

GUÍA PRÁCTICA PARA

LA OPERACIÓN DE LA

PLATAFORMA TELEMETA

DE ARCHIVOS SONOROS DEL

CNRS-MUSÉE DE L'HOMME

<https://archives.crem-cnrs.fr/>



LESC • CREM
Centre de recherche
en ethnomusicologie
UMR 7186 • Université Paris Nanterre • CNRS

Huma-Num

Soutenu
par

**MINISTÈRE
DE LA CULTURE**
*Liberté
Égalité
Fraternité*

telemeta

parisson.com

FONOTECA NACIONAL

ESTAS ACTIVIDADES FORMAN PARTE DEL PROYECTO DGAPA UNAM PAPIIT IT400118

Documentación profesional de la plataforma Telemeta, Versión CREM

V.6

**Joséphine Simonnot, Ingeniera en investigación,
LESC CNRS, 12 de mayo 2020.**

**Traducción en español: Sibylle Hayem, Perla Olivia Rodríguez Reséndiz
y Dafne Alí Martínez, con el auspicio del proyecto DGAPA UNAM PAPIIT
IT400118 y de la Fonoteca Nacional de México.**

Índice

1. Interfaz (usuario y administrador)

- 1.1 Inicio
- 1.2 Escritorio
- 1.3 Archivos
- 1.4 Geonavegador
- 1.5 Búsqueda avanzada
- 1.6 Terminología
- 1.7 Ayuda
- 1.8 Admin

2. Organización de la base de datos

- 2.1 El modelo de datos: histórico
- 2.2 La base de datos
- 2.3 Niveles de catálogo
- 2.4 Plan de asignación: sistema de código
- 2.5 Descripción de los campos documentales
 - 2.5.1 Fondos
 - 2.5.2 Corpus
 - 2.5.3 Colecciones
 - 2.5.4 Ítems
- 2.6 Descripción de las listas de vocabulario
- 2.7 Funcionalidades de modificaciones e incorporación de elementos en la base de datos
 - 2.7.1 Importación de medios
 - 2.7.2 Incorporación y modificación de los metadatos

- 2.7.3 Regla para la captura de datos
- 2.7.4 La estructuración de las Colecciones
- 2.7.5 Modificación de las listas de vocabulario
- 2.7.6 Modificación masiva (automática)
- 2.8 Interoperabilidad
 - 2.8.1 Recolección vía el Dublin Core
 - 2.8.2 Mapeo en el formato CSV

3. Funcionalidades de búsqueda y selección

- 3.1 Búsqueda sencilla
- 3.2 Búsqueda avanzada
- 3.3 Visualización de los resultados (búsqueda sencilla y avanzada)
 - 3.3.1 Visualización por lista
 - 3.3.2 Filtro de resultados
- 3.4 Búsqueda en “Colecciones” e “Ítems” en el menú “Archivos”
- 3.5 Búsqueda por países
- 3.6 Búsqueda por listas / enumeraciones

4. Funcionalidades de navegación y visualización

- 4.1 En un Fondo o en un Corpus
- 4.2 En una Colección
- 4.3 En un *Ítem*

5. Funcionalidades de gestión de accesos a los datos

- 5.1 Gestión de la conexión
 - 5.1.1 Creación de la cuenta de usuario
 - 5.1.2 Conexión a la plataforma

5.2 Tipos de acceso a las Colecciones e Ítems (derechos de uso)

5.3 Perfiles de usuarios

5.4 Funciones de acceso a los medios (sonido, video e imágenes)
y metadatos

6. Funcionalidades de gestión de medios/media (sonido y video)

6.1 Datos producidos por *Time Side*, transcodificación (análisis de audio)

6.2 Lector de *streaming*

6.3 Representación gráfica de los medios

6.3.1 Con una cuenta de usuario

6.3.2 Sin cuenta de usuario

6.3.3 Los algoritmos del ANR DIAMENS

6.4 Anotación temporal

6.5 Ilustración de registros con una imagen o una liga externa

6.6 Reutilización de un medio

6.6.1 Enlace *iframe*.

6.6.2 Libro electrónico (*export EPUB*)

7. Varios

8. Conclusión

Bibliografía

Introducción

Telemeta es un sistema de gestión para contenidos audiovisuales en línea, concebido y desarrollado desde 2007. En este documento se expone su funcionamiento y se expresan los objetivos formulados por nuestra antigua directora, Rosalía Martínez. En ese entonces era urgente salvar los metadatos de la antigua base 4D, amenazada por la obsolescencia.¹ Además, se requería facilitar el acceso y el trabajo de los investigadores. Este patrimonio cultural audiovisual, mundial, frágil, inaccesible en su totalidad al público, cuyo destino era quedarse confinado por décadas, debía ser compartido ampliamente a través del internet.

Contexto 2

En el mercado se carecía de plataformas de gestión documental que no fueran propietarias, así, el objetivo de este proyecto fue proponer una plataforma *open source* adaptada a las necesidades profesionales de los archivos sonoros, ofreciendo las posibilidades de gestión, anotación y uso compartido en línea, de archivos sonoros y audiovisuales (video).³ Actualmente el CNRS (Centro Nacional de Investigación Científica) propone soluciones de almacenamiento y acceso a datos digitales, pero carece de un sistema en línea de gestión de bases de datos audiovisuales.

La plataforma Telemeta es colaborativa y está orientada al público en general, aún cuando no tenga conocimientos especiales en informática, con el propósito de implicar directamente a los productores de datos y de conocimientos (investigadores, depositantes diversos, estudiantes, etc) en el enriquecimiento de los metadatos, con el fin de generar plusvalía científica para determinado equipo de investigación.

El desarrollo del *software* Telemeta empezó en 2007, luego de la participación de Guillaume Pellerin en el LAM, quien recibió dos encomiendas principales:⁴

- Migrar la base de datos 4D que contiene miles de registros documentales capturados (gracias a la intervención de la Sociedad Samalyse) hacia un sistema *open source*, libre y compatible con los estándares del internet.
- Entregar herramientas de análisis y transcodificación en tiempo real que permitan a los profesionales valorar los datos.

1 La base 4D funcionaba únicamente con Mac OS9, posteriormente su actualización no se pudo llevar a cabo.

2 Ver artículo del boletín de l'AFAS sobre el contexto histórico: <https://journals.openedition.org/afas/4056>

3 Portal tipo Alexandria

4 Equipo Lutería, Acústica, Música en el Instituto d'Alembert (CNRS, Universidad de la Sorbona): <http://www.lam.jussieu.fr/>

En efecto, el análisis de la base de datos realizada por la Sociedad Samalyse en aquellos tiempos fue el siguiente:

Parece que, en este momento, el fondo documental de esta base de datos no se está aprovechando. Esta base de datos es poco utilizada, inclusive por los investigadores, pues algunos ni la usan. La razón principal de esta situación se debe a que el entorno técnico y la estructura de la base fueron concebidos para inscribirse en un medio documental tradicional, donde en general se contacta a un documentalista para una investigación, quien posee los conocimientos técnicos necesarios para atender las consultas. Pero en la era del internet este proceso de investigación es demasiado pesado e inadecuado, en una palabra: obsoleto⁵

Las funciones desarrolladas en esta plataforma posibilitaron el acceso, la interoperabilidad y la disponibilidad a los archivos sonoros del CNRS-Musée de l'Homme.⁶ Desde 2011, la plataforma está disponible en línea, mucho antes que se promovieran los principios del FAIR.⁷

Lenguajes y formatos

La aplicación web se apoya en bloques y depende al 100% del *open source*. Además de su modelo de datos específico dedicado a la etnomusicología, su peculiaridad reside en la gestión automática de las transcodificaciones (cambio de formatos) y en el análisis de los audios de la plataforma. La base de datos relacional, MySQL, gestiona los metadatos (o documentación descriptiva).

La totalidad de la documentación informática y el programa completo son accesibles en GitHub:

<https://github.com/Parisson/Telemeta>

<https://github.com/Parisson/TimeSide/>

Features:

<https://github.com/Parisson/Telemeta#overview>

Changelog:

<https://github.com/Parisson/Telemeta#news>

Documentación técnica de instalación y de configuración (dependencias del *software*):

<https://github.com/Parisson/Telemeta-doc>

5 Cf. O. Guiyardi, Docref-1.1-stable, Mycore

6 Dicha institución, a partir de este momento y a lo largo del texto, será mencionado por su traducción en español: Museo del Hombre (N. del Ed.).

7 FAIR: Fácil de encontrar, Accesible, Interoperable y Reutilizable.

Observación: gracias a esta documentación, el CMAM en Túnez y el Instituto de Investigaciones Bibliotecológicas y de la Información de la Universidad Nacional Autónoma de México instalaron Telemeta en su servidor para su propio uso.⁸ Los archivos sonoros de acústica musical de Francia también utilizan Telemeta, al igual que la Universidad de Poitiers, la cual se apoya en la tecnología TimeSide.

La aplicación se basa en el *framework* Django del tipo MVC (Modelo, Vista, Controlador) escrito en lenguaje Python. Esta elección permite que evolucione el modelo de datos y que se incluyan métodos de análisis provenientes de laboratorios de investigación, tanto franceses como extranjeros.⁹ El *frontend* (interfaz web) está ubicado de manera clásica en HTML5, CSS y JavaScript. Se puede ejecutar en cualquier servidor gracias a Docker, que permite empaquetar la aplicación y sus dependencias en un contenedor.

Composición Docker (módulo de instalación multi-sistema):

<https://github.com/Parisson/Telemeta/blob/master/docker-compose.yml>

Lista de las dependencias *backend*:

<https://github.com/Parisson/Telemeta/blob/master/setup.py#L51>

Lista de las dependencias *frontend*:

<https://github.com/Parisson/Telemeta/blob/master/app/settings.py#L349>

Entorno técnico

El proyecto dispone de una máquina virtual (VM) de producción, basada en OpenStack y desplegada sobre la infraestructura del CNRS Huma-Num8.¹⁰ Sistema de explotación (OS):

Debian Stretch (V.10).

Servidor web / *load balancing*: Nginx

Gestión de los contenedores: Docker

Memoria RAM: 80 Go

Almacenamiento de datos del CREM (excepto cálculo para la NYU): 5To memoria asignada a la VM de desarrollo principio del 2020: 160 G

8 Telemeta del CMAM, Túnez: <http://phonotheque.cmam.tn>.

9 *Framework*: infraestructura de software concebida para modelar la arquitectura de los softwares aplicativos.

10 Muy Grande Estructura de Investigación del INSHS del CNRS: www.huma-num.fr.

Formatos de los datos audio / video

Los formatos utilizados son los más generalizados y aceptados para el archivo a largo plazo (a excepción del MP3):

- Archivos no comprimidos de audio para su conservación: WAV (48kHz/24bit; 44.1kHz/16bit), FLAC.
- Archivos comprimidos audio: MP3 y OGG Vorbis.
- Archivos comprimidos video: MP4 y WebM.

Las opciones de formatos se detallan en el capítulo 6.1. La compresión de los archivos fuentes audio (WAV) y el cálculo para la visualización gráfica (forma onda espectral predeterminada) es automática.

La interoperabilidad con otras plataformas, como Isidore, está asegurada por la compatibilidad con la norma Dublin Core y el protocolo OAI-PMH (*Open Archive Initiative-Protocol for Metadata Harvesting*).¹¹ La compatibilidad de la base de datos documental con la norma de descripción Dublin Core se efectúa para los campos principales en la versión del CREM (ver capítulo “Interoperabilidad”).

Observación: dado que existen diferentes aproximaciones comerciales, las instrucciones de uso se repiten para que cada apartado esté completo.

Agradecimientos

La idea de esta plataforma nació gracias a Hugues Genevois (LAM 10) y su concepción es el resultado de una colaboración entre el CREM y varios socios:¹² la empresa emergente Parisson, el IRCAM (con G. Pellerin y T. Fillon), la MMSH (estructuración de los metadatos) y los laboratorios de investigación en tratamiento informático de la señal (IRIT, LIMSI, LaBRI, Universidad Queen Mary de Londres, Universidad de New York).¹³

También agradecemos a nuestra antigua directora, Rosalía Martínez, por haber sostenido este proyecto a pesar de los numerosos desafíos y de la ausencia de financiamientos, así como a los participantes activos, Aude Da Cruz Lima y Sara Tandar, así como a los investigadores que nos tuvieron confianza. Igualmente hay que hacer énfasis en el apoyo del TGE Adonis (Stéphane Pouyllau y Richard Walter) y de Caroline Bodolec para el colectivo Cristal del CNRS. Las contribuciones de Nedra Mellouli-Nauwynck (París 8) y de los pasantes del IUT Montreuil que permitieron mejorar el motor de investigación

11 Isidore: <https://isidore.science/>

12 Equipo de Luteria,, Acustica, Musica del Instituto d'Alembert (CNRS Paris-Sorbonne): <http://telemeta.lam.jussieu.fr>, Telemeta LAM : <http://telemeta.lam.jussieu.fr>

13 Ver. ANR y Labex cap. VI.

(módulos Haystack y ElasticSearch):¹⁴ Killian Mary, Angy Fils-Aimé, Raouf Benmansour, Mathieu Boyé, Nicolas Neto, Thomas Balastegui. Un vasto número de generosos participantes contribuyeron en la traducción de la interfaz en varios idiomas (Anas Ghrab al árabe, Susanne Fürniss al alemán, Carlos Sandroni al portugués, Zaichao Xia al chino).

Pero nada se hubiera logrado sin el entusiasmo y las competencias de Guillaume Pellerin, el principal arquitecto de este proyecto, que pone las nuevas tecnologías al servicio del patrimonio cultural inmaterial. También rendimos homenaje a los fundadores de esos archivos: André Schaeffner y Gilbert Rouget, cuya voluntad fue transmitir este patrimonio a las generaciones futuras. Esperamos que sus deseos se hagan realidad.

14 Seguimiento de los desarrollos y orientación benévola de los pasantes en informática.

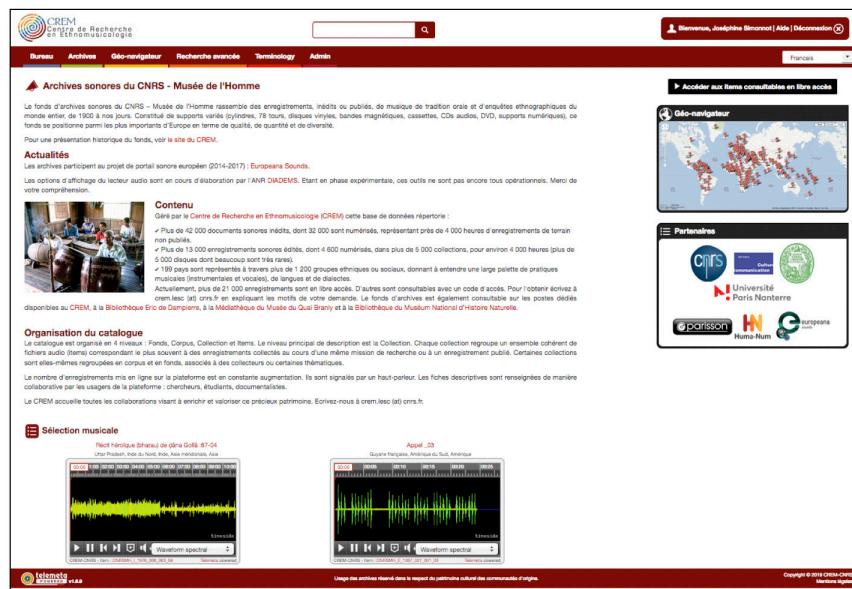
1. INTERFAZ (USUARIO Y ADMINISTRADOR)

1. Interfaz (usuario y administrador)

Este capítulo describe de manera general la página del inicio de la sesión. Cada elemento se detalla en los diferentes capítulos.

1.1 Inicio

La página de inicio se presenta, de arriba hacia abajo y de izquierda a derecha, con los elementos siguientes:



- Logo del equipo.
- Motor de búsqueda sencilla (botón “lupa”).
- Botón conexión / ayuda: ventana de *login*, visualización del nombre del usuario.
- **Barra horizontal con 6 menús: Escritorio, Archivos, Geo-navegador, Búsqueda avanzada, Menú Terminología, Administrador.**
- Menú “Idioma” de la interfaz (7 idiomas para escoger: francés, inglés, alemán, chino, árabe, español, portugués). Selección automática del idioma de acuerdo con el navegador del usuario.

1. INTERFAZ (USUARIO Y ADMINISTRADOR)

- Un texto introductorio con una foto (modificable en GitHub).¹⁵ El texto se traduce en inglés, alemán y portugués.
- Un botón “Acceder a los Ítems en acceso abierto”.
- Una ventana “Geonavegador” con un clic para la geolocalización por país.
- Una ventana con los logos de los socios.
- Una selección musical con 2 reproductores incrustados con el título del ítem, lugar y liga hipertexto, hacia la fuente (selección aleatoria para invitar al usuario a consultar los archivos).
- *Banner* con el enlace hacia el wiki telemeta.org, mención para las comunidades originarias y menciones legales.

1.2 Escritorio

La página “Escritorio” es accesible con una cuenta usuario. Éste es un espacio personal (privado).

The screenshot displays the Telemeta desktop application interface. At the top, a navigation bar includes links for Bureau, Archives, Géo-navigateur, Recherche avancée, Terminology, Admin, and a language switcher set to Français. Below the navigation bar are four main panels:

- Mes listes de lecture**: A sidebar listing various collections and exports, such as "Collections Inédites", "Digitized collections (Éditées)", and "MP3 export SET1, SET2, SET3".
- Mes recherches**: A search history panel showing queries like "loncke", "Exposition Coloniale de Paris, 1931", and "sardigne 1979".
- Mes dernières modifications**: A list of recent item modifications by user nyuad, including entries for "Chant masculin, flûte, clarinette et calebasse".
- Toutes les dernières modifications**: A broader list of all recent item modifications, including entries for "Suite d'airs damâlî" and "Séance guilli-damâlî: a".

At the bottom of the interface, there is a table titled "Moussali, Published records (Éditées)" showing digitization proposals. The columns include Titre, Type, Cote, Collecteur, Période d'enregistrement, Numérisé, and Action. Several rows are listed, such as "Les grands maltres de la chanson tunisienne Vol. II : Cheikh Hamada" and "Musique traditionnelle et religieuse des Comores".

1. INTERFAZ (USUARIO Y ADMINISTRADOR)

Listas de lecturas

El usuario puede crear una o varias listas personales de lectura a través de la selección de elementos de la base de datos, gracias al botón “Aregar”, que incorpora un título y una descripción. En las listas de lectura se visualizan las listas de las Colecciones, de Ítems o de marcadores temporales seleccionados por el usuario en la base de datos con el botón “Listar” o al marcar en una lista los resultados de búsqueda con “Aregar a la lista”. Cada lista puede ser minimizada o desplegada para aligerar la visualización.

Cada lista cuenta con los siguientes botones:

- “Editar”: para modificar un título y una descripción de la lista.
- “CSV Colecciones”: Exportación de los metadatos del registro “Colección” en el formato CSV (tabulador).
- “CSV Ítems”: Exportación de los metadatos del registro Ítem en el formato CSV (tabulador).
- “Eliminar”: suprimir las listas.

