



HAL
open science

Partager les archives sonores du musée de l'Homme sur le web avec la plateforme Telemeta

Josephine Simonnot

► To cite this version:

Josephine Simonnot. Partager les archives sonores du musée de l'Homme sur le web avec la plateforme Telemeta. *Sonorités*, 2020, Les archives de l'ethnomusicologie en France : le cas du MNATP et du MH, 46, pp.88-101. 10.4000/afas.4056 . hal-02860091

HAL Id: hal-02860091

<https://hal.science/hal-02860091v1>

Submitted on 16 Oct 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

N° 46

AVRIL

2020

Éditorial

Florence Descamps

DOSSIER LES ARCHIVES DE L'ETHNO- MUSICOLOGIE EN FRANCE : LE CAS DU MNATP ET DU MH

Marie-Barbara Le Gonidec, Martine Sin Blima-Barru
Pascal Riviale, Aude Julien Da Cruz Lima,
Joséphine Simonnot, Marie-France Mifune,
Florabelle Spielmann, François Gasnault,
Florence Neveux, Marie-Charlotte Calafat, Thomas Henry

PORTRAIT

*Du musée instrumental du Conservatoire
de Paris au musée des Arts et Traditions
populaires*

Entretien avec Florence Gétreau

Véronique Ginouvès, Florence Descamps

RECHERCHE

*Oral History / Oral Sources - Polish
Historians' Dilemmas*

Marta Kurkowska-Budzan, Marcin Stasiak

SONORITÉS

Éditorial

*Pourquoi consacrer un dossier thématique
aux archives de l'ethnomusicologie ?*

Florence Descamps

4-6

DOSSIER
LES ARCHIVES DE L'ETHNO-
MUSICOLOGIE EN FRANCE :
LE CAS DU MNATP ET DU MH

*Que sont les archives devenues ?
Actualité des fonds de l'ethnomusicologie
institutionnelle en France*

*Du musée des Arts et Traditions populaires
aux Archives nationales, du musée de l'Homme
à l'université de Nanterre*

Marie-Barbara Le Gonidec

8-14

LES PHONOGRAMMES
DANS LES MUSÉES ETHNOGRAPHIQUES,
QUESTIONS PATRIMONIALES
ET JURIDIQUES : LE CAS DU MNATP

*Les fonds sonores de l'ancien musée
national des Arts et Traditions populaires :
collections ou archives ?*

Marie-Barbara Le Gonidec

16-28

*Les enregistrements sonores du musée
national des Arts et Traditions populaires
aux Archives nationales*

Questions juridiques d'hier et d'aujourd'hui

Martine Sin Blima-Barru

30-43

SONORITÉS

**VERSEMENT ET TRAITEMENT DES FONDS :
LES RESSOURCES DEVIENNENT-ELLES
NÉCESSAIREMENT DES ARCHIVES ?**

***Des Sablons à Pierrefitte, du Musée
aux Archives***

*Versement et traitement des archives
ethnomusicologiques du musée national
des Arts et Traditions populaires*

Pascal Riviale 46-58

***Les « archives sonores » du CNRS
et du musée de l'Homme***

*Retour sur un ensemble documentaire
emblématique et enjeux actuels*

Aude Julien Da Cruz Lima 60-85

**PROJETS DE VALORISATION :
LES NOUVELLES TECHNOLOGIES,
UN « UNIVERS » DE POSSIBLES**

***Partager les archives sonores du musée
de l'Homme sur le web avec la plateforme
Telemeta***

Joséphine Simonnot 88-101

***Les archives sonores ethnomusicologiques
du CNRS-musée de l'Homme***

*Quelles implications dans des projets
d'indexation et d'analyse du contenu audio*

Marie-France Mifune, Florabelle Spielmann 102-112

***Du « corpus des musiques ethniques
de la France » au projet Les Réveillées
Éditer les enquêtes ethnomusicologiques
de la « maison Rivière »***

François Gasnault 114-120

***Conception et mise en œuvre
de la plateforme du projet Les Réveillées
De la valorisation de données de recherche
à l'édition de contenus***

Florence Neveux 122-129

LES COLLECTIONS PHONOGRAPHIQUES
DES DEUX ANCIENS MUSÉES :
DES COMPLÉMENTS INDISPENSABLES

*La collection de disques édités du Mucem,
un bel héritage aujourd'hui valorisé*

Marie-Charlotte Calafat 132-138

*Le fonds de catalogues de marques
phonographiques du CNRS-musée
de l'Homme*

Thomas Henry 140-145

Bibliographie générale du dossier 146-152

PORTRAIT

*Du musée instrumental du Conservatoire
de Paris au musée des Arts et Traditions
populaires*

Entretien avec Florence Gétreau

Véronique Ginouvès, Florence Descamps 154-165

RECHERCHE

*Oral History / Oral Sources - Polish
Historians' Dilemmas*

Marta Kurkowska-Budzan, Marcin Stasiak 168-179

Auteur·e·s 182-188

Partager les archives sonores du musée de l'Homme sur le web avec la plateforme Telemeta

Joséphine Simonnot

Ingénieure de recherche CNRS

Depuis 2011, le CREM, centre de recherche en Ethnomusicologie (<http://crem-cnrs.fr>) basé à l'université de Paris Nanterre, équipe du laboratoire d'Ethnologie et de Sociologie comparative du CNRS, diffuse en ligne les archives sonores de l'ancien musée de l'Homme grâce à une plateforme web intitulée Telemeta. Joséphine Simonnot qui en est la principale instigatrice, explique la genèse de la plateforme qui huit ans plus tard, ne cesse d'évoluer et de s'enrichir.

Since 2011, the "Centre for Research in Ethnomusicology", attached to the Laboratory of Ethnology and Comparative Sociology of the University of Paris Nanterre (<http://crem-cnrs.fr>) broadcasts online the sound archives of the former Museum of Man (musée de l'Homme, Paris) thanks to a web platform called Telemeta. Joséphine Simonnot, who is the main instigator of Telemeta, explains the genesis of the platform eight years later and how it continues to evolve and develop.

