# A 2 Million Commercial Song Interactive Navigator


Michel Buffa
Jerome Lebrun
Université Côte d'Azur
CNRS, INRIA
{buffa, lebrun}@i3s.unice.fr

Johan Pauwels
Queen Mary University of London
j.pauwels@qmul.ac.uk

Guillaume Pellerin
IRCAM
guillaume.pellerin@ircam.fr



## ABSTRACT

In this paper, we present a web-based interactive tool for exploring a collection of two million commercially released songs. It gathers song information from a large number of heterogeneous sources, web-based and audio-based, and integrates work from multiple research groups. The resulting tool can be used to request information about a specific song such as lyrics, metadata and chords; to navigate further on to linked external resources such as Discogs, AllMusic, MusicBrainz or a number of streaming providers; or to browse the collection by artist's discographies or band membership. Several Web Audio applications are integrated and use the dataset to enrich the experience.


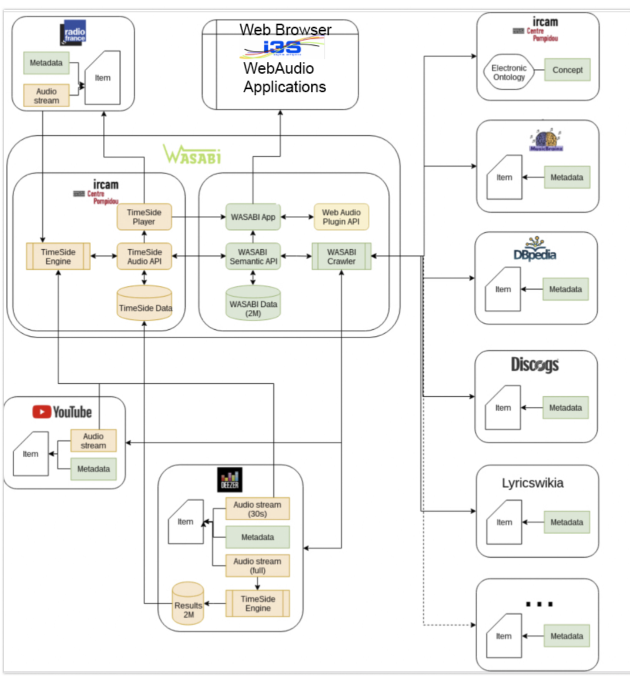


**Figure 1: a schema of the WASABI network architecture and workflow.**




*Web Audio Conference WAC-2019, December 4–6, 2019, Trondheim, Norway.*




## 1. INTRODUCTION

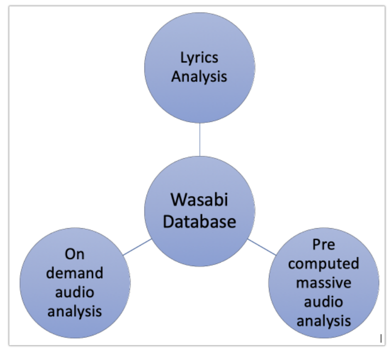


**Figure 2: the metadata of the WASABI database is complemented by computational analysis of lyrics and audio, both on-demand and precomputed.**

Since 2017, a 2M song database consisting of metadata collected from the Web of Data [4] has been constructed in the context of the WASABI research project[1]. The heterogeneous sources that have been consulted are displayed on the right of Figure 1.

These metadata include the identifiers of the corresponding audio on a variety of audio platforms, which allowed to enrich the database with computational analyses of the lyrics and audio data (see Figure 2). Several research groups have contributed to the analysis and have built interactive Web Audio applications on top of the output. For example, IRCAM linked their TimeSide analysis and annotation tool to make on-demand audio analysis possible. Queen Mary University of London, through the FAST project[2], performed an offline chord analysis of 442k songs, and built an online enhanced audio player [15] and chord search engine [16] around it. The I3S laboratory, responsible for the design and implementation of the database, extracted song structure based on lyrics analysis [10,12]. Furthermore, they have developed a rich set of Web Audio applications and plugins [6,7] that allow, for example, songs to be played along with sounds similar to those used by artists.

These metadata, computational analyses and Web Audio applications have now been gathered in one easy-to-use web interface, the WASABI Interactive Navigator, which is presented in this paper.

[1] http://wasabihome.i3s.unice.fr/
[2] http://www.semanticaudio.ac.uk

## 2. RELATED WORKS

Other research projects aimed to collect metadata on a large set of commercial songs, such as The Million Song Dataset project from 2011, which is mainly based on audio data [3] and did not exploit the availability of large structured data sources from the Web of Data to the full extent.

