by the Community programme eContentplus

Porto, Sibelius Academy, and Association Europeenne of Conservatoires, Academies de Musique et Musikhochschulen) and technical partners (Germinus XXI, Rigel Engineering, Exitech, Universitat Pompeu Fabra and Università degli Studi di Firenze).
The purpose of Variazioni is to provide a Content Enrichment Portal where users and musical institutions can publish, annotate and access musical contents, including its protection. In order to validate its approach, the project will provide a minimum of 700 audiovisual hours, 1000 audio hours and 2000 written documents.

3. Content Enrichment Techniques

In the context of Variazioni project, content enrichment is defined as the process of adding new metadata to contents. In order to bring the Library 2.0 concept into the musical assets, the project has analyzed the requirements of users and musical content providers in order to publish and enrich their assets. Some of these requirements are (i) *easiness*, content should be easy to add and enrich; (ii) *specialization*, specific metadata should be defined in order to provide efficient retrieval and accurate cataloguing; (iii) *security*, since some of the contents are copyright protected securing mechanism should be provided in order to make them available online. Several content enrichment techniques have been identified: manual enrichment, automatic enrichment, social enrichment, and repurposing contents into different contexts.

Manual enrichment is the process of adding metadata by users according to a predefined metadata model. This is the traditional way of cataloguing items by librarians, with cataloguing standards and rules, such as AACR [AacrURL], MARC [MarcURL] or MODS [ModsURL],

Automatic enrichment adds new metadata based on the content characteristics (textual or audiovisual) or pre-existing metadata.

Social enrichment consists of adding annotations and comments by the users. The annotations given by the users constitute a free text taxonomy that is called folksonomy (also known as collaborative tagging, social classification, social indexing or social tagging). There are several important differences between this approach and the manual one. Firstly, this is a bottom-up cataloguing process, since there is no predefined taxonomy, but it emerges from the individual free annotations. In addition, social enrichment metadata is generated not only by experts, but also by content creators and consumers. Moreover, since the same item is annotated per different users, it is possible to improve the annotation of contents using the most popular annotations.

Repurposing contents consists of reusing a content in a different context, and adding metadata in this process. In the traditional libraries, items are classified in an aseptic way. Social tagging has brought contextual tagging, since users can add tags to a content depending of their current interests.

Variazioni combines these approaches taking in mind a user orientation approach, and are described below: manual enrichment (section 4), social enrichment (section 5), automatic enrichment, which is carried out in two different ways: based on audio properties of the contents (section 6) and on pre-existing metadata (section 7).; and repurposing, thanks to the usage of a standard content management system, where users can reuse contents for building new contents, such as articles, critics, etc.

4. Manual Content Enrichment

Variazioni proposes that manual content enrichment can be done in a collaborative way, and both content providers and content consumers can help in the cataloguing process.

4.1. Variazioni Content Metadata Model

Variazioni Content Model is based on FRBR (*Functional Requirements for Bibliographic Records*) [IF98]. FRBR is a conceptual entity relationship model developed by the International Federation of Library Associations and Institutions (IFLA), which has supposed a paradigm shift in cataloguing, since it considers the requirements of the users for searching, identifying, selecting and obtaining details of the bibliographic record.

The FRBR model is structured in three groups with different entities for each group. The first group comprises the core entities that represent the products of the artistic or intellectual model: work, expression, manifestation and digital item. A *work* is a distinct intellectual or artistic creation, for example a composition. An *expression* is the intellectual or artistic realization of a work, for example, the interpretation of a composition in a concert. A *manifestation* is the physical embodiment of a work, for example a CD production with the recording of the concert. A *digital item* is a single exemplar of a manifestation, for example, one CD bought at a shop with a serial number.

Some of the advantages of this model is that it is easy to establish relationships between different digital

items. One can catalogue that different manifestations (i. e. video, book and audio recording) correspond to the same expression (i.e. musical event), or define that one digital item complements another one, for example. In addition, FRBR metadata has been defined taking into account the users' needs.