Cada lista muestra la siguiente información por columnas:

- Título.
- Tipo (colección / Ítem).
- Persona que deposita.
- Fecha (para los Ítems).
- Digitalizado (para los Ítems).
- Eliminación (papelera): elimina la selección únicamente de la lista.

Existe la posibilidad de ordenar cada columna por orden alfabético, ascendente / descendente.

El escritorio muestra también las funciones siguientes:

- Botón “Aregar”: añadir una *lista de reproducción* con un título y una descripción.
- “Mis búsquedas”: muestra las últimas búsquedas del usuario junto con la fecha.
- “Mis últimas modificaciones”: muestra las últimas 25 modificaciones del usuario con el título, el tipo (Colección / Ítem / Marcador), la fecha y el flujo RSS.
- “Todas las últimas modificaciones”: muestra las últimas 100 modificaciones hechas por todos los usuarios con el título, el tipo (Colección / Ítem / Marcador), la fecha y el flujo RSS.

1. INTERFAZ (USUARIO Y ADMINISTRADOR)

1.3 Archivos

El menú “Archivos” permite seleccionar un nivel del catálogo. El catálogo tiene 4 niveles: Fondo, Corpus, Colección e Ítem (ver el capítulo 2). El Ítem corresponde a un documento sonoro o audiovisual y a su documentación descriptiva (metadatos) y debe ser parte de una Colección. Una Colección es el conjunto de grabaciones y puede eventualmente ser parte de uno o varios Corpus, mientras que un Corpus puede eventualmente ser parte de uno o varios Fondos.

En cada nivel del catálogo, la página indica una lista de elementos con:

- La visualización del número de páginas de los resultados, flechas de salto de página.
- El título.
- La presencia de la media digitalizada (audio o video).
- El acopiador.
- Año de la grabación.
- El código.
- El botón “Aregar” permite añadir una referencia en el nivel seleccionado.
- El botón “Dublin Core” permite la visualización de los metadatos en este formato.
- El botón “Eliminar” permite suprimir el registro.

Nota: El botón para “Eliminar” únicamente es accesible desde las cuentas de Admin / Documentalistas. En el registro “Colección” este botón elimina tanto a la colección como a sus Ítems; no hay que confundir este proceso con los recuadros de selección de los Ítems de las listas de reproducción, pues no sirven para eliminar los Ítems.

En la navegación en los niveles Colecciones e Ítems se encuentra:

- El botón “Filtros”, arriba a la derecha, el cual se usa para seleccionar todo, sean materiales inéditos, editados, o bien, las referencias con el medio digitalizado dentro de las Colecciones o de los Ítems.

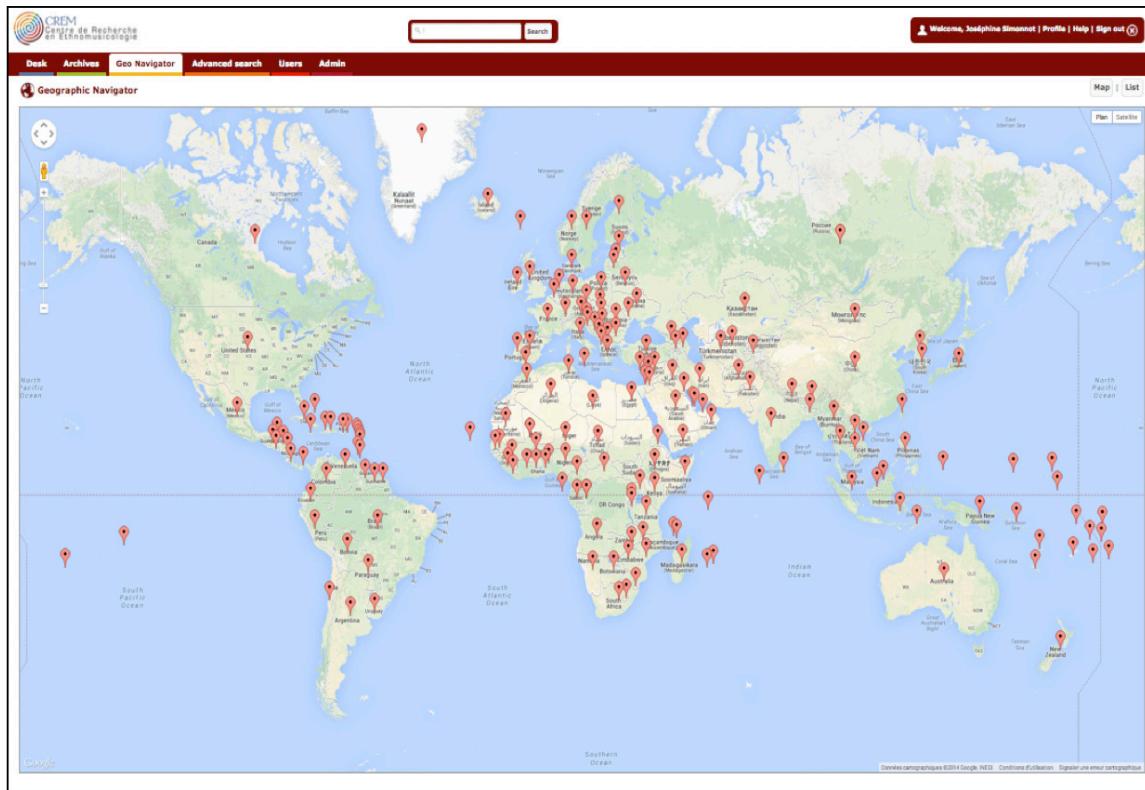
Para el nivel Ítem únicamente:

- El menú de selección del número de resultados por página: 20, 50, 100, 250, 500, 1000 y 10000.
- La columna de selección: toda la página, o bien, Ítem por Ítem (a seleccionar) para agregar a una *lista de reproducción* personal previamente elaborada (“Escritorio”) con el botón de abajo “Aregar a la lista”, el cual abre un menú desplegable para escoger entre las listas.

1. INTERFAZ (USUARIO Y ADMINISTRADOR)

1.4 Geonavegador

El “Geonavegador” permite visualizar las Colecciones por países en un mapamundi (opción predeterminada), o bien, en las listas de los países por continente, gracias al menú “Listar”, arriba a la derecha (ver capítulo 3.5).



1.5 Búsqueda avanzada

Existe un motor de búsqueda avanzado que permite afinar las búsquedas de acuerdo con 13 criterios. Este motor de búsqueda fue desarrollado entre 2015 y 2016 bajo las necesidades prioritarias definidas por los archivistas del CREM (ver capítulo 3.2).

1.6 Menú Terminologías

Este menú propone la lista de los instrumentos (términos genéricos y términos vernáculos) accesibles a todos los usuarios. Las demás enumeraciones (listas de vocabulario) son accesibles a través de las cuentas “súper usuarios”, para dar acceso a los ítems indexados según los términos de una lista (ver capítulo 3.6).

1. INTERFAZ (USUARIO Y ADMINISTRADOR)

1.7. Ayuda

Esta página indica a quién pueden contactar los usuarios en caso de problemas o preguntas. La página se puede modificar vía GitHub, versión de prueba (ver capítulo 7).

1.8. Admin

Esta página está reservada para los administradores de la base de datos. Su papel principal es gestionar y crear las cuentas de los usuarios, atribuir los perfiles y otorgar los permisos “a la carta” (ver capítulo 5). También permite que los administradores accedan a toda la base de datos. Con el motor de búsqueda sencillo se puede seleccionar el Ítem, la Colección, el Corpus o el Fondo para modificarlo y guardar el cambio. También se puede localizar a un usuario y consultar los datos de su cuenta.

2. ORGANIZACIÓN DE LA BASE DE DATOS

2. Organización de la base de datos

2.1. El modelo de datos: histórico

El modelo de datos es el resultado de reflexiones de etnomusicólogos que empezaron, desde 1930, a describir adecuadamente los archivos de música tradicional del mundo. La información actual de la base de datos se rige por ese modelo. Los descriptores, o campos, están detallados en el capítulo “Descripción de los campos documentales”. El modelo contiene numerosas listas, o tablas, de vocabularios específicos para la materia descrita más adelante. Antes de la migración, los metadatos se estructuran en 3 niveles heredados del sistema de gestión de las Colecciones que aplica el Museo del Hombre: Colección, Soporte físico e Ítems. El modelo se enriqueció en 2011 con nuevos campos que serán descritos más adelante.

Nota: La migración de este modelo, a uno más sencillo, podría facilitar la gestión del conjunto de los metadatos, pero implica la pérdida de información y una baja en el nivel de granularidad de la descripción científica o en la contextualización. La decisión de suprimir algunos campos tiene que tomarse de acuerdo con la utilidad efectiva de los metadatos.

2.2 La base de datos

La base de datos actual es una herencia de la base de datos de Pribislav Pitoëff bajo 4D, un formato propietario.¹⁶ El esquema simplificado del UML (*Unified Modeling Language*) de la antigua base, describe el conjunto de las relaciones internas recomendadas para la base de datos actual.

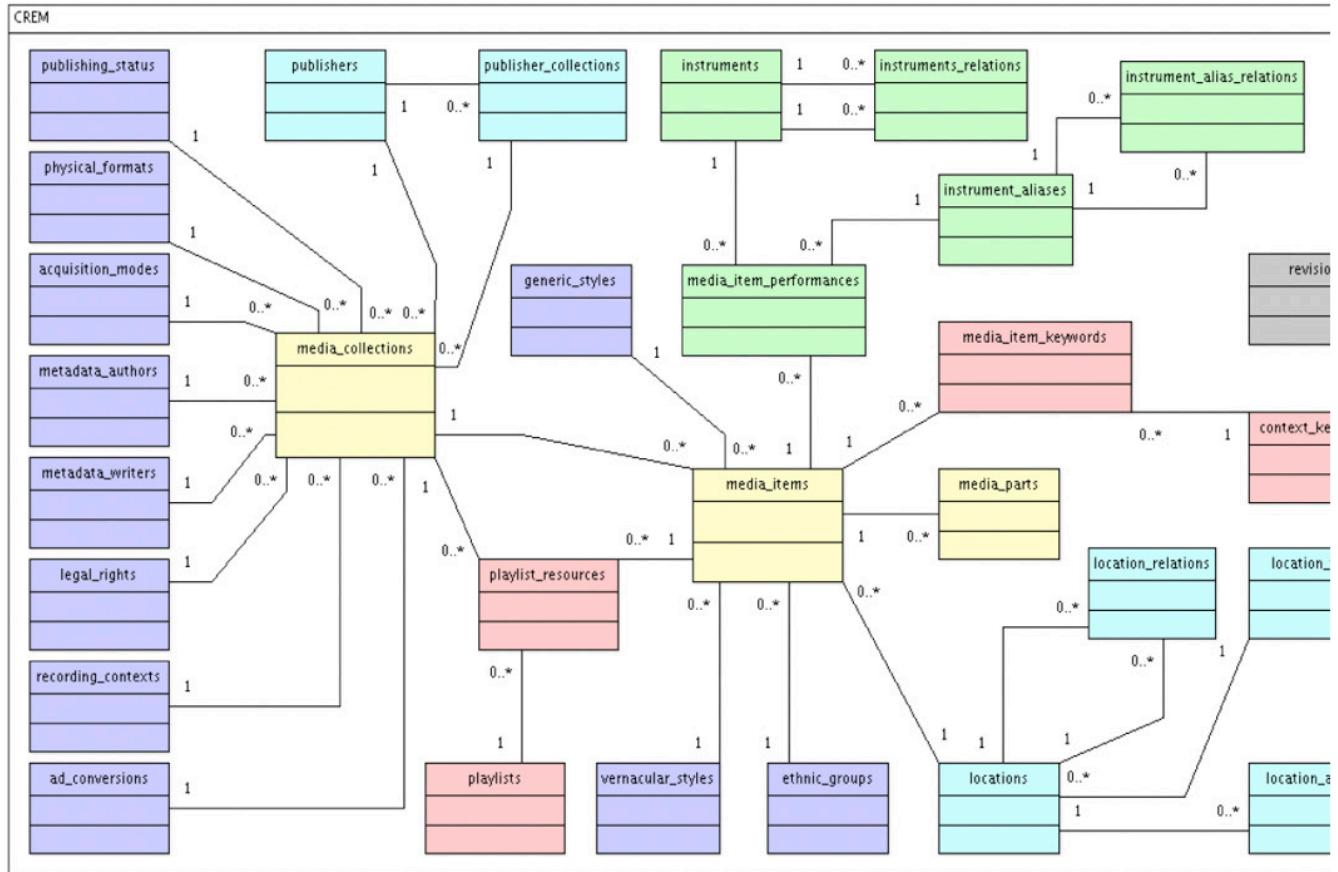
Para más información, se puede consultar el documento referencia del 2009, con las recomendaciones de reorganización de la base de datos del CRE,M producido por Olivier Guilyardi, Société Samalyse.

Este documento junto con los esquemas técnicos puede consultarse en:<https://mycore.core-cloud.net/index.php/s/48dedcecb574f9cf6c38f66f365b26f2>

16

Ingeniero CNRS responsable de los archivos sonoros del Laboratorio de Etnomusicología hasta 2008.

2. ORGANIZACIÓN DE LA BASE DE DATOS

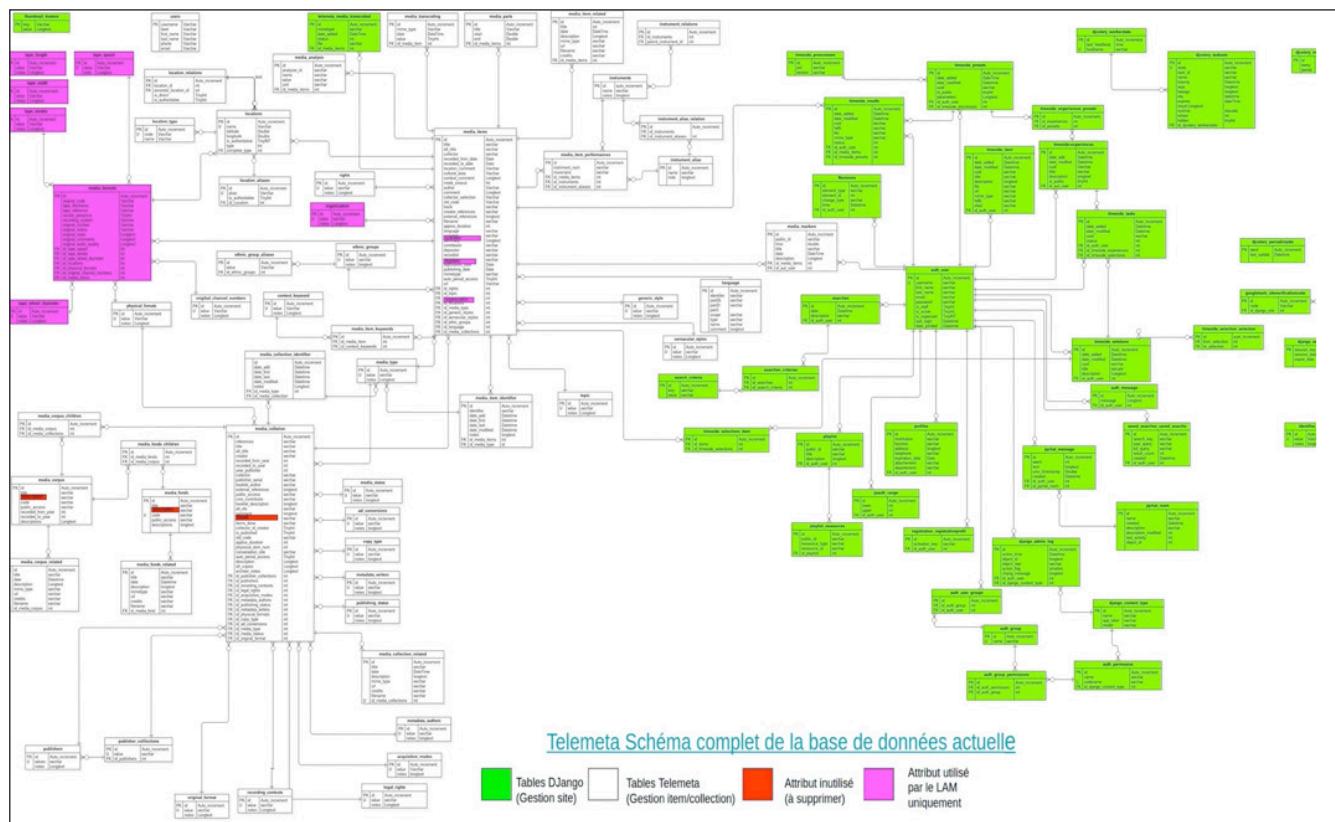


Esta migración ofreció la oportunidad para llevar a cabo una reorganización estructural y semántica profunda. La operación, sumamente compleja, fue el objeto de una especificación por separado (recomendaciones y “motor” de exportación de la base de datos desarrollado entre 2009 y 2011, con el fin de recuperar los registros Ítems y Colecciones capturadas en 4D).

Los principales metadatos, anteriores a 2011, proceden de una migración hecha a partir de la antigua base de datos hacia una estructura SQL. Esa migración fue la oportunidad para poner en marcha una reorganización estructural y semántica profunda de la base de datos, a fin de evitar redundancias de informaciones. El motor de exportación, desarrollado entre 2009 y 2011, será objeto de una especificación aparte. Entre 2011 y 2015 se agregaron nuevas especificaciones.

2. ORGANIZACIÓN DE LA BASE DE DATOS

El esquema UML actual (consultable en MyCore: <https://mycore.corecloud.net/index.php/s/48dedcecb574f9cf6c38f66f365b26f2>), completado por Nicolas Neto en 2017 como pasante en informática en el CREM, es el siguiente:



La plataforma respeta los estándares de respaldo y transmisión de los datos:

- Lenguaje de las páginas web.
 - **HTML**: lenguaje de hipertexto para representar las páginas web.
 - **CSS**: estilo de las páginas web.
 - **XML**: formato para el respaldo en serie de la base de datos e intercambio de metadatos.
 - Base de datos.
 - **SQL**: lenguaje de la base de datos relacional.

2. ORGANIZACIÓN DE LA BASE DE DATOS

- Descripción de los metadatos y de las ontologías.
- **Dublin Core**: formato de datos que permite el acopio del conjunto de la base de datos gracias a entidades externas (Isidore).
- **OAI-PMH**: protocolo de transferencia de los metadatos integrado a Dublin Core (acopio).
- **OWL**: (Web Ontology Language): estructuras semánticas de datos (tesauros y antologías).

2.3 Niveles del catálogo

Desde 2011, el catálogo está estructurado en 5 niveles para una mejor contextualización de los datos en conjuntos temáticos.

- Fondos: el primer nivel jerárquico de los metadatos incluye datos textuales y un conjunto coherente de Corpus.
- Corpus: el segundo nivel jerárquico de metadatos que incluye datos textuales y un conjunto coherente de Colecciones.
- Colección: el tercer nivel jerárquico de metadatos que incluye datos textuales y un conjunto coherente de Ítems.
- Ítems: el cuarto nivel jerárquico de metadatos que incluye datos textuales y la media. Esta media puede contener partes (Marcadores).
- Parte (marcador): el quinto nivel jerárquico que incluye los mismos metadatos del Ítem asociado y el intervalo de tiempo (secuencia).