MusicWeb and its successor MusicLynx [1] link music artists within a Web-based application for discovering connections between them and provides a browsing experience using extra-musical relations. The project shares some ideas with WASABI, but works on the artist level, and does not perform analyses on the audio and lyrics content itself. It reuses, for example, MIR metadata from AcousticBrainz.

The WASABI project has been built on a broader scope than these projects and mixes a wider set of metadata, including ones from audio and natural language processing of lyrics. In addition, as presented in this paper, it comes with a large set of Web Audio enhanced applications.

## 3. THE WASABI INTERACTIVE NAVIGATOR HOMEPAGE

The primary starting point for the interactive navigator is the home screen as seen in Figure 3. From here, a textual search can be performed for one of the 2 million songs, 200k albums or 77k artists included in the database [4]. The resulting page displays information aggregated from a huge set of online sources, as shown in Figure 4.

A set of tabs at the top link directly to the tools based on the computational analysis of audio and lyrics, which will be discussed next.

## 4. SONG STRUCTURE DERIVED FROM LYRICS AND AUDIO

On the pages for individual songs, the musical structure is indicated by grouping the lyrics in blocks, as shown in Figure 5. Song lyrics contain repeated patterns that facilitate automated segmentation, with the detection of constitutive elements of a song text (e.g., intro, chorus) as goal.

We proposed to segment lyrics by applying a convolutional neural network to a synchronized audio-text representation of a song. First, we created a corpus projecting the segmentation of the lyrics of the WASABI corpus onto a synchronized lyric-audio corpus (DALI corpus[3]). We have shown that the information in the text enriched with the characteristics of the audio signal allows our segmentation model to surpass the state of the art method, which is based solely on textual characteristics [10, 12].

## 5. CHORD PLAYER AND SEARCH

Chord transcriptions can be requested from the song pages, and are presented as an interactive player showing the chords in sync with the music, such that musicians can play along with them (see Figure 6, top). As an alternative entrypoint to the WASABI database, a chord searching interface is available to find specific songs that contain a certain set of chords (bottom of Figure 6). This type of search interface has been previously tested with the Jamendo music collection [16].

[3]https://github.com/gabolsgabs/DALI

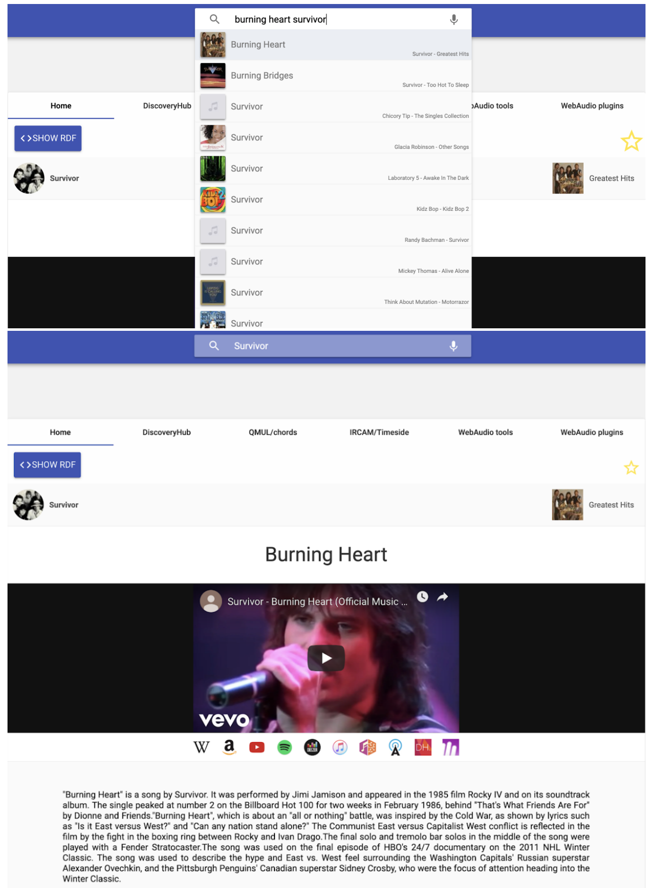


**Figure 3: main page of the WASABI navigator interface.**

For the search interface to work, the songs need to be indexed offline. So far, 442k files have been analysed with the chord analysis algorithm proposed in [14]. In order to comply with copyright requirements, a remote processing toolchain has been set-up in which algorithms packaged as a Docker container are sent to Deezer for processing on their servers. The output is then returned to us and stored in the WASABI database. Calculation of further music descriptors such as tempo and key is ongoing.

