



HAL
open science

HI-AUDIO ONLINE PLATFORM: OPPORTUNITIES AND CHALLENGES OF COLLECTING VARIED MUSIC DATA ON THE WEB

José Manuel Gil Panal, Aurélien David, Gael Richard

► **To cite this version:**

José Manuel Gil Panal, Aurélien David, Gael Richard. HI-AUDIO ONLINE PLATFORM: OPPORTUNITIES AND CHALLENGES OF COLLECTING VARIED MUSIC DATA ON THE WEB. International Society for Music Information Retrieval Conference (ISMIR), Late-Breaking Demo, 2024, San Francisco CA, United States. hal-04809004

HAL Id: hal-04809004

<https://telecom-paris.hal.science/hal-04809004v1>

Submitted on 28 Nov 2024

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.

HI-AUDIO ONLINE PLATFORM: OPPORTUNITIES AND CHALLENGES OF COLLECTING VARIED MUSIC DATA ON THE WEB

José Manuel Gil Panal, Aurélien David, Gaël Richard
LTCI, Telecom Paris, Institut polytechnique de Paris, France

ABSTRACT

We present in this paper the extended online HI-AUDIO platform which relies on a distributed and iterative music recording paradigm to asynchronously record musicians localised at different remote individual sites. The major goal of this platform is to become a key enabling tool for building a large, varied, multi-genre, multi-track, multi-instrument music dataset, to be ultimately publicly distributed for MIR research purposes. We describe in this paper the main characteristics of the web platform and discuss some of the major challenges for collecting music data on the web. The platform will be demonstrated on site with local and distant access and illustrate its merits for recording collaborative compositions.

1. INTRODUCTION

There is a growing interest for online Digital Audio Workstations (DAWs) [1,2] which are essential tools for recording music through the web. They open new opportunities [3,4] for musicians and composers: online DAWs are now sufficiently powerful to represent a decent alternative to traditional DAWs and can potentially be used anywhere, on any platform and any type of hardware. Though, most of the existing DAWs remain private [5–10], and only a few are distributed in open source [11,12]. They also face some important challenges in terms of robustness to the various browsers and operating systems, and in terms of flexibility or performance.

The common use of DAWs is to record and mix music in a potentially collaborative fashion. With the HI-AUDIO online platform, we open new opportunities, in particular, for building novel varied music databases for research purposes. Such databases are essential for building efficient machine learning models. In Music Information Retrieval (MIR), numerous datasets exist for a wide range of applications, including for example Freesound [13], ENST-Drums [14], DALI [15], MUSDB18 [16] or the Million Song Dataset [17] (see also [18] for an extensive list of datasets for MIR applications). While valuable, these datasets of-

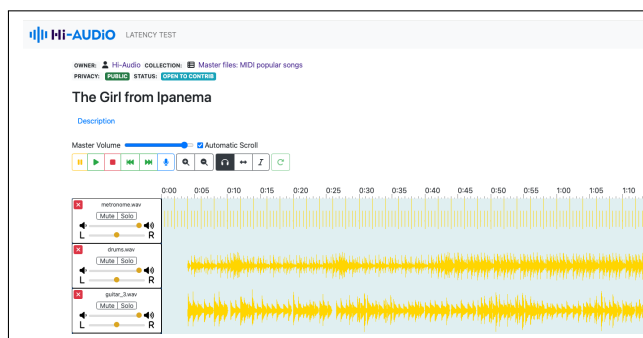


Figure 1. The HI-AUDIO platform UI multitrack editor based on waveform-playlist [20]

ten face limitations in terms of volume, annotation quality, variety in genre and style, or public accessibility. For some applications such as music source separation or music style transfer, it is of particular interest to have access to multi-track recordings without any bleeding between tracks, in a variety of styles or versions.

The major promise of the HI-AUDIO platform is to become a key tool, distributed in open source, for building a large, varied, multi-genre, multi-track, multi-instrument publicly available music dataset. To that aim, the core idea is to rely on a distributed and iterative music recording paradigm where the individual tracks are recorded at different remote sites by volunteer musicians.

The aim of this paper is to describe the main advances of the HI-AUDIO platform¹. We also briefly describe how to use the platform including details on the mechanism implemented for the automatic synchronisation of tracks played asynchronously from different locations.