In order to apply this model in Variazioni, several assumptions have been taken. Firstly, since Variazioni is a digital library, the notion of digital item has not been included. Secondly, since our scope is a musical library, the notion of work has been mapped onto musical composition. Users only catalogue manifestations (called musical contents), but the model makes the difference between expression and manifestation. In this way, if for example, a user has catalogued a video of a concert, he/she can add an audio of the same concert, without having to introduce again all the metadata of the musical event. The reason of this is to simplify the way uses introduce the contents.

In order to adapt the general model to musical contents, several expressions (content types) have been defined with their associated metadata (concert, score, libretto, class) and several manifestations with their associated metadata (video, audio, paper and image). The user interface makes this differentiation transparent. In addition, metadata has been organized in aspects, in order to promote its reuse for different content types.

4.2. Content Enrichment Workflows

Variazioni has defined several workflows in order to assist in the enrichment process, which is described below.

The first workflow is *Content Enrichment Review*. When a user different from the content owner enriches a content, there is the possibility that the content owner do not agree with this contribution. It is needed some quality assurance mechanism in order to ensure metadata quality. Several strategies could be applied in this case. An approval of changes could be defined, and each content owner would validate the content contributions. Another alternative is notification of changes, the contributions are accepted automatically but the user receives notification of these changes. The latter is the one selected currently for Variazioni, taking into account mainly the overhead that this approval could mean for content providers.

The second workflow is *Content Protection*. When content owners select that the content should be protected, there is a workflow which starts the protection of the media file with Axmedis P2P Network and synchronizes all the metadata between the content management system and the Axmedis MPEG21 database. Once the media is protected, the media is available through the portal.

Translation workflow. It would be feasible to include a translation workshop, inside a musical institution, that assigns translation tasks to translators when a content is created. Nevertheless, translation is treated as a standard enrichment workflow, since there are no special requirements.

Finally, there are workflows for *automatic enrichment*. Depending on the nature of the content, a workflow can be started in order to produce a thumbnail of the content, or invoke automatic enrichment as described in section 6 and 7.

5. Social Annotation

VARIAZIONI adopts and refines emerging collaborative practices for content enrichment based on web2.0 concepts, leveraging the use of folksonomies , focusing on user participation, and exploiting the architecture of the web as a platform.

In particular, it developed tools and facilities for exploiting socially derived taxonomies (i.e. folksonomies). This classification schema has proven to be very accurate when communities tag the same resources, and in addition it provides feedback for improving quality of tagged resources.

Such tools enable the participation of user communities in the classification of existing content they are interested in, as well as content they create/integrate. Hence, the tools developed support the creation of user generated tags (in form of folksonomies) which can complement and enrich the existing metadata. The tools consist of a rich multifaceted user interface (UI) adapted to the specific VARIAZIONI content and metadata, wizards for tags selection and insertion, powerful functions for simultaneous tagging operations on multiple content objects, support for quality assurance mechanisms (support for evaluation, revisions and modifications of tags).

VARIAZIONI tools allow users to tag content objects with a descriptive word, expressing a characteristic of the content or associated meaning. They represent folksonomies as a tag cloud, which displays the most popular tags;

When a user creates a tag for a specific content object, the tag is stored in the database and associated with that object (by its ID). The system keeps track of all the tags that users have entered and the number of times that they have entered the same tag.

Each tag is visualized with a font size based on the popularity of that tag. This allows users to browse the content by way of a user-driven categorization of that content. When a user clicks on one of the tags in the tag cloud, the application retrieves a list of the associated contents. searching by tags (folksonomies) or by metadata.

Like most advanced digital users, VARIAZIONI users are increasingly interested in accessing all the aspects of a digital content, like user-generated video, photos, podcasts, music, games and more. They want access to all available data, all in real-time. VARIAZIONI is leveraging the users themselves to help organize the content and make it accessible and searcheable, using the "wisdom of the crowds". As VARIAZIONI CEP support the use of tags, it is able to assemble collections of social media based on the interests of VARIAZIONI users.

Significant effort went in making the user interface simpler, clean and more intuitive, doing user testing, performing validation sessions and listening to VARIAZIONI users, collecting and prioritizing what they wanted, liked and disliked.

