

UNVEILING SOUTHEAST ASIAN MUSICAL DATA IN EUROPE WITH THE PRATINADA PLATFORM: FUNCTIONS, ORIGINS AND CULTURAL PERSPECTIVES

Joséphine Simonnot, *Research Engineer, Centre National de la Recherche Scientifique, France*¹

Dana Rappoport, *Director of Research, Centre National de la Recherche Scientifique, France*²

DOI: 10.35320/ij.178

Abstract

Colonial-era sound recordings from Southeast Asia remain largely housed in European archives, historically inaccessible to the societies from which these archives come. The Pratinada website addresses this colonial legacy by democratizing access to collections held in France, the Netherlands and the United Kingdom. This online portal consolidates multiple European-held Southeast Asian sound databases, making thousands of historical recordings publicly available. However, the digital repatriation process raises complex cultural, ethical, and technological considerations that warrant careful examination.

Keywords: Sound archives, aggregated database, colonial heritage, Southeast Asia, web platform, decolonization

Introduction

The digital revolution has transformed the landscape of archival preservation and access, notably in the realm of sound recordings. While this transformation offers unprecedented opportunities for expanding access to cultural heritage, it also brings to the surface complex challenges at the intersection of colonialism, cultural ownership, and digital humanities ethics. The Pratinada platform,³ a pioneering initiative to aggregate and provide online access to Southeast Asian music collections held in European institutions, exemplifies these tensions. Through the Pratinada website, historically overlooked sound archives are now available to international audiences, including originating communities previously disconnected from their cultural heritage. However, this digital repatriation occurs without direct consultation with the original performers, creating an ethical paradox: the imperative to provide access must be balanced against risks of cultural misappropriation and misuse. This dilemma raises fundamental questions about the cultural, ethical and technological considerations that should guide the digitization and dissemination of colonial-era sound archives.

The Pratinada website is hosted by the Centre National de la Recherche Scientifique (Délégation de Provence) in France and its database by the University of Amsterdam, which is also responsible for scientific and IT management. It is the result of a collective team managed by ethnomusicologist Dana Rappoport, sound database engineer Joséphine Simonnot and IT specialist Anas Ghrab. The present article first examines Pratinada's architecture and functionality, situating it within the broader historical context of French archival practices, before critically analyzing the cultural, ethical and technical implications of this digital platform.

1 Joséphine Simonnot is a Senior Sound Database Expert at Perception, Representation, Image, Sound, Music (PRISM), Aix Marseille Univ, CNRS, Marseille.

2 Dana Rappoport is a Senior Ethnomusicologist at the Centre Asie du Sud-Est (CASE), Centre National de la Recherche Scientifique (CNRS), Ecole des Hautes Etudes en Sciences Sociales (EHESS), and Institut National des Langues et Civilisations Orientales (INALCO).

3 <https://pratinada.net/>

General Information

Pratinada provides access to 7,304 recordings from Southeast Asia, including field recordings, radio programs and published recordings. These materials, all held in European archives, represent more than a century of sound recordings (1900–2023), now made accessible through a multilingual interface in nine languages, seven of which are Southeast Asian. Researchers and sound archive specialists from France, the Netherlands and the United Kingdom collaborated with academics and Southeast Asian partners to improve access to European sound collections from Southeast Asia, which have been largely inaccessible to the public. Historically, barriers to access have been numerous. Among them are language, lack of digitalization, physical distance, restrictive access policies, institutional staffing problems, legal issues and bureaucratic resistance. The creators of this website have attempted to overcome all these obstacles.

Launched in May 2024, Pratinada is the result of years of experimentation with ethnomusicological sound data on the internet. It was made possible thanks to the support of the European Hub Joint Program Initiative on Cultural Heritage and Global Change,⁴ which funded a three-year project called *Decolonizing Southeast Asia Sound Archives* (DeCoSEAS).⁵ This project involved the collaborative efforts of curators, ethnomusicologists, collectors, database managers, IT specialists and Southeast Asian partners. The French team⁶ oversaw the creation of the Pratinada digital platform in a participatory way, aiming to design a new digital curational framework by aggregating various sound databases in dialogue with Southeast Asian partners.

Content

The Pratinada website aggregates the contents of three major European sound archives: Southeast Asian sound archives hosted in France; the Jaap Kunst Collection of wax cylinders hosted in Germany and managed by the University of Amsterdam,⁷ and a set of radio programs broadcast in Southeast Asia from BBC Empire Service Broadcasts in the United Kingdom.⁸ Thanks to this project, the archives of the Jaap Kunst and BBC sound collections are available online for the first time, and have been integrated into a single database with the French collections, which were already accessible but spread across different institutions with only French-language metadata.

The contents of these aggregated archives differ in various respects, including the nature of the sounds, quantity of recordings, and access rights. The sound archives provided by France offer a diverse range of collections comprising 6,500 mainly music items from the CNRS - Musée de l'Homme in Paris,⁹ managed and disseminated online by the Centre de Recherche en Ethnomusicologie (CREM),¹⁰ the French National Library

4 <https://www.heritageresearch-hub.eu/joint-programming-initiative-on-cultural-heritage-homepage/>. Each partner country finances its own team. The French team is funded by the Agence Nationale de la Recherche (ANR-21-CHIP-0001).

5 <https://www.decoseas.org>

6 In addition to the project managers, Pratinada benefited from the help of many French collaborators from Institut National des Langues et Civilisations Orientales (INALCO), the French Center for Southeast Asia Studies (CASE/CNRS-EHESS-INALCO) and Southeast Asian partners.

7 <https://jaapkunst.org>

8 <https://omekas.seasia-hearing.org/s/bbc-sea/page/welcome>

9 CNRS - Musée de l'Homme Sound Archives: <https://archives.crem-cnrs.fr/>; Centre National de la Recherche Scientifique (CNRS): <https://www.cnrs.fr/>.

10 <https://lesc-cnrs.fr/fr/laboratoire/lesc-crem>

(BNF-Gallica),¹¹ the Phonobase¹² and the Office de Radiodiffusion-Télévision Française (ORTF).¹³ The sound archive at the University of Amsterdam consists of a collection of 600 wax cylinders collected by Jaap Kunst in Indonesia in the 1930s. 78 broadcast recordings from the BBC are also included, which are managed by the School of Oriental and African Studies (SOAS) and hosted at the British Library, though very few recordings are accessible because of copyright issues. This disparity between collections reflects divergent histories of archival work in Western Europe and the various levels of access provided to digital sounds in the international academic world.

The Pratinada sound archives primarily consist of field recordings of traditional Southeast Asian music, from 1900 to the present day, recorded in rural and urban areas across the region, from Myanmar to Papua New Guinea, including Thailand, Vietnam, Laos, Cambodia, Taiwan, Malaysia, Indonesia, the Philippines and East Timor.¹⁴ The main categories of content are vocal music (3,300), instrumental (1,500) and mixed vocal-instrumental (1,000).¹⁵ The extensive presence of vocal music may come as a surprise, in view of the prevalence of instrumental ensembles in the published recordings from this part of the world. Only 341 of the recordings in Pratinada were produced professionally.¹⁶ Most of the recordings were gathered as part of field research, with the goal of preserving cultural heritage for future generations that was either undergoing profound changes or at risk of vanishing. The majority of collectors were ethnomusicologists, anthropologists, or linguists. Some were explorers or radio program producers, while a few were filmmakers, composers and missionaries.