Sauver les archives sonores du musée de l'Homme

Depuis les années 2010, le partage des données audiovisuelles de la recherche est une idée qui a progressé notablement dans les esprits, mais au début de ce siècle, c'était encore un sujet sensible parmi les ethnomusicologues. Les débats sur la propriété intellectuelle des données collectées étaient très vifs¹. Les chercheurs étaient recrutés pour faire des recherches, publier des résultats mais, pour la majorité d'entre eux, ils ne se considéraient pas investis d'une mission de service public pour la diffusion des données qu'ils collectaient. C'est ainsi que les choses m'ont été présentées à mon arrivée au musée de l'Homme en 1999². Quant à pérenniser ces archives collectées pendant près de 100 ans dans le monde entier, les moyens humains et financiers n'étaient pas prévus dans les budgets du laboratoire à cette période. Le CNRS n'avait pas non plus de programme de numérisation. À quoi bon numériser des milliers de vieilles bandes, représentant à l'époque environ 5 000 supports analogiques (fig. 1, en ligne) alors que les meilleurs morceaux avaient été publiés en disque depuis l'après-guerre et largement commentés dans divers ouvrages ?

De fait, il faut reconnaître que l'apport des publications de disques fut indéniable pour la diffusion des connaissances des musiques de l'humanité. L'importante production discographique du musée de l'Homme de 1947 à 2001, aujourd'hui accessible en ligne³, témoigne de l'investissement de Gilbert Rouget⁴. Le disque a été longtemps le seul moyen de faire entendre au grand public ce patrimoine musical, en dehors des concerts bien entendu. Mais, lorsque la production de CD a cessé d'être viable économiquement pour l'éditeur *Le Chant du Monde*, notre préoccupation a été de trouver un nouveau moyen d'illustrer les publications scientifiques avec des enregistrements de terrain, de donner accès aux archives musicales du musée de l'Homme, de faciliter l'étude de ces documents mais surtout, d'assurer la survie de ce patrimoine en danger d'obsolescence et de dégradation. La mission initiale fixée par André Schaeffner⁵ dans les années 1930, reprise par Gilbert Rouget après la guerre, était d'étudier et de conserver ces musiques en voie de disparition et de les transmettre aux générations futures. Mais les temps avaient changé et à la fin des années 1960, le département d'ethnologie musicale

1 Voir à ce sujet le programme des journées de la Société Française d'Ethnomusicologie en 2009 et leur captation. <http://www.ethnomusicologie.fr/activites/jetu/jetu-2009-video/>

2 J'interviens dans cet article en tant qu'ingénieur de recherche au CNRS responsable des productions audiovisuelles, des techniques de l'information et de la communication au CREM-LESC. J'ai publié au tout début du projet Telemeta un article dans le *Bulletin de l'AFAS. Sonorités* (n° 36, 2011), une présentation de cette plateforme [<https://journals.openedition.org/afas/2621>] qui permet d'évaluer le chemin parcouru.

3 http://archives.crem-cnrs.fr/archives/fonds/CNRSMH_Editions/

4 Gilbert Rouget (1916-2017) était un ethnomusicologue français, directeur de recherche et fondateur de l'équipe de recherche CNRS.

5 André Schaeffner (1895-1980), anthropologue et ethnomusicologue français, directeur du Département d'ethnomusicologie du Musée de l'Homme (1929-1965).

du musée était devenu, sous l'impulsion de Rouget, le laboratoire d'Ethnomusicologie du CNRS, avec une approche plus anthropologique que muséographique. Or, l'anthropologie de la musique étudie avant tout les aspects socioculturels des pratiques musicales, elle ne s'intéresse pas à la conservation des données elles-mêmes. L'ethnomusicologie qui s'intéresse à « l'œuvre plutôt qu'à l'ouvrier qui la rend possible », selon les mots de Bernard Lortat-Jacob et Miriam Roving Olsen⁶, fait débat. Dans les années 1970, les travaux sur les collections d'instruments et les archives sonores, activités phares du département à l'époque de Schaeffner et Rouget deviennent secondaires, comme si la « gestion » de ce patrimoine, presque suspecté de néocolonialisme, n'était plus du ressort de la nouvelle génération de chercheurs. D'ailleurs, les collections d'instruments de musique restent sous la responsabilité du MNHN, le muséum national d'histoire naturelle, avec le recrutement de Geneviève Dournon⁷ en 1967 puis de Lucie Rault⁸ en 1993, marquant ainsi la répartition des « tâches » entre l'équipe de recherche et le département du musée. Une nouvelle étape est franchie au début des années 2000 lorsque la tutelle historique du MNHN sur l'équipe de recherche CNRS prend fin. La mission de conservation des archives sonores, prérogative du MNHN, se trouve donc en suspens et le numérique est surtout apprécié pour l'apport de la vidéo devenue plus accessible et considérée comme un nouveau « carnet de terrain ». Autrement dit, en l'an 2000, la numérisation systématique de ce patrimoine n'était pas la priorité du musée ni de l'équipe de recherche.

6 Bernard Lortat-Jacob et Miriam Roving Olsen, « Musique, anthropologie : la conjonction nécessaire », *L'Homme*, 2004, 3-4, p. 11.

7 Geneviève Dournon, née en 1928, a été ethnomusicologue du CNRS, conservateur de musée et chargée des collections d'instruments de musique du département d'ethnomusicologie du musée de l'Homme.

8 Sinologue et ethnomusicologue, Lucie Rault a été maîtresse de conférences au muséum national d'histoire naturelle de Paris

9 Sur Florence Gétéreau, consulter le portrait qui lui est dédié dans ce numéro du *Bulletin de l'AFAS. Sonorités* 46.

10 Jean-Pierre Dalbéra a été chef de la Mission de la recherche et de la technologie au Ministère de la culture et de la communication, Paris entre 1990 et 2004 puis en poste au MNATP (Paris) en 2004.