## 6. ON-DEMAND AUDIO ANALYSIS

On-demand audio analysis is complementary to precomputed analysis in the sense that it avoids large, upfront computational costs and scales easily to changes in data and algorithms. It comes with its own challenges, however. Building a web platform that depends on various external services requires that the underlying software architecture and data model must be robust against disappearing sources. For example, if a YouTube video gets removed, we still want its computational analysis to remain available in case any of the web services built on top refer to the track and its related metadata. The TimeSide framework[4] has been designed to provide a RESTful API as well as plugin based core library dedicated to audio processing that can

[4]https://github.com/Parisson/TimeSide/


**Figure 4: a large set of online databases has been exploited for building the metadata part of the WASABI database.**

provide this resilience.

In the context of the WASABI project, we demonstrate how the service (see Figure 1) can be used on demand by a master semantical application that will feed TimeSide with URLs coming from various providers (YouTube, Deezer, etc.) and then dynamically return the analysis data to the client application to feed Web Audio based tools. A number of audio-based characteristics can be requested from a song's page in this way.

## 7. MULTITRACK PLAYER AND EFFECTS

In order to assist music schools, which is one of the target audiences of the WASABI project, some Web Audio tools have been integrated into the Navigator. One possible scenario is that music teachers pick a song from the database for their students to learn. An enhanced multi-track player (Figure 7) is presented to the users, which displays the song's sheet music and allows to play back or selectively mute the different instruments in the song. Furthermore, a simplified DAW is available in the browser for recording and playing back student or teacher performances. It includes real-time audio effects, with ready to use presets, to attain a realistic and attractive studio sound.

The in-browser audio effects have been implemented using a Web Audio plugin standard [8, 9] (including SDK, online validation tools and examples of host applications), for which a large set of plugins[5] already exists. An online IDE for designing these plugins [11, 13] and a guitar tube amp simulator designer [6, 7] are also available for developers.

## 8. FUTURE ENHANCEMENTS

Fine-tuning the parameters in many of our Web Audio applications (tube amplifiers, pedals, etc. . . ) has been quite tedious and time-consuming with a high level of expertise