Los 4 niveles principales del catálogo son visualizados en el menú “Archivos” en la página de inicio: Fondos, Corpus, Colección e Ítem. La organización del catálogo está descrita en la página de inicio de los usuarios.

La Colección es un conjunto básico heredado del Museo del Hombre. Agrupa un conjunto coherente de registros audio y video (Ítems) de una misma misión, de un mismo disco o de una misma temática, recopilado por una misma entidad moral o física. A su vez, algunas Colecciones están agrupadas según ciertas temáticas étnicas o acopiadores emblemáticos. Los niveles Fondos, Corpus son opcionales pero el nivel Colección es necesario para agregar datos en la plataforma.

El Ítem corresponde a un documento audiovisual y a su documentación descriptiva (metadatos) y debe ser parte de una colección. Eventualmente, una colección es parte de uno o varios Corpus y un Corpus a su vez puede ser parte de uno o varios Fondos. Un Ítem, forzosamente, es parte de una

2. ORGANIZACIÓN DE LA BASE DE DATOS

colección y su código de identificación es resultado de la clave de su colección y tiene que ser única. El ítem se puede duplicar con una nueva clave en su colección o en una colección diferente (escogiendo otra colección en el campo “Código de la colección”).

2.4 Plan de asignación de los archivos (sistema de identificación)

Los códigos de identificación tuvieron que ser modificados en el momento de la migración de la base 4D según un nuevo sistema de *Under score* (“_”) para sustituir los caracteres prohibidos (« . » ; « / » « : »). En el fondo, copropiedad del CNRS y del Museo del Hombre, CNRSMH fue elegido como encabezado para cualquier referencia del catálogo.

Los antiguos códigos de la migración del 2011 se conservaron en el campo “código antiguo” (*c code*). El antiguo sistema era demasiado heterogéneo para poder realizar una conversión automática. Poco a poco los antiguos códigos fueron reemplazados tomando en cuenta que cada modificación de un registro antiguo exige la atribución de un nuevo código para su ingreso. Hoy en día todavía existen unos 10.000 registros ítem con el código anterior.

Los registros de campo inéditos y los discos publicados comercialmente resaltan con los códigos:

CNRSMH_I_: registros de campo, antiguamente, BM.

CNRSMH_E_: archivos publicados comercialmente (editados) antiguamente DI.

Los códigos tienen que respetar la clave alfanumérica de procedencia de la colección. Ejemplo con el código del ítem: CNRSMH_I_9999_999_999_99.

- La primera serie de números: año de depósito de la colección.
- La segunda serie de números: no. de colección.
- La tercera serie de números: no. de soporte.
- L+La cuarta serie de números: no. ítem.

A veces los discos editados son agrupados en cajas o series, entonces el código de la colección tiene 3 números demás de acuerdo con el antiguo sistema heredado del Museo del Hombre.

Ejemplo de un código de identificación de un disco: CNRSMH_E_1963_007_001.

El motor de búsqueda avanzada permite la selección del tipo de documento buscado (Inédito / Editado) de acuerdo con el estatuto de la colección indicado en el código (CNRSMH_I_ o CNRSMH_E_).

2. ORGANIZACIÓN DE LA BASE DE DATOS

2.5 Descripción de los campos documentales

Algunos campos se forman por una cadena de caracteres de áreas de textos extensos o listas cerradas. Cada nivel del catálogo cuenta con un formulario de captura al que se accede con el botón “Editar”, siempre y cuando se posea una cuenta de usuario. En cada nivel del catálogo, aparece el botón “Dublin Core”, abajo y a la izquierda de la página, éste permite visualizar en una ventana nueva la información del registro en el formato metadatos Dublin Core sencillo y extendido.

Los niveles *Fondos* y *Corpus* tienen un número de campos limitados.

2.5.1 Fondos

- Título.
- Descripción.
- Código de identificación.
- Duración total disponible alfanumérica (calculada por el sistema: HH:MM:SS).
- Última modificación del registro: fecha, hora y usuario (información automática).
- Media asociada desplegable.
- Lista de Corpus del Fondo.

The screenshot displays a detailed record for the 'Fonds Dana Rappoport'. The top navigation bar includes 'Accueil', 'Archives', 'Géo-Navigateur', 'Recherche avancée', 'Aide | Connexion', and a search bar. The main content area is titled 'Fonds : Fonds Dana Rappoport'. It contains sections for 'Titre', 'Description', 'Cote', 'Durée totale disponible', 'Dernière modification de la fiche', and 'Média associés'. Below this is a 'Méta' section with fields for 'Titre' (Biographie de D. Rappoport), 'Description' (Elements de biographie: Dana Rappoport est ethnomusicologue au Centre national de la Recherche Scientifique. Ses recherches portent sur les musiques austroasiatiques de l'archipel indonésien, étudiées à la fois sous l'angle de la musicologie formelle, de l'anthropologie de la religion et de l'organisation sociale.), and 'Crédits' (CNRS). Further down are sections for 'Prévisualisation' (with a thumbnail image of Dana Rappoport), 'Fiche de présentation des travaux de Dana Rappoport' (listing publications from 1995 to 2011), 'Crédits' (CNRS), and 'Corpus' (listing various collections: Sulawesi-Selaiyer-Togian, Flores-Solor-Adonara-Lembata, Sumbawa, and Kaimana). At the bottom, there are buttons for 'Dublin Core', 'Telemetg v1.6.2', and copyright information.

2. ORGANIZACIÓN DE LA BASE DE DATOS

2.5.2 Corpus

- Título.
- Descripción.
- Código de identificación-
- Año de grabación (desde): Formato: AAAA.
- Año de grabación (hasta): formato: AAAA.
- Duración total disponible (calculada): alfanumérica, HH:MM:SS.
- Última modificación del registro: fecha, hora y usuario (información automática).
- Media asociada desplegable.
- Fondo asociado.
- Lista de las Colecciones del Corpus.

2.5.3 Colecciones

El registro “Colección” contiene toda la información válida para todos los Ítems de esta Colección. Algunos campos agregan automáticamente la información precedentes de los metadatos de los Ítems, como los lugares y el nombre de las poblaciones. Esos campos, calculados o visualizados automáticamente por la aplicación son los siguientes:

- Estados / naciones: concatenación de los países Ítems.
- Población / grupos sociales: concatenación de las poblaciones de los Ítems.
- Última modificación del registro.
- Duración calculada.
- Volumen de la colección.
- Cantidad de Ítems.

El registro “Colección” está organizado por bloques temáticos para agilizar su lectura.

2. ORGANIZACIÓN DE LA BASE DE DATOS

Bloque central:

Campos	Descripciones	Observaciones
Número de Ítems: lista	Número de Ítems en la colección. Liga dinámica hacia la lista de Ítems	Cálculo automático
Título		Cadena de caracteres
Título original / traducción	Título original si el título fue traducido o traducción del título si necesario	Cadena de caracteres
Depositante	Nombre del depositante de la colección en los archivos	Cadena de caracteres. Formato de captura: apellido, nombre
Estado del documentodocumento	Información automática en función del código colección (CNRSMH_I = Inédito ; CNRSMH_E = publicado)	Cálculo automático
Descripción	Texto descriptivo	Zona de texto
Contexto de grabación	Contexto técnico	Enumeración
Periodo de la grabación	Año de principio y año de fin de grabación	Formato de captura AAAA en los 2 campos concatenados “año de grabación”
Fecha de publicación	Fecha de publicación para un documento registrado	Formato de captura AAAA
Tipo de acceso	Condición de acceso libre, restringida, reservada, mixta	Enumeración. Modificable por el Admin / Documentalista
Corpus	Ligas hacia los corpus parientes	Visualización automática

Bloque de indicaciones geográficas y culturales:

Estados / naciones	Recopilación de los países por Ítems	Visualización automática
Poblaciones / grupos sociales	Recopilación de los nombres de las poblaciones, etnias	Visualización automática

2. ORGANIZACIÓN DE LA BASE DE DATOS

Bloque de Información Legal:

Recopilador	Recopilador, colector (persona moral o física)	Cadena de caracteres. Formato de captura: apellido, nombre
Editor	Editor de la versión comercial	La enumeración no se visualiza si no existe
Colección del editor	Colección del editor del disco	Enumeración
Número de la serie	Número en la colección del editor	Cadena de caracteres
Autor del documento publicado	Autor del documento	Cadena de caracteres. Formato de captura: apellido, nombre
Referencia del editor	Número de referencia del editor del disco	Cadena de caracteres
Referencias bibliográficas	Bibliografías sobre la colección archivada	Zona de texto
Acceso automático cuando expiran los derechos de autor (barrera móvil)	Casilla que se desactiva si el acceso al público sigue restringido después de 50 años	
Derecho de uso	Instrucciones para la difusión y utilización	Enumeración

Bloque de datos de archivo, plegados por defecto para aligerar la visualización, datos de archiveros:

Código	Referencia única	Cf. sistema de denominación
Código antiguo	Código de la antigua bases de datos (BM o DI)	
Modo de adquisición	Condición del depósito, donación, compra o intercambio	Enumeración
Depositario CNRS	Pertenencia al CNRS del depositario.	
Tipo de ejemplar	Ejemplar principal o secundario (versión del documento)	Enumeración
Documentalista / Redactor de la ficha	Autor de la descripción	Enumeración
Documentación asociada	En relación con la colección archivada	Zona de texto

2. ORGANIZACIÓN DE LA BASE DE DATOS

Reedición	Tipo de reedición de publicación	Enumeración
Estado de la colección	Original o copia	Enumeración
Otros ejemplares	Indicaciones sobre los otros ejemplares de la colección en los archivos	Zona de texto
Comentarios	Comentarios varios	Zona de texto
Captura de la ficha	Nombre de la persona que hizo la captura en la base de datos	Enumeración
Última modificación de la ficha	Nombre y fecha de la última modificación	Visualización automática
Nota del archivista	Notas sobre el archivado	
Lista de elementos completados	Estado de la documentación	Cadena de caracteres
Colector idéntico al depositante	Marcar casilla	
Editado	Marcar casilla si el documento fue publicado comercialmente	
Lugar de conservación	Ubicación para conservación en el archivo	Cadena de caracteres

Bloque de datos técnicos (plegado por defecto para aligerar la pantalla):

Tipo de medio	Audio / Video	Enumeración
Duración estimada	Duración indicada en la documentación. No se visualiza si el campo "Duración calculada" tiene información	Formato de Captura "HH:MM:SS"
Duración estimada	Duración total de los Ítems	Cálculo automático
Volumen de la colección	Volumen total de la colección en MB	Cálculo automático
Cantidad de componentes: (soportes / pieza)	Cantidad de soportes o de piezas en la colección	Alfanumérico
Cantidad de Ítems	Cantidad de Ítems en la colección	Cálculo automático
Soporte original	Tipo de soporte audiovisual original	La enumeración no se visualiza si no existe
Soporte archivado	Tipo de soporte audiovisual archivado	La enumeración no se visualiza si no existe
Digitalización	Operador de la digitalización	La enumeración no se visualiza si no existe

2. ORGANIZACIÓN DE LA BASE DE DATOS

Bloque Medios asociados:

Título	Título del documento asociado	Cadena de caracteres
Descripción	Descripción del documento asociado	Zona de texto
URL	Enlace URL a un recurso externo	Cadena de caracteres
Créditos	Responsabilidad jurídica	Cadena de caracteres
Expediente	Liga hacia el documento / viñeta de visualización (JPEG)	Descarga

**Bloque identificador (introducido en una pestaña especial).
Este bloque se ha añadido para dar cabida a un identificador perenne generado por la institución:**

Identificador	Referencia	Cadena de caracteres
Tipo de identificador	Tipo ARK o Handle o DOI	Lista 3 entradas
Fecha de primera atribución	Fecha de atribución del identificador	Alfanumérica. "AAAA-MM- DD"
Fecha de última atribución	Fecha de atribución del identificador	Alfanumérica. "AAAA-MM- DD"
Notas	Observaciones	Zona de texto

2.5.4 Ítems

Algunos campos en blanco no se visualizan para aligerar la consulta. El registro Ítem está organizado por bloques temáticos para facilitar su lectura.

Bloque central:

Campos	Descripción	Observaciones
Título	Título del documento	Cadena de caracteres
Título original / traducción	Título original si el título fue traducido o traducción del título si es necesario	Cadena de caracteres
Depositante	Nombre de la persona que deposita el ítem	Cadena de caracteres. Formato de captura: apellido, nombre
Colección	Enlaces a la Colección "madre"	Visualización automática

2. ORGANIZACIÓN DE LA BASE DE DATOS

Fecha de captura (desde)	Fecha	Formato de captura AAAA-MM-DD
Fecha de captura (hasta)	Fecha	Formato de captura AAAA-MM-DD
Tipo de acceso	Condición para la difusión: libre, restringida, reservada, mixto	Lista modificable por admin / documentalista

Bloque de Indicaciones geográficas y culturales:

Lugar	Lugar de la grabación	Enumeración
Detalles de la ubicación	Lugar preciso de la grabación	Cadena de caracteres
Área cultural	Zona de influencia geo-cultural	Cadena de caracteres
Idioma	Idioma del medio	Si ausente de la lista ISO
Idioma (norma ISO)	Idioma del medio	Enumeración
Población / grupos sociales	Población / grupos sociales de los intérpretes	Enumeración
Contexto etnográfico	Contexto etnográfico de la grabación	Cadena de caracteres
Palabras claves	Temática	Enumeración (pestaña de captura especial)

Bloque de Información musical:

Modalidad de ejecución	Modo de ejecución musical	Cadena de caracteres
Estilo vernáculo	Estilo musical vernáculo	Enumeración
Estilo genérico	Estilo musical genérico	Enmeración
Autor / compositor	Autor de la música o del texto	Cadena de caracteres
Captura en una pestaña especial		
Nombre	Nombre del elemento.	Alfanumérica
Voces / instrumentos	Nombre genérico (voices o instrumentos)	Lista
Nombre vernáculo	Término vernáculo del instrumento	Lista
Intérpretes	Nombre del músico	Cadena de caracteres: apellido, nombre

2. ORGANIZACIÓN DE LA BASE DE DATOS

Bloque Datos de archivo:

Código	Nuevo código	Cadena de caracteres
Código original	Código antiguo	Cadena de caracteres
Número del ítem	Número del soporte y/o de la plataforma	Cadena de caracteres
Selección del depositante	Indicación cualitativa del colector	Cadena de caracteres
Depositante / Coleccionista idéntico a la colección	Depositante/ Coleccionista idéntico al de la colección	Casilla a marcar
Referencia del depositante	Nomenclatura / Numeración del depositante	Cadena de caracteres
Referencia editada	Referencias discográficas o bibliográficas	Zona de texto
Acceso automático después de la fecha móvil	Desactivar la casilla si el acceso al público sigue restringido después de 50 años	Casilla a marcar predeterminada
Observaciones	Observaciones varias	Zona de texto
Última modificación de la ficha	Nombre y fecha de la última modificación	Visualización automática

Bloque de Datos técnicos (plegado por defecto):

Tipo de medio	Audio o video	Enumeración
Tipo MIME	Formato digital del medio (norma web)	Visualización automática
Duración estimada	Duración indicativa proporcionada por el archivo	Cadena de caracteres
Duración calculada (visualización "duración")	Duración calculada por Timeside	Calculo automático
Título	Título del documento asociado	Cadena de caracteres
Descripción	Descripción del documento asociado	Zona de texto
URL	Enlace URL hacia una fuente externa	Cadena de caracteres
Créditos	Responsabilidad jurídica	Cadena de caracteres
Archivo	Enlace al documento / Viñeta de visualización (JPEG)	Descarga

2. ORGANIZACIÓN DE LA BASE DE DATOS

Bloque Medios asociados (introducido en una pestaña especial):

Título	Título del documento asociado	Cadena de caracteres
Descripción	Descripción del documento asociado	Zona de texto
URL	Enlace URL a una fuente externa	Cadena de caracteres
Créditos	Responsabilidad jurídica	Cadena de caracteres
Archivo	Enlace al documento / viñeta de visualización (JPEG)	Descarga

Bloque Marcadores (incorporado en una etiqueta especial)

Número del marcador	Número en el orden cronológico del medio	Alfanumérico
Time code	Posición temporal del marcador	HH:MM:SS:FF
Título	Título de la secuencia	Cadena de caracteres
Comentario	Comentario sobre la secuencia	Zona de texto
Autor del marcador	Nombre de la cuenta del usuario	Visualización automática

Bloque Identificación (introducido en una pestaña especial). Este bloque se ha añadido para dar cabida a un identificador perenne generado por la institución:

Identificador	Referencia	Cadena de caracteres
Tipo de identificador	Tipo ARK o Handle o DOI	Lista con 3 entradas
Fecha de primera atribución	Fecha de atribución del identificador	Alfanumérica. AAAA-MM-DD
Fecha de última atribución	Fecha de atribución del identificador	Alfanumérica. AAAA-MM-DD
Notas	Observaciones	Zona de texto

2.6 Descripción de las listas de vocabulario

Las Colecciones y los Ítems tienen tablas externas que en su mayoría son enumeraciones sencillas y las otras más complejas están relacionadas en listas.

Esas son las enumeraciones sencillas:

- Contexto de grabación: campo, espectáculo, radio, etc.
- Derecho de utilización: condición de comunicación (libre, restringida, prohibida, etc.).
- Editor: casa editorial de discos.

2. ORGANIZACIÓN DE LA BASE DE DATOS

- Modo de adquisición: donación, depósito, intercambio, compra, etc.
- Palabra clave palabras del vocabulario de la disciplina..
- Digitalización: operador de la digitalización.
- Población / grupo social: nombre de las poblaciones / etnias.
- Redactor del registro: persona que escribió el documento.
- Reedición: Tipo de reedición de la publicación (parcial, aumentada,etc.).
- Captura del registro: persona que capturó el documento en la base de datos.
- Estilo genérico: término general describiendo el tipo de archivo (popular, religioso,etc.).
- Estilo vernáculo: términos vernáculos describiendo el archivo.
- Soporte archivado: formato del soporte archivado físicamente por el equipo.
- Soporte original: formato original del soporte archivado por el equipo.

Las listas siguientes están incluidas en una tabla (Media_ítem_performance/Formation) que contiene cuatro elementos:

- Instrumentos: nombre de los instrumentos de música (términos genéricos) y de los tipos de producción vocal, 534 términos en 2020.
- Instrumentos (vernáculos): nombre de los instrumentos de música (términos vernáculos) 2020, términos en 2020.
- Nombre: nombre de los instrumentos en cada categoría o nombre de los intérpretes.
- Intérpretes: apellido, nombre del (los) intérprete(s).

En el menú “Admin” hay una tabla denominada “Instrumento alias relación” para indicar las relaciones entre dos listas. Por ejemplo: Ampongavilany > Timbale.

2. ORGANIZACIÓN DE LA BASE DE DATOS

Las listas siguientes fueron agregadas en la nueva base de datos en 2011:

- Tipo de ejemplar: versión del ejemplar, principal o secundario.
- Estatuto de la colección: copia o original.
- Idioma ISO: Lista norma «ISO 639-3» (versión 2010). De no existir en la lista “Idioma ISO”, se lo puede indicar en el campo texto “Idioma”.
- Tipo de identificador: ARK, Handle o DOI (en previsión de la atribución de un identificador permanente). El campo “Tipo de identificador” no está activo actualmente.
- Tipo de media: audio o video.