6. Audio-based Content Enrichment

In order to automatically describe musical content by analyzing audio material, software has been developed and customized for the Variazioni project. In this software, music is described according to different meaningful facets:

1. Timbre, related to its instrumentation.
2. Tonality, related to its harmony and melody.
3. Rhythm and structure, related to the temporal location of events.
4. Dynamics, related to loudness and expressivity.

The result of this automatic tagging is integrated into the metadata description scheme defined in WP2, more specifically in D2.2 VARIAZIONI Musical Metadata Definition.

For this task, existing technology related to audio description of popular music has been adapted to the user profile of music professional and music lovers and specialized in classical music and genres included in the VARIAZIONI collection, which are mainly classical and folk music.

The software is based on an analysis of different levels of abstraction:

- Low-level features are closely related to the audio signal.
- Mid-level features use statistics and machine learning to create descriptors that are semantically meaningful.
- High-level descriptors provide relevant meaning to human users. They often require a modeling process.

7. Smart Clipping Content Enrichment

In order to supply some metadata without user intervention, existing sources of information are used to produce an initial set of descriptive tags for the given content.

In particular, Wikipedia serves as a basis for selecting an initial set of potential labels, taken from the Wikipedia page for an artist/composer or a specific work. A special algorithm is used to robustly match the available metadata (composer name, title of the work) to the pages contained in Wikipedia, allowing for variations in spelling, etc.

The candidate labels are then weighted, ranked, and filtered based on a measure of mutual information with the work or composer in question. The Yahoo Search web services are used to obtain the necessary statistics about the co-occurrence of different terms on the web, which we take to be representative of topical correlation.

The enrichment component also leverages existing folksonomies by connecting the content added to the Variazioni library to annotations provided by users in music related communities on the web. These annotations are filtered to reduce the noise typically found in uncontrolled open sources of user supplied information, and then included in Variazioni.

The use of tags extracted purely from web searches was also investigated (e.g. using the Yahoo term extraction APIs), but has proven to not be feasible with sufficient accuracy.

8. Conclusions and Future Work

This paper has shown different strategies for content enrichment in the project VARIAZIONI. Currently these strategies are being testing with real users. A real time monitoring system has been developed in order to adapt these strategies to users' interests and improve its effectiveness. In particular, next phases of the project will explore the ability to create user communities for attracting users to the system and improving the metadata of digital items according to their interests and peculiarities. The usage of web2.0 strategies for content enrichment can help to provide accurate cataloguing with a sustainable financing model.

9. References

[AxmURL] Axmedis Project Official Web Site, available at http://www.axmedis.org.

[AacrURL] Anglo-American Cataloguing Rules (AACR) official web site, available at http://www.aacr2.org/.

[Ca06] Casey, Michael and Savastinuk, Laura. Library 2.0. Service for the next-generation library. Library Journal, January 2006

[Cr06] Crawford, Walt. "Library 2.0 and Library 2.0". Cites & Insight: Crawford at Large, 6, number 2, 2006. Available at: http://citesandinsights.info/v6i2a.htm

[IF98] International Federation of Library Associations and Institutions, IFLA, Functional Requirements for Bibliographic Records, 1998. Available at: http://www.ifla.org/VII/s13/frbr/frbr.pdf

[La06] Martin LaMonica, "Web 2.0 meets the enterprise", ZDNet News, April, 2006.

[MarcURL] MAchine Readable Cataloguing, developed by the Library of Congress, official web site available at http://www.loc.gov/marc/.

[ModsURL] Metadata Object Description Schema (MODS), official web site. Available at http://www.loc.gov/standards/mods/

[WiURL] Entry Folkonomy at Wikipedia, available at http://en.wikipedia.org/wiki/Folksonomy

[1] A.B. Smith, C.D. Jones, and E.F. Roberts, "Article Title", *Journal*, Publisher, Location, Date, pp. 1-10.

[2] Jones, C.D., A.B. Smith, and E.F. Roberts, *Book Title*, Publisher, Location, Date.