Pratinada's primary mission is to reconnect Southeast Asian communities with sonic heritage currently housed in European institutions. This material is preserved in Europe for two main reasons: it was either recorded during the colonial period or collected later by European researchers as part of scientific investigations. These archives preserve irreplaceable cultural expressions, including musical traditions that have since vanished or evolved. For instance, the Toraja ritual music of Indonesia documented in the 1990s captures ceremonial ancient performances that were subsequently forbidden by Christian missionary influence.¹⁷ Similarly, recordings in the Boutary collection¹⁸ constitute the sole surviving documentation of extinct cultural practices and dialects, such as those of Vietnam's Lac people. These examples underscore the archives' significance not only as historical documents, but as vital cultural patrimony.

Historical Background and Context

From the 16th century onwards, Southeast Asia became a strategic region coveted by European powers for its natural riches and geographical position. The Portuguese were the first to settle there, notably in Malacca in 1511 and in East Timor, which was of-

11 <https://gallica.bnf.fr/accueil/en/html/accueil-en>

12 <https://www.phonobase.org/>

13 <https://www.ina.fr/ina-eclaire-actu/1964-1974-naissance-et-disparition-de-l-ortf>

14 One might question the inclusion of Taiwan and New Guinea in Southeast Asia. Much of the region shares a common Austronesian linguistic heritage, originating from Taiwan—where the oldest Austronesian languages are found—and extending westward to Madagascar and eastward to Easter Island, spanning maritime Southeast Asia, Melanesia, Micronesia, and Polynesia. Since Papuan languages are present in the Malay Archipelago, we include New Guinea in Pratinada as it can be considered as a transitional zone, even though it is not traditionally considered part of Southeast Asia.

15 Approximation is based on Pratinada search results.

16 Exhibitions, performances, studio recording sessions and radio programs comprise 5% of total recordings available.

17 Toraja-Sulawesi collection: <https://pratinada.net/item/14486>.

18 <https://pratinada.net/item/20444>

ficially colonized in 1702. The Spanish controlled the Philippines from 1565 until 1898, when the United States took over after the Spanish-American War; the Philippines then became an American colony until independence in 1946. The Dutch took control of the Malay Archipelago (the so-called Dutch East Indies) in the 17th century, consolidating their hold until Indonesian independence in 1945. The United Kingdom colonized Burma, Malaysia and Singapore, which it controlled until the 20th century. French Indochina was established from 1862, bringing together Cochinchina, Annam and Tonkin (now parts of Vietnam); Cambodia and Laos, all of which became independent between 1953 and 1954. Portugal finally relinquished control of East Timor in 1975, which was then invaded by Indonesia nine days later. It was only after a United Nations-backed referendum that the country achieved definitive independence in 2002. Thus, the region, long divided between colonial empires, gradually became independent over the course of the 20th century, often after protracted struggles.

After decolonization, certain territories with high natural resource potential became the focus of predation from other Southeast Asian countries. This led to the political division of some of Southeast Asia's islands. Borneo is now divided in three: Malaysia controls the north with the states of Sabah and Sarawak, Indonesia administers most of the south under the name of Kalimantan, and the small sultanate of Brunei is an independent state. In the far east, the island of New Guinea is divided into two countries: Indonesia annexed the west of the island in 1963 (now called Papua), while the eastern half of New Guinea has been a state in its own right since 1975, when it officially separated from Australia, which had administered it since 1920 under a mandate from the League of Nations, then from the UN after the Second World War.

On Pratinada's map¹⁹ (Figure 1), the geographical distribution of recordings reveals a significant concentration from insular Southeast Asia; with Indonesia, Papua-New Guinea, the Philippines and East Timor collectively accounting for 4,720 sound items, more than half of the total archive. This large quantity may be explained by many studies of the region and by the central role of the Ethnomusicological Department of the Musée de l'Homme in archiving sound recordings. Vietnam accounts for the second-largest concentration of recordings (1,925 sound items), reflecting intensive ethnographic fieldwork conducted in the post-independence period after 1954. Notable ethnographers and missionaries during Vietnam's post-colonial transition, including the team of Georges Condominas, Geneviève de Chambure and Marius Boutary (from 1958), Anne de Hauteclocque (from 1961) and Jacques Dournes (from 1963), documented oral traditions that would later face crisis and decline due to prolonged conflict. Myanmar, Cambodia, Thailand, and Malaysia are underrepresented, reflecting a lack of sound collections from those countries in Europe.

19 <https://pratinada.net/search>

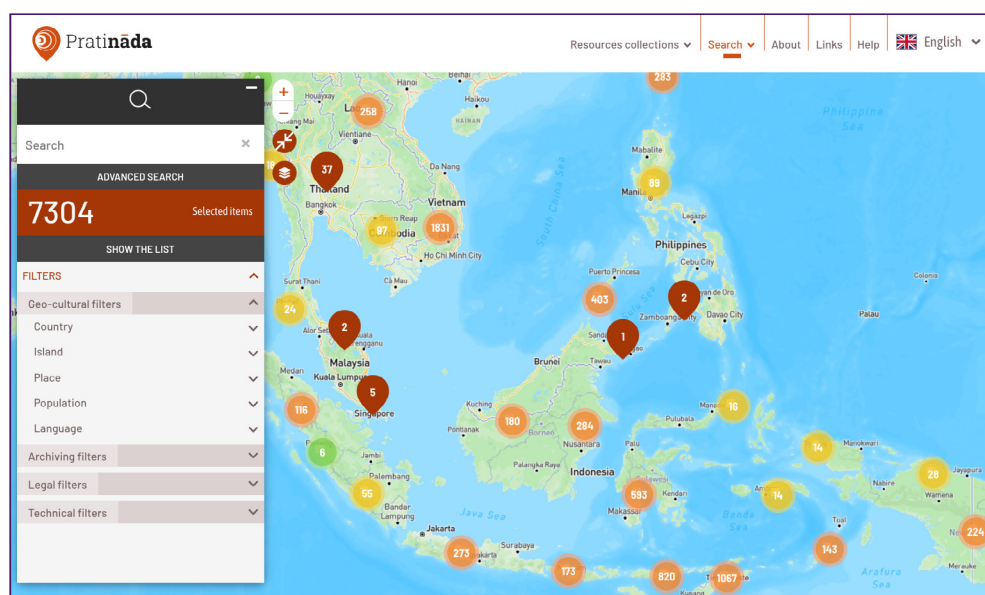


Figure 1. Map with geo-cultural (country, island, place, population, language) and other filters.

Database Structure

The database, built from three European archives, has a basic 3-level internal structure of archive/collection/item. Each collection contains items (recordings) from the same collecting context. One way of browsing recordings is via the ‘resources collections’ tab²⁰ in the main menu bar, where users can access the three main archives of Jaap Kunst, Musée de l’Homme and BBC. Access to ‘all collections’²¹ is also available from the resources collections tab. The collection titles are presented in chronological order and grouped by country.

Individual collectors are also accessible through this menu.²² One collector may be responsible for several collections (for example, by depositing field recordings from different research areas). Each item contains a link to the basic description of the collection, the collector, and the code numbers of the original holding institution for traceability.

Currently in 2025, the database contains 72 collections, systematically named according to country, region, date and collector (Figure 2). The collection sheet provides a general description of the form of the collection, such as fieldwork or concert performance, accompanied by photos. A brief summary gives historical context and situates the collection. After the collection title, the total number of items is indicated in brackets.²³ All the descriptions of collections and are available in English, and sometimes have been translated into English, French, and Indonesian.²⁴

²⁰ <https://pratinada.net/collections>

²¹ <https://pratinada.net/collection>

²² <https://pratinada.net/collectors>

²³ As an example: ‘Vietnam: Centre, Highland, Hauteclacque, 1961’ (83 items).

²⁴ E.g., Dutch collector Johannes Anceaux (1920–1988): see <https://pratinada.net/item/1679>, or French collector Louis Berthe (1927–1968): <https://pratinada.net/item/1676>.



Figure 2. Collection view: Jaap Kunst's collection from Papua.