Los metadatos no siempre están presentes en estas nuevas listas para las colecciones anteriores a 2011.

Nota: El LAM puede utilizar ciertas enumeraciones y pueden ser ocultadas por los usuarios.¹⁷

Están vacías en la versión del CREM:

- Diámetro del carrete
- Disciplina
- Derechos
- Número de pistas
- Organización
- Velocidad de la cinta

¹⁷ Telemeta del LAM: <http://telemeta.lam.jussieu.fr>

2. ORGANIZACIÓN DE LA BASE DE DATOS

Lista alfabética de las enumeraciones/listas

Nombre de las tablas	Origen	Tipo	Nombre interno	Añadido	Nivel catálogo
Colección Editor	Interno	Lista	publisher_collection	4D	Colección
Contexto de grabación	Interno	Enumeración	recording_context	4D	Colección
Derecho de utilización	Interno	Enumeración	legal_rights	4D	Colección
Editor	Interno	Enumeración	publisher	4D	Colección
Instrumentos	Interno	Lista	instrumentos	4D	Ítem
Instrumentos (vernáculos)	Interno	Lista	Instrumentos Alias	4D	Ítem
Instrumentos Alias Relación	Interno	Lista	Instrumentos Alias Relación	4D	
Instrumentos Relación	Interno	Lista	Instrumentos Relacion	4D	
Idiomas ISO	Norma ISO 639-3	Lista	Idiomas_iso	2011	Ítem
Lugares (geográficos)	GéoEthno	Lista	Ubicación	2011	Ítem
Modo de adquisición	Interno	Enumeración	adquisición_modo	4D	Colección
Palabra llave	Interno	Enumeración	keywords	4D	Ítem
Digitalización	Interno	Enumeración	ad_conversion	4D	Colección
Población/ grupo social	Interno	Enumeración	ethnic_group	4D	Ítem
Documentalista/ redactor de la ficha	Interno	Enumeración	metadata_author	4D	Colección
Reedición	Interno	Enumeración	publishing_status	4D	Colección
Captura de la ficha	Interno	Enumeración	metada_writer	4D	Colección
Estado de la colección	Interno	Enumeración	status_id	2014	Colección
Estado del documento	Interno	Enumeración	doc_status	2014	Colección
Estilo genérico	Interno	Enumeración	generic_style	4D	Ítem
Estilo vernáculo	Interno	Enumeración	vernacular_style	4D	Ítem
Soporte archivado	Interno	Enumeración	physical_format	4D	Colección
Soporte original	interno+ LAM		original_format	2014	Colección
Tipo de medio	Interno	Enumeración	media_type	2014	Colección
Tipo de acceso	Interno	Enumeración	public_access	2011	Colección/ Ítem
Tipo de ejemplar	Interno	Enumeración	copy_type	2014	Colección
Tipo de identificador	Interno	Enumeración	archive_id	2014	Colección
Tipo MIME	norma IANA	Variable	mime_type	2011	Ítem

16 <http://www.openarchives.org/OAI/1.1/openarchivesprotocol.htm>
 17 <https://www.iri.centre Pompidou.fr/projets/anthroponet/>

2. ORGANIZACIÓN DE LA BASE DE DATOS

2.7 Funcionalidades de modificación e incorporación de elementos en la base de datos

2.7.1 Importación de medios

La importación manual de la media o medios (audio y video) se realiza en las fichas Ítem a través de la descarga individual en el servidor del archivo digital (sin compresión para audio) gracias al botón “Navegar”, ubicado debajo de la pestaña principal, y en “Registrar”.

La importación de una imagen (JPEG o PDF) se hace en la pestaña “Medios asociados” del Ítem, Colección, Corpus o Fondo de la misma manera. El nombre del archivo de la media tiene que respetar, en la medida del posible, el código de identificación del ítem. De lo contrario, el archivo descargado por el administrador tendrá el nombre del ítem (código del ítem).

2.7.2 Incorporación y modificación de los metadatos

En cada nivel del catálogo se puede agregar un elemento con el botón “Aregar” y duplicarlo con el botón “Duplicar”. Durante una duplicación, la información que se encuentran en la parte “Voces / Instrumentos” se conserva de manera predeterminada. Si se requiere modificarla, hay que registrar y rehacer “Editar” y así modificar los elementos. Al agregar el nombre de una población o de un lugar este queda automáticamente en el registro Colección.

Entonces, la base de datos se puede modificar sencilla y progresivamente, aún sin tener conocimientos especiales en informática. Basta con respetar las convenciones de captura y el sistema de códigos de identificación definido por los archivistas (se verá más adelante) para conservar un sistema homogéneo de identificación de los archivos digitalizados. Es importante subrayar que los metadatos existentes o entregados por el acopiador muy raras veces están completos. La disponibilidad simultánea de la media y su documentación, aunque sea parcial, permite el enriquecimiento en tiempo real, conforme se llevan a cabo las búsquedas del archivista o del acopiador; así como gracias a la disposición de escucha y las contribuciones externas.

En caso ser necesario, algunos campos se pueden duplicar: Palabras claves, Voces / Instrumentos, Media asociada. Para obtener campos adicionales de este tipo, es necesario dar de alta el registro y luego “Editar” nuevamente con el fin de tener nuevos campos disponibles para la captura.

El menú “Revisión” de la página “Admin” permite ver el historial de las modificaciones y el ID de referencia desde la apertura de la plataforma (el 18 de mayo del 2011). Había 280, 867 modificaciones en marzo 2020.

2. ORGANIZACIÓN DE LA BASE DE DATOS

Los usuarios pueden estar al tanto del historial de las modificaciones. Estas informaciones están disponibles en los flujos tipos RSS.

También se pueden modificar los metadatos a través de PhpMyAdmin con una contraseña: https://archives.crem--cnrs.fr/phpmyadmin/db_structure.php?server=1&db=telemeta&token=f448e42ec974fd94167bd6c548_76ce6f. Con esta herramienta, los datos se enlistan y, en ciertos casos, la modificación en serie es más rápida o práctica. Cabe señalar que las modificaciones masivas también son posibles (ver capítulo 2.7.6).

2.7.3 Regla para la captura de datos

Los códigos de identificación:

Cada código de identificación debe ser único, por lo tanto, se requiere uno para cada nuevo elemento y se respeta el modelo existente.

- La clave Fondos es semejante a: CNRSMH_Berthe.
- La clave Corpus es semejante a: CNRSMH_Berthe_001.
- La clave Colección es semejante a: CNRSMH_I_ 9999_999 (Inédito / Campo) o del tipo CNRSMH_E_ 9999_999_999 (Editado / Publicado).

En particular, el código del ítem tiene que ser conforme al de la colección:

Ejemplo: los ítems de la colección CNRSMH_I_ 9999_999 tendrán la forma CNRSMH_I_ 9999_999_999_99 o CNRSMH_I_ 9999_999_999_999.

- Los campos “Fechas”, “Nombre” y “n.º de Ítem”:

Hay que respetar las convenciones de captura para las fechas y el nombre de las personas:

- Fecha: AAAA-MM-DD.
- Nombre de las personas: nombre, apellido.
- Nombre de varias personas: nombre, apellido; nombre, apellido.
- Para los discos: cara y número de la pista (Cara A - Pista 1).
- Para las cintas: número de cinta y número de la pista (:02-15).

Los formularios de captura están organizados por pestañas. Algunos campos tienen una lista de vocabularios o una cadena de caracteres con una función de autocompletar para facilitar la captura de una ficha o duplicarla. Para cada modificación se debe “Grabar” y luego pasar a el registro siguiente.

2. ORGANIZACIÓN DE LA BASE DE DATOS

2.7.4 La estructuración de las colecciones

La estructura de un fondo, al momento de su depósito, se hace en concertación con el depositante, el investigador y los archivistas con el objetivo de organizar y contextualizar de la mejor forma posible las colecciones de acuerdo con la trayectoria científica del coleccionista (o de la institución) y así dar cuenta de la coherencia de los datos recaudados. Las *Colecciones* se pueden agrupar en corpus y los corpus se pueden repartir entre fondos.

Agregar una Colección en un Corpus o un Corpus en un Fondo es una operación sencilla y accesible, que no requiere contar con competencias especiales. Este proceso se realiza desde una cuenta de Admin / Documentalista. (Con una cuenta Admin / Documentalista).

El formulario de edición (botón Editar) permite seleccionar elementos para agregarlos en un Corpus o un Fondo. La lista de los elementos del nivel inferior aparece y se debe seleccionar uno o más elementos ya sea con la barra de desplazamiento de la lista o con el filtro / motor de búsqueda. Una flecha permite la transferencia de su selección en la lista “Selección de colección” o “Selección de Corpus”. Esta estructura puede modificarse en cualquier momento por los administradores y los documentalistas, lo cual permite reagrupar y contextualizar las colecciones y los corpus (campo, título, descripción y fechas).

2.7.5 Modificación de las listas de vocabulario

La mayoría de las listas pueden ser modificadas por los Administradores y los Documentalistas en el menú “Terminología” de la página de inicio. Agregar una palabra de vocabulario se logra con el botón “Aregar”. La modificación de una palabra del vocabulario se logra seleccionando un término en una lista y registrando la modificación.

La normalización de dos palabras de vocabulario redundantes se logra seleccionando un término en “Reemplazar por” y al registrar la modificación. Si se marca “Borrar el valor después de reemplazar”, el término reemplazado por otro desaparece. De este modo, todas las fichas son indexadas sin necesidad de cambiar cada ficha. Se puede agregar una nota o un comentario para cada término.

2.7.6 Modificación masiva (automática)

A través de la herramienta PhpMyAdmin y la contraseña se puede hacer una modificación masiva tomando con las precauciones necesarias y hacer verificaciones de forma sistemática sobre la instancia de pruebas. Este tipo de modificación lleva riesgos porque las funcionalidades de protección y de seguridad presentes en Telemeta no se aplican en esta herramienta.

2. ORGANIZACIÓN DE LA BASE DE DATOS

Instancia de prueba:

<http://sandbox.crem.telemeta.org/phpmyadmin/>

Instancia de producción:

https://archives.crem-cnrs.fr/phpmyadmin/db_structure.php?server=1&db=telemeta&token=f448e42ec974fd94167bd6c54876ce6f

A continuación, se presenta un ejemplo de consulta SQL para modificar el nombre del colector en los ítems de una colección:

UPDATE media_items SET collector = 'Nuevo colector' WHERE code LIKE 'Cote collection%';

UPDATE media_items SET collector = 'Beaudet, Jean-Michel' WHERE code LIKE 'CNRSMH_I_2011_005%'

En el siguiente documento se comparte una lista de consultas típicas:

https://docs.google.com/document/d/1saMv6_Ff3WUR3c67onl6-F_bq5VMS72XgOWhxgB3Ur8/edit

2.8 Interoperabilidad

2.8.1 Cosecha vía el Dublin Core

Para facilitar la cosecha se usan funciones de correspondencia entre los datos de la base y el formato destino, por ejemplo, el Dublin Core. Estas funciones proponen metadatos ejes en unos cuadros de formato CSV.

La correspondencia de los metadatos se visualiza en el formato Dublin Core para cada nivel del catálogo: Fondos, Corpus, Colección, Ítem gracias al botón Dublin Core. El formato Dublin Core es compatible con el protocolo del Open Archives Initiative (OAI-PMH).¹⁸ Esta función se desarrolló gracias al apoyo financiero del MNHN en 2009 (proyecto Anthroponet).¹⁹ El motor de búsqueda del INSHS del CNRS permite la cosecha a través de Isidore:

<https://isidore.science/organization/http://archives.crem-cnrs.fr/>

18 <http://www.openarchives.org/OAI/1.1/openarchivesprotocol.htm>

19 <https://www.iri.centre Pompidou.fr/projets/anthroponet/>

2. ORGANIZACIÓN DE LA BASE DE DATOS

Las equivalencias Dublin Core se visualizan en el cuadro siguiente (colección e ítem)

- Vista Dublin Core de la colección

Element	Refinement	Valeur	Remarque
identifier		media_items.old_code OU media_items.code si présent	Ancienne ou nouvelle cote
type		"Sound"	
title		media_items.title si présent, OU media_collections.title de la collection liée, éventuellement suivi d'un suffixe (n° d'item, etc..)	Titre de la collection ou de l'item si disponible
title		media_items.alt_title	Traduction du titre
creator		media_collections.creator de la collection liée	Auteur / Cédant de la collection
contributor		media_items.collector si présent ou media_collections.collector de la collection liée	Collecteur de l'item ou de la collection
contributor		metadata_authors.value référencé par media_collections.metadata_author_id, de la collection liée	Rédacteur fiche
subject		« Ethnologie »	Valeur constante
subject		« Ethnomusicologie »	Valeur constante
subject*		context_keywords.value liés à cet item via media_item_keywords	Mots clefs du contexte ethnographique
description	abstract	media_item.context_comment	Commentaire contexte ethnographique
publisher		publishers.value référencé par media_collections.publisher_id, de la collection liée	Editeur
publisher		« CNRS – Musée de l'homme »	Valeur constante
date	created	Par priorité, selon disponibilité : 1. media_items.recorded_from_date et media_items.recorded_to_date au format: start=AAAA-MM-JJ; end=AAAA-MM-JJ; 2. media_items.recorded_from_date au format AAAA-MM-JJ 3. (collection liée) media_collections.recorded_from_year et media_collections.recorded_to_year au format : start=AAAA; end=AAAA; 4. (collection liée) media_collections.recorded_from_year au format AAAA	Année ou période d'enregistrement de l'item ou de la collection (conforme à W3C-DTF et DCMI Period)
date	issued	media_collections.year_published au format AAAA, de la collection liée	Année de publication de la collection (conforme à W3C-DTF)
coverage	spatial	Pays, région (si disponible), ville/village (si disponible) du lieu référencé par media_items.location_name (exemple: Sarawak, Malaisie)	Nom complet du lieu
rightsHolder		media_collections.creator de la collection liée	Détenteur des droits
rightsHolder		media_collections.collector de la collection liée si différent de media_collections.creator	Détenteur des droits
rightsHolder		publishers.value référencé par media_collections.publisher_id de la collection liée si éditée	Détenteur des droits
rightsHolder		media_item.performances.musicians si renseignés	Détenteur des droits (interprètes)
rights	license	legal_rights.value référencé par media_collections.legal_rights_id de la collection liée	Statut juridique
rights	accessRights	selon la valeur de media_items.public_access : 4. full : public 5. metadata : restricted 6. none : private	Droits d'accès
format	extent	media_items.approx_duration ou durée calculée de l'item, au format HH:MM:SS	Durée

2. ORGANIZACIÓN DE LA BASE DE DATOS

- Vista Dublin Core del ítem

Referencias:

- DCMI Metadata Terms : <https://www.dublincore.org/specifications/>
- W3C-DTF : <http://www.w3.org/TR/NOTE-datetime>
- DCMI Period : <https://www.dublincore.org/specifications/dublin-core/dcmi-period/2006-04-10/>

2.8.2 Mapeo en el formato CSV

Los metadatos se pueden exportar en formato CSV, lo cual permite hacer un *data mapping* de los metadatos. Esta fue la forma en la que el proyecto Europeana Sounds, recolectó los datos en la página de inicio Europeana: www.europeana.eu/

Entonces se pueden reutilizar los metadatos de esta manera: una o varias colecciones son seleccionadas en una playlis *lista de reproducción* reproducción (escritorio con cuenta de usuario). La exportación de las fichas Colección o fichas Ítems en el formato abierto CSV se describen en el esquema del documento compartido a continuación:

<https://docs.google.com/spreadsheets/d/1Hpvk-5ZvKTr7AOTdrtIt0FkFjrw4lFIMKTYCi0neyOo/edit#gid=359255607>

3. FUNCIONALIDADES DE BÚSQUEDA Y SELECCIÓN

3. Funcionalidades de búsqueda y selección

La mayoría de los usuarios que consultan los Archivos del CNRS-Musée de l'Homme llegan directamente después de una búsqueda en Google. Una vez en el sitio, el visitante puede hacer una búsqueda sencilla o avanzada gracias a un motor interno específico. Todas las búsquedas del usuario con una cuenta son guardadas y pueden recargarse (visualizadas en la página "Escritorio" en "Mis búsquedas"),

3.1 Búsqueda sencilla

Se trata de una búsqueda de texto completo en los campos “texto” de los niveles *Ítem*, *Colección*, *Corpus* y *Fondos*. Los resultados son clasificados de acuerdo con seis campos listados en los dos bloques de filtros, tal como se indica más adelante (ver capítulo 3.3 “Visualización de los resultados”).

3.2 Búsqueda avanzada

Cuando se hizo la puesta en línea en 2011, la búsqueda avanzada ofrecía siete criterios. Actualmente se pueden cruzar 13 criterios.

The screenshot shows the advanced search interface of the CREM (Centre de Recherche en Ethnomusicologie) website. The interface is in French and includes the following fields:

- Rechercher** (Search):
 - titre : (Title) - Text input field
 - collecteur : (Collector) - Text input field
 - lieu : (Place) - Text input field containing "Indonésie"
 - population / groupe social : (Population / Social group) - Radio buttons: "no preference" (selected), "audio", "video"
 - instruments : (Instruments) - Text input field
 - type de média : (Media type) - Radio buttons: "no preference" (selected), "audio", "video"
 - Available for consultation media : (Available for consultation media)
 - no preference
 - online and public
 - online (account required)- Date d'enregistrement:** (Registration date):
 - 1920
 - 1969
- Statut du document :** (Document status):
 - no preference
 - Édités
 - Inédits
- année de parution :** (Year of publication):
 -
 -
- contexte d'enregistrement :** (Recording context):
 - Terrain
- nature du support original :** (Original support nature):
 - Bande(s) magnétique(s)
- cote :** (Call number): Text input field

At the bottom of the form is a "Rechercher" (Search) button with a magnifying glass icon.

At the very bottom of the page, there are footer links: "telemeta POWERED v1.6", "Usage des archives réservé dans le respect du patrimoine culturel des communautés d'origine.", "Copyright © 2015 CREM-CNRS", and "Mentions légales".

3. FUNCIONALIDADES DE BÚSQUEDA Y SELECCIÓN

En 2015, el CREM eligió los criterios de búsqueda con base en los descriptores que consideró prioritarios y gracias a los medios de desarrollo de varios pasantes en informática bajo la tutela de Nédra Mellouli Nauwynck (IUT de informática de Montreuil). Así este motor se convirtió en un módulo independiente (Haystack y ElasticSearch). Algunos criterios recurren a listas con demasiada información (poblaciones e instrumentos).