À la même époque cependant, le ministère de la Culture prône la numérisation des données de la recherche, au même titre que celle des collections des musées. C'est ainsi que les fonds sonores du musée national des arts et traditions populaires, les plus anciens d'abord pour des raisons de sauvegarde, ont bénéficié du plan de numérisation lancé par la mission recherche et technologie (MRT). L'opération a été mise en place sous la houlette avisée de Florence Gétéreau⁹, alors conservatrice de la phonothèque, ce qui a permis d'envoyer les collections dans les régions correspondantes. Le responsable de la MRT, Jean-Pierre Dalbéra¹⁰, annonce alors clairement les enjeux : les financements de la numérisation doivent s'inscrire dans le cadre de projets d'une diffusion au public, totale ou partielle. Et c'est bien cette notion de diffusion au public qui nous a interpellés : cette incitation de la MRT fut décisive au cours des années suivantes. L'accord-cadre CNRS-ministère de la Culture mis en place à cette époque fut crucial pour la sauvegarde et la valorisation de ce patrimoine vers un public plus large, grâce à un soutien financier récurrent.

Méthodologiquement, toute mise à disposition commence d'abord par la sauvegarde et donc par la numérisation des anciens supports car les machines de lecture se font rares et sont de plus en plus difficiles à réparer. Il faut également rappeler qu'un document numérisé

est plus facile à consulter et ne risque aucune dégradation à la lecture. La duplication est également beaucoup plus rapide et sans perte de qualité, contrairement aux supports analogiques. Lors de mon recrutement dans l'équipe en 1999 comme responsable de l'audiovisuel, j'étais fascinée par toutes ces bandes magnétiques accumulées pendant tant d'années. Ayant enregistré moi-même sur le terrain dans l'est indonésien¹¹ en 1992 et 1993, j'étais consciente de la difficulté de la collecte sur le terrain, surtout à l'époque pionnière et « héroïque » de l'enregistrement analogique sur des supports tels que le cylindre, le disque puis la bande. C'est pourquoi un fonds d'une telle richesse est si rare aujourd'hui. Préoccupée par la fragilité des bandes dont les caractéristiques physiques et magnétiques se dégradent au fil du temps, j'ai commencé la numérisation de certaines collections emblématiques telles que celles de Gilbert Rouget (fig. 2, en ligne), Mireille Helffer¹² ou Hugo Zemp¹³. Je traitais également l'archivage des nouveaux dépôts, numériques ou analogiques, soit environ une quinzaine de collections par an entre 1999 et 2009. Un rapide calcul me fait entrevoir qu'une trentaine d'années à temps plein me serait nécessaire pour numériser et segmenter toutes les bandes des collections du musée de l'Homme.

Au début des années 2000, une solution providentielle de numérisation apparaît grâce au musée du quai Branly (MQB) : le nouvel établissement propose la numérisation des archives sonores si l'équipe CNRS rejoint ses rangs. Mais à cette période, il régnait au musée de l'Homme une opposition farouche au départ des collections ethnographiques vers Branly. Contrairement au transfert des instruments de musique au MQB, les bandes magnétiques étaient restées dans leurs placards, entre les squelettes du département d'anthropologie biologique et les pointes de flèches du département de préhistoire, nos voisins au musée de l'Homme. Quelques années plus tard, trois cents bandes magnétiques bénéficient du programme de numérisation du ministère de la Culture, ce qui a constitué un premier pas dans l'externalisation de cet énorme chantier. À partir de 2005, les budgets de numérisation de la MRT sont réservés aux langues de France et le laboratoire ne peut plus en bénéficier mais heureusement l'accord-cadre CNRS-ministère de la Culture a pris le relais. En 2009, le musée de l'Homme est fermé pour rénovation. Les bandes, dont un quart est numérisé, sont transférées à la Bibliothèque nationale de France (BnF) à la demande du MNHN, où elles sont conservées dans un environnement climatisé. La numérisation se poursuit en partie par la BnF et en partie en interne. Dix ans plus tard, en 2019, 90 % des collectes de terrain sont numérisées, dont 30 % par la BnF. Quant à la collection de disques commerciaux (78 tours, 33 tours, 45 tours, CD, cassettes), toujours conservée par l'équipe de recherche car moins fragile, c'est une autre histoire... Environ 10 % des disques sont numérisés et consultables en ligne actuellement.

11 Collectes aux Moluques, Indonésie. http://archives.cremcnrs.fr/archives/corpus/CNRSMH_Simonnot_001/

12 Mireille Helffer est une ethnomusicologue française née en 1928, spécialiste des musiques himalayennes.

13 Ethnomusicologue franco-suisse né en 1937, Hugo Zemp est spécialiste des îles Salomon et de Côte d'Ivoire.

Vers une base de données enrichie sur le web : la création de Telemeta

La documentation d'une collection sonore, souvent faite en collaboration avec les collecteurs au moment du dépôt ou ultérieurement, est aussi importante que la numérisation du contenu sonore lui-même. La documentation associée s'enrichit progressivement depuis plus de 50 ans grâce à la participation des chercheurs, des étudiants ainsi que des retraités. L'expertise des membres du CREM et la présence d'ingénieurs spécialisés permettent de procéder, jour après jour, à l'indexation de ces contenus sonores d'origines diverses, au gré de la disponibilité des collecteurs, de la documentation disponible sur place ou à la Bibliothèque du MNHN¹⁴. Consignée dans des registres d'inventaires depuis les années 1930, la documentation est peu à peu complétée par des fiches descriptives dactylographiées indexant les enregistrements de terrain ainsi que les disques 78 tours et 33 tours (fig. 3, en ligne).