2. MAIN ADVANCES

To address some of the challenges discussed above, the HI-AUDIO platform was updated with some new features:

Secured/private registering: Besides the initial possibility to sign in with a Google account, it is now possible to register with any email provider following a *two steps authentication*. The user receives a 6 digits code which then needs to be introduced in a verification screen. For both sign-in processes, we do not need to store any user's password.

Compression on server-side: While the original audio files are saved without compression on the server, they are

¹ A very preliminary version of the platform was demonstrated in the 2023 ISMIR's Late-Breaking Demo (LBD) session [19].



downloaded to the web client in a compressed format (with AAC [21]). This significantly enhances the loading speed.

Annotations: An intuitive user interface to annotate track information at different levels is now provided.

Latency estimation which is one of the major challenges addressed. In fact, the latency highly depends on the web browsers and operating systems, and their current workload, capabilities, memory or battery consumption. See section 4 for further details.

Data privacy, copyright/Intellectual Property: which are major issues especially in music. See section 5 for details.

3. USING THE PLATFORM

We describe in this section the main steps for creating or contributing to a composition. The platform is publicly accessible at <https://hiaudio.fr> but it is necessary to register to contribute to the platform.

Opening a composition and recording a track. Once a composition is opened, the platform displays a classic, although simple, multitrack web-based DAWs interface (see Figure 1). If playing the current file is readily possible at this stage, the key step to allow recording is to evaluate the round trip latency between the playback and record actions (see section 4). Then, when the overall delay is known, it is possible to play the master file (potentially along with other previously recorded tracks) and to record a new track which is automatically synchronised and saved along the other existing tracks.

Creating master files. To gather the music database, we rely on the concept of Master files that are by default the track on which a musician can play along. Master files can be of different types including for instance the recording of a bass line with a metronome, the recording of a rhythm section, a full music recording that can serve as a guide for re-recording each track or a MIDI version based of such a recording. Initially, the platform integrates a small but varied set of master files, but any registered user can expand this initial list and create new master files, in any style.

4. LATENCY TEST

As recalled above, to record a new track, the musician directly plays on the master file. Then, one of the major challenges in this distributed and asynchronous recording setup is to assess the delay of the current recording. This round-trip delay must be estimated precisely, to within a few milliseconds. The estimation requires using the audio input (microphone) to record the output (loudspeakers/headphones) assuming that they are very close to each other and that the volume is sufficiently high.

Unlike some other platforms which use audio clicks or sinusoidal beeps as test signals to estimate the overall latency [22–24], we opt for a novel and more robust approach exploiting Maximum Length Sequences (MLS)². Such sequences are pseudo-random noises which have the

²<https://github.com/gilpanal/weblatencytest>

property that their linear autocorrelation approximates a Kronecker delta. They can, therefore, be used to robustly estimate the round trip delay [25]. The confidence of the estimation is assessed by computing the following ratio:

$$R = 10 \log_{10} \left(\frac{\max_{\tau} [C(\tau)^2]}{\frac{1}{N} \sum_1^N C(\tau)^2} \right) \quad (1)$$

where $C(\tau)$ is the crosscorrelation between the input and output signals at time lag τ and N is the number of samples of the crosscorrelation (typically $N = 22050$ for signals at $44.1kHz$). In practice, a pre-defined fixed threshold set to $+18dB$ is defined to consider the test successful.

5. PRIVACY, COPYRIGHTS AND TERMS OF USE

Specific care has been taken to handle privacy and data copyrights. The HI-AUDIO does collect and processes (1) profile data, (2) data related to the music uploaded and (3) data created while contributing to the community. This data is collected for the purpose of creating the user account and to provide the user with our service. The security and confidentiality of users data are maintained in accordance with applicable law and will not make personal data public other than on the site (which is limited to the automatically generated username).

By completing and submitting personal data to the platform, the user declares to have read and accepted the terms of use and of the privacy policy directly accessible on the platform. The user may revoke consent at any time. Except for some master files, everything that is recorded or uploaded to our platform is registered as Creative Commons [26] with CC-BY 4.0 license (see the intellectual property section in the terms of use [27]).

As long as the master files are concerned, it is worth mentioning that the aim is to only have copyright free master files. Nevertheless, even for copyrighted master files, their use can be considered to be within the legal doctrine of *fair use*: they are only used for research purpose, are not distributed but only used as a guide for the recording of new tracks, cannot be downloaded from the platform and are on the platform in a degraded audio quality.