Los criterios de búsqueda son:

- El título: (Colección o ítem)
- Acopiador: (con función de autocompletar): nombre de la persona moral o física que colectó el archivo.
- Lugar: (con función de autocompletar): lugar de grabación
- Población / grupo social: lista y autocompletado
- Instrumentos: instrumentos de música, nombre genérico y vernáculo (con función de autocompletar)
- Tipo de media: Audio /Video /Sin preferencia
- Disponibilidad del medio: en línea (digitalizado y en acceso libre) / accesible con una cuenta (digitalizado pero no en acceso libre) / sin preferencia.
- Periodo de grabación: es la fecha de grabación de la media, la entrada es la fecha de inicio y la fecha del final, con captura sencilla o con un menú desplegable con las fechas posibles de la base de datos.
- Estado del documento: editados / inéditos / sin preferencia. La selección se hace según si el código de identificación contiene CNRSMH_I_ (inédito/campo) o CNRSMH_E_ (editado/publicado).
- Año de publicación: para un documento editado (disco, casete), entrada con una fecha de principio y una fecha de final, por captura sencilla o con el menú desplegable con las fechas posibles de la base de datos.
- Contexto de grabación: captura sencilla o con el menú desplegable.
- Naturaleza del documento archivado: formato del soporte físico o digitalizado archivado lo cual puede diferir del formato original.
- Código de identificación: los códigos de identificación siempre serán así: CNRSMH_I_AAAA_NNN (inéditos) o CNRSMH_E_AAAA_NNN (editados). Este campo es muy útil para los archivistas (Ver plan de asignación Capítulo 2.4)

3. FUNCIONALIDADES DE BÚSQUEDA Y SELECCIÓN

3.3 Visualización de los resultados (Búsqueda sencilla y avanzada)

3.3.1 Visualización por lista

- Los resultados de la búsqueda aparecen en la lista Colecciones, predeterminada con el número de resultados por nivel del catálogo. También existe la posibilidad de escoger la visualización de los resultados en las listas de ítems, corpus o.
- Se puede organizar cada columna por orden alfabético, creciente/decreciente o hacer una selección y marcar la lista de la izquierda para agregar a una *Lista de reproducción* (se escoge en el menú “Selecciona su *playlist*”). La lista de los ítems seleccionados se exhibe para su verificación antes de “Agregar a la lista”, nuevamente. Se avisa al usuario si el elemento ya fue seleccionado. También uno puede regresar a la lista de resultados de la búsqueda escogiendo “Regreso a los resultados de la búsqueda”.
- Al marcar “Selecciona todo”, todos los ítems mostrados de la búsqueda aparecen y se pueden agregar a una *playlist* de reproducción.
- En cada nivel del catálogo se puede seleccionar el número de resultados por página : 20, 50, 100, 250, 500, 1000.

3.3.2 Filtro de los resultados

Con el objetivo de afinar los resultados de la búsqueda, se propone un filtro de seis criterios, dividido en dos bloques temáticos. Al seleccionar varios criterios, los filtros se multiplican y el resultado toma en cuenta todos los criterios activados.

- Bloque 1, se refiere a la accesibilidad de los resultados de búsqueda (desplegado predeterminado).
- Criterios de acceso: los resultados inclusive los inéditos o los editados.
- Criterios de acceso: los resultados con la media digitalizada (sonido).
- Tipo de acceso: los resultados con la media de libre acceso sin cuenta. (visible).
- Bloque 2, se refiere a las características de la media (se pliega de forma predeterminada para aligerar la ventana)
- Tipo de media: los resultados con el audio o el video.

3. FUNCIONALIDADES DE BÚSQUEDA Y SELECCIÓN

- Contexto de la grabación: Campo o Estudio o Radio.
- Naturaleza del soporte archivado: CDR o disco o cilindro.

Nota: La velocidad de acceso es más lenta porque el módulo ElasticSearch-Haystack necesita una actualización para mantener su compatibilidad con la última versión de Telemeta 1.7 que está en Django 1.8 (en 2020).

3.4 Búsqueda en “Colecciones” e “Ítems” en el menú “Archivos”

En el menú “Archivos” en la página de inicio, se puede hacer una selección en los niveles Colecciones e Ítems gracias al botón “Filtros”, arriba a la derecha de la pantalla dentro de todas las colecciones o todos los ítems. Este botón selecciona las referencias inéditas, las referencias editadas (discos), las referencias que tienen un medio (digitalizado). Este botón de filtro resulta muy útil sobre todo para los archivistas que necesitan acceder a informaciones cuantitativas: ¿Cuántos ítems digitalizados existen? y ¿cuántas colecciones inéditas hay? El botón de filtro existía previamente como mejora para el motor de búsqueda avanzada. A partir de 2014, tiene doble uso y su acceso es más directo cuando uno navega en el menú “Archivos” y por eso se conservó.

En el nivel Ítem del menú “Archivos”, se puede seleccionar de la misma manera el número de resultados por página: 20, 50, 100, 250, 500, 1000 ... para facilitar la lectura de los resultados y marcar las selecciones en una lista personal.

3.5 Búsqueda por países

El Geo-navegador es accesible para todos los usuarios y facilita la visualización de las colecciones por país, en un mapamundi o en listas de país por continente en el menú “Listar”, ubicado arriba y a la derecha de la pantalla.

En apoyo a la localización de las colecciones se puede combinar el tesoro Géo Ethno²⁰ del LESC (que toma en cuenta los términos históricos de los nombres de los lugares) y Geonames²¹(que proporciona las coordenadas geográficas). La cantidad de colecciones e ítems por país y el enlace del resultado aparecen seleccionados con un cursor.

20 GéoEthno: <http://www.mae.u-paris10.fr/dbtw-wpd/bed/index-lesc.html>

21 <https://www.geonames.de/>

3. FUNCIONALIDADES DE BÚSQUEDA Y SELECCIÓN

Nota:

- Por ejemplo, utilizar GéoEthno permite asociar en una búsqueda a Burkina Faso y Alta Volta como un mismo lugar. Este tesauro está en francés.
- La actualización de GéoEthno tendrá que efectuarse a partir de la integración de nuevos términos capturados en el campo “Precisión lugar” (ver consignas capítulo 7).
- En la actualidad la plataforma utiliza Google Map. Se preve el uso de Open Street Map que es gratuito y colaborativo. El propósito en futuros proyectos es utilizar una herramienta más abierta y afinar la geolocalización de las fuentes.

3.6 Búsqueda por listas/enumeraciones

La selección de una lista en el menú “Terminología” en la página de inicio, permite acceder a todos los términos de una lista y buscar todos los ítems indexados a cada término de una lista. En cada lista, la cantidad de ítems se visualiza frente a cada término correspondiente. Por ejemplo, en la lista “Palabra clave”, se encuentran 201 canciones de cuna. Al hacer clic en el número “201”, se muestra la lista de los 201 ítems de las canciones de cuna, se pueden seleccionar y agregar en una *lista de reproducción*. Esta funcionalidad es útil para los documentalistas y para la gestión de las listas.

Se accede a las enumeraciones siguientes a través del perfil administradores y documentalistas quienes tienen el estatus de “superusuario”.

- Contexto de la grabación
- Derecho de utilización
- Editor
- Modo de adquisición
- Palabra clave
- Digitalización
- Población / grupo social
- Redactor de el registro
- Reedición
- Captura ficha

3. FUNCIONALIDADES DE BÚSQUEDA Y SELECCIÓN

- Estado de la colección
- Estilo genérico
- Estilo vernáculo
- Soporte archivado
- Soporte original
- Tipo de ejemplar
- Tipo de identificado
- Tipo de media

Las listas de “Instrumentos” (términos genéricos y términos vernáculos) se pueden consultar por todo el público en el menú “Terminologías” lo cual permite hacer una investigación por instrumento de música del mismo modo que en las enumeraciones mencionadas arriba.

Nota: La consulta de las listas para cada perfil de usuario debería ser agregada en los permisos por perfil con la finalidad que estas listas se visualizan (modo sólo lectura) por los perfiles diferentes a los “superusuario”. Este permiso no fue especificado al momento del desarrollo del acceso a las listas en la página de inicio por los pasantes en informática.

4. FUNCIONALIDADES DE NAVEGACIÓN Y VISUALIZACIÓN

4. Funcionalidades de navegación y visualización

4.1 En un Fondo o en un Corpus

Cuando se accede a través del menú “Archivos” las posibilidades de navegación son las siguientes:

- Visualización por listas de títulos de *Fondos o de Corpus*,
- Títulos con ligas en hipertexto (en rojo) para acceder al *Fondo y al Corpus*
- Visualización de los números de páginas de las listas
- Flechas para salto de página.

→ Navegación en un Fondo:

Al seleccionar la ficha Fondos, los títulos de los Corpus “hijos” aparecen en la lista. El registro puede incluir recursos asociados visibles en el apartado “Media Asociada” (imagen, ligas, bibliografía, etc.) y la lista de los Corpus se sitúa abajo del registro. Al igual que en los otros sitios, las ligas hacia fuentes exteriores y los Corpus “hijos” están en rojo. La lista de los Corpus muestra los títulos, la descripción, la existencia de medios/media digitalizados y el código.

→ Navegación en un Corpus:

En el registro Corpus seleccionado, los títulos de las Colecciones “hijas” se muestran en listas. El registro puede contener fuentes asociadas visibles en la sección “Media asociada” (imagen, ligas, bibliografía, etc.) y la lista de las Colecciones se sitúa abajo del registro. El o los fondos “padre(s)” están marcados en rojo para facilitar el rastreo en el catálogo. La lista de las Colecciones muestra los títulos, la descripción, la presencia de los medios/ media digitalizados y el código.

El botón **Dublin Core** permite visualizar en una nueva ventana los resultados de cada registro descriptivo Fondo o Corpus en el formato de los metadatos Dublin Core sencillo y extendido. Así se puede acceder a los valores recogidos por el protocolo OAI-PMH .

4. FUNCIONALIDADES DE NAVEGACIÓN Y VISUALIZACIÓN

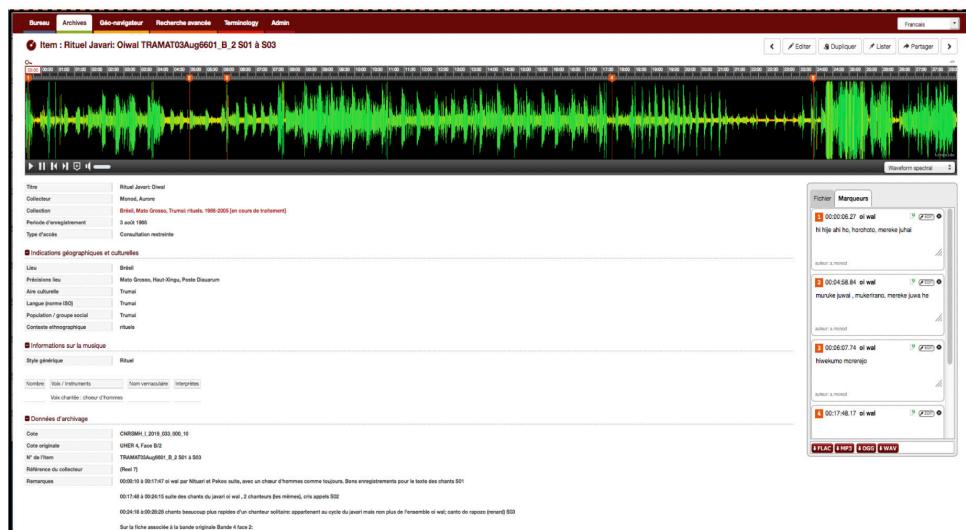
4.2 En una Colección

La colección es un conjunto básico que recopila los ítems de una misma misión, de un mismo disco o de una misma temática recopilados por la misma entidad moral o física.

Una vez seleccionada la colección, el registro “Colección” aparece con el número de ítems. Una liga dinámica a la lista de los ítems permite el acceso directo a los ítems localizados abajo del registro Colección. El o los Corpus “padre(s)” están marcados en rojo (liga hipertexto) para facilitar el rastreo en el catálogo. Así se puede navegar en el catálogo en los corpus y fondos “parientes”. Se pueden visualizar los siguientes elementos:

- La lista de los ítems muestra los títulos, la descripción, la presencia de los medios/media digitalizados (audio y video) y el código.
- La presencia de la medios/media de audio o video indicada por una marca verde.
- Algunos campos en blanco no se visualizan para aligerar la consulta.
- De acuerdo con el estatus del acceso de la colección y de los ítems y según su perfil, el usuario puede escuchar directamente un ítem en la lista (botón “lectura / pausa “a la izquierda del título) o seleccionar el registro ítem para consultar el registro documental y el sonido / video en streaming (si la media existe en la base de datos).
- El botón “Dublin Core” permite la visualización en una nueva ventana con las equivalencias en el formato, sencillo y extendido, de los metadatos Dublin Core.

4.3 En un Ítem



4. FUNCIONALIDADES DE NAVEGACIÓN Y VISUALIZACIÓN

Una vez que el ítem fue seleccionado, el registro descriptivo aparece. El enlace hacia la Colección “madre” está en rojo (liga hipertexto) y permite rastrear el catálogo. Para aligerar la consulta algunos campos en blanco no se visualizan. El sonido o el video aparecen junto con la descripción documental. (ver TimeSide capítulo 6).

Las funciones de navegación del registro ítem son las siguientes:

- Paso al ítem siguiente o anterior (flechas arriba a la derecha)
- Visualización del reproductor en pantalla grande (Lupa)
- Pestaña “Marcadores”: lista de marcadores temporales eventuales
- Pestaña “Archivo”: metadatos técnicos calculados
- Navegación en la media: lectura, pausa, regreso al marcador anterior/siguiente, marcador (bandera) y nivel de escucha.
- Cursor móvil de lectura con visualización del Time Code (MM.SS)
- Botón “Marcador” para posicionar una anotación (título y comentario) en el tiempo.
- Muestra en rojo la posición del marcador de tiempo con números y la anotación del mouse durante la lectura.
- Menú opción de visualización gráfica: forma de ondas (Waveform predeterminada).
- Liga hipertexto en rojo para llegar al nivel Colección
- Liga hipertexto en rojo para acceder al “Media asociada”

El botón Dublin Core permite visualizar en una nueva ventana las equivalencias con el formato sencillo y extendido de metadatos Dublin Core.

5. Funcionalidades de gestión de accesos a los datos

Con Telemeta se pueden gestionar los accesos a los datos de acuerdo con el estatus de las colecciones y el perfil del usuario.

Esta función mejoró en 2014. En aquellos tiempos, la legislación sobre los derechos de los artistas-intérpretes abogaba por una restricción de uso de 50 años. El Código del Patrimonio²² prevé igualmente un plazo de 50 años para que los archivos públicos puedan ser difundidos (categoría cida privada). Esta regla se modula de acuerdo con las consignas del depositante y de las comunidades de los pueblos originarios, caso por caso según las condiciones éticas. En la página de inicio se invita a los usuarios académicos a pedir una cuenta usuario para acceder al conjunto de las fuentes audiovisuales digitalizadas que se pueden consultar en línea. Los datos descriptivos (metadatos) están en acceso libre, con la excepción de las fichas ítems de las colecciones disponibles “sobre pedido”. (ver preguntas jurídicas, capítulo 7).

5.1 Gestión de la conexión

La aplicación habilita una restricción de uso (depósito) de 50 años (o más) sobre colecciones de la base de datos. La cuenta usuario permite acceder a todos los metadatos (con excepción de los ítems consultados sobre demanda) y a todos los sonidos/videos en depósito (con excepción de los ítems en consultados bajo demanda). Es posible elegir una duración distinta a 50 años, para ello, es necesario desmarcar la casilla del acceso automático cuando expiren los derechos de autor y elegir la fecha en que la colección tendrá acceso automático de acuerdo con la legislación de cada país.

5.1.1 Creación de la cuenta usuario

Esta cuenta es asignada por los administradores del sitio (en la página “Admin” / Aut / Usuarios) y la creación de la cuenta necesita como mínimo el nombre del usuario y su dirección email.

Una cuenta puede estar asignada a una serie de direcciones IP proporcionadas por una institución que desea dar un acceso en sus instalaciones sin pedir una cuenta para cada usuario (solamente consulta).

A cada cuenta, el administrador puede asignar un perfil de entre cinco grupos: Administrador, Documentalista, Investigador, Miembro, Visitante.

5. FUNCIONALIDADES DE GESTIÓN DE ACCESOS A LOS DATOS

Hay 275 permisos diferentes asignados al Administrador, Documentalista y para cada cuenta, el administrador escoge un perfil dentro de 5 grupos: Administrador, Documentalista, Investigador, Miembro, Visitante.

Existen 275 permisos distintos asignados a los grupos Administrador, Documentalista e Investigador. Los permisos (lectura, escritura y eliminación) pueden ser administrados para cada cuenta de usuario, así como la cantidad de permisos de grupo. Estos permisos se pueden consultar en PhpMyAdmin²³ (contraseña):

https://archives.crem-cnrs.fr/phpmyadmin/sql.php?server=1&db=telemeta&table=auth_permission&pos=0&token=a53d4d8e81e82df4dd6e7ced8ea322ca

Todas las cuentas de usuario pueden ser consultadas en (contraseña):

https://archives.crem-cnrs.fr/phpmyadmin/sql.php?server=1&db=telemeta&table=auth_user&pos=0&token=a53d4d8e81e82df4dd6e7ced8ea322ca

Los administradores pueden escoger una fecha de caducidad para cada cuenta y verificar la fecha de la última conexión. También pueden atribuir uno o varios niveles a la cuenta del usuario:

- Activo: especifica si el usuario debe considerarse activo. Desactiva la casilla en lugar de suprimir la cuenta.
- Estado del equipo: precisa si el usuario se puede conectar al sitio de administración.
- Estado de superusuario: precisa si el usuario tiene todos los permisos sin permisos específicos.

5.1.2 Conexión a la plataforma

La conexión a la plataforma se realiza en la página de inicio con el botón “Conexión”. La elección del idioma de la interfaz es automática de acuerdo con el idioma utilizado en el navegador. El usuario coloca el nombre de su cuenta (generalmente en minúscula, por ejemplo: d.abad) y su contraseña. En caso de olvido, puede hacer clic sobre “Contraseña olvidada”. En este caso, el usuario recibe una nueva contraseña a la dirección email que utilizó para la creación de su cuenta.

23 https://archives.crem-cnrs.fr/phpmyadmin/db_structure.php?server=1&db=telemeta&token=f448e42ec974fd94167bd6c54876ce6f

5. FUNCIONALIDADES DE GESTIÓN DE ACCESOS A LOS DATOS

5.2 Tipos de acceso a las colecciones e ítems (restricción de uso)

Los administradores pueden escoger una duración de restricción de uso de 50 años para el acceso público a la media. El acceso está en constante evolución, cada año gracias a la utilización de una “fecha móvil”. Actualmente, la mitad de los datos no están en acceso abierto. Las colecciones se pueden configurar por ítem y por colección en el campo Tipo de acceso al nivel Colección y al nivel Ítem (con una cuenta administrador o documentalista). El estado de la colección es prioritaria sobre el del ítem. En caso de querer cambiar el estado de todos los ítems, basta cambiar el de la colección y el tipo de acceso a los ítems será el mismo que el de la colección. De manera predeterminada, una colección tiene acceso restringido al igual que los ítems. Si la colección se vuelve accesible para todos (“Consulta en acceso abierto”) no es necesario modificar el estado de cada uno de los ítems de la colección únicamente el estado de la colección. El “Acceso restringido” de los ítems sigue visible pero este estado está desactivado y el acceso se vuelve abierto.

Los ítems de una misma colección pueden necesitar estados de acceso distintos (abierto, restringido, reservado). En este caso, hay que escoger “Acceso parcial a los ítems” para la colección.