14 Voir dans ce dossier la contribution de Aude Julien Da Cruz Lima.

15 Monique Desprez était alors technicienne en documentation au CNRS jusqu'au début des années 2000.

Dans les années 1990, à la demande de la directrice du laboratoire qui est alors Mireille Helffer, une base de données informatique a été conçue spécialement pour les archives sonores sous la direction de Pribislav Pittoëff, ingénieur chargé des ressources documentaires. Tran Quang Hai, ingénieur responsable de la phonothèque, assisté de Monique Desprez¹⁵ ainsi que divers vacataires, participaient également à la saisie des données descriptives. Celle-ci s'effectuait en deux étapes : les collecteurs rédigeaient un brouillon de fiche pour chaque item sous un format de type Word, et celui-ci était saisi ensuite dans la base de données informatisée. La saisie par le collecteur lui-même était impossible en raison de la complexité de l'interface du système. Il faut préciser que les fiches descriptives des archives sonores n'étaient que partiellement informatisées : environ 20 000 fiches, niveau collections et niveau items, étaient saisies dans la base de données existante, gérée par le système propriétaire 4D avant la migration (à titre de comparaison, la nouvelle base en compte 78 000 en 2019). Lors du déménagement à Nanterre, il a fallu insister fortement auprès du MNHN pour conserver ces précieuses fiches dactylographiées qui consignaient la documentation réalisée jusqu'en 1994.

Après une tentative infructueuse de mise à jour de cette base 4D, son évolution, et donc sa survie, était compromise : la version de l'application ne fonctionnait que sous Mac OS 9 et dépendait de la durée de vie de l'ordinateur où elle était installée. Le manque d'accessibilité à la documentation par toute autre personne que les deux ingénieurs et la technicienne de l'équipe décourageaient les non-initiés cherchant à consulter les archives. D'autre part, toute ouverture en réseau était évidemment impossible. Une migration des données s'est alors avérée indispensable pour ne pas perdre les métadonnées et leur structuration complexe (tables, énumérations thématiques, liste de mots-clés...) saisies

pendant plus de quinze ans. Un poste de chargé des ressources documentaires a été accordé en 2009 par le CNRS et Aude Julien Da Cruz Lima a été recrutée (en remplacement des 3 personnes citées plus haut), ce qui a permis de renforcer le traitement documentaire des archives et de poursuivre sa modernisation.

Le besoin d'améliorer la gestion des archives sonores s'est imposé rapidement : il fallait simplifier la saisie et la consultation, tout en partageant les fruits d'une coûteuse numérisation avec un plus large public. Mais pour concevoir un nouvel outil de travail adapté, il fallait d'abord bien connaître toute la chaîne de l'archivage et ce qu'on appelle le « scénario d'usage ». L'opération de numérisation elle-même, notamment celle effectuée avec Gilbert Rouget ou Mireille Helffer, m'avait fait découvrir l'analyse documentaire que l'on pratique pour l'archivage en ethnomusicologie. L'identification du contenu musical pour la segmentation numérique des séquences audio et la captation des informations données par le collecteur se faisaient simultanément pendant le transfert des supports analogiques : notes de terrain, mémoire réveillée à l'écoute, données techniques... Certaines informations implicites sont rarement écrites et doivent être consignées à « chaud » : type d'instruments, effectifs, langue, contexte... Pendant près de 10 ans, j'ai joint ces informations sur les pochettes des CD où étaient transférées les séquences numérisées pour l'archivage (fig. 4, en ligne). Cette opération permettait au chercheur de s'y retrouver dans sa copie sur CD et constituait une « sécurité » pour conserver les informations descriptives en attendant leur saisie définitive dans la base de données. Mais après quelques années de pratique, nous nous sommes rendu compte que le décalage entre l'analyse documentaire et la saisie informatique était source d'erreur, de perte de temps et même d'oubli, à force de reporter à plus tard ce fastidieux travail de recopie du brouillon dans la base 4D.

À cette époque, il n'existait que des systèmes propriétaires pour la gestion d'archives sonores, comme par exemple la plateforme de la phonothèque de la maison méditerranéenne des sciences de l'Homme (MMSH), de l'institut national de l'audiovisuel (INA) ou de la médiathèque de la Cité de la musique (actuelle Philharmonie de Paris). Un état de l'art fut entrepris pour explorer les modèles de données descriptives utilisés pour les archives sonores à la BnF et au musée du quai Branly. Je compris rapidement que notre modèle de métadonnées était très spécifique et différent des standards en vigueur. Mais le système de description des bibliothèques n'était pas pertinent non plus pour nos données en ethnomusicologie.

C'est alors que les nombreux échanges que j'avais avec le laboratoire d'acoustique musicale de l'université de Jussieu (LAM), dépendant du CNRS, de l'université de Pierre et Marie Curie et du ministère de la Culture, ont été profitables : Hugues Genevois, alors

responsable du LAM, a proposé l'idée de réaliser une base de données accessible en ligne incluant des fichiers audio, réalisée par Guillaume Pellerin, ancien doctorant en acoustique. Un système de gestion de contenu *open source* fut alors envisagé pour exploiter la base de données et les enregistrements associés, les enrichir progressivement et les diffuser grâce à une compression automatique des fichiers *wav* et un lecteur de streaming permettant la visualisation et l'annotation dans la temporalité du document. C'était, en un mot, réaliser un rêve : avoir toutes les données dans le même outil, travailler à plusieurs et à distance. La plateforme Telemeta était née (fig. 5, en ligne).

Il faut préciser que notre équipe ne bénéficiait d'aucun support humain en informatique, ce qui est toujours le cas actuellement. Un prototype en ligne est alors mis au point progressivement avec la société Parisson¹⁶ entre 2007 et 2009, ainsi qu'une première migration de la base 4D, migration représentant déjà en soi un gros défi... Cette étape a permis de tester l'outil, mais également de procéder à la saisie des nouvelles collections et donc d'enrichir la base de données documentaires à plusieurs, sans passer par l'ancienne interface. C'était déjà une révolution dans le quotidien de l'archivage des collections.

D'une certaine manière, la création de cette interface web collaborative constitue une nouvelle étape dans le processus de diffusion et d'adaptation progressive aux nouvelles technologies depuis l'invention de l'enregistrement. Cette adaptation constante aux nouvelles techniques était chère à Gilbert Rouget, toujours curieux des nouveautés : en 1946, il avait personnellement insisté pour emporter une machine de gravure de disque lors de la Mission Ogooué-Congo en Afrique¹⁷ afin d'avoir une meilleure qualité que les cylindres proposés par le responsable Noël Ballif¹⁸.