6. CONCLUSION

The HI-AUDIO platform, which shall be distributed in open source, is still in an early stage of development. Nevertheless, it is now ready for initiating the collection of the targeted multi-track, multi-style dataset for MIR research. The platform will be demonstrated live at the late-breaking demo session, with local and distant connections.

7. ACKNOWLEDGEMENT

This work was 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.

8. REFERENCES

- [1] C. Holbrow, "Turning the daw inside out," in *Audio Engineering Society Convention 146*. Audio Engineering Society, 2019.
- [2] S. Knapp, "Browser-based engine prototype for a digital audio workstation," 2024. [Online]. Available: <https://hdl.handle.net/20.500.12738/16337>
- [3] M. Buffa, J. Lebrun, S. Ren, S. Letz, Y. Orlarey, R. Michon, and D. Fober, "Emerging w3c apis opened up commercial opportunities for computer music applications," in *The Web Conference 2020-DevTrack*, 2020.
- [4] J. F. M. Sánchez-Jara, S. G. Gutiérrez, S. O. Migueláñez, and M. A. G. Malheiro, "Digital platforms for music production (daw): educational innovation from music teacher training: Plataformas digitales de producción musical (daw): innovación educativa desde la formación docente del profesorado de música," *Electronic Journal of Music in Education*, no. 52, pp. 53–72, 2023.
- [5] 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.
- [6] Kompoz. Kompoz. [Online]. Available: <https://www.kompoz.com/>
- [7] A. Studio. Amped studio. [Online]. Available: <https://ampedstudio.com/>
- [8] Bandlab. Bandlab. [Online]. Available: <https://www.bandlab.com>
- [9] AudioTool. Audiotool. [Online]. Available: <https://www.audiotool.com/>
- [10] Soundation. Soundation. [Online]. Available: <https://soundation.com/online-studio>
- [11] M. Buffa and S. Demont, "Can you daw it online?" in *IEEE International Symposium on the Internet of Sounds - Workshop on the Musical Metaverse*, 2024.
- [12] M. D. Thomas Tortorini. Gridsound. [Online]. Available: <https://github.com/gridsound>
- [13] F. Font, G. Roma, and X. Serra., "Freesound technical demo," in *Proceedings of the 21st ACM international conference on Multimedia*. , 2013.
- [14] 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>
- [15] 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.
- [16] Z. Rafii, A. Liutkus, F.-R. Stöter, S. I. Mimilakis, and R. Bittner, "The MUSDB18 corpus for music separation," Dec. 2017. [Online]. Available: <https://doi.org/10.5281/zenodo.1117372>
- [17] 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.
- [18] Resources/datasets listed from the mir-datasets repository. [Online]. Available: <https://ismir.net/resources/datasets/>
- [19] J. M. Gil-Panal, A. David, and G. Richard, "The HI-AUDIO online platform for distributed music crowdsourcing database collection," in *Late Breaking Demo - International Society for Music Information Retrieval Conference (ISMIR)*, Milan, Italy, 2023.
- [20] N. Aro. Waveform playlist. [Online]. Available: <https://github.com/naomiaro/waveform-playlist>
- [21] I. J. S. 29, *ISO/IEC 14496-3:2009, Information technology – Coding of audio-visual objects – Part 3 : Audio*. ISO/IEC, Sep 2009, vol. 4th Ed.
- [22] Y. Wang, R. Stables, and J. Reiss, "Audio latency measurement for desktop operating systems with onboard soundcards," in *128th Convention of Audio Engineering Society, London*, 2010.
- [23] 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, pp. 543–548.
- [24] U. Hammarqvist. Audio latency in browser-based daws. [Online]. Available: <https://www.w3.org/2021/03/media-production-workshop/talks/ulf-hammarqvist-audio-latency.html>
- [25] D. Sarwate and M. Pursley, "Crosscorrelation properties of pseudorandom and related sequences," *Proceedings of the IEEE*, vol. 68, no. 5, pp. 593–619, 1980.
- [26] A. Xambó, F. Font, G. Fazekas, and M. Barthet, "Leveraging online audio commons content for media production," in *Foundations in Sound Design for Linear Media*. Routledge, 2019, pp. 248–282.
- [27] Terms and conditions of the HI-AUDIO platform. [Online]. Available: <https://hiaudio.fr/static/terms.html>