- El tipo de acceso de una colección

Existen cuatro tipos de acceso que pueden ser seleccionados con el propósito de responder a todas las posibilidades de comunicabilidad de los contenidos:

- Consulta abierta: el acceso a los metadatos y al sonido está autorizado al público (full). Estado predeterminado.
- Consulta restringida: solamente el acceso a los metadatos está autorizado al público (metadata).
- Consulta sobre pedido: el acceso a los metadatos de los ítems y a la media está reservado para los perfiles administrador y documentalista (none).
- Acceso parcial a los ítems: los ítems de la colección pueden tener estados diferentes (mixto) dentro de los tres otros estatutos (Full, Metadata, None). El estado de cada ítem tendrá que ser configurado de acuerdo con las necesidades.

La casilla “Acceso automático después de que expiran los derechos de autor (barrera móvil)” determina cuándo pasará la colección a estar en acceso abierto. En el caso de Francia son 50 años después de la fecha de registro. Esta fecha puede ser parametrizada en función de la legislación de cada país. Si la colección tiene varias fechas de registro se tomará en cuenta, de manera automática, la fecha

5. FUNCIONALIDADES DE GESTIÓN DE ACCESOS A LOS DATOS

del ítem más reciente. El estatus “Consulta restringida” sigue visible a pesar de que la colección se volvió en acceso abierto de manera automática según la fecha móvil. Cada año, las colecciones que cumplen 50 años pasan al acceso abierto. Si la casilla no está activada la colección no se vuelve en acceso abierto después de 50 años, por lo tanto sigue en consulta restringida. Esta funcionalidad alienta a numerosos investigadores a depositar y archivar sus recopilaciones de campo, aun cuando esos datos son “sensibles” lo cual es el caso más frecuente.

El tipo de acceso de un ítem

Existen tres tipos de acceso:

- Consulta en acceso abierto: el acceso abierto a los metadatos y al sonido está autorizado al público (acceso total). Estado predeterminado.
- Consulta restringida: solamente el acceso a los metadatos está autorizado al público. (metadata)
- Consulta sobre pedido: El acceso a los metadatos de los ítems y a la media está reservado a los perfiles administrador y documentalista (none). El registro Colección continúa siendo accesible (metadatos).

Cuando no se puede consultar la media del ítem, el usuario recibe un mensaje que le invita a contactar el equipo para obtener una autorización con una liga de hipertexto hacia la página de avisos legales.

El siguiente cuadro muestra todas las posibilidades que deben ser tomadas en cuenta para brindar acceso a los datos tal como se estableció en 2014.

5. FUNCIONALIDADES DE GESTIÓN DE ACCESOS A LOS DATOS

Tabla resumida de metadatos y acceso a los medios

Statut Collection	Statut Item	Statut prioritaire	Soumis à date glissante	Accès Admin Doc	Accès Chercheur Membre	Accès Public
Full	Full, metadata ou none	Collection	oui	full	full	full
metadata only	Full, metadata ou none	Collection	oui	full	full	metadata (pendant 50 ans)
metadata only	Full, metadata ou none	Collection	non	full	metadata	metadata
none	Full, metadata ou none	Collection	oui	full	none (pendant 50 ans) Mais fiche coll. accessible	none (pendant 50 ans) Mais fiche coll. accessible
none	Full, metadata ou none	Collection	non	full	none	none
mixte	full	Item	oui	full	full	full
mixte	metadata		oui	full	full	Metadata (pendant 50 ans)
mixte	metadata		non	full	metadata	metadata
mixte	none		oui	full	none (pendant 50 ans) Mais fiche coll. accessible	none (pendant 50 ans) Mais fiche coll. accessible
mixte	none		non	full	none mais fiche coll. accessible	none mais fiche coll. accessible

Admin= administradores + documentalistas.

Durante 50 años = estatus temporal (depende de la fecha de expiración de 50 años).

5. FUNCIONALIDADES DE GESTIÓN DE ACCESOS A LOS DATOS

5.3 Perfiles de usuarios

Los derechos de acceso varían en función de los perfiles usuario.

Se pueden consultar y modificar los cinco perfiles en el menú “admin”:

- Admin
- Documentalista
- Investigador
- Miembro
- Visitante

La lista de los permisos se indica en cada perfil (página Admin). Las principales características son las siguientes:

- **Perfil Admin**
 - Acceso a todos los datos y a todas las opciones de visualización gráfica audio.
 - Permiso para modificar todos los datos (Creación, Supresión, Duplicación)
 - Exportación de fichas en ZIP o en EPUB
 - Exportación de sonidos (WAV, FLAC, MP3, OGG)
 - Gestión de tesauros: modificar, suprimir, fusionar y agregar nuevas entradas.
 - Gestión de los accesos (permisos), creación de cuentas usuarios y atribución de los códigos de acceso (*login*).
 - Acceso al escritorio (*playlist*, últimas modificaciones)
 - Exportación de metadatos en formato CSV (a través de una *playlist*)
 - Modificación y supresión de todos los marcadores
 - Incorporación de “una media asociada”: PDF, JPG, PNG, URL

5. FUNCIONALIDADES DE GESTIÓN DE ACCESOS A LOS DATOS

- **Perfil Documentalista**

- Acceso a todos los datos y a todas las opciones de visualización gráfica audio.
- Permiso para modificar todos los datos (Creación, Supresión, Duplicación)
- Exportación de registros en zip o en EPUB
- Exportación de sonidos (WAV, FLAC, MP3, OGG)
- Gestión de tesauros: modificar, suprimir, fusionar y agregar nuevas entradas
- Acceso al escritorio (*playlist*, últimas modificaciones)
- Exportación de metadatos en formato CSV (a través de una *playlist*)
- Modificación y supresión de todos los marcadores
- Incorporación de “una media asociada”: PDF, JPG, PNG, URL

- **Perfil Investigador**

- Acceso a todos los datos (menos al estado None /Consulta sobre demanda) y a todas las opciones de visualización gráfica audio.
- No puede hacer descargas
- Permiso para modificar datos de ítems, corpus y fondos (Creación, Duplicación, sin supresión)
- Sin creación de ficha colección
- Acceso al escritorio (*playlist*, últimas modificaciones)
- Exportación de metadatos en formato CSV (a través de una *playlist*)
- Colocación de “marcadores” y captura de comentarios
- Incorporación de una media asociada: PDF, JPG, PNG, URL
- Firma del marcador (automático)

5. FUNCIONALIDADES DE GESTIÓN DE ACCESOS A LOS DATOS

- **Perfil Miembro (cuenta institucional con dirección IP)**
 - Acceso a los metadatos y a la media (menos al estatuto None /Consulta sobre demanda)
 - No hay descarga
 - Escucha de los archivos de sonido comprimidos en monoaural
 - Acceso al escritorio (*playlist*, últimas modificaciones)
 - No puede colocar marcadores
 - Búsqueda sencilla y multicriterio
 - Sin modificación posible, consulta sencilla
- **Perfil Visitante**
 - Acceso a los metadatos (menos None) y media en acceso libre
 - No hay descarga
 - Búsqueda sencilla y multicriterio
 - Liga con el sitio web del Laboratorio y de los sitios socios

5.4 Funciones de acceso a los media y metadatos

Hay diferentes funcionalidades previstas para acceder a los datos y metadatos. Todo el catálogo es accesible a todo el público (excepto los *ítems* de las colecciones en “Acceso sobre pedido”), incluso si no se tiene una cuenta de usuario. Los medios son accesibles según el estatus de las colecciones (y de los *ítems*).

Se puede oír un *ítem* directamente en la lista de los *ítems* (botón “Lectura” a la izquierda de los títulos) o seleccionar el registro *ítem* para consultar el registro documental al mismo tiempo que la visualización del media en streaming (si el media existe en la base de datos).

5. FUNCIONALIDADES DE GESTIÓN DE ACCESOS A LOS DATOS

También los metadatos son accesibles a través de la exportación y por lo tanto reutilizables: en la colección con el formato EEPUB (libro digital con los metadatos y los medios) o en una *lista de reproducción* (escritorio) con el formato CSV (con cuenta de usuario). La exportación de los registros Colección o las fichas *Ítems* de la “*lista de reproducción*” con el formato abierto CSV se logra gracias al esquema descrito en el siguiente documento:

<https://docs.google.com/spreadsheets/d/1Hpvk-5ZvKTr7AOTdrtIt0FkFjrw4IFIMKTYCi0neyOo/edit#gid=359255607>

Los medios (audio y video) de una colección son accesibles al descargarlos a través del botón “Zip” en el registro colección (cuenta Admin y Doc). Las exportaciones bajo diversos formatos se pueden realizar en la aplicación web de gestión de la base de datos PhpMyAdmin (ver capítulo 2.8).

La “media asociada” (imágenes, PDF, JPG, PNG) son accesibles en una viñeta previsualizada o por descarga.

6. Funcionalidades de gestión de medios/media (sonido y video)

TimeSide²⁴ es un *framework*²⁵ que gestiona el análisis de audio, la representación gráfica, la transcodificación, el streaming de audio y video y la anotación temporal con marcadores editables. Es una API en lenguaje Python que permite tener una interfaz web con visualización dinámica de la media en el registro descriptivo. Es gratuita (desarrollada con el ANR Wasabi desde 2016), descargable y que se puede instalar a través de Docker en todos los sistemas de explotación.

La ventaja para gestión de archivos es la transcodificación en tiempo real y sobre pedido de un archivo audio para consulta en línea de forma comprimida (OGG). La consultan de un sonido es casi inmediata después de su descarga en el servidor y así no es necesario esperar el cálculo de la compresión ni gestionar el respaldo de dos archivos de sonido para cada ítem (comprimido y no comprimido).

La lectura utiliza un player HTML 5 extensible, lo cual permite una ampliación para una visualización en pantalla completa. El análisis audio permite extraer características sonoras y técnicas gracias a varias “librerías” audio, como Aubio, Yaafe o VAMP. Todos los plug-in o algoritmos fueron desarrollados bajo licencia libre y se pueden reutilizar. (ver: <https://github.com/Parisson/TimeSide/>)

6.1 Datos producidos por TimeSide transcodificación (análisis de audio)

Una de las funcionalidades más importante de TimeSide para archivar es la compresión de los sonidos. La compresión de los archivos de sonido se calcula de forma automática lo cual permite consultar un sonido comprimido más adaptado al streaming en el momento de su descarga en la plataforma. La operación de compresión y la de doble backup en el servidor (el sonido no comprimido WAV y el comprimido OGG) no se realiza por el archivista, lo cual ahorra tiempo. En cada cálculo el archivo comprimido automáticamente por TimeSide es almacenado en la memoria del servidor. Los usuarios no pueden acceder al sonido alta definición (estéreo, WAV, 48 Hz o 44 Hz, CREM). Solamente los administradores lo pueden descargar. El formato WAV es el más común, bien documentado y no comprimido. Permite el almacenamiento a largo plazo (CINES, Archivos Nacionales).

24 <https://github.com/Parisson/TimeSide/>

25

6. FUNCIONALIDADES DE GESTIÓN DE MEDIOS/MEDIA (SONIDO Y VIDEO)

TimeSide administra automáticamente los siguientes formatos audiovisuales:

- Archivos no comprimidos para la conservación: WAV (48kHz / 24 bit; 44.1 kHz / 16 bit), FLAC
- Archivos comprimidos audio: MP3 y OGG Vorbis
- Archivos comprimidos video: MP4 y WebM

Nota: El formato OGG es un formato abierto mientras que el MP3, utilizado ampliamente, es un formato propietario (con licencia). Además, a la misma velocidad de flujo, es de mejor calidad. Además, a igual velocidad de flujo tiene mejor calidad.

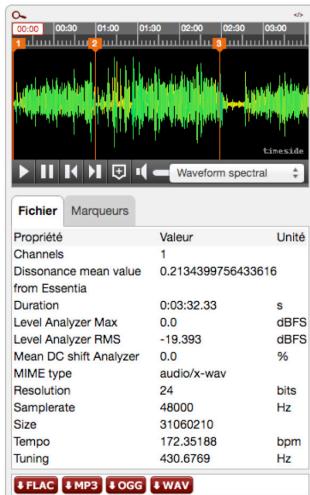
Las características técnicas de los datos de audio gestionados por TimeSide son las siguientes:

WAV: Formato privilegiado para almacenamiento y respaldo audio, este formato sin pérdida tiene una calidad casi igual a la de un CD audio. N = 8, 16, 24, o 32 bits. Frecuencia de muestreo (simple rate) de 8,000 a 192,000 Hz.

MP3: Utilizado para transmisión en streaming y lectura para los lectores en línea, comprime con pérdidas para disminuir el flujo. Ampliamente utilizado por su portabilidad, puede encapsular parcialmente los metadatos. Este formato no es código abierto (las regalías son para el encodificador, no para el decodificador) N = 8 o 16 bits, fe de 8,000 a 48,000 Hz, k de 0,05 a 0,2.

OGG Vorbis: Calidad superior para un flujo igual al MP3 y código abierto (*open source*), encapsula todos los metadatos. N = 8 o 16 bits, fe de 11, 050 a 96,000 Hz, k de 0,01 a 0,3.

FLAC: Compresión sin pérdida, calidad igual al WAV en la escucha, multipista, *open source*, encapsula totalmente los metadatos. N = 16 o 24 bits, fe de 11, 050 a 96, 000 Hz, k =0,6. El CREM no utiliza este formato.



6. FUNCIONALIDADES DE GESTIÓN DE MEDIOS/MEDIA (SONIDO Y VIDEO)

El reproductor incluye una ventana de visualización y de navegación en el medio así como una pestaña “Archivo” con los datos técnicos calculados.

Esta pestaña es predeterminada. Algunos algoritmos (Disonancia, Tempo y Tuning) se calculan al momento de la descarga del sonido en el servidor, desde 2018 fueron integrados en TimeSide para el proyecto DaCaRyH (ver tabla página siguiente).

Nota: En la versión del CREM, la descarga de los formatos de los archivos de sonido visualizada aquí es posible únicamente para los perfiles de “Administradores” (WAV, FLAC, OGG, MP3) pero este permiso se puede ampliar en el menú Admin (Perfiles de usuarios).

Tres de los algoritmos del proyecto DaCaRyH funcionan directamente en esta ventana, los demás están disponibles en el menú desplegable del reproductor con una cuenta (ver capítulo 6.3).

El proyecto “DaCaRyH, Data science for the study of Calypso Rhythm through History” (2016-2017) reúne a un conjunto de etnomusicólogos e ingenieros del CREM-LESC, la Sociedad Parisson y del Centro de Música Digital (C4DM) de la Universidad Queen Mary de Londres financiado por el LABEX “Los Pasados en el Presente” y su homólogo británico l’AHRC Care for the Future.

Los datos calculados visualizados en la ventana “Archivo” del reproductor son los siguientes:

Nombre/valor calculado	Significado	Unidad	Observaciones
Channels	Cantidad de canalews		1 o 2
Duración	Duración	S	
Level Analyzer Max	Nivel máximo	dBFS	
Level Analyzer RMS	Nivel promedio	dBFS	
Mean DC shift Analyzer	Diferencia entre los canales	%	
MIME type	Tipo MIME		
Resolución	Resolución audio	bits	16 o 24
Samplerate	Tasa de muestuario	Hz	
Size	Tamaño archivo digital (Mb)	Bit	
Valores musicales:			
Tempo	Tempo promedio	bpm	Oscilaciones por mn. Algoritmo (DACARYH)
Tuning	Frecuencia del LA3	Hz	Algoritmo (DACARYH)
Dissonance mean value from Essentia	Tasa de disonancia	entre 0 y 1	Algoritmo Essentia (DACARYH)

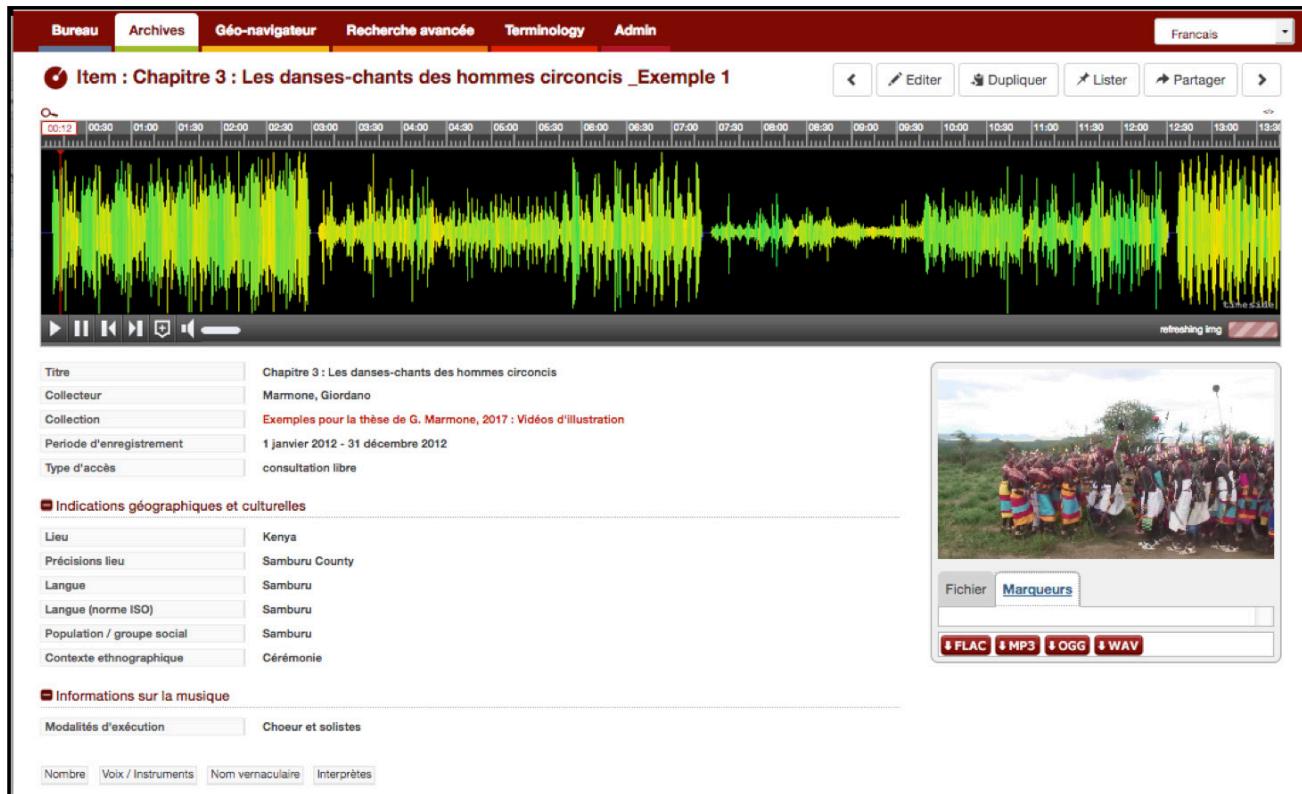
El tamaño del archivo (en Mb) se ve también en el registro ítem y en la suma de los archivos ítems en colección (duración y tamaño). La suma de las duraciones se muestra en los Corpus y los Fondos.