Sur Telemeta, la visualisation de la vidéo et des données sonores, sous forme d'onde et de spectrogramme se fait grâce au moteur multimédia *TimeSide*¹⁹ qui gère les flux permettant l'affichage, la compression automatique des sons et la lecture. Il permet aussi l'annotation au moyen de marqueurs temporels (fig. 6, en ligne) et l'accès simultané aux métadonnées qui apportent un confort de travail à l'équipe tout en motivant les contributions diverses, celles des chercheurs, déposants, étudiants, collaborateurs étrangers, etc. La visibilité par la communauté scientifique et la sécurisation des données sur un serveur sont également des arguments de poids pour convaincre les chercheurs à participer à cette activité qui demeure malgré tout très chronophage. L'attractivité et la simplicité d'utilisation de l'interface sont aussi cruciales dans la réussite de ce genre d'outil dédié à la communauté SHS, non spécialisée en informatique.

16 <http://parisson.com/>

17 Exposition virtuelle sur la mission Ogooué-Congo de 1946 : <http://www.europeana.eu/portal/fr/exhibitions/1946-ogooe-congo-mission>

18 Sur l'histoire des collections et du catalogue, voir Brice Gérard, « Ethnomusicologie et archives sonores, entretiens avec Pribislav Pitoëff » et « Présentation du projet Telemeta, entretiens avec Joséphine Simonnot », *Transposition, musique et sciences sociales*, 2011, 1. <https://transposition.revues.org/112> et Gilbert Rouget, « Le Département d'ethnomusicologie du musée de l'Homme », *L'Homme*, 2004, p. 171-172, p. 513-523. <http://journals.openedition.org/lhomme/24998>

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

Le partage entre les besoins de la recherche scientifique et les défis patrimoniaux est souvent inextricable et le soutien du ministère de la Culture est, dans ce cas, pleinement justifié. Depuis la fin des années 2000, cette spécificité est prise en compte par le CNRS avec la création du TGR (très grand équipement) Adonis, devenu ensuite Huma-Num. L'exposition en ligne de données audiovisuelles de la recherche via un système normalisé et interopérable répond tout à fait aux directives du ministère de la Culture comme du CNRS (normes *Dublin Core* et OAI-PMH)²⁰. Et le fait d'être en ligne, c'est-à-dire d'offrir une nouvelle visibilité, a provoqué une ouverture sans précédent tout en créant une nouvelle dynamique qui perdure encore aujourd'hui. Mais l'investissement humain et financier dans de nouvelles technologies n'aurait pu être possible sans le soutien financier significatif apporté par l'accord-cadre entre le CNRS et le ministère de la Culture²¹. Le CNRS n'ayant pas de vocation patrimoniale à la différence du ministère, cet accord a pour vocation de soutenir la valorisation des données patrimoniales de la recherche. En 2014, cette volonté a été confirmée dans le blog *Culturelabs*²², où il est écrit que « l'innovation numérique [... est soutenue] pour développer des usages culturels numériques innovants qui s'appuient sur des contenus numérisés disponibles ». Ce soutien régulier constitue la clé de voûte de l'existence de la plateforme Telemeta. Il permet également des opérations de numérisation et d'informatisation des métadonnées, indispensables pour alimenter le catalogue en contenus de qualité, sachant que la majorité du budget est dépensée pour l'analyse musicologique et documentaire des archives, faite par des spécialistes jour après jour pour indexer au mieux les données.

Côté CNRS, le projet a été soutenu en 2009 par le TGEA (très grand équipement Adonis) du CNRS devenu aujourd'hui le dispositif Huma-Num, la TGIR (très grande infrastructure de recherche) des humanités numériques, en tant que projet pilote innovant. La structuration des données en quatre niveaux de catalogage a été réalisée sur les conseils de Véronique Ginouvès²³. Diverses fonctionnalités ont été développées pour répondre aux besoins des utilisateurs : géolocalisation, moteur de recherche simple et avancé, c'est-à-dire avec des critères adaptés à la discipline, listes de lecture, annotations avec des marqueurs temporels, listes de vocabulaires adaptées aux données, etc. Un moteur d'exportation de l'ancienne base a été développé par Samalyse, prestataire spécialisé, et la migration d'environ 25 000 fiches descriptives a pu être effectuée en 2010 vers une base de données en ligne de type relationnel, MySQL, un des systèmes de gestion de bases de données les plus utilisés dans le monde. La version de production de l'application a été lancée en mai 2011, donnant déjà accès à 7 000 enregistrements sonores.

20 Abréviation de *Open Archive Initiatives Protocol for Metadata Harvesting*. <http://www.openarchives.org/>

21 <http://www.culturecommunication.gouv.fr/Thematiques/Enseignement-superieur-et-Recherche/La-recherche/Accord-cadre-avec-le-CNRS>

22 <http://culturelabs.culture.fr/objetsifs.html>

23 Responsable de la phonothèque de la maison méditerranéenne des sciences de l'Homme, Aix-en-Provence. <http://phonothèque.mmsh.huma-num.fr/>

Aujourd'hui, les archives numérisées ainsi que leur catalogue sont donc consultables sur le web, avec le grand avantage que la vérification des métadonnées peut se faire par les chercheurs à distance ; certains font eux-mêmes la saisie du descriptif de leurs collections de terrain, grâce à un compte utilisateur. Les résultats sont éloquentes : depuis son lancement, Telemeta a permis de multiplier par quatre les données du catalogue du CREM. Divers profils d'utilisateurs permettent de gérer les accès aux collections. La consultation des sons peut être restreinte aux détenteurs d'un mot de passe ou totalement réservée aux administrateurs du site, sachant qu'actuellement, plus de la moitié des documents audio ou vidéo sont consultables librement en streaming et sans mot de passe. Il faut également préciser que les sons ne sont pas téléchargeables et que les utilisateurs sont invités à nous contacter pour toutes demandes d'utilisation. Depuis 2013, les utilisateurs peuvent également trouver le fonds²⁴ sur *Isidore*, le moteur de recherche de l'INSHS du CNRS²⁵.