Sound Access Functionalities: Map, Text Research, Filters

Pratinada offers a range of features that support exploration through both textual and geographical interfaces. A key priority from the outset was developing a responsive web application, recognizing that most users in Asia access the internet with mobile devices. Consequently, two distinct interfaces have been developed: one tailored specifically for mobile phones (Figure 3) and another designed for desktop users (Figure 4).²⁵

25 For optimal adaptation to user needs, the Pratinada user interface was designed using the Agile method for software development which includes improvements through a collaborative effort and a cross-functional team : https://en.wikipedia.org/wiki/Agile_software_development.



Figure 3. Pratinada mobile interface (left).

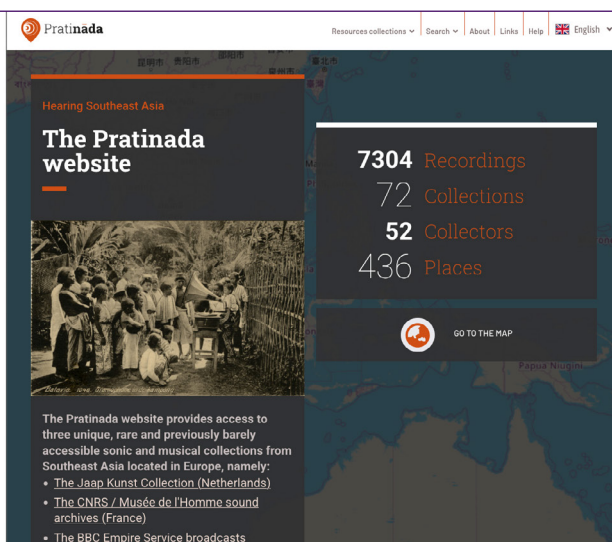


Figure 4. Pratinada desktop interface (right).

The audio recordings can be accessed by way of several geographical and cultural parameters. Ordinary consultation of the catalogue is possible via the Resources Collections menu, but the most direct and intuitive way to listen to recordings is to browse the map or to use text search and filters.

Geographical access

Each sound file's recording location is geolocated on a map (Figure 5). Users may search for locations in Southeast Asia and also in Europe (76 recordings of Southeast Asian music were made in France or the United Kingdom during exhibitions or broadcasts). A flag on the map indicates the number of recordings collected in a given location. Users may choose from three kinds of maps: an Open Street Map, with place names written in local alphabets or characters; a map with all place names in the Latin alphabet; or a satellite view.

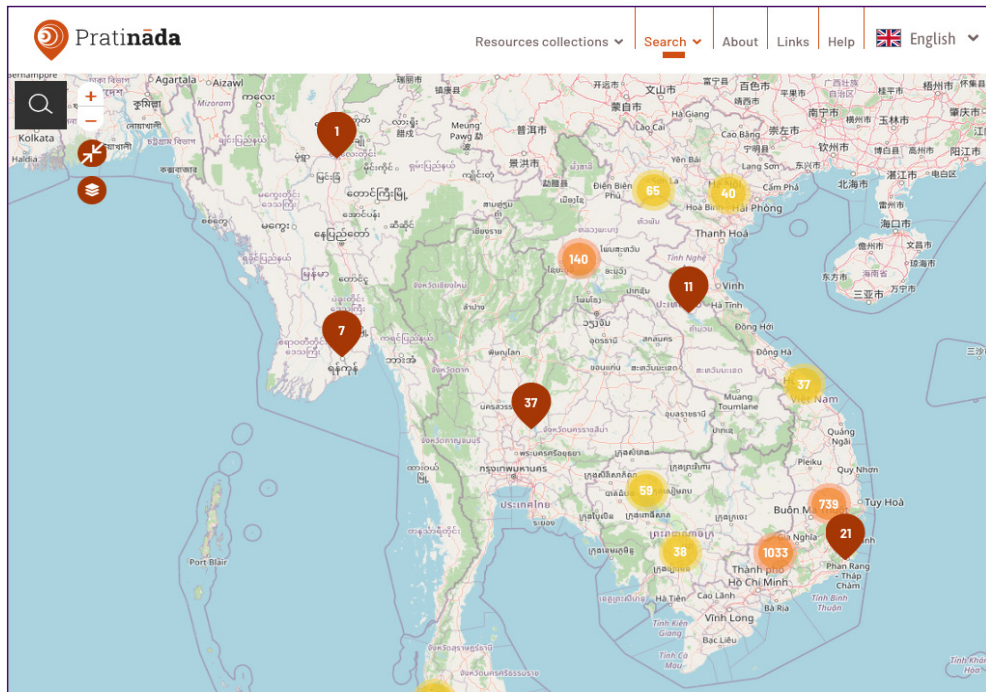


Figure 5. Open Street Map view of Pratinada map.

Search Filters

Full-text search and advanced search with Boolean operators are also possible. Specific filters were designed to refine searches with 17 different parameters,²⁶ which are divided into four groups. The platform's filters are a distinctive strength for enhancing data findability, developed on the basis of a thorough assessment of user needs. A video tutorial has been produced to explain how to use the interface and is available online on the Pratinada 'links' page.²⁷

The first group of filters is based on relevant 'geo-cultural' features (Figure 1). Sound items can be searched by country, island, place, population group and language. Pratinada provides access to music, spoken voices and soundscape from 14 countries. The 'country' filter selects items according to nation, displaying the number of sound items available for each one. The majority of sound files originate from Indonesia (2,634 items) and Vietnam (1,923 items). Using the 'island' filter, users can explore recordings from 38 different Southeast Asian islands. In some cases, determining the appropriate level of granularity islands and island groups, especially smaller ones, proved challenging. Because of the sizable number of recordings made on individual islands of Indonesia's Tanimbar archipelago, a decision was made to name the island group itself along with the names of individual islands to improve navigation. Another advantage of displaying islands separately from countries is that users can choose geographic rather than political borders. For example, the islands of Timor and Borneo are divided among multiple nations (Indonesia/Timor-Leste and Indonesia/Malaysia/Brunei, respectively).

²⁶ <https://pratinada.net/search>

²⁷ <https://sharedocs.huma-num.fr/wl/?id=s8ic1MI1risZZOSnO323hFJZ9zWqV308>

Users can search the entire island of Timor or the entire island of Borneo, without having to select a country.

The ‘place’ filter allows users to access sound recordings based on their specific recording locations. As of 2025, 436 place names are displayed on the map, including cities, villages, regions, districts, and even sites along rivers or river basins with no formal names. In many cases, existing metadata lacked detail; in others, place names had changed over time or were difficult to trace due to issues with transliteration.²⁸ Not all villages were included on Google Maps or Open Street Map, political events sometimes caused places where sounds were recorded to literally disappear,²⁹ and vernacular names at times differed from those used by geographers. While the geolocated place name is indicated in the field ‘place’, all location information available in the source metadata is included in the field ‘recording place’. For example, a recording collected in Kalimantan, Indonesia³⁰ is geolocated as ‘Kenyah Umaq Taw - Tanjung Manis’ and the description ‘Kalimantan Timur, Long Segar sur la rivière Telen’ is found in ‘recording place’. The project team also had to choose which location to include when recordings were made outside of their original cultural context. The decision was made to prioritize cultural origin over recording site. For example, a Vietnamese singing performance recorded at the 1900 Universal Exhibition in Paris was not mapped to Paris, but to the singers’ place of origin in Vietnam.

The ‘population’ filter provides access to 152 ethno-linguistic groups. The selection process raised difficult questions, particularly about naming conventions. Should historical, often colonial-era names be used; or current, self-identified terms? Should only the main language groups be considered or subgroups as well? Subgroups often have distinct local names and speakers may have their own music. Ultimately, the current names of populations were chosen rather than their colonial-era labels, making it easier for contemporary users to search for and identify groups.