6. FUNCIONALIDADES DE GESTIÓN DE MEDIOS/MEDIA (SONIDO Y VIDEO)

6.2 Lector de streaming

En cada registro *Ítem* es posible (pero no obligatorio) asociar un elemento audio o video al registro descriptivo, lo cual permite verificar la documentación durante la descripción o consulta del archivo audiovisual al mismo tiempo que la lectura del media en la misma página. De hecho, la descripción de un sonido musical es compleja (menos inmediata que una foto) y algunos detalles organológicos, por ejemplo, escapan a menudo a la memoria del coleccionista. Sus notas de campo son generalmente incompletas y solamente la escucha permite validar o corregir los metadatos que varían de una pieza a otra. En cada una de las sesiones de trabajo de archivo con un investigador el sonido escuchado despierta la memoria casi sistemáticamente. Por lo tanto, los datos suelen enriquecerse gradualmente.

Para eso, una visualización dinámica del registro descriptivo del medio resulta muy útil. Esto es posible gracias al lenguaje informático “orientado a objetos” utilizado por TimeSide (Python). Esta característica aporta comodidad y es muy apreciada por los usuarios. Permite igualmente “navegar” en la situación actual del documento e identificar visualmente las diferentes secuencias del material sonoro.



6. FUNCIONALIDADES DE GESTIÓN DE MEDIOS/MEDIA (SONIDO Y VIDEO)

El reproductor (audio y video) propone las siguientes funcionalidades:

- Visualización del reproductor en la pantalla completa en el registro ítem para el audio (botón lupa en la barra superior)
- Visualización del reproductor en pantalla completa en el registro ítem para el video (botón pantalla completa en sobreimpresión del lector video)
- Un Time Line con un medidor de tiempo para el audio y el video (sincrónico)
- Cursor móvil de lectura con la visualización del Time Code (MM:SS)
- Navegación en la media: lectura (>), pausa (II), regreso al marcador anterior / siguiente, volumen de escucha.
- </>: Botón “i-frame” “arriba a la derecha, para insertar el enlace HTML del reproductor en otro sitio web (muy utilizado).
- Botón “Marcadores” para posicionar una anotación (título y comentario) en el tiempo. Visualización en rojo de la posición del marcador temporal con su número y de la anotación al pasar el mouse o durante la lectura.
- Pestaña “Marcadores”: lista de los marcadores eventuales con el Time Code.
- Pestaña “Archivo”: metadatos técnicos calculados.

6. FUNCIONALIDADES DE GESTIÓN DE MEDIOS/MEDIA (SONIDO Y VIDEO)

6.3 Representación gráfica de los medios

TimeSide determina la visualización del sonido directamente en el servidor (sin pasar por otra herramienta local). Facilita la consulta del documento con su duración y la colocación de marcadores temporales para anotaciones. Permite detectar las diferentes secuencias en el documento y sus patrones melódico-rítmicos.



TimeSide calcula la visualización en el servidor, esta tarea puede ser tardada si el archivo es pesado (superior a 10 minutos).

La siguiente ilustración muestra la visualización con “Spectrogram Log” que permite la detección de las partes del canto.

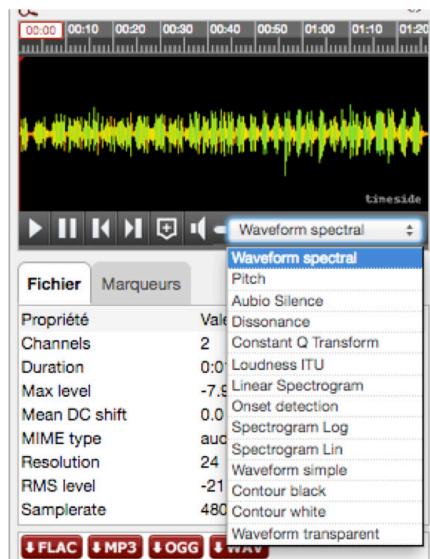
La visualización gráfica es predeterminada, de forma de ondapectral (waveform spectral) tal como se ve [página 68](#). Es la variación de la intensidad en función del tiempo con una variación de los colores en función de la frecuencia. Un menú desplegable en la barra de navegación audio propone otras representaciones gráficas del sonido. Para limitar la saturación del servidor, unas opciones son de acceso libre, otras disponibles con una cuenta.

6. FUNCIONALIDADES DE GESTIÓN DE MEDIOS/MEDIA (SONIDO Y VIDEO)

6.3.1 Con una de cuenta usuario

Todas las herramientas de visualización (o su opción de visualización gráfica) y de análisis audio disponibles en TimeSide son accesibles con una cuenta usuario (versión 2020):

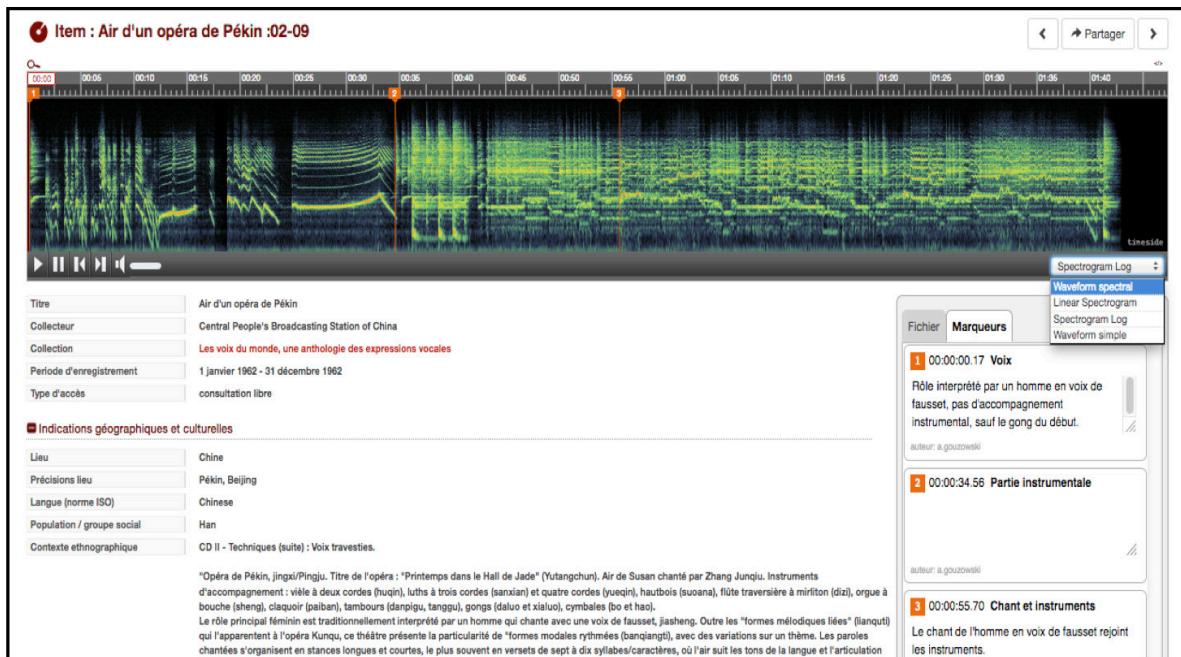
- Waveform spectral: forma de onda espectral
- Pitch: visualización de alturas relativas (utilizable en archivos cortos)
- Audio Silence: detección de los silencios audio
- Dissonance: cálculo de la disonancia armónica en función del tiempo
- Constant Q Transform: otra fórmula para visualizar el espectro
- Loudness: cálculo de la energía en función del tiempo
- Spectrogram Lin: espectrograma (tiempo / frecuencia, escala lineal)
- Spectrogram Log: espectrograma (tiempo / frecuencia, escala logarítmica)
- Onset detection: detección de los ataques, utilizables en archivos cortos
- Waveform simple: forma de onda, color negro
- Contour black: contorno de forma de onda, color negro
- Contour white: contorno de forma de onda, color blanco
- Waveform transparent: forma de onda



6. FUNCIONALIDADES DE GESTIÓN DE MEDIOS/MEDIA (SONIDO Y VIDEO)

6.3.2 Sin una cuenta de usuario

Las herramientas de visualización accesibles sin cuenta de usuario en el menú del reproductor son las siguientes: waveform spectral, spectrogram log, spectrogram lin y waveform simple.



6.3.3 Los algoritmos del ANR DIADEMS²⁶:

Las herramientas de detección (algoritmos de análisis audio) implementados al momento del ARN Diadems (2013-2016) están en espera de reintegración en el órgano de producción (abril 2020). Son visibles a partir del 2020 con una cuenta usuario en el órgano Telemeta : <http://diadems.telemeta.org>

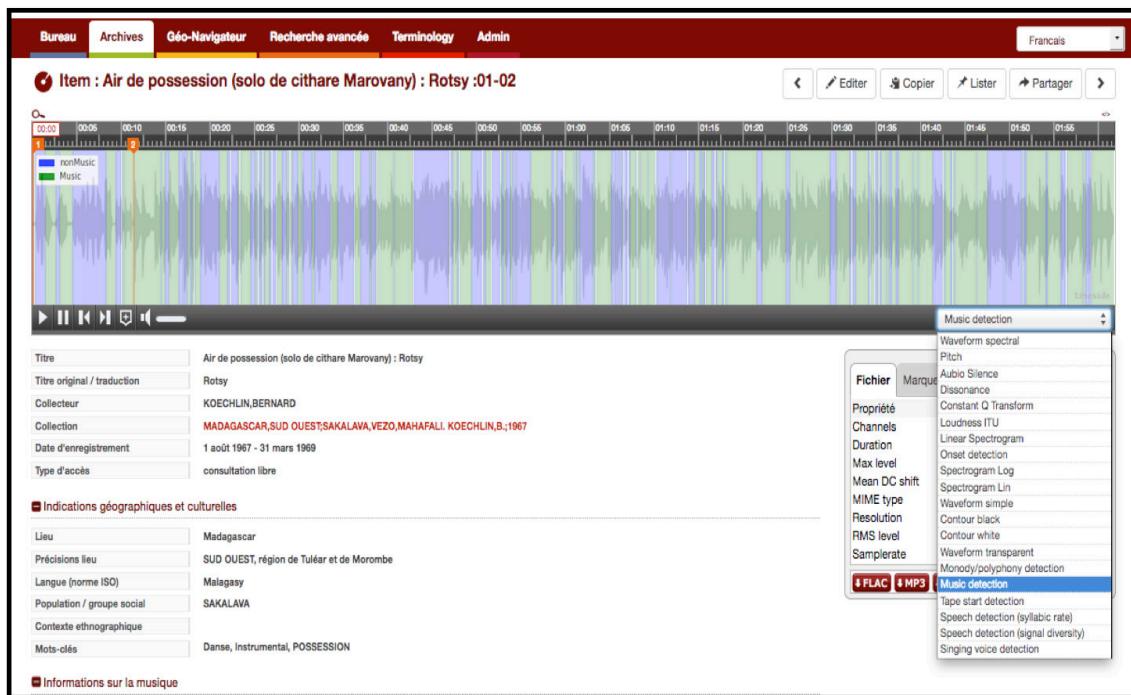
Esta instancia contiene 16,000 archivos.

6. FUNCIONALIDADES DE GESTIÓN DE MEDIOS/MEDIA (SONIDO Y VIDEO)

El siguiente cuadro define los algoritmos desarrollados:

Nombre	Definición
<i>Monody / Polyphony detection</i>	Mono / polifónico
<i>Music detection</i>	Detección de zonas de música
<i>Tape start detection</i>	Arranque de cinta
<i>Speech detection (syllabique rate)</i>	Detección palabra / no-palabra
<i>Speech detection (signal diversity)</i>	Detección palabra / no-palabra (según contexto, variación, etc.).
<i>Singing voice detection</i>	Detección voz cantada

La visualización gráfica de los análisis se presenta en el reproductor con un juego de colores en sobreimpresión sobre la forma de onda y los marcadores de anotación manual. A continuación se ve un ejemplo del resultado del algoritmo “Music detection” proporcionado por el LaBRI dentro del marco del ANR DIAMENS:



Reconocimiento de música/no música, menú desplegable, instancia de Diadems.

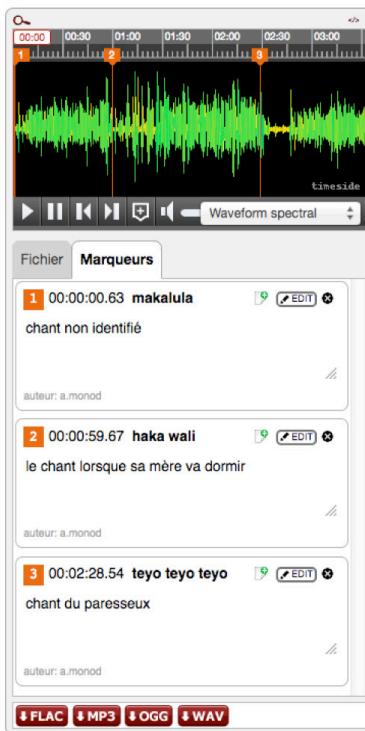
6. FUNCIONALIDADES DE GESTIÓN DE MEDIOS/MEDIA (SONIDO Y VIDEO)

6.4 Anotación temporal

Los usuarios con una cuenta pueden hacer anotaciones temporales en el medio de su elección. Esta funcionalidad propuesta por TimeSide fue concebida para resolver los problemas de indexación de los datos de campo que no se pueden segmentar debido a que a veces hay que conservar la continuidad de un evento ritual.

Por otra parte, el archivista no dispone siempre de las informaciones culturales y etnográficas o no tiene las competencias lingüísticas para identificar y segmentar las secuencias de un archivo de campo. La plataforma permite entonces, colocar marcadores sobre las representaciones gráficas del sonido gracias al botón Marcador (bandera con una cruz) situado al lado de la bocina para anotar el contenido sonoro. El usuario puede poner el marcador en tiempo real durante la lectura o colocarlo (y moverlo con un clic-jalar sencillo). La anotación la puede realizar el archivista en el momento en que carga en línea el audio, cuando el investigador lo revisa o por un usuario con las competencias culturales o musicales necesarias.

Una ventana de color naranja se abre entonces en la pestaña “Marcadores” con la lista de los marcadores, como se observa en la siguiente imagen.



El usuario puede capturar un título, un comentario, grabar o suprimir su marcador. El Time Code aparece y también el nombre de su cuenta de usuario para poder “firmar” cada marcador.

6. FUNCIONALIDADES DE GESTIÓN DE MEDIOS/MEDIA (SONIDO Y VIDEO)

Durante la reproducción aparecen un cursor en rojo y el número del marcador. El título y el comentario se visualizan en sobreimpresión durante la lectura o pasando el mouse.

- El botón “+” permite agregar el marcador a una “Lista de reproducción”
- El botón “Editar” permite modificar su marcador como usuario o administrador.
- El botón “x” permite la supresión del marcador y de esta manera se incrementa la numeración de los marcadores automáticamente. Nadie más que el usuario o un administrador puede suprimir un marcador.

Todo el público puede ver los marcadores y varios usuarios pueden anotar el mismo documento. Sin embargo, es raro que varios usuarios realicen anotaciones en el mismo documento.

6.5 Ilustración de registros con una imagen o una liga externa

Es posible integrar una o varias imágenes para ilustrar un registro en cualquier momento con una cuenta usuario. Por otro lado, a veces algunos metadatos aparecen sobre el soporte físico del archivo y la incorporación de una imagen es valiosa para identificar su contenido durante la búsqueda.

Por lo tanto, cada nivel del catálogo se puede ilustrar con un documento imagen o por otra fuente web (liga URL) gracias al menú “Medios asociados” en el formulario de captura. Este menú es accesible en modo “Edición” (Ver el botón “Editar”).

Durante la captura se puede describir una imagen o la fuente externa con título, descripción, derechos y descargar un documento imagen (PDF, JPG y PNG).

Se puede acceder a los “Medios asociados” en una viñeta de previsualización (imágenes en formato JPEG) o en descarga (con o sin cuenta usuario).

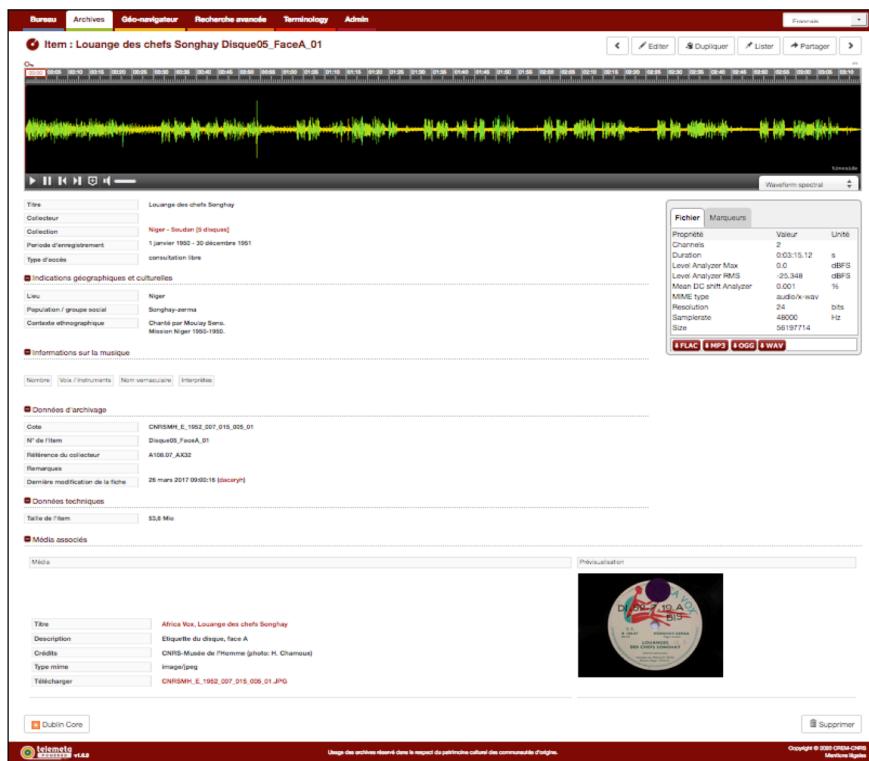
Ejemplo de registro Fondo de Mireille Hellfer:

The screenshot shows a detailed catalog record for the 'Fonds Mireille Hellfer (1928-)'. At the top, there's a navigation bar with links like 'Bureau', 'Archives', 'Géo-navigateur', 'Recherche avancée', 'Terminology', and 'Admin'. Below the title, there's a brief description of the collection: 'Fonds Mireille Hellfer (1928-)'. It includes fields for 'Titre', 'Description', 'Créateur', 'Date totale disponible', and 'Dernière modification de la fiche'. A 'Média associés' section is present, with a 'Médiathèque' link. The main content area contains a large amount of descriptive text in French, detailing Hellfer's work in India and her research interests. There are also sections for 'Photo' (with a thumbnail of Hellfer), 'Extrait de l'article' (with a link to a PDF), and 'Bibliographie' (with a link to a PDF). The entire page has a clean, professional layout with a white background and black text.

6. FUNCIONALIDADES DE GESTIÓN DE MEDIOS/MEDIA (SONIDO Y VIDEO)

Ejemplo de registro ítem de un disco Africa Vox (ilustración con la cara del disco):

CUADRO p. 48



6.6 Reutilización de un medio

La reutilización de un medio es la parte más delicada de los principios “FAIR” para los datos antropológicos que se encuentran muy raras veces se publican en línea. Los datos de los fondos de acceso abierto son reutilizables de varias maneras.

