

AN OPEN WEB AUDIO PLATFORM FOR ETHNOMUSICOLOGICAL SOUND ARCHIVES MANAGEMENT AND AUTOMATIC ANALYSIS

Thomas Fillon
PARISSON / LAM,
Institut Jean Le Rond d'Alembert,
UPMC Univ. Paris 06,
UMR CNRS 7190
thomas.fillon@parisson.com

Guillaume Pellerin, Paul Brossier
PARISSON
16 rue Jacques Louvel-Tessier
Paris, France
guillaume.pellerin@parisson.com

Joséphine Simonnot
CREM, LESC, UMR CNRS 7186
MAE, Université Paris Ouest
Nanterre La Défense

ABSTRACT

Since 2007, ethnomusicologist and engineers have joint their effort to develop a scalable and collaborative web platform for management of and access to digital sound archives. This platform has been deployed since 2011 and hold the archives of the *Center for Research in Ethnomusicology*, which is the most important collection in Europe. This web platform is based on *Telemeta*, an open-source web audio framework dedicated to digital sound archives secure storing, indexing and publishing. It focuses on the enhanced and collaborative user-experience in accessing audio items and their associated metadata and on the possibility for the expert users to further enrich those metadata.

Telemeta architecture relies on *TimeSide*, an open audio processing framework written in Python which provides decoding, encoding and streaming methods for various formats together with a smart embeddable HTML audio player. TimeSide also includes a set of audio analysis plugins and additionally wraps several audio features extraction libraries to provide automatic annotation, segmentation and musicological analysis.

1. INTRODUCTION

In social sciences like anthropology and linguistics, researchers have to work on multiple types of multimedia documents such as photos, videos, sound recordings or databases. The need to easily access, visualize and annotate such materials can be problematic given their diverse formats, sources and given their chronological nature.

In the context of ethnomusicological research, the Research Center on Ethnomusicology (CREM) and Parisson, a company specialized in big music data projects, have been developing an innovative, collaborative and interdisciplinary open-source web-based multimedia platform since 2007.

This platform, *Telemeta* is designed to fit the professional requirements from both sound archivists, researchers and musicians to work together on big music data. The first prototype of this platform has been online since 2010 and is now fully operational and used on a daily basis for ethnomusicological studies since 2011. A description of these archives and some use cases are given in Section 4.

The benefit of this collaborative platform for ethnomusicological research has been described in several publications (Simonnot, 2011; Julien Da Cruz Lima, 2011; Simonnot et al., 2014).



Figure 1: Screenshot excerpt of the Telemeta web interface

Recently, an open-source audio analysis framework, TimeSide, has been developed to bring automatic music analysis capabilities to the web platform and thus have turned Telemeta into a complete resource for *Computational Ethnomusicology* (Tzanetakis et al., 2007; Gómez et al., 2013).

2. THE TELEMETA PLATFORM

2.1 Web audio content management features and architecture

The primary purpose of the project is to provide the ethnomusicological communities with a scalable system to access, preserve and share audio research materials together with their associated metadata, as these data provide key information on the context and significance of the recording. Telemeta¹, as a free and open source², is a unique scalable web audio platform for backuping, indexing, transcoding, analyzing, sharing and visualizing any digital audio or video file in accordance with open web standards.

The time-based nature of such audio-visual materials and some associated metadata as annotations raises issues of access and visualization at a large scales. Easy and on demand access to these data, as you listen to the recording, represents a significant improvement.

An overview of the Telemeta's web interface is illustrated in Figure 1. Its flexible and streaming safe architecture

¹ <http://telemeta.org>

² Telemeta code is available under the CeCILL Free Software License Agreement

This work is partly supported by a grant from the french National Research Agency (ANR) with reference ANR-12-CORD-0022.