The ‘language’ filter contains a list of 122 languages. Some of these have not yet been described by linguists and for others, their very naming is still open to discussion (for example: Dadu’a or Du’a for the name of subgroup language in Atauro Island, East Timor). As some languages collected have not been inventoried in common linguistic thesauri, two linguists specializing in this geographic area compiled their own list for Pratinada. Many Austronesian languages are ordered in groups and subgroups. In some cases, the subgroup was named (e.g. Sikka language group, Tana ‘Ai dialect). But most of the time, only the large groups and not the subgroups were retained. For example, the Toraja Mamasa subgroup has been merged into the Toraja group. Similarly, the Tanjung Bunga subgroup has been merged into the Lamaholot language group. Local names of the populations and their languages were respected as much as possible. However, language is currently less precisely identified in the original databases than population names.

In a second group, ‘archiving filters’, users can search for files according to usual archiving parameters: the collection, collector, recording date, recording context, or the category of content. Users can browse all items grouped under a single ‘collection’, which

28 When possible, collectors who had knowledge of their deceased colleagues’ fieldwork were interviewed to help identify recording locations.

29 For example, the Vietnam War destroyed all traces of the village of Sar Luk, where Georges Condominas conducted most of his fieldwork. See also Paul Levy’s ‘Les Mnong Gar du Centre Viêt-Nam et Georges Condominas’, p. 82 (https://www.persee.fr/doc/hom_0439-4216_1969_num_9_1_367023).

30 Udoq dance: <https://pratinada.net/item/24982>.

primarily represents fieldwork recordings, but may also include published records or radio broadcasts such as *BBC: London Calling Asia*. When a collector worked in a single location, all their recordings from that place are grouped together as one collection unit. In the CNRS - Musée de l'Homme Sound Archive web platform, the Nicole Revel Corpus contains seven sound collections, each corresponding to a field expedition on Palawan Island in the Philippines.³¹ In Pratinada, however, all of these recordings have been grouped into one collection.³² If a collector recorded in different places and at different times, several collections are listed.³³

Currently, the database features 52 collectors, with recordings dating from 1900 to 2023. The 'collector' filter allows users to select all recordings collected by a specific person (or organization), regardless of location or date. For instance, the renowned Filipino composer and ethnomusicologist Jose Maceda recorded in the Philippines and Indonesia between 1955 and 1977. His four collections available in Pratinada are grouped together under his collector sheet.³⁴

The filter 'recording date' is divided from 1900 to 2030 in ten-year increments. The 1960–1970 decade is the richest, with 2,131 items. The user can select several years in the same decade but cannot choose more than one decade at a time (Figure 6). Therefore, recordings can be explored by decade or by selecting specific years to see the historical evolution of collections during the colonial era. 'Advanced search' allows search over wider date ranges.

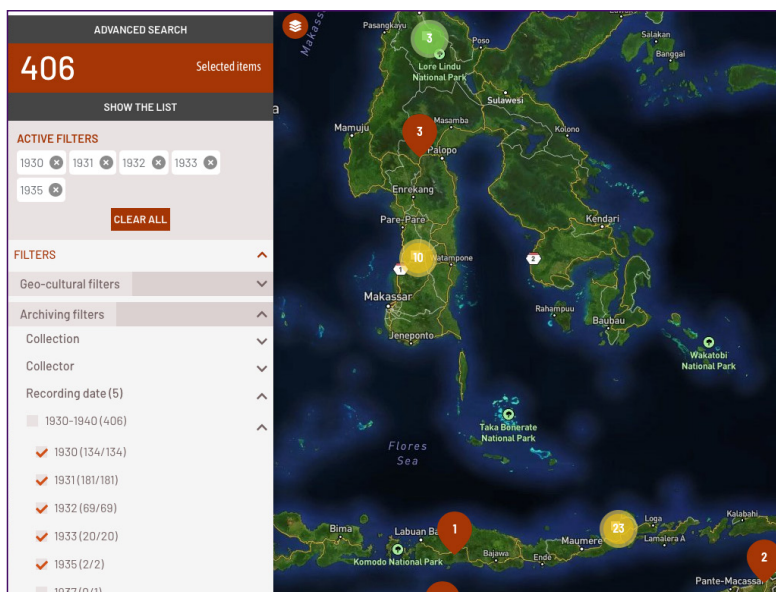


Figure 6. Recording date selection in the filter and the results on the map.

31 Nicole Revel Corpus: https://archives.crem-cnrs.fr/archives/corpus/CNRSMH_Revel_001/.

32 'Philippines, Palawan population, N. Revel, Ch. Macdonald, 1970–1999': <https://pratinada.net/item/20428>.

33 This is the case for Jaap Kunst's recordings, which have been divided according to location.

34 <https://pratinada.net/item/1671>

‘Recording context’ allows the filtering of content according to the collecting situation, whether it is a field recording, interview, performance, published recording, radio program, exhibition, social media content or studio recording. As stated above, most of the recordings are field recordings (6,728 items).

Sound databases may be used by many sorts of users, including musicologists, ethnomusicologists, composers, teachers, or filmmakers, who might look for generic content such as the sound of a drum, a spoken voice, or a soundscape. However, granular meta-data for medium of performance or genre is not present in the hosting institutions’ databases. Two parameters (‘instrument’ and ‘category’) requiring further development will be displayed at a later date.

In a third group of ‘legal filters’, parameters are linked to the legal status and ownership of these archives, with the aim of maximizing transparency for indigenous societies’ rights. Users may choose recordings from a specific provider: the holding institution that stores the physical object, the accessing institution which manages the archive, or those who manage copyright and licensing issues. The purpose of these filters is to clearly identify each legal responsibility, which may lie with one institution or be shared among several.

A fourth group, ‘technical filters’, helps to inform users about the sound quality and usability of each recording. Items can also be filtered based on the presence or absence of sound files, as some media remain offline due to a lack of digitization or pending legal authorization. Most of the archived items were originally recorded on magnetic tape (4,600 items from France) and were professionally digitized over the past two decades (Simonnot, 2020, pp. 3-5). Published recordings, including 78 rpm discs, LPs and CDs, are also available in Pratinada (185 items). BBC recordings, mainly 78 rpm discs, were digitized by the British Library according to their own technical criteria. The earliest recordings in the Pratinada database were made on wax cylinders: 652 items recorded between 1900 and 1933. Despite their poor sound quality, this format continued to be used for field recordings by the French collectors until 1936, primarily because it required no electricity.³⁵ The Jaap Kunst cylinders (1929-1933) were digitized by the Berliner Phonogramm Archiv³⁶ in the early 2000s but have been largely inaccessible until now.³⁷

Display of Results

Users can browse the results of their filtering either on an interactive map or as a list and can switch between the two views at any time. The list view allows users to sort search results according to parameters including title, recording date, location, and population name. This sorting function is particularly useful when a search yields a large number of results, as it enables users to quickly locate the most relevant items. Overall, the ability to toggle between map and list views, combined with customizable sorting, makes the filtering and browsing experience more tailored and efficient.

These filters demonstrate Pratinada’s potential and highlight its key advantages over commercial platforms and social networks that rely on basic keyword searches. Many users prioritize relevant, precise results and prefer not to be steered by algorithms de-

35 Musée de l’Homme’s last field collection recorded on cylinder: ‘Mission Algérie – Aurès 1936’: https://archives.crem-cnrs.fr/archives/corpus/CNRSMH_Cylindres_003/.

36 <https://www.sammlungen.hu-berlin.de/sammlungen/phonogramm-archiv/>

37 <https://jaapkunst.org/about/>

signed to capture attention for commercial purposes. Of course, the quality of the results generated by user queries depends on the quality of the metadata.

Sound Visualization

The Pratinada interface embeds the audio players from different content providers' interfaces alongside each item's metadata. As a result, the site presents several distinct audio players, each with its own playback controls, waveform display, and navigation features. The platform offers two different interaction paths to an audio player: from the map view, clicking on a location loads a list of titles, selecting one item then loads its basic metadata and an audio player (Figure 7); alternately, users can browse the results list from a simple search, scroll through the list and click on an item, thus displaying its metadata and audio player. From either path, the user can also expand the item to view its full metadata and embedded sound player.

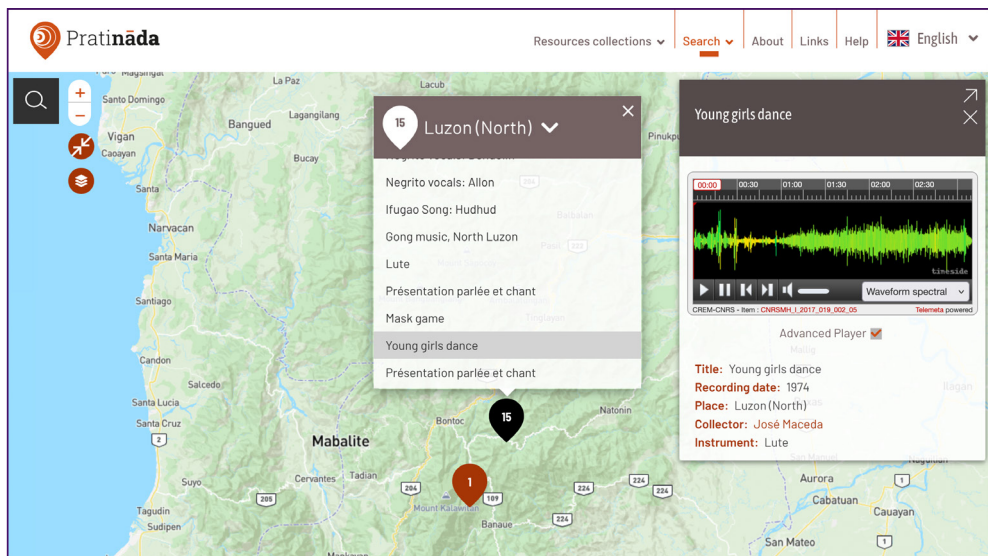


Figure 7. Basic metadata, geolocation and audio player for a J. Maceda recording in Luzon, Philippines.

The 'advanced player', which shows a visualized wave form, can be toggled on or off for any item with an available sound file. Today in Pratinada, most items utilize the player from the CNRS - Musée de l'Homme Sound Archives web platform,³⁸ which can include temporal markers. Experimentation with sound visualization tools in online databases, such as waveform and spectral displays, has proven helpful for sound archivists. The CNRS player has also encouraged greater engagement from other stakeholders who add time-coded annotations.

38 CNRS - Musée de l'Homme Sound Archives: <https://archives.crem-cnrs.fr/>. CREM is part of the Laboratoire d'ethnologie et sociologie comparative (LESC), Nanterre University.

Multilingualism

The interface’s front end is translated into nine languages, seven of which are Southeast Asian: Indonesian, Thai, Lao, Khmer, Viet, Tetun and Filipino, in addition to English and French (Figure 8). The main tabs and control buttons have also been translated, as has the home page text.



Figure 8. Interface with choice between 9 languages.

In selecting the Southeast Asia national languages to be included on the site, the needs of users from more recently established countries were considered. Pratinada contains 796 sound files from East Timor; thus Tetun, a language that was elevated to the status of an official language early in the 21st century, was included for the East Timor collections (Figure 9).

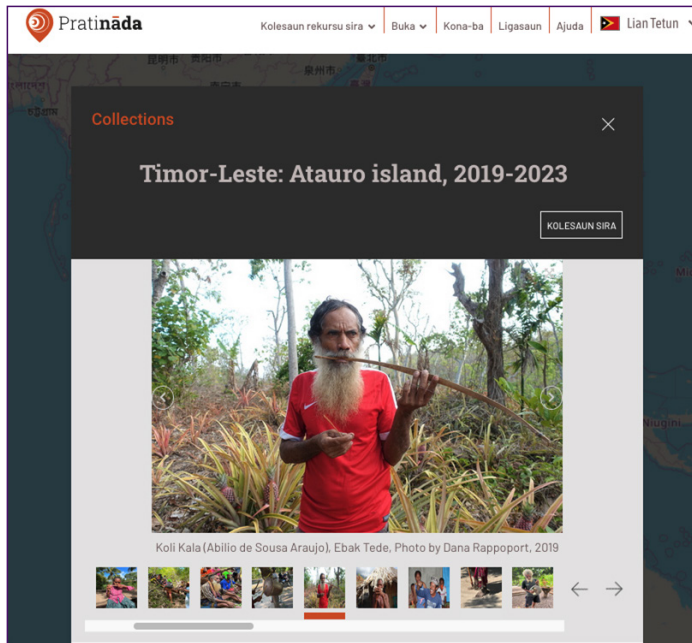


Figure 9. Timor-Leste Atauro Island collection, shown in Tetun language.

While some metadata (primarily titles) have been translated into the site's official languages,³⁹ this is not yet the case for most of the database. Translations were possible thanks to the work of ethnomusicologists involved in musical revitalization efforts who maintain local contacts to assist with descriptions.⁴⁰ In the future, full translations for more collections are planned.

Help and Legal and Ethical Usage Pages

The Help page⁴¹ provides advice on search and site functionality and also defines the metadata fields used on the website. The Pratinada team's experience managing ethnological databases with input from many different contributors has shown the importance of providing explicit definitions of fields. Clear metadata definitions will also streamline future database migrations and facilitate the reuse of information in other contexts.

Details about data management, intellectual property and data usage are given in the Legal and ethical usage page,⁴² which also alerts users that metadata may contain errors due to lack of information collected or changes in terminology or spelling over the last century. Descriptions are constantly evolving, thanks to contributions and corrections from researchers, local societies and partners in Southeast Asia, and to the ongoing

39 The items in some collections (e.g. Kalimantan, Flores, Sulawesi) are translated into Indonesian. Some, including those from Timor-Leste, Atauro, are given both in local languages and Tetun.

40 Pratinada's translators include Pierre Prouteau (Thai), Sèng Aloune Keovanthin (Lao), Michel Antelme (Khmer), Ignatius Aditya and Citra Aryandari (Indonesian), Y-Lin Lê (Vietnamese), Romeu Silva (Tetun), Philip Yampolsky (English), Dana Rappoport (French), Verne de la Peña, Elizabeth Enriquez, Roan May Opiso, Sol Maris Trinidad and Jose Mirabueno (Filipino).

41 <https://pratinada.net/help>

42 <https://pratinada.net/legal>

ing process of translating metadata into Southeast Asian languages. Translations into English or local languages may also contain errors.

Origin of the Pratinada Website

The Pratinada database is the culmination of a long development process and each of its component collections has its own unique story. The Jaap Kunst wax cylinder collection, physically located in Berlin, is managed by the University of Amsterdam, but the cylinders were not available online. In contrast, CNRS operates under open data policies (CNRS, 2020, pp. 6–8) and its curators have made web access to its sound archives a priority. The University of Amsterdam entrusted the CNRS team with the task of creating a common web sound database because of its previous experience with the Musée de l'Homme sound archives web platform, which runs on Telemeta software⁴³ (Simonnot, 2017, pp. 3–7). Telemeta was created in 2011 for the Musée de l'Homme, one of the world's richest collections of sound archives.

Since 1928, the Phonothèque of the Ethnographic Museum in Paris (which became the Département d'ethnologie musicale of Musée de l'Homme in 1937 and then the Centre de Recherche en Ethnomusicologie in 2009) has hosted a large number of sound collections deposited by French and foreign collectors. From the outset, the department housed a sound library with a large collection of 78 rpm records, augmented by cylinders from the 1900 Universal Exposition and from early field expeditions. As sound recording techniques evolved from magnetic tapes to digital files, an ever-growing amount of material collected in the field has been added to the collection, along with commercial LPs and CDs containing field recordings from researchers (Rouget, 2004, pp. 513–523; Gérard, 2014, pp. 192–215).

For decades, French and foreign ethnographers have carried out long-term research with communities in the Southeast Asian archipelago. Many researchers have donated their sound archives to be preserved at the French Musée de l'Homme, among them José Maceda, Christian Pelras, Rosario de Santos, Georges Condominas, Nicole Revel, Paul Wirtz and the group of French ethnographers known collectively as L'Equipe Timor (which included Louis Berthe, Claudine Friedberg, Brigitte Clamagirand and Henri and Maria-Olímpia Lameiras-Campagnolo). Between 1957 and 1970, L'Equipe Timor conducted monthslong fieldwork among the Buna', Ema, and Fataluku ethno-linguistic groups in Timor. Even though no one on this team was a musicologist, they recorded music as one of the important activities of the people they were studying. At that time, societies had not yet been changed by the arrival of electricity and the dissemination of amplified recordings. These ethnographers witnessed breathtaking artistic vitality that is preserved in their recordings.

The sound archives of the CNRS - Musée de l'Homme were not open to the public for many years, but only to researchers. Moreover, there was no plan to digitize the collections, which consisted of around 5,000 reel-to-reel tapes containing tens of thousands of audio tracks. Finally in 2000, a two-stage work plan was instituted. The first stage (2000–2009) was to digitize collections in urgent need due to the fragility of their carriers. Tapes were assigned metadata based on the paper descriptions in the archive. During the second stage (2009–2021) an online database was created, and metadata was integrated item by item, while the French National Library continued to digitize less fragile recordings.

43 Telemeta musicology media asset management system: <https://github.com/Parisson/Telemeta>; <https://parisson.github.io/Telemeta/>.

Convincing collectors to allow access to their archives, initially only dedicated to scientific research, was challenging. Some were reluctant due to the sensitivity of the content of spoken or sung words, while others hesitated because parts of the collections lacked proper metadata and were not ready for public release. Documenting the context of the collections in detail and extracting tracks for each piece of music was a time-consuming and painstaking process. Managing thousands of sound items and their associated documentation required an asset management system that linked each digital file to its metadata. To devise a new tool, first data needed to be migrated from obsolete proprietary database software.⁴⁴ The CNRS - Musée de l'Homme web platform was launched in 2011, the culmination of a long-term effort⁴⁵ to fund and develop a user-friendly interface dedicated to audiovisual data of traditional music and suitable for various types of users and contributors, as well as archivists.⁴⁶

As of 2025, the CNRS - Musée de l'Homme sound archives platform⁴⁷ contains 57,600 sound items, including 33,400 with open access. This database contains field recordings and published recordings from around the world, as well as their accompanying documentation. Access rights are managed according to profiles of users and contributors, who can add audio files and metadata themselves with a user account. Similarly, the ability for contributors to annotate sound with temporal markers (Figure 10) has also brought greater precision to the description of often complex content, attracting researchers in data sciences and Artificial Intelligence (Simonnot, 2020, pp. 26–28).

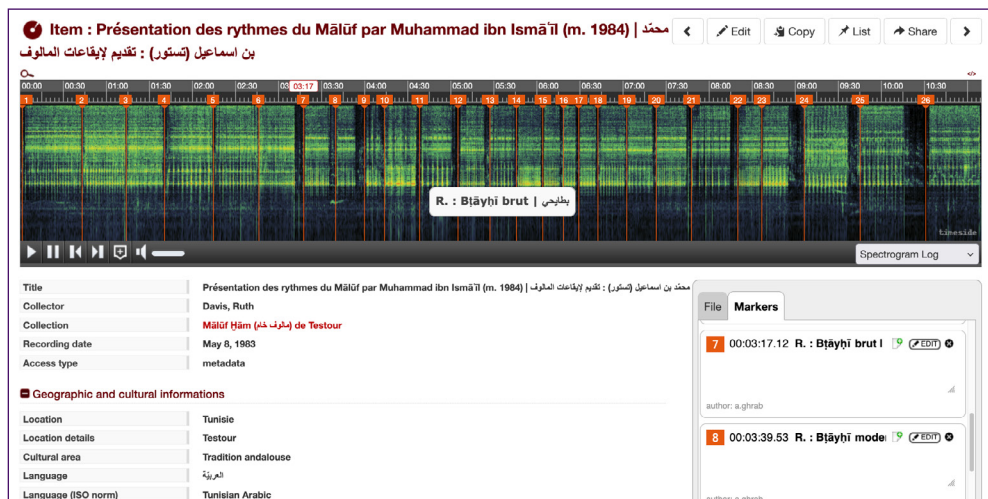


Figure 10. The CNRS - Musée de l'Homme Sound archives platform with annotated markers (orange numbers) for Tunisian Ma'luf music.

44 At the time, all the metadata had been entered in 4D Software, a proprietary format. Rosalia Martinez, director of CREM, advocated for a solution to the problem of retrieving and preserving the archive's metadata.

45 Led by the Centre de Recherche en Ethnomusicologie with Joséphine Simonnot, Aude Julien Da Cruz Lima, Guillaume Pellerin (Parisson startup).

46 Archiving platforms from this institution have mainly been designed for archivists, not for researchers or musicians.

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

Since 2011, the CNRS - Musée de l'Homme platform has proactively addressed the expectations of users from formerly colonized countries. Within the framework of Open Science in France, it has initiated a gradual restitution of its sound archives and attracts over 30,000 visitors annually, 70% of whom are from abroad (Simonnot, 2020, p. 97). Ongoing curation of the archives, the hosting of foreign students and dissemination of scientific papers have led to the establishment of contacts from other countries. This enhanced visibility has facilitated the repatriation of numerous collections by sending sound files to originating communities, exemplified by the return of archival content to Burkina Faso (Simonnot & Koudougou, 2019, pp. 4–5).

However, usage was constrained by using French as the primary language for description, despite the platform being localized into seven international languages including Arabic. As of 2025, metadata is still only available in French, limiting accessibility for international societies and researchers who may struggle to understand it. Although the site presents language barriers, users still manage to make meaningful use of it by identifying and listening to recordings relevant to their communities. For example, Marsel Lembang, a Toraja singer from Indonesia, found the D. Rappoport collection through an internet search. He used her archived recordings made in the 1990s to learn sung texts and revitalize Toraja rituals in 2015 (Rappoport, 2024).

Pratinada Methodology

Pratinada represents a new stage in the mission to preserve and share audio cultural heritage. Most of its sound data have been extracted from the CNRS - Musée de l'Homme sound archives web platform. To improve access to sound items, existing descriptions were enriched by the addition of geolocation and metadata in new languages. Like the Musée de l'Homme platform, Pratinada employs a multilingual interface, which includes several Southeast Asian languages. Access to the sound files depends on the functionality of the original databases, whereas the metadata are site-specific.

Joséphine Simonnot, who worked on the Europeana Sounds project from 2014–2017,⁴⁸ managed the process of aggregating metadata from each provider and the development of a common metadata model for Pratinada. This metadata model was created in partnership with Southeast Asian curators to follow the FAIR (findable, accessible, interoperable and reusable) principles,⁴⁹ and to be suitable for all providers' data. Cultural, geographical and linguistic criteria were emphasized. Schema.org's types and properties were utilized,⁵⁰ and several metadata templates were created in an Omeka S⁵¹ platform based on types including CreativeWorks, AudioObject, Person, Place, Organization and Collection. It is advantageous to utilize existing metadata schemas rather than creating new ones, and schema.org has additional benefits of greater expressivity and specificity than the Dublin Core metadata terms. Therefore, a clear and unambiguous distinction exists between concepts like 'accountable person', 'author' and 'performer'; and between 'country of origin' and 'location created'. Ambiguity in field names can be a source of confusion and error for contributions made by specialists who are rarely themselves archivists.

Another important phase in the work was to classify information into the appropriate descriptive fields. Collectors often compile various information into only one field to

48 <https://pro.europeana.eu/organisation/europeana-sounds>

49 https://en.wikipedia.org/wiki/FAIR_data

50 <https://schema.org/>

51 <https://omeka.org/s/>

save working time. Information parsing was essential to enabling filtering tools for the site. This substantial operation occupied the project team for two years.

Cultural and Ethical Issues in Pratinada

One of the project's goals was to decolonize European music archives by making them accessible to the public while balancing the concerns of the holding institutions. From a broader perspective, the aim was to share music archives recorded by Europeans in Southeast Asia during and after Colonialism. Most of the recordings in Pratinada were made after the respective countries became independent. After decolonization, Western collectors, ethnographers and ethnomusicologists came to Southeast Asia to record music that was rapidly vanishing.

Another aim was to advocate for a duty of restitution. During database development, legal and ethical issues were discussed with Southeast Asian partners who were part of a network established during the project.⁵² One result was a joint declaration on the decolonization of sound archives (DecoSEAS, 2024, p. 1). This text calls for, among other things, improvement of access to sound archives and diversification of dialogue about curation.

Pratinada's network has discussed many cultural and ethical issues in the development of this project. Who owns these recordings: the individuals recorded, their descendants, their community, their ethnic group, the local government, the provincial government, the national government? What rights are held by the researchers who collected them, or the institutions that preserve them? How can copyright and dissemination rights be managed fairly? Who should have access to these archives? How should the communities of origin have a say in the use of these recordings? Did the recorded individuals give informed consent? Was the collection context ethical, or was there an unequal power relationship? The network continues to contemplate these issues and to seek just policies that further the project's mission.

Data Protection

Pratinada's recordings, which are stored in public institutions in France, the Netherlands, and the United Kingdom, are managed by collectors, musicologists and/or archivists. Many of these recordings were created during the British, French and Dutch colonial periods and some of them are in the public domain. Content on the Pratinada website is also safeguarded by intellectual property laws.

In 2011, the European Union Directive (European Parliament and Council, 2011, article 1) extended the protections from 50 to 70 years for rightsholders including performers and collectors. As a result, the rightsholder's consent must be given for broadcast within 70 years of a work's initial performance or recording date. Recordings made before 1963 are in the public domain, according to European law. Pratinada also considers ethical issues, and access restrictions may be placed on recordings of sacred music or where private personal information is shared. Online access is based on the agreement of performers, researchers and depositors, and on the laws in force concerning performers' rights. Even though many recordings are technically in the public domain, project participants support the right of indigenous people to maintain, control, protect and develop their cultural heritage, traditional knowledge and traditional cultural expressions (DOCIP, 2025, Article 31:1).

52 The list of the partners can be consulted at: <https://www.decoseas.org/people/>.

According to the spirit of the Berne Convention, moral rights, such as the right to object to distortion prejudicial to honour or reputation, are generally interpreted to include performers and recordings, including in digital/online contexts (WIPO, 1979, Article 6bis). Copying and commercial reuse of content on Pratinada are prohibited unless permitted by rightsholders or the institution. In all cases, users are invited to contact the archival institution managing the content for any reuse requests, as the holding institution manages contacts with the rightsholders and collectors who produced the data.

In the context of research data, the content providers are responsible for data licensing and Pratinada data adheres to their choices for included collections. Creative Commons licenses are widely used when an agreement exists between the holding institution, the collectors and the performers. However, in many cases, anthropological research collections lack contact information for rightsholders, or the original rightsholders may have died, making it difficult to apply Creative Commons licenses. To facilitate the fair reuse of cultural data, standardized rights statements from RightsStatements.org are provided in the 'license' field. Twelve different options for materials that are more or less restrictive about data reuse are available. These statements are both human- and machine-readable (Rightsstatements.org, n.d.).

CREM has chosen to adopt rights statements that will be applied on a case-by-case basis over the next few years for the CNRS - Musée de l'Homme sound archives. Some collectors have chosen the 'In Copyright–Non-Commercial Use Permitted' statement to avoid improper reuse. RightsStatements.org offers additional options to clarify the use of online digital heritage, for example the 'No Copyright–Other Known Legal Restriction' statement,⁵³ which implies that the user must obtain permission from the responsible organization because although the recording is not restricted by copyright and/or related rights, other laws are known to impose restrictions on its usage.

In the United Kingdom, British partner SOAS obtained a specific agreement in 2023 from the BBC for this project which states that the data are under BBC copyright. Therefore, permission is necessary for any other use than consultation in Pratinada. Despite their recording dates (1932–1961), the recordings held by the BBC are not in the public domain.

For each item from the University of Amsterdam's Jaap Kunst collection, a copyright notice is provided.⁵⁴ Because the recordings were made between 1925 and 1933, they are in the public domain, but their use is subject to Indonesian laws including the constitutional recognition of Indigenous peoples' rights, legislation protecting Indigenous and customary law communities, and national cultural heritage laws aligned with the UNESCO World Heritage Convention. Thus, the rights statement 'No Copyright–Other Known Legal Restriction' was applied to communicate the need to respect the heritage of traditional societies.

In the absence of any commercial exploitation, the legal risks of granting open access are minimal for project archiving institutions, because there is no threat of financial loss for the rightsholders. Moreover, the public can only access compressed MP3 versions; uncompressed archival files are never made available, which helps limit the risk of com-

⁵³ <https://rightsstatements.org/page/NoC-OKLR/1.0/?language=en>

⁵⁴ An example of legal restrictions mentioned in the 'copyright notice' field: <https://pratinada.net/item/543>.

mercial misuse. If need be, users can request a high-resolution file from the rightsholder for a record or film production. Furthermore, a right of withdrawal of the sound files from free online access can be exercised by the performers recorded or their descendants by making a request for suspension of streaming to the institution providing access.

Finally, newly created metadata are under the responsibility of the project partners involved in putting them online. Consequently, available content is under the shared responsibility of the collectors, the archive holder and the project partners, including the Southeast Asian partner organizations involved. The field 'SEA annotation' (short for Southeast Asia annotation) is included in the metadata model to denote descriptions added by Southeast Asian contributors, and free comments may be added by a wider array of contributors.

Towards the Decolonization of Southeast Asian Sound Archives

Since UNESCO established a committee in 1978 to address the restitution of cultural property, cultural heritage decolonization has become an international political issue (UNESCO, Intergovernmental Committee for Promoting the Return of Cultural Property to its Countries of Origin or its Restitution in case of Illicit Appropriation, 1978). For Pratinada, the decolonization of archives depends above all on making data accessible: technically, by ensuring it can be discovered, viewed, and reused without technical barriers; and linguistically, by providing descriptions and interfaces in languages understood by the communities concerned. Secondly, the presentation and design of the web platform seek to take user needs into account. To this end, Southeast Asian partners were consulted through numerous meetings and workshops, including discussions about specific website features. Step by step, the interface functionalities were validated collaboratively: these included the website logo, the search index, the criteria for sound description and the name of the website. The name 'Pratinada' was chosen by vote from among seven Southeast Asian words meaning 'echo'. The Sanskrit word prevailed, a fitting choice given Sanskrit's historical influence throughout the Indonesian archipelago.