24 <https://isidore.science/collection/10670/2.hjtany>

25 <https://www.rechercheisidore.fr>

26 <http://www.iasa-web.org/>

27 <http://www.europeansounds.eu/fr/>

28 « Les Archives sonores de la Belle époque en ligne ». <http://www.phonobase.org/>

29 <http://www.europeana.eu/>

En 2013, la *British Library*, ayant découvert notre base de données en ligne lors de conférences à l'*International Association of Sound and Audiovisual Archives*²⁶, nous a proposé de participer au projet *Europeana Sounds*²⁷ (2014-2017) en tant que fournisseur de contenu afin de contribuer au partage des données sonores de la recherche au sein de la bibliothèque numérique européenne. À cette occasion, le CREM s'est associé à la phonothèque de la MMSH d'Aix, au LARHRA (laboratoire de recherche historique Rhône-Alpes) gérant la Phonobase²⁸ et au CRESSON (centre de recherche sur l'espace sonore et l'environnement urbain). Cette collaboration permet de mettre à disposition 37 000 documents sonores de la recherche en libre accès sur *Europeana*²⁹. Le CNRS est ainsi devenu le premier contributeur français à la bibliothèque européenne en nombre de documents sonores.

Grâce à cette nouvelle manne financière, le traitement des collections librement consultables s'est accéléré. De manière générale, le fonds s'enrichit chaque année d'environ 6 000 items audio et vidéo supplémentaires. En 2019, 70 600 items, répartis dans 7 100 collections, sont catalogués dans cette base de données qui comprend plus de 52 900 fichiers audio et vidéo, parmi lesquels environ 28 400 accessibles en streaming sans compte utilisateur. À titre indicatif, les enregistrements librement consultables sur la plateforme étaient deux fois moins nombreux au début du projet européen en 2013.

Les usages de Telemeta par des publics diversifiés

Telemeta est un outil de travail quotidien de l'équipe du CREM pour la gestion des fonds sonores et visuels numérisés, tout en assurant leur conservation. En 2015, la plateforme a migré de Nanterre vers le serveur d'Huma-Num à Lyon, situé à l'IN2P3 (institut national de physique nucléaire et de physique des particules). La sauvegarde des données et des

métadonnées est donc assurée ainsi que la maintenance des machines. L'amélioration et la maintenance de l'application restent à la charge du CREM mais les frais de développements informatiques sont en passe d'être mutualisés avec d'autres structures, en France et à l'étranger. En effet, de nouvelles fonctionnalités sont sans cesse demandées, comme ces derniers temps, l'utilisation de thésaurus partagés et gérés à plusieurs comme « Open théso » ou d'un système de géolocalisation collaboratif et ouvert tel qu'« *Open Street Map* » déjà intégré dans le Telemeta de l'université de Poitiers. Quant aux données elles-mêmes, la prochaine étape dans le processus de pérennisation sera l'archivage à long terme aux Archives Nationales, missionnées par l'État pour la sauvegarde des données de la recherche publique.

Nous pouvons décompter 3 000 visiteurs uniques chaque mois (37 000/AN), et 400 000 pages sont consultées par an. Il s'agit d'un public français et étranger, respectivement 30 % et 70 % environ. Comparées aux quelques visiteurs du laboratoire du musée de l'Homme d'autrefois, la diffusion et l'accessibilité ont « explosé ». Mais surtout, même si c'est une évidence aujourd'hui, il est possible de connaître le contenu des collections, à n'importe quelle heure et de n'importe quel endroit. La couverture géographique de diffusion a changé d'échelle. À titre d'exemple, le deuxième pays après la France à consulter les archives sonores du CREM en 2019 est le Gabon, phénomène impensable quelques années plus tôt. Les visiteurs étrangers proviennent d'origine géographique toujours plus diverses : États-Unis, Vietnam, Côte d'Ivoire, Maroc, Allemagne, etc. (fig. 7, en ligne). Pour la première fois depuis le lancement de la plateforme, des pays du continent africain figurent parmi les 10 premiers pays consultant les archives du musée de l'Homme.

Autre surprise : la réutilisation des archives par les blogueurs. Certains mois, 80 % des visiteurs transitent par les réseaux sociaux qui intègrent le lecteur de l'application Telemeta pour illustrer une recherche anthropologique, une identité culturelle ou des recettes de cuisine africaine... Cette technique de partage grâce à la copie d'un lien « i-frame » dans un site web permet l'utilisation des documents audio et vidéo, tout en évitant les téléchargements abusifs. Cette plateforme simplifie également les opérations de « retour au pays d'origine ». L'accès gratuit à ces données constitue une forme de restitution d'un patrimoine culturel et le son est plus facile à restituer, surtout avec internet, qu'une momie égyptienne ou une sculpture antique... Les échanges se multiplient ainsi avec les pays d'origine des enregistrements : Ladakh, Indonésie, Burkina-Faso, Tchad, Guyane, etc.

Le témoignage de Marsel Lembang³⁰, utilisateur indonésien rencontré par Dana Rappoport, directrice de recherche au CNRS, est éloquent. Pour apprendre le culte de ses ancêtres et le pratiquer lui-même, ce Toraja en exil consulte les archives sonores documentées

30 http://archives.crem-cnrs.fr/archives/items/CNRSMH_I_2016_007_001_01/

31 http://archives.crem-cnrs.fr/archives/fonds/CNRSMH_Rapport/

32 Terme venant du monde de l'entreprise, document de présentation de la stratégie financière de l'entreprise.

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

par cette ethnomusicologue, dont l'intégralité des enregistrements de terrain est en libre accès sur Telemeta³¹. Dana Rappoport a pleinement pris la mesure de l'impact d'une telle diffusion auprès des communautés concernées et a intégré dans sa démarche la perspective d'un large partage du patrimoine culturel toraja en voie de disparition. L'objectif du partage des données de la recherche est donc atteint au-delà de nos espérances.

Il est clair qu'aujourd'hui, ces archives n'intéressent plus seulement quelques chercheurs ou visiteurs du musée de l'Homme. De nombreux pays sont intéressés par ces témoignages de leur patrimoine oral, surtout quand leurs institutions locales n'ont pas les moyens de l'archivage ou lorsque l'*open data* n'est pas prévu dans leur *business model*³², comme c'est souvent le cas en Europe ou ailleurs. Pour l'heure, les messages d'encouragement du monde académique et des pays d'origine se multiplient, par courriel ou sur Facebook. Personne ne remet en cause une telle diffusion mise en place avec l'accord des chercheurs et des déposants, et selon les lois en vigueur pour les droits des artistes-interprètes. La gestion des droits d'accès permet d'ailleurs de moduler au cas par cas la diffusion des données selon des considérations éthiques. Bien heureusement, en France, le CNRS et le ministère de la Culture encouragent le libre accès aux données patrimoniales, dans les limites du cadre légal du Code de la propriété intellectuelle.

Dans la grande majorité des cas, il n'y a pas d'auteur ou de compositeur déclarés ou connus en musique traditionnelle. Parfois les noms des interprètes sont connus, mais les personnes sont difficiles à joindre. Nous référant aux écrits d'Anne-Laure Sterin³³ qui définissent le cadre éthique et juridique pour les enregistrements sonores, les enregistrements antérieurs à 1963, sauf cas particuliers que nous respectons pour des questions d'éthique, sont librement consultables. Si leur contenu n'a pas de caractère secret ou « risqué » pour les communautés d'origine, et que le collecteur souhaite diffuser ses enregistrements postérieurs à 1963, nous considérons que c'est la personne la mieux placée pour respecter la volonté des interprètes et de leurs descendants. S'il souhaite garder une restriction d'accès, nous respectons sa volonté, et l'enregistrement est réservé pendant 70 ans après sa collecte. Dans le cas où la mise en ligne est libre d'accès (c'est-à-dire sans mot de passe), les témoignages nombreux et enthousiastes des communautés nous encouragent à poursuivre dans cette voie. En cas d'avis contraire, la diffusion en ligne peut être arrêtée à tout moment, ce qui n'est pas possible dans le cas d'un disque. En l'absence d'exploitation commerciale et donc d'éventuel préjudice financier pour les ayants droits, les risques sont nuls pour l'institution selon les juristes spécialisés. Par ailleurs, les fichiers haute définition ne sont jamais accessibles, contrairement à un CD, ce qui limite les exploitations commerciales.

Depuis le début du ^{xxi}^e siècle, le domaine des humanités numériques est en pleine expansion. Tout est à inventer et à expérimenter. Les bases de données sonores bien documentées sont rares et s'avèrent précieuses pour les chercheurs qui développent l'intelligence artificielle dans le domaine de l'image et du son. Dans ce contexte, divers partenaires à la recherche de contenus de qualité nous contactent pour participer à des projets innovants et patrimoniaux. Une série de projets de recherche et de valorisation s'appuyant sur l'accès en ligne des archives du CREM se mettent ainsi en place, tant au niveau national qu'international. Peu après le lancement de la plateforme en ligne, le CREM-LESC et le laboratoire d'Eco-anthropologie du MNHN ont collaboré à un projet de recherche *Contenus et Interactions* financé par l'agence nationale de la Recherche (ANR). L'aventure a débuté en 2013 avec trois laboratoires de recherche en informatique du CNRS (IRIT, LIMSI, LaBRI) et l'équipe du LAM, notre partenaire habituel. Ce programme a permis de renouer des relations scientifiques avec le musée de l'Homme, partenaire historique des archives sonores. Intitulé DIADEMS (acronyme pour « description, indexation, accès aux documents ethnomusicologiques et sonores »)³⁴, ce projet a pour objectif de faciliter l'accès et l'indexation semi-automatique des contenus sonores grâce aux outils d'analyse du son développés par les spécialistes du traitement acoustique (fig. 8, en ligne).

Le CREM-LESC participe également depuis 2012 au Labex *Les Passés dans le présent*³⁵, avec le projet *Les sources de l'ethnomusicologie* en collaboration avec la BnF et le musée du quai Branly, afin de créer un catalogue commun relatif aux documents d'archives et les instruments de musiques répartis dans les trois institutions. Financé conjointement par le Labex et par son homologue britannique l'AHRC *Care for the Future*, le projet DaCaRyH³⁶ associe des ethnomusicologues et des ingénieurs du CREM-LESC, la société Parisson et le centre de Musique numérique (C4DM) de l'université Queen Mary de Londres. Ce projet propose d'analyser le calypso à travers l'histoire. L'objectif est de développer des modules algorithmiques d'analyse musicale des archives sonores et vidéo du répertoire de *steelbands* de Trinidad dans une perspective diachronique (fig. 9, en ligne).

En 2018, un projet équivalent s'est réalisé avec l'université de New York pour l'analyse des similarités rythmiques sur une large couverture géographique autour du monde arabe. Grâce à ce financement, de nombreuses collections³⁷ ont été numérisées et mises en ligne (environ 400 heures de contenu).

On peut donc distinguer deux types de projets : la valorisation basée sur l'exploitation des techniques de l'information et de la communication avec une forte composante en ingénierie documentaire (tel que le portail *Europeana*) et les programmes de recherche en traitement du signal pour développer le domaine de l'intelligence artificielle, ou plutôt

34 Voir dans ce dossier la contribution de Marie-France Mifune et Florabelle Spielmann.

35 <http://passes-present.eu>

36 Voir dans ce dossier la contribution de M.-F. Mifune et F. Spielmann.

37 Voir quelles ont été les collections bénéficiaires sur : https://archives.crem-cnrs.fr/archives/fonds/CNRSMH_NYUAD

dans notre cas, l'oreille musicale artificielle. L'investissement initial est donc globalement rentable en terme d'augmentation des données disponibles, même si les budgets ne sont toujours pas explicitement orientés vers la numérisation des données de la recherche. Une réflexion devrait cependant être menée pour envisager une mutualisation à un niveau plus large de la gestion de ces projets qui se multiplient.

Pour conclure, on peut dire que trois « bonnes fées » se sont penchées sur le berceau respectable des archives sonores du musée de l'Homme. Tout d'abord, une équipe de recherche avec du personnel spécialisé pour gérer et documenter un patrimoine immatériel mondial exceptionnel. Ensuite, ce fonds a bénéficié de la volonté du ministère de la Culture de sauvegarder et valoriser les données de la recherche. Ce soutien indispensable est accompagné d'une incitation par toutes les tutelles à l'utilisation de logiciels libres, ou *open Source* et aux données ouvertes, ou *open data*. Enfin, l'existence de compétences en analyse informatique du signal audio au CNRS et la possibilité de collaborations interdisciplinaires entre ses équipes de recherche a constitué un grand avantage sur les grandes institutions de la conservation dans le monde. La combinaison de ces trois facteurs a permis de mettre en œuvre le partage des données audiovisuelles en ethnomusicologie et d'entrer de plain-pied dans la révolution numérique, chère à nos tutelles. Lors des rencontres avec nos homologues internationaux, on se rend compte aisément que cette aventure est assez unique à l'échelle d'une équipe de recherche dont la taille est modeste et les moyens financiers limités. L'investissement initial est cependant largement compensé, et donc récompensé par les projets induits et il profite à d'autres équipes françaises et étrangères utilisant le même outil tels que le LAM, qui est un des plus anciens partenaires dans notre aventure, le centre des Musiques arabes et méditerranéennes³⁸, l'université de Poitiers, l'université de Mexico et d'autres (fig. 10).

38 <http://phonotheque.cmam.tn>

Est-ce la mission d'une équipe de recherche de supporter une telle aventure, surtout sans personnel informatique en interne ? Ce manque de moyens est le principal frein pour l'avenir du projet et nécessite une course permanente aux financements. Certains chercheurs prennent conscience des apports innovants d'une telle mission de partage des connaissances mais l'interdisciplinarité nécessaire à ce type de projet est un exercice difficile qui demande de la curiosité, de la patience et beaucoup de pragmatisme. En pratique, rien n'est donc jamais gagné. Une chose est sûre : un outil doit répondre avant tout aux besoins spécifiques de l'équipe qui produit et utilise ces données. Il doit donc permettre de les traiter et de les enrichir de manière efficace, tout en étant ouvert au public. L'outil parfait ou la base de données universelle n'existent pas, et le « sur-mesure » a un prix, financier et humain. Cependant, l'existence à ce jour de 70 000 fiches documentaires accessibles à tous grâce à cet outil est déjà en soi en succès. Ces collections sont restées « vivantes »,

ce qui n'était pas acquis il y a 20 ans. Le pari de l'ouverture est gagné dès lors que ces données sont visibles sur deux autres plateformes, Isidore et Europeana dont nous avons parlé plus haut. Les demandes d'utilisation « académique » des données, en dehors de la simple consultation, sont en progression mais restent en quantités raisonnables. Elles sont examinées par le CREM, sachant que toute demande d'exploitation commerciale est rejetée. De manière plus pragmatique, les projets qui se succèdent apportent un complément d'activité appréciable pour les jeunes diplômés tels que les post-doctorants qui cherchent à s'intégrer au marché du travail qui demeure très restreint dans cette discipline. Enfin, les efforts des ingénieurs collaborant à ce projet innovant ont été récompensés en 2018 par la remise du « cristal collectif » du CNRS.

Aujourd'hui, la numérisation et la publication en ligne permettent l'échange des données à grande échelle, ce qui était totalement inimaginable auparavant. Grâce à Internet, de nouveaux usages sont offerts aux chercheurs mais aussi aux artistes-interprètes et à leur communauté très motivés pour accéder aux patrimoines oraux de leurs ancêtres. Les nouveaux mécanismes de partage des archives créent des approches de recherche novatrices et présentent de nouvelles opportunités de collaborations entre les disciplines autour de l'oralité. De nouveaux défis apparaissent également, comme les questions de propriété intellectuelle et d'éthique, dans un monde de *Big Data* où les enjeux commerciaux ne sont jamais très loin, même lorsque le patrimoine et la connaissance sont considérés comme des biens communs. S'il est difficile d'imaginer l'avenir de ce nouvel écosystème du web, ce patrimoine musical de l'humanité n'est plus limité à quelques spécialistes parisiens. Chacun peut désormais découvrir ces ressources historiques, dont les voix reprennent vie pour faire écho parmi les générations futures.

Mars 2019



Association française
des archives, sonores,
orales & audiovisuelles

L'AFAS [Association française des détenteurs de documents audiovisuels et sonores] est une association créée en 1979 sur l'initiative de responsables de fonds d'archives sonores ou de phonogrammes inédits ou édités à vocation de recherche, d'études ou d'usage professionnel. Son objectif est au départ de réunir institutions et personnes qui s'intéressent à la sauvegarde, au traitement et à la communication des documents sonores et audiovisuels inédits ou édités. En 2014, sur l'impulsion de sa présidente Florence Descamps, l'AFAS conserve son sigle mais change son intitulé pour devenir l'Association française des archives orales, sonores et audiovisuelles et s'ouvrir plus largement aux sciences humaines et sociales. L'association souhaite également reprendre pour son Bulletin l'ancien titre de sa revue thématique créée en 1980 *Sonorités*, qui devient ainsi *Bulletin de l'association française des archives sonores, orales et audiovisuelles. Sonorités*.

Version électronique
ISSN : 2109-9537 ISBN : 979-10-97301-44-8

Version imprimée
ISSN : 1246-7529 ISBN : 979-10-97301-05-7