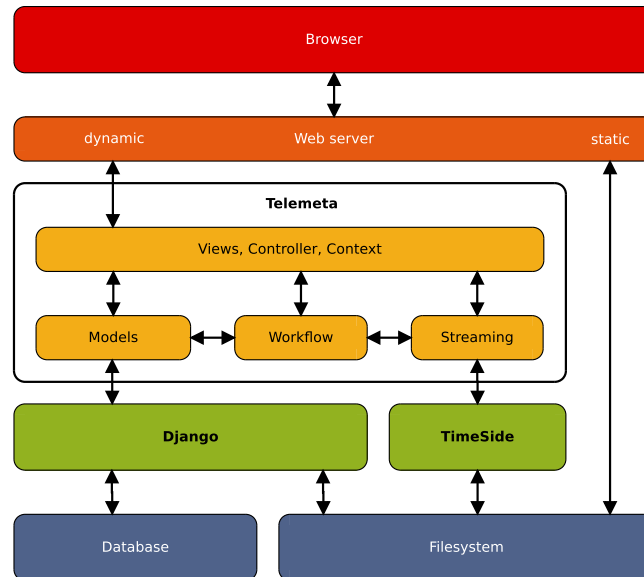


Figure 2: Telemeta architecture

ture is represented in Figure 2.

The main features of *Telemeta* are:

- Pure HTML5 web user interface including dynamical forms
- On the fly audio analyzing, transcoding and metadata embedding in various formats
- Social editing with semantic ontologies, smart workflows, realtime tools, human or automatic annotations and segmentations
- User management with individual desk, playlists, profiles and group access rights
- High level search engine (geolocation, instruments, ethnic groups, etc...)
- Data providers : DublinCore, OAI-PMH, RSS, XML, JSON and other
- Multi-language support (now english and french)

Beside database management, the audio support is mainly provided through an external component, TimeSide, which is described in Section 3.

2.2 Metadata

In addition to the audio data, an efficient and dynamic management of the associated metadata is also required. Consulting metadata provide both an exhaustive access to valuable information about the source of the data and to the related work of peer researchers. Dynamically handling metadata in a collaborative manner optimises the continuous process of knowledge gathering and enrichment of the materials in the database. One of the major challenge is thus the standardization of audio and metadata formats with the aim of long-term preservation and usage of the

different materials. The compatibility with other systems is facilitated by the integration of the metadata standards protocols *Dublin Core*³ and *OAI-PMH* (Open Archives Initiative Protocol for Metadata Harvesting)⁴.

Metadata provide two different kinds of information about the audio item: contextual information and annotations.

2.2.1 Contextual Information

In ethnomusicology, contextual information could be geographic, cultural and musical. It could also store archive related information and include related materials in any multimedia format.

2.2.2 Annotations and segmentation

Metadata also consist in temporally-indexed information such as a list of time-coded markers associated with annotations and a list of time-segments associated with labels. The ontology for those labels is relevant for ethnomusicology (e.g. speech versus singing voice segment, chorus, ...).

Ethnomusicological researchers and archivists can produce their own annotations and share them with colleagues. These annotations are accessible from the sound archive item web page and are indexed through the database.

It should be noted that annotations and segmentation can also be produce by some automatic signal processing analysis (see Section 3).

3. TIMESIDE, AN AUDIO ANALYSIS FRAMEWORK

One specificity of the Telemeta architecture is to rely on an external component, TimeSide⁵, that offers audio player

³ Dublin Core Metadata Initiative, <http://dublincore.org/>

⁴ <http://www.openarchives.org/pmh/>

⁵ <https://github.com/yomguy/TimeSide>

web integration together with audio signal processing analysis capabilities.

TimeSide is an audio analysis and visualization framework based on both python and javascript languages to provide state-of-the-art signal processing and machine learning algorithms together with web audio capabilities for display and streaming. Figure 3 illustrates the overall architecture of TimeSide together with the data flow between TimeSide and the Telemeta web-server.