6.6.1 Enlace “i-frame”

Por medio de el enlace “i-frame” propuesto en el reproductor de TimeSide, es posible encapsular el reproductor y el medio seleccionado en otro sitio web, sin necesidad de descargarlo. La plusvalía científica y técnica de este tipo de reutilización es triple:

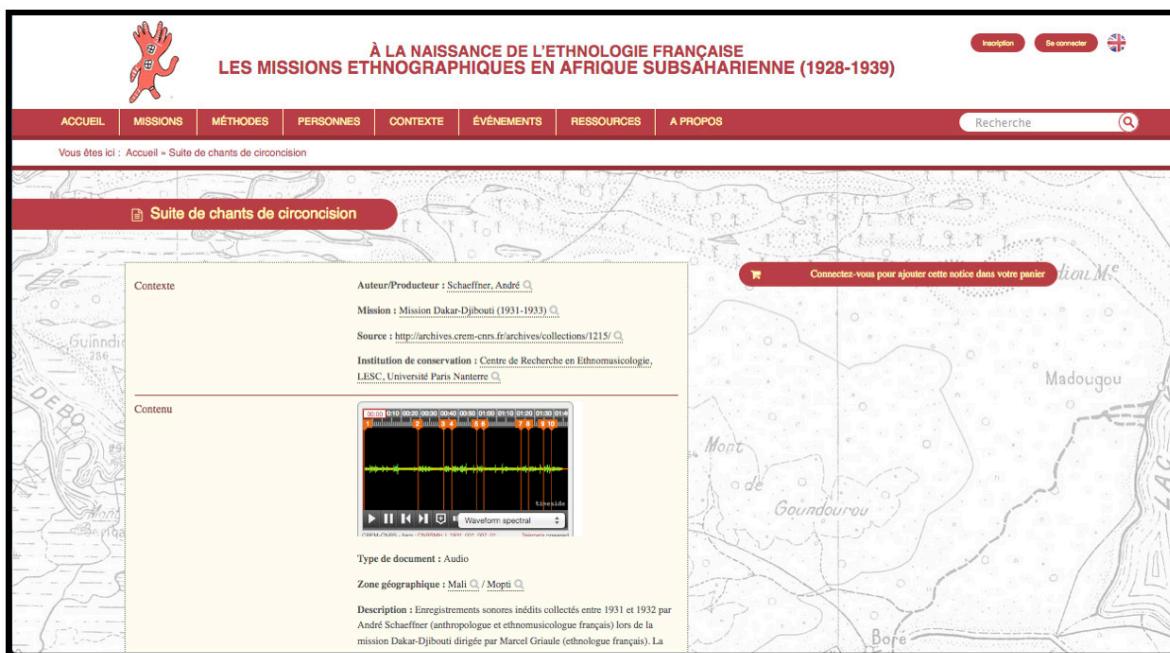
- Los medios son reproducibles y los marcadores visibles
- Las funcionalidades de navegación del reproductor son utilizables
- El acceso a la documentación completa es posible a través del enlace incluido en el lector.

6. FUNCIONALIDADES DE GESTIÓN DE MEDIOS/MEDIA (SONIDO Y VIDEO)

El usuario evita entonces la descarga abusiva de sonidos gracias a estas múltiples ventajas. Éstas se complementan con el clásico enlace URL. Los administradores de los archivos conservan de este modo la trazabilidad del uso del medio con herramientas de estadísticas tales como Google Analytics o Matomo de Huma-Num.

Se pueden ver ejemplos de reutilización en el sitio “Al nacimiento de la etnología francesa”:

<http://naissanceethnologie.fr/items/show/20351>



El sitio del ANR *Islamic Learning in Morocco*²⁷ utiliza grabaciones de Anis Fariji depositados en la plataforma Telemeta. Desde el sitio de esta ANR ILM (2017-2020) 3,000 usuarios accedieron a sus grabaciones.

La mayor cantidad de conexiones por “i-frame” proviene del sitio de Gabón, Dungandzi (cuyo propósito es promover el idioma yipunu) con 18,000 visitas en tres años: <http://bajag-mujabitsi.blogspot.com/>

27

ANR Islamic Learning in Morocco http://anrlim.cnrs.fr/presentation_anr_ilm/

6. FUNCIONALIDADES DE GESTIÓN DE MEDIOS/MEDIA (SONIDO Y VIDEO)

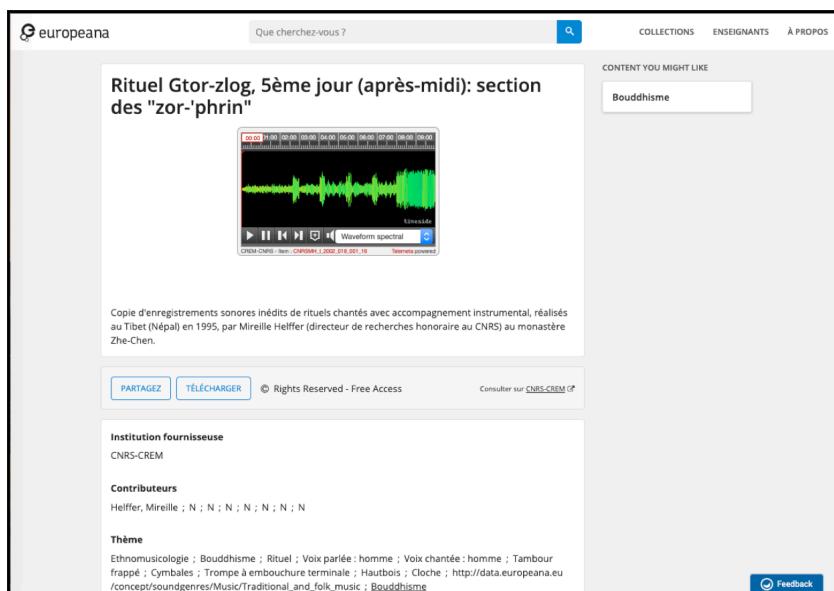
En el nuevo sitio web Watau del proyecto SAWA “Conocimiento indígena Wayana-apalai (Guyana)” también se utiliza el reproductor TimeSide de los archivos del CREM.

(<https://watau.fr/s/watau-fra/page/accueil>)



El reproductor de TimeSide también es manejado por Europeana:

http://www.europeana.eu/fr/item/2059208/data_sounds_http_archives_crem_cnrs_fr_archives_items_10296



6. FUNCIONALIDADES DE GESTIÓN DE MEDIOS/MEDIA (SONIDO Y VIDEO)

Gracias a esta funcionalidad sitios web del mundo de la investigación y el público en general, a través de redes sociales utilizan ampliamente estos archivos audiovisuales (30% de las conexiones) en proporciones mucho más superiores a las referenciadas en las plataformas de cosecha (Isidore o Europeana) con 1% o 2% de conexiones. Las conexiones procedentes de las redes sociales, muy mayoritarias, se dividen así: 65% de blogs, 33% de Facebook.

6.6.2 Libro electrónico (export EPUB)

Esta funcionalidad permite exportar una colección completa bajo la forma de un libro electrónico (e-book) para consultar los contenidos en una tableta o un lector electrónico. El botón “EPUB” se localiza arriba, a la derecha del registro

Colección. Esta función fue desarrollada para el libro de Michèle Castellengo con el fin de escuchar los ejemplos sonoros durante la lectura: “Écoute musicale et acoustique, avec 420 sons et leurs sonagrammes décryptés”.²⁸ Se pueden descargar los ejemplos sonoros desde la plataforma Telemeta del LAM. Esta funcionalidad no ha sido explotada, según tenemos entendido, para los datos gestionados por el CREM.

28

<https://www.editions-eyrolles.com/Livre/9782212138726/ecoute-musicale-et-acoustique>

7. Varios

→ Consigna para la gestión de la geolocalización con GeoEthno

Las recomendaciones de la Sociedad Samalyse en cuanto a las informaciones de los lugares fueron las siguientes en 2009:

En la base 4D, las informaciones geográficas estaban contenidas en la tabla de *ítems*. Los antiguos campos Continente, Estado, y Región_Pueblo indicaban la zona de cobertura del ítem, los dos primeros de esos campos eran enumeración en el 4D. Para modernizar la gestión de las informaciones geográficas, fundamentales en la base de datos del CREM y así permitir evoluciones interesantes más o menos a largo plazo (geolocalización, cálculo de distancias, representación en un mapa, etc.) prevaleció la utilización de un tesauro geográfico. De esta manera, se necesita un solo campo ‘Lugar’ en la tabla de *ítems*, este campo hace referencia a un elemento del tesauro que es la arborescencia de lugares . El tesauro escogido es GeoEthno, mantenido por el LESC en la Maison de l’Archéologie et de l’Ethnologie de Nanterre, del cual depende el CREM. GeoEthno está disponible en un formato de datos HML.

Estas tablas incluyen también unos indicadores booleanos (*is_authoritative*) para permitir a los usuarios, en cuanto a la concepción de la nueva aplicación del CREM d introducir nuevos nombres de lugares si no existían en el tesauro GeoEthno. Luego, periódicamente, los responsables del CREM y de GeoEthno se reunirán para decidir cuáles de estos nuevos lugares podrán integrar oficialmente GeoEthno. Gracias a los llamados indicadores booleanos se podrán encontrar lugares creados en el CREM pero que todavía no son partes de GeoEthno.

Cabe subrayar que la utilización de un tesauro evitará errores de captura (como marcar un país en un continente equivocado, un pueblo en una región equivocada,etc.) y permitirá una normalización de las regiones y pueblos. Además, el tesauro tendrá que ser actualizado regularmente en la base del CREM, para incorporar los eventuales cambios de nombre que surjan por razones políticas u otras. En el momento de la migración de la base será muy simple determinar el nombre del lugar (término descriptor GeoEthno) de los países.

En cuanto a las regiones y pueblos (campo antiguo), es difícil determinar el término descriptor GeoEthno, por lo tanto, se recomienda migrar esta información en crudo en el campo ‘Precisión lugar’. Sin embargo, se intentará poner en correspondencia el campo Región_Pueblo de la tabla Fono y esta se considerará exitosa si el término exacto se encuentra una sola vez en GeoEthno. Luego, y a medida del mantenimiento de la nueva base,los documentalistas podrán seleccionar términos descriptores GeoEthno más precisos, las informaciones del campo ‘Precisión lugar’, lo que lo vuelve poco a poco inútil.» (Ver O. Guilyardi, Docref-1.1-stable, en Mycore).

7. VARIOS

→ Modificación de contenidos textuales (página de inicio, ayuda, menciones legales)

Los cambios de los contenidos textuales (página de inicio, ayuda, menciones legales, etc.) se hacen en Github en la rama CREM, versión de prueba, menú “páginas”:

<https://github.com/Parisson/Telemeta-pages/blob/crem/fr/home.rst>

La nueva versión se controla en la instancia de prueba (contenido y paginación) antes de que se integre en la versión de producción.

→ Cuestiones jurídicas y éticas

En la gran mayoría de los casos, no hay un autor o un compositor declarado o conocido en música tradicional. A veces los nombres de los intérpretes son conocidos pero las personas son de difícil acceso. De acuerdo con los textos de la jurista Anne-Laure Sterin²⁹ que definen el marco ético y jurídico para los archivos sonoros, las grabaciones anteriores a 1963 son consultables libremente salvo algunos casos particulares por cuestiones éticas. Después de 1963, la mayoría de los depositarios necesitan una restricción de uso sobre sus datos de campo con restricción de uso por 70 años después de la recopilación. Pero si el contenido no tiene un carácter secreto o “sensible” para las comunidades de origen, y el depositario desea difundir sus grabaciones posteriores a 1963, consideramos que la persona es la más indicada para respetar la voluntad de los intérpretes y sus descendientes. Cuando se publican en línea en acceso abierto (sin contraseña) los testimonios de las comunidades son numerosos y entusiastas, lo cual alienta a continuar en esta dirección. En caso de que suceda lo contrario, la difusión en línea se puede limitar en cualquier momento, lo cual no es posible en el caso de un disco. En ausencia de explotación comercial y por lo tanto de un eventual perjuicio financiero para los derechohabientes, los riesgos jurídicos son inexistentes para la institución, según juristas especializados. Además, los archivos en alta definición nunca son accesibles, a diferencia de un CD, lo cual limita las explotaciones comerciales.

29

<http://ethiquedroit.hypotheses.org/1135>

8. Conclusión

Ante la falta de una herramienta web adaptada en nuestras instituciones y de medios informáticos internos, la colaboración primero con Parisson y después con el IRCAM fue una oportunidad inesperada para valorar un fondo de archivos sonoros de una gran riqueza, pero poco accesible³⁰. Hay que rendir homenaje a los numerosos colaboradores y contribuyentes que participaron en este proyecto, de manera progresiva (o “ágil”), a veces de benévolos y motivados por su dinámica innovadora. También hay que subrayar el impulso que dio el TGE Adonis³¹ (convertido en Huma-Num) y el apoyo del Ministerio de la Cultura.

La plataforma colaborativa Telemeta corresponde a las necesidades definidas en 2009 por el equipo del CREM, en coherencia con los medios financieros de una pequeña estructura de investigación en cuanto a la gestión, la difusión y la anotación de datos de audio y video. Esta interfaz multiusuario respeta el modelo de datos específicos de la etnomusicología y las necesidades particulares en materia de gestión de acceso. La gestión de la restricción de acceso sobre los datos alienta a numerosos investigadores a depositar sus recopilaciones, a pesar de que estos datos no pueden ser consultados por el público en general, lo cual es a menudo el caso.

El modelo de los datos definido y enriquecido poco a poco, desde la creación del equipo del Museo del Hombre, sigue las necesidades de la disciplina. Podrían simplificarse (algunos campos son poco utilizados) las listas de vocabularios pueden mejorar internamente y unas cuantas pueden estar visibles al exterior, al considerar el balance beneficio-restricciones de depender de tesoros externos.

El rendimiento de la herramienta es satisfactorio si se toma en cuenta el volumen de datos en línea. La implicación de los usuarios es eficiente, especialistas o público amplio (con un promedio de 350,000 páginas visualizadas por año y 48,000 visitas por año). La velocidad de acceso es limitada por la velocidad del servidor Huma-Num (el nuevo sistema de almacenamiento Active Circle no está concebido para un acceso ultra-veloz) pero esta lentitud relativa se compensa por la oferta de servicios de la infraestructura (alojamiento, resguardo, mantenimiento, máquina, etc.).

La base de datos ya no está reservada a unos cuantos archivistas iniciados, el enriquecimiento de la base es constante desde 2011 (54,200 medios hoy) y hoy en día han surgido numerosos proyectos de investigaciones innovadoras. La disponibilidad y flexibilidad de los datos de investigación en una interfaz ergonómica para la reutilización de los medios significa una prueba de concepto que podría ser útil en el equilibrio entre la ergonomía y el grado de complejidad de las funcionalidades esperadas para una base de datos audiovisuales.

30 Article du Bulletin de l'AFAS: <https://journals.openedition.org/afas/4056>

31 Très Grand Équipement Adonis, INSHS, CNRS, 2007-2010

8. CONCLUSIÓN

La herramienta no tiene vocación editorial, pero se pretende que sea útil sobre todo para la preservación y, por lo tanto, esté disponible por los coleccionistas y usuarios de estos archivos. Las mejoras son posibles (licencia *open source*) y dependen de los recursos financieros y humanos disponibles. Desde 2017, se crearon propuestas que aún están en espera como son, la modularización con el fin de mutualizar funcionalidades comunes a diferentes instancias y dotar de autonomía al *framework* TimeSide. Ninguna solución técnica que incluya el conjunto de funcionalidades de esta herramienta ha sido propuesta por Huma-Num, los especialistas en este dominio no están a favor de una herramienta en lenguaje PHP. Además, la solución de almacenamiento del CNRS (Nakala) no propone la gestión de una restricción de uso sobre los medios.

El proyecto Telemeta / TimeSide suscitó otras plataformas en Francia y en el extranjero y contribuyó a dar a conocer la riqueza de este patrimonio científico y cultural del CNRS, así como la excelencia de los laboratorios de investigación en el ámbito de los Data Ciencia.

Bibliografía

- » Simonnot, Joséphine, 2008, Telemeta: un projet pour les archives sonores de la recherche. Traitement, valorisation, diffusion, en ligne: <http://www.imageson.org/document1045.html>, mise en ligne le 31 octobre 2008. - Simonnot, Joséphine, 2008, Telemeta: un projet d'application collaboratif dédiée à l'archivage des documents sonores, La lettre d'information du programme ARSHS,http://mshdijon.ubourgogne.fr/msh_cnr/arshs/Lettre_arshs_2.html, mise en ligne en novembre 2008.
- » Simonnot, Joséphine, 2011, Telemeta, un projet Web pour les archives sonores de la recherche, Lettre de l'INSHS, en ligne: <http://www.cnrs.fr/inshs/recherche/telemeta2.htm>, mise en ligne le 24 février 2011
- » Fillon Thomas, Simonnot Joséphine, Mifune Marie-France, Khoury Stéphanie, Pellerin Guillaume, Lecozi Maxime, Amy de La Bretèque Estelle, Doukhan David, Fourer Dominique, Rouas Jean-Luc, Pinquier Julien, Mauclair Julie and Barras Claude, 2014. "Telemeta: An open-source web framework for ethnomusicological audio archives management and automatic analysis", working paper for the Digital Library for Musicology workshop, London, Septembre 2014 - Simonnot Joséphine, Khoury Stéphanie, "Applications and implications of digital audio databases for the field of ethnomusicology: A discussion of the CNRS — Musée de l'Homme sound archives", First Monday, Volume 19, Number 10 - 6 octobre 2014 - <http://firstmonday.org/ojs/index.php/fm/article/view/5554/4133>
- » SIMONNOT, Joséphine, "The CNRS - Musée de l'Homme Sound Archives from 1900 to the present: a long way between heritage, knowledge and technologies", in ZIEGLER, Susanne, Ingrid AKESSON et LECHLEITNER, Gerda (eds.), Historical Sources of ethnomusicology in contemporary debate, Cambridge Scholars Publishing (Cambridge, s.n.): 103-116,2017 en ligne: <http://www.cambridgescholars.com/historical-sources-ofethnomusicology-in-contemporary-debate>
- » SPIELMANN, Florabelle, HELMLINGER Aurélie, FILLON Thomas, PELLERIN Guillaume, STURM Bob L., BEN-TAL Oded, QUINTON Elio et SIMONNOT Joséphine 2017, Zoom arrière. L'ethnomusicologie à l'ère du Big Data, Cahiers d'ethnomusicologie. Anciennement Cahiers de musiques traditionnelles, 30: 9-28.
- » Simonnot Joséphine, «Partager les archives sonores du musée de l'Homme sur le web avec la plateforme Telemeta », Bulletin de l'AFAS [En ligne], 46 | 2020, mis en ligne le 14 mars 2020, consulté le 07 mai 2020. URL: <http://journals.openedition.org/afas/4056>
- » Marie-France Mifune et Florabelle Spielmann, «Les archives sonores ethnomusicologiques du CNRSmusée de l'Homme. Quelles implications dans des projets d'indexation et d'analyse du contenu audio», Bulletin de l'AFAS [En ligne], 46 | 2020, mis en ligne le 14 mars 2020, consulté le 07 mai 2020. URL: <http://journals.openedition.org/afas/4070>; DOI: <https://doi.org/10.4000/afas.4070>

Pedro Sergio Montes de Oca Quiroz
Corrección de estilo

Andreina Villaseñor Espinosa
Diseño editorial