



HAL
open science

THE HI-AUDIO ONLINE PLATFORM FOR DISTRIBUTED MUSIC CROWDSOURCING DATABASE COLLECTION

Jose Manuel Gil Panal, Aurélien David, Gael Richard

► **To cite this version:**

Jose Manuel Gil Panal, Aurélien David, Gael Richard. THE HI-AUDIO ONLINE PLATFORM FOR DISTRIBUTED MUSIC CROWDSOURCING DATABASE COLLECTION. Late Breaking Demo - International Society for Music Information Retrieval Conference (ISMIR), Nov 2023, Milan (Italie), Italy. hal-04265346

HAL Id: hal-04265346

<https://telecom-paris.hal.science/hal-04265346v1>

Submitted on 30 Oct 2023

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.

THE HI-AUDIO ONLINE PLATFORM FOR DISTRIBUTED MUSIC CROWDSOURCING DATABASE COLLECTION

José Manuel Gil Panal, Aurélien David, Gaël Richard
LTCI, Telecom Paris, IP Paris

ABSTRACT

We present in this paper the recent development of an on-line platform for musicians, researchers and an open community of enthusiasts of audio and music with a view to build a public database of music recordings from a wide variety of styles and different cultures. The data generated and collected will primarily be audio data, coming from various sources, including field recordings, existing datasets, and users' collaboration. The platform aims at gathering a distributed music crowdsourcing database collection where each music piece is built from asynchronous recordings of different tracks at remote sites. The complete tool and databases generated will be openly distributed for research purposes.

1. INTRODUCTION

Access to data, and in particular annotated data, is essential to learn the appropriate statistical models for the targeted applications. Many such datasets do exist for nearly all applications in Music Information Retrieval (MIR). They include for instance ENST-Drums [1], DALI [2], the Million song data set [3], Freesound [4], The Open Multitrack Testbed [5], MUSDB18 [6] or QUASI [7]. However, these datasets, while valuable, are either too small, lack annotation, are only weakly-annotated, or, in many cases, are only partially publicly available.

One of the major goals of the Hi-Audio online platform is to build a large, varied, multi-genre, multi-channel, multi-instrument publicly available music dataset for research that is suitable for a wide variety of MIR applications (such as music source separation, automatic track generation and mixing, cover version detection, music style transfer, ...). The Hi-Audio online platform permits to rely on a distributed and iterative music recording paradigm where the final music recordings will be semi-automatically mixed from individual tracks recorded at different remote individual sites by volunteer musicians. It is also important to note that it will be possible using the platform to straightforwardly extend an existing database with new tracks or recordings with a rapid impact on research in

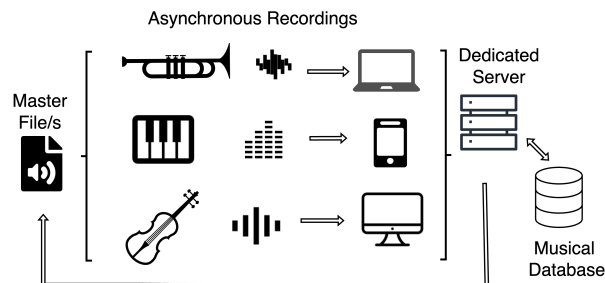


Figure 1. Overall recording scheme. The recording is conducted in a distributed way in different locations and the information is stored and annotated in a dedicated server. A master file, extracted from an existing recording can serve as a main synchronizing track with a count-down at the beginning or with a metronome signal

MIR. The second major goal of the Hi-Audio platform is to provide musicians, researchers and an open community of enthusiasts of audio and music with a virtual meeting point to practice and record music jointly with others (in an asynchronous mode) as experienced by many ensembles during the Coronavirus crisis in creating music videos from individual tracks recorded at different remote individual sites [8–10]. We also believe that the platform could be of particular interest for further developing the concept of “music repurposing” as discussed in [11] where individual musicians/composer add their own recorded track to a given music excerpt that is then gradually enriched with each new contribution.

Although the market for Digital Audio Workstations (DAWs) and online platforms for music creation and collaboration offers a wide range of possibilities and a rich and varied selection, with a somewhat smaller number of open source alternatives [12], we believe that our proposal goes one step further. In fact, it will allow free access to content and code and provide a connection point such as the public API, becoming an entirely new open ecosystem, sharing the same social view as Freesound for audio clips [13]. The overall recording scheme is illustrated on Figure 1.

2. ARCHITECTURE

To allow centralization of access to resources, a client-server architecture has been designed. The client and the server technologies are depicted on Figure 2.



© F. Author, S. Author, and T. Author. Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

Attribution: F. Author, S. Author, and T. Author, “The Hi-Audio Online Platform for distributed music crowdsourcing database collection”, in *Extended Abstracts for the Late-Breaking Demo Session of the 22nd Int. Society for Music Information Retrieval Conf.*, Online, 2023.

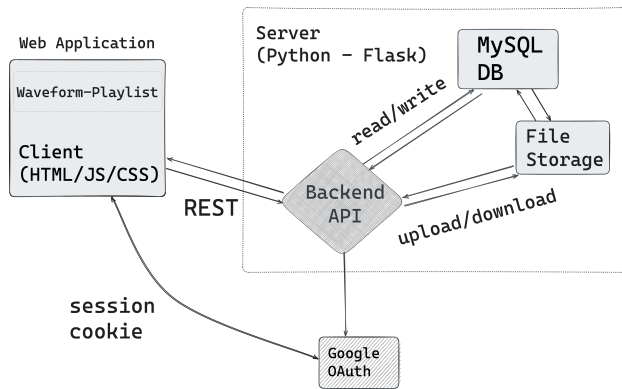


Figure 2. The diagram shows the different blocks and parts involved in the system interaction. The web application developed using JavaScript, HTML and CSS communicates with a dedicated server for data management and also with Google OAuth service to generate new user accounts and valid sessions within the platform. The back-end is used as an endpoint to store and retrieve information and files from the database.

2.1 Client

The client, an online web application, allows musicians to record and upload music files and collaborate with others along the internet by creating collections and compositions. The application, developed using standard web technologies such as Javascript, HTML and CSS, provides a friendly user interface, compatible with major modern browsers. It consists of a multitrack editor with basic capacities similar to traditional DAWs. Although there are many web services regarding music production workstations (e.g. Ampedstudio, Bandlab, Soundtrap [14]), most of the existing software is proprietary and its usage is limited. We then chose to rely on open source [15], to contribute and promote collaboration following the FAIR principles¹ of open science. Several existing open source libraries for audio interaction were considered. For instance, Gridsound² was pre-selected but finally not retained especially due its intricate user interface to display compositions. Instead, we finally chose a simple library called waveform-playlist³ [16, 17] which has the main features required for the audio functionalities (recording, playing, uploading and displaying tracks). We have tailored it to improve server communication and to allow users' interconnection.

2.2 Server

The aforementioned cloud solutions available in the market are not compatible with the idea of providing an online system to researchers to develop bespoke tools and transparent access to the information [18]. This motivated us to create our own server instance and an endpoint to facilitate an API that will be publicly available for anyone

willing to contribute, making the server module a key and differentiating component of our proposal. The aim of the developed ecosystem focuses on data collection applied to music creation, as in Beat Byte Bot [19]. The server side is built using the Python programming language and it is based on Flask⁴, a web application framework written in Python. The back-end provides a RESTful API that allows any client to connect to the server infrastructure and retrieve or exchange information via JSON data format, which is one of the most common formats for data exchange [20]. Regarding authentication and registration of users within the system, OAuth 2.0 flow is used through Google OAuth 2.0 and the requests are authorized in the API by a cookie session stored in the browser. The access to the information is also controlled by different roles and levels of privacy and users can decide which information they want to share and with which audience (from a single invited collaborator to everyone).

3. DISCUSSION AND FUTURE WORK

The Hi-Audio platform is still in an early stage of development and only part of the envisaged features will be demonstrated at the late breaking demo session. If the current platform is functional, important developments are still needed in particular regarding the user interface or the control of the recording latency, for example by replacing the MediaRecorder⁵ API with more efficient approaches like WebAssembly or AudioWorklet. The integration of classic DAWs features such as editing, sound effect and the integration of automatic annotation tools will be a major asset for building MIR research databases and is clearly on the development roadmap. Our main interest in this demo will be to gather first feedback from the MIR community and to identify potential collaborators for the platform development as amateur users, beta-testers or co-developers.

The platform will be temporally accessible using a dedicated login account during the conference at <https://hiaudio.fr>

4. CONCLUSION

In this paper we have presented the Hi-Audio Music Platform, our future open source online tool for researchers and performers in the music field. We explained the main goal of the developed platform which is to collect a large number of compositions in a multi track structure and distribute both the code and the data produced in a public way to share it with the audio and MIR community. We provided some details about the client-server architecture and technology stack chosen. We then discussed some of the immediate future work related to the web application that we plan to start intensively testing in the coming months.

¹ <https://www.go-fair.org/fair-principles/>

² <https://github.com/gridsound/daw>

³ <https://github.com/naomiaro/waveform-playlist>

⁴ <https://flask.palletsprojects.com/en/2.3.x/>

⁵ <https://developer.mozilla.org/en-US/docs/Web/API/MediaRecorder>

5. ACKNOWLEDGEMENT

This work was partly funded by the European Union (ERC, Hi-Audio, 101052978). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

6. REFERENCES

- [1] O. Gillet and G. Richard, “ENST-Drums: an extensive audio-visual database for drum signals processing.” in *Proceedings of the 7th International Conference on Music Information Retrieval*. Victoria, Canada: ISMIR, Oct. 2006, pp. 156–159. [Online]. Available: <https://doi.org/10.5281/zenodo.1415902>
- [2] G. Meseguer-Brocal, A. Cohen-Hadria, and G. Peeters, “Creating dali, a large dataset of synchronized audio, lyrics, and notes,” *Trans. of the Int. Soc. for Music Information Retrieval (TISMIR)*, vol. 3, no. 1, 2020.
- [3] T. Bertin-Mahieux, D. P. Ellis, B. Whitman, and P. Lamere, “The million song dataset,” in *Proceedings of the 12th International Conference on Music Information Retrieval (ISMIR 2011)*, 2011.
- [4] F. Font, G. Roma, and X. Serra, “Freesound technical demo,” in *Proceedings of the 21st ACM international conference on Multimedia*. , 2013.
- [5] B. De Man, J. D. Reiss *et al.*, “The open multitrack testbed: Features, content and use cases,” in *Proceedings of the 2nd AES Workshop on Intelligent Music Production*, vol. 13, 2016.
- [6] Z. Rafii, A. Liutkus, F.-R. Stöter, S. I. Mimitakis, and R. Bittner, “The MUSDB18 corpus for music separation,” Dec. 2017. [Online]. Available: <https://doi.org/10.5281/zenodo.1117372>
- [7] A. Liutkus, R. Badeau, and G. Richard, “Gaussian processes for underdetermined source separation,” *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3155–3167, 2011.
- [8] “<https://www.youtube.com/watch?v=mtjrnxdyr6e>.”
- [9] “<https://www.youtube.com/watch?v=nbfslttoxy>.”
- [10] “<https://www.youtube.com/watch?v=pkexicvotq>.”
- [11] M. Goto, *Grand Challenges in Music Information Research*. Multimodal Music Processing, Dagstuhl Follow-Ups, 2012.
- [12] M. Buffa and A. Vidal-Mazuy, “Wam-studio, a digital audio workstation (daw) for the web,” in *Companion Proceedings of the ACM Web Conference 2023*, 2023, pp. 543–548.
- [13] E. Fonseca, J. Pons Puig, X. Favory, F. Font Corbera, D. Bogdanov, A. Ferraro, S. Oramas, A. Porter, and X. Serra, “Freesound datasets: a platform for the creation of open audio datasets,” in *Hu X, Cunningham SJ, Turnbull D, Duan Z, editors. Proceedings of the 18th ISMIR Conference; 2017 oct 23-27; Suzhou, China.[Canada]: International Society for Music Information Retrieval; 2017. p. 486-93*. International Society for Music Information Retrieval (ISMIR), 2017.
- [14] F. Lind and A. MacPherson, “Soundtrap: A collaborative music studio with web audio,” in *Proceedings of the International Web Audio Conference*, ser. WAC ’17, F. Thalmann and S. Ewert, Eds. London, United Kingdom: Queen Mary University of London, August 2017.
- [15] M. Buffa, A. Hallili, and P. Renevier, “Mt5: a html5 multitrack player for musicians,” in *WAC 1st Web Audio Conference January 26-28, 2015-IRCAM & Mozilla Paris, France*, 2015.
- [16] W. Cai, A. Doshi, C. L. Miller, and S. Zhou, “Musi: Facilitating asynchronous practice between classical musicians,” in *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–6.
- [17] J. Pauwels and M. B. Sandler, “pywebaudioplayer: Bridging the gap between audio processing code in python and attractive visualisations based on web technology,” in *Proceedings of the International Web Audio Conference*, ser. WAC ’18, J. Monschke, C. Gut-tandin, N. Schnell, T. Jenkinson, and J. Schaedler, Eds. Berlin, Germany: TU Berlin, September 2018.
- [18] A. Srinivasamurthy, S. Gulati, R. C. Repetto, and X. Serra, “Saraga: Open datasets for research on indian art music,” *Empirical Musicology Review*, vol. 16, no. 1, pp. 85–98, 2021.
- [19] G. Panal and L. Arandas, “Beat byte bot: A chatbot architecture for web-based audio management,” in *Proceedings of the 11th Workshop on Ubiquitous Music (UbiMus 2021)*. g-ubimus, 2021, pp. 72–82.
- [20] A. Neumann, N. Laranjeiro, and J. Bernardino, “An analysis of public rest web service apis,” *IEEE Transactions on Services Computing*, vol. 14, no. 4, pp. 957–970, 2018.