To ensure this audio heritage remains accessible, it is essential to raise its profile across Southeast Asia and collaboratively enhance any incomplete descriptions.⁵⁵ Encouraging active local involvement is key to this effort. Another aggregation operation is in progress with MyArchives (Malaysian Audio-Visual Archives), a project developed by the International Council for Traditions of Music and Dance's National Committee of Malaysia, based at Sunway University in Kuala Lumpur.

Deeper engagement with local stakeholders is also needed to gather meaningful feedback. At a workshop at the University of Jakarta in July 2024, Indonesian researchers testified to the importance of this data for their own research. Another event was organized in September 2024 in Kupang, Indonesia by the University of Amsterdam with the Institute for Resource Governance and Social Change to digitally repatriate the Jaap Kunst collections.⁵⁶ Additional social events may be held to encourage people to reuse historic cultural heritage, revitalize oral practices and stimulate artistic creation.

55 As an example, during the DeCoSEAS project, a board was created during a 2024 workshop in Denpasar to manage the database collectively by adding documentation and to encourage contributors to import new collections, in particular from Asia.

56 <https://www.decoseas.org/events/digital-repatriation-of-jaap-kunst-collection-to-east-nusa-tenggara-archive-and-library-service/>

Sustainability

The sustainability of the Pratinada platform relies on the long-term maintenance of its technical infrastructure, the active participation of its user community, and continuous development of its content. In the years ahead, a dedicated technical team at the French CNRS research center will support the platform's technical maintenance, ensure that the website is updated, that the server remains operational and data security is maintained.

The Pratinada website is more likely to persist if new users host and describe their data there. Additionally, maintaining database consistency with the collectively developed level of description needs to be upheld by future contributors. Importantly, the large volume of new metadata created since 2022 is easily exportable from the Omeka database and can be secured without incurring additional costs, a feature that is not available in some content management systems. Long-term preservation of audio files is managed by the holding institutions and this content will exist in the future regardless of Pratinada. However, as previously mentioned, the enhanced metadata created by users exists only within Pratinada, potentially creating an information gap in the holding institutions' databases. A future project could be initiated to update the original metadata in collaboration with curatorial institutions. Description is much more detailed and granular than what is available on platforms like YouTube, and the sovereignty of the originating archives is ensured. Maintenance and development costs, amounting to a few thousand Euros annually, are handled by partner institutions, such as university or museum libraries.

Conclusion

The Pratinada platform represents a significant breakthrough in cultural heritage preservation, making Southeast Asian sound archives held in European institutions accessible to international audiences. European funding has enabled the dissemination of these recordings under open data policies and FAIR principles, resulting in a predominantly open access collection. However, preserving this intangible heritage for future generations presents ongoing challenges. The sustainability of this comprehensive work requires attention at multiple levels—technical infrastructure, institutional support, and most critically, meaningful engagement with the local communities to whom this heritage belongs. The Pratinada team has identified further development tasks to strengthen the platform's capabilities. Translating all metadata into relevant local languages would significantly improve accessibility for Southeast Asian communities. Additionally, developing a multicultural thesaurus to manage diverse terminologies for instruments, sounds, and genres would create a more inclusive and comprehensive resource. The platform's future depends on securing continued technical maintenance, institutional backing, and community involvement. Southeast Asia's growing urban populations represent a crucial audience for these preserved sounds. By providing accessible digital access, Pratinada can facilitate artistic creation, educational applications, and historical research. The platform thus serves as more than an archive—it becomes an active tool for cultural transmission, helping to sustain endangered traditions that face erosion from rapid globalization.

Acknowledgements

We would like to thank Philip Yampolsky for his invaluable help with an earlier version of this paper.

References

- CNRS (2020). *CNRS Research Data Plan*. Paris: Centre national de la recherche scientifique. Available at: https://www.science-ouverte.cnrs.fr/wp-content/uploads/2021/04/Cnrs_Research-Data-Plan_mars21.pdf [Accessed: 2 August 2025].
- DeCoSeas (2024). *Joint Declaration – Decolonizing Southeast Asian Sound Archives*. Available at: <https://www.decoseas.org/manifesto/> [Accessed 2 August 2025].
- DOCIP (Indigenous Peoples' Centre for Documentation, Research and Information) (n.d.). *Culture and Education*. Available at: <https://www.docip.org/en/rights-and-freedoms/culture-and-education> [Accessed: 2 August 2025].
- European Parliament and Council (2011) *Directive 2011/77/EU of 27 September 2011 amending Directive 2006/116/EC on the Term of Protection of Copyright and Certain Related Rights*. *Official Journal of the European Union*, L 265, 11 October, pp. 1–5. Available at: <https://eur-lex.europa.eu/eli/dir/2011/77/oj> [Accessed: 2 August 2025].
- Gérard, B. (2012). 'Gilbert Rouget and the Ogooué-Congo Mission (1946). Institution and Epistemology in the History of Ethnomusicology in France', *Gradhiva* 16. Available at: <https://journals.openedition.org/gradhiva/2538> [Accessed 2 August 2025].
- Rappoport, D. (2024). *Entretien avec Pa' Marsel Lembang, Apprenti Chanteur Originaire de Toraja (Sulawesi, Indonésie)*. Available at: <https://nakala.fr/10.34847/nkl.d2c53o23> [Accessed 27 July 2025].
- RightsStatements.org (n.d.) *RightsStatements.org – Rights Statements (version 1.0)*. Available at: <https://rightsstatements.org/page/1.0/?language=en> [Accessed: 2 August 2025].
- Rouget, G. (2004). 'Le Département d'Ethnomusicologie du Musée de l'Homme', *L'Homme* 171–172. Available at: <https://journals.openedition.org/lhomme/24998> [Accessed 2 August 2025].
- Simonnot, J. (2017). 'The CNRS - Musée de l'Homme Sound Archives from 1900 to the Present: A Long Way Between Heritage, Knowledge and Technologies', in Ziegler S., Akesson I. and Lecheitner G. (eds.), *Historical Sources of Ethnomusicology in Contemporary Debate*. Newcastle: Cambridge Scholars Publishing, pp. 103–116. Available at: <https://shs.hal.science/hal-02638852v1> [Accessed 2 August 2025].
- Simonnot, J. (2020). 'Partager les Archives Sonores du Musée de l'Homme sur le Web avec la Plateforme Telemeta', *Bulletin de l'AFAS*, 46. Available at: <http://journals.openedition.org/afas/4056> [Accessed 1 August 2025].
- Simonnot, J., Koudougou, J.-P. (2019). 'La Restitution des Œuvres Culturelles Immatérielles: l'Exemple des Archives Sonores du Burkina Faso Conservées au Musée de l'Homme', *Patrimoines Africains: Réussir Ensemble Notre Nouvelle Coopération Culturelle*. Institut de France, 4 July. Available at: <https://shs.hal.science/halshs-02879952v1> [Accessed 1 August 2025].
- UNESCO, Intergovernmental Committee for Promoting the Return of Cultural Property to its Countries of Origin or its Restitution in case of Illicit Appropriation (1978). *Resolution Adopted at the First Session of the Intergovernmental Committee*, Paris, 4–7 December. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000114032> [Accessed: 2 August 2025].
- WIPO (1979). *Berne Convention for the Protection of Literary and Artistic Works (Paris Act of 24 July 1971, as amended on 28 September 1979)*. Geneva: World Intellectual Property Organization. Available at: <https://www.wipo.int/wipolex/en/text/283698> [Accessed: 2 August 2025].