3.1 Audio management

TimeSide provides the following main features:

- Secure archiving, editing and publishing of audio files over internet.
- Smart audio player with enhanced visualisation (waveform, spectrogram)
- Multi-format support: reads all available audio and video formats through Gstreamer, transcoding with smart streaming and caching methods
- "On the fly" audio analyzing, transcoding and meta-data embedding based on an easy plugin architecture

3.2 Audio features extraction

In order to provide Music Information Retrieval analysis methods to be implemented over a large corpus for ethnomusicological studies, TimeSide incorporates some state-of-the-art audio feature extraction libraries such as Aubio⁶ (Brossier, 2006), Yaafe⁷ (Mathieu et al., 2010) and Vamp plugins⁸.

As an open-source framework and given its architecture and the flexibility provided by Python, the implementation of any audio and music analysis algorithm can be considered. Thus, it makes it a very convenient framework for researchers in computational ethnomusicology to develop and evaluate their algorithms.

Given the extracted features, every sound item in a given collection can be automatically analyzed. The results of this analysis can be stored in a scientific file format like Numpy and HDF5 and serialized to the web browser through common markup languages: XML, JSON and YAML.

3.3 Automatic Analysis of ethnomusicological sound archives

Ongoing works lead by the DIADEMS project consist in implementing advanced classification, indexation, segmentation and similarity analysis methods dedicated to ethnomusicological sound archives.

Besides music analysis, such automatic tools also deal with speech and noises classification and segmentation to enable a full annotation of the audio materials.

In the context of this project, both researchers from Ethnomusicological, Speech and Music Information Retrieval

communities are working together to specify the tasks to be addressed by automatic analysis tools.

4. SOUND ARCHIVES OF THE CNRS - MUSÉE DE L'HOMME

Since June 2011, the Telemeta platform has been deployed to hold the *Sound archives of the CNRS - Musée de l'Homme*⁹ and is managed by the CREM (Center for Research in Ethnomusicology). The platform aims to make these archives available to researchers and to the extent possible, the public, in compliance with the intellectual and moral rights of musicians and collectors.

4.1 Archiving research materials

The archives of CREM, the most important in Europe, are distinguished by their wealth:

- Nearly 3,500 hours of recordings of unpublished field.
- Approximately 3700 hours of material published (more than 5000 discs, many of which are very rare).

The collection is sustained by the field missions of researchers on all continents.

Through this platform, archivists can properly ensure the long-term preservation of data and continuously maintain and enrich the associated metadata.

Accessing the collections aid laboratory research, diachronic and synchronic comparisons, the preparation of new fieldwork and the training of PhD students.

Publishing collections also helps researchers making their work more visible. Besides making available and listenable the related corpora, researchers can also append related academic publications and provide temporal annotations to further illustrate their work.

4.2 A collaborative platform

Given the collaborative nature of the platform, both research and archivist can cooperate with colleagues to continuously enrich metadata associated to a sound item or a collection.

Collaborative tools like markers and comments enable researchers from different institutions to work together on some common audio materials. It also allows researchers to exchange data online with communities producing their music in their home countries.

5. CONCLUSION

The Telemeta open-source framework provides the researchers in musicology with a new platform to efficiently distribute, share and work on their research materials. The platform has been deployed since 2011 to manage the *Sound archives of the CNRS - Musée de l'Homme* which is the most important European collection of ethnomusicological resources.

⁶ <http://aubio.org/>

⁷ <https://github.com/Yaafe/Yaafe>

⁸ <http://www.vamp-plugins.org>

⁹ <http://archives.crem-cnrs.fr>

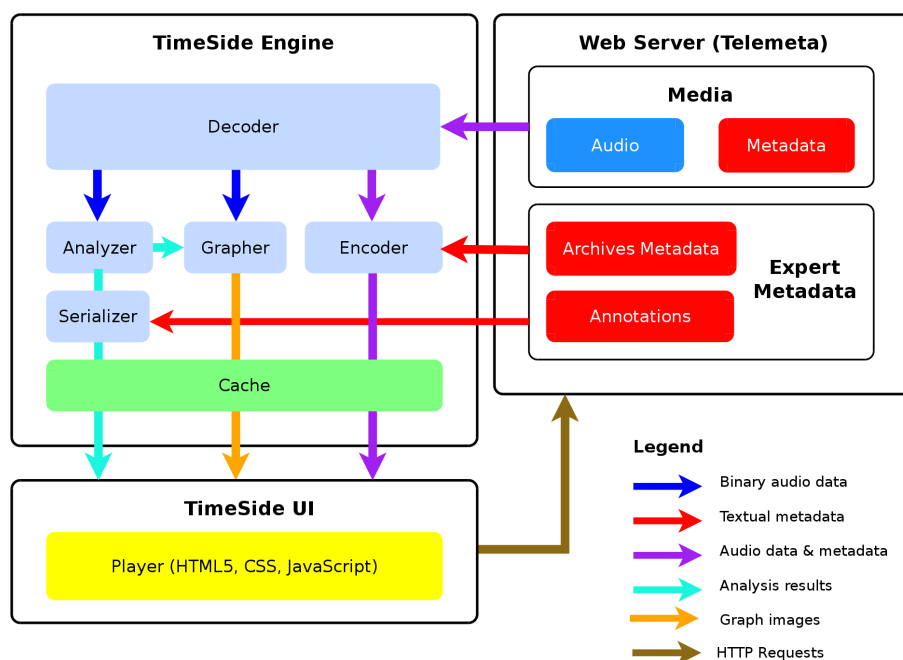


Figure 3: TimeSide engine architecture and data flow with Telemeta web-server

Furthermore, this platform is offered automatic music analysis capabilities through an external component, TimeSide that provides a flexible computational analysis engine together with web serialization and visualization capabilities. As an open-source framework TimeSide could be an appropriate platform for researchers in computational ethnomusicology to develop and evaluate their algorithms

Further works on the user interface will enhance the visualization experience with time and frequency zooming capabilities and will thus improve the accuracy and the quality of time-segment base annotations.

Acknowledgments

The authors would like to thank all the people that have been involved in Telemeta specification and development or have provide useful input and feedback. The project has been partially funded by the French National Centre for Scientific Research (CNRS), the French Ministry of Culture and Communication, the TGE Adonis Consortium, and the Centre of Research in Ethnomusicology (CREM).

6. REFERENCES

- Brossier, P. (2006). *Automatic annotation of musical audio for interactive systems*. PhD thesis, Centre for Digital music, Queen Mary University of London, UK.
- Gómez, E., Herrera, P., & Gómez-Martin, F. (2013). Computational ethnomusicology: perspectives and challenges. *Journal of New Music Research*, 42(2), 111–112.
- Júlien Da Cruz Lima, A. (2011). The CNRS — Musée de l’Homme audio archives: a short introduction. *International Association of Sound and Audiovisual Archives journal*, 36.
- Mathieu, B., Essid, S., Fillon, T., Prado, J., & Richard, G. (2010). Yaafe, an easy to use and efficient audio feature extraction software. In *Proc. of ISMIR 2010, Utrecht, Netherlands*, (pp. 441–446). International Society for Music Information Retrieval.
- Simonnot, J. (2011). TELEMETA: an audio content management system for the web. *International Association of Sound and Audiovisual Archives journal*, 36.
- Simonnot, J., Mifune, M.-F., & Lambert, J. (2014). Telemeta: Resources of an online archive of ethnomusicological recordings. Panel accepted at ICTM Study Group on Historical Sources of Traditional Music, Aveiro, Portugal, May 12–17 2014.
- Tzanetakis, G., Kapur, A., Schloss, W. A., & Wright, M. (2007). Computational ethnomusicology. *Journal of Interdisciplinary Music Studies*, 1(2), 1–24.