The unpublished records also cover a wide range of carriers now preserved by the Audiovisual Department of the French National Library (Bibliothèque nationale de France). These include our oldest records — 400 cylinders recorded by Léon Azoulay (from the Anthropological Society of Paris) at the Universal Exhibition in Paris in 1900 —, the lacquers discs recorded directly in the field during the 1940s and 1950s, and of course magnetic tapes, recorded from the 1950s to the 1990s.



**Figure 1**. Published series from the CREM collections © CREM – Le chant de monde



### *TELEMETA: an audio Content Management System for the Web*
*Joséphine Simonnot, research engineer, CNRS, Research Centre for Ethnomusicology*[12]

Researchers in the field of humanities disciplines such as anthropology and linguistic, work with a wide variety of documents including pictures, sound recordings and videos. The time-based nature of these audiovisual materials raises issues of access and visualization. As these resources are research materials, it is important to allow their access and management, as well as their preservation and distribution. For sound recordings, it is essential to manage the sounds together with their associated metadata, to enrich them and to facilitate access to them.

As there was no open source application available on the market, the CNRS Research Center for Ethnomusicology (CREM), the Laboratory of Musical Acoustics (LAM), and the Sound Archive of Aix-en-Provence (MMSH), have been working together since 2007 on the design of an innovative, collaborative and interdisciplinary tool. For these teams, whose core activity is to work on and manage sound records, the project had a double mission: to meet their specific management needs and also to satisfy the demands of the research sector.

The multimedia Web application TELEMATA has been on line since 2008. It is specifically designed to give access to audio archives and their associated metadata, to facilitate the work of researchers and to enhance the availability of data from the database catalog, according to common standards and interoperability.

**Edit view:**



The main features of TELEMETA are:

- User-focused web front-end supporting workflows,
- Advanced search methods with thesaurus,
- Dynamic audio player,
- Immediate audio analysis, transcoding and metadata embedding, based on easy plug-in architecture,
- Multi-format support: FLAC, OGG, MP3, WAV,
- XML serialized back-up, SQL back-end,
- Long-term preservation of audio files and metadata.

TELEMATA can be easily adapted to the specific needs of each archive and is compatible with other systems through the integration of standard protocols Dublin Core and OAI-PMH.

It also offers user-friendly functions to navigate inside the recordings such as:

- Visualization of the wave form resized in full screen,
- Navigation within audiovisual files (dynamic head player),
- Acoustic analysis tools (spectral view),
- "Markers" with annotation to identify special events in the audio files (segmentation).

**Spectral view with markers:**
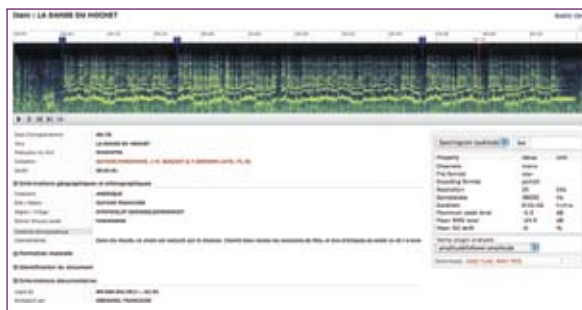


The online sharing of data and annotation allows all the different persons involved in a specific research project to collaborate and to optimize the enrichment of the metadata.

TELEMETA also allows the geolocation of records through the integration of the GéoEthno and GeoNames thesauri.



The application supports export and sharing of data sources, compressed (MP3 & OGG) or not (WAV, FLAC), and the management of access profiles.

The next step is to include analysis tools in order to improve the semantic search, including the detection of drop, tone, rhythm, speech-music segmentation, speech-to-text transcription, recognition of musical instruments family, etc.

The National Museum of Natural History (MNHN) is also involved in the project, as well as the French National Library (BNF) specifically with regards to the preservation of old media.

Prototypes are already online for the CREM, the MMSH, MuCEM and MNHN:

http://crem.telemeta.org
http://mmsh.telemeta.org
http://mnhn.telemeta.org
http://mucem.telemeta.org

Finally, TELEMETA is supported by a national infrastructure for digital humanities (TGE Adonis) and its new search engine, "ISIDORE", specializing in humanities sources, will harvest the data (http://www.rechercheisidore.fr).

For more details, see the wiki of the TELEMATA project: http://telemeta.org

The technologies involved are:

1. **Open Source**
   License CeCILL
   The sharing of resources to ensure long-term development (wiki)

2. **Software**
   Python, Django, TimeSide, MySQL
   Linux, OSX, Windows

3. **Formats and standards**
   Web: HTML, CSS, SQL
   Audio: WAV, MP3, OGG Vorbis, FLAC
   Metadata: Dublin Core and OAI-PMH