[5] https://github.com/micbuffa/WebAudioPlugins

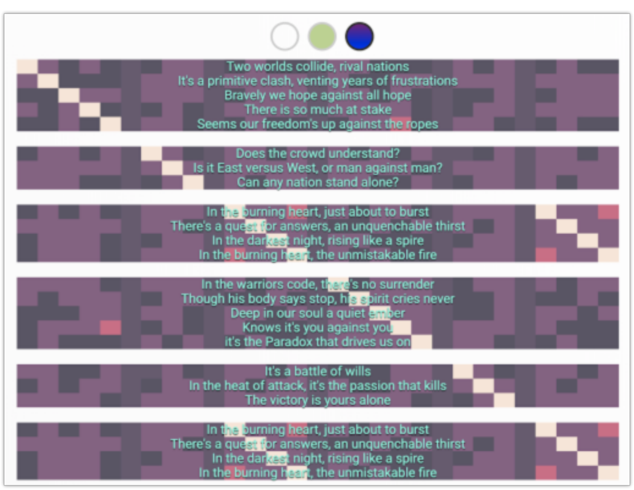


**Figure 5: song lyric structure detection [10].**

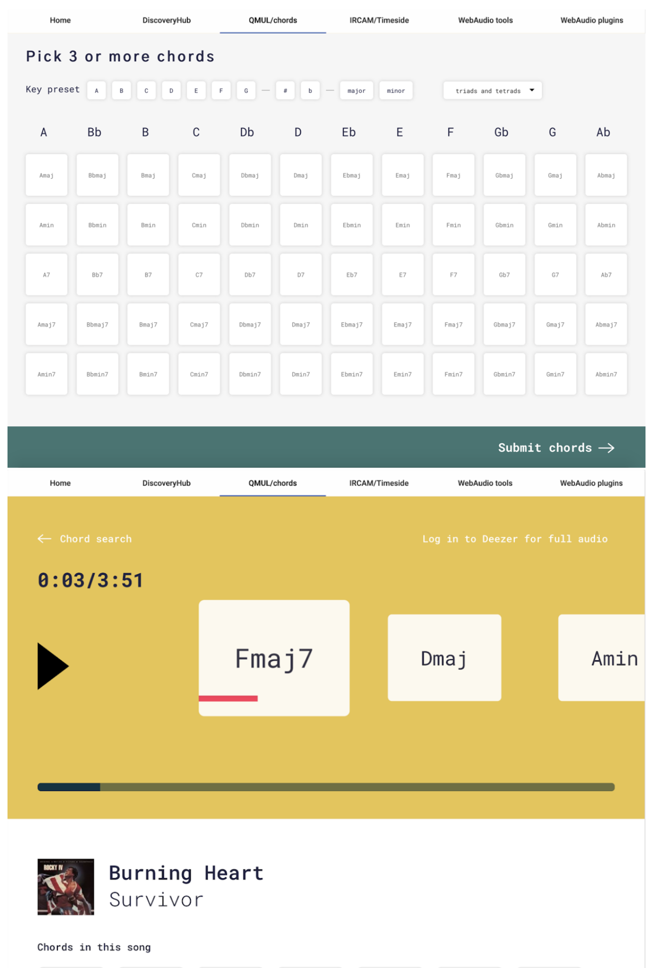


**Figure 6: augmented player and search engine, integrated in the navigator interface.**

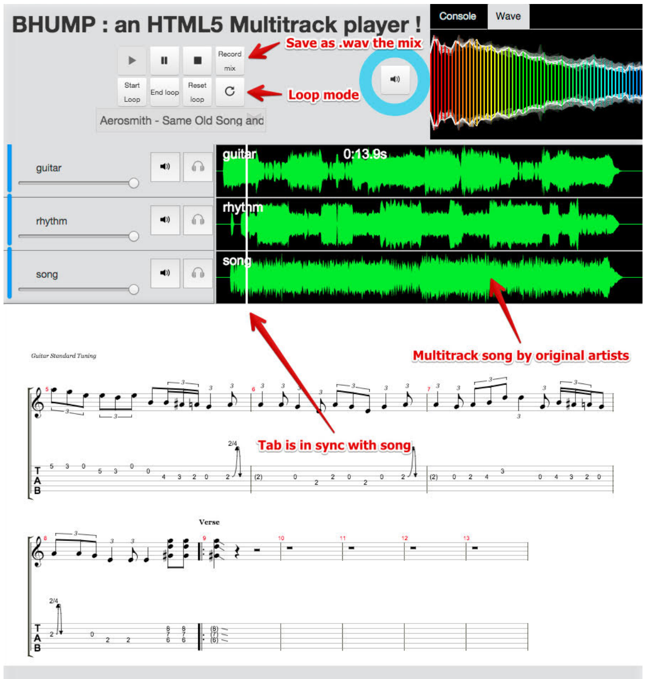


**Figure 7: a multitrack player capable of displaying Guitar Pro music tabs in sync with the audio [5].**

required [6, 7]. To ease this process of presetting for music school students, we are currently exploring machine-learning based approaches (with scattering wavelet based convolutional neural networks [2]) to learn and extract the relevant features from large datasets of songs and to match these with presets leading to similar subjective timbre in simulated instruments and also to automatically classify songs in the WASABI database.

Furthermore, offline indexation of more audio-based characteristics is ongoing.

## 9. ACKNOWLEDGMENTS

This work was supported by the French Research National Agency (ANR) and the WASABI team (contract ANR-16-CE23-0017-01). Part of this work has been funded by UK EPSRC grant EP/L019981/1.

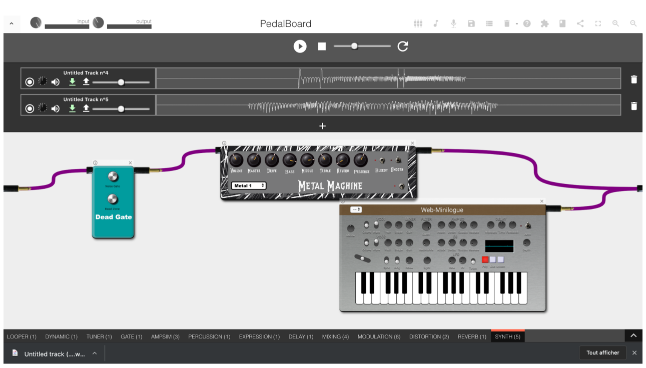


**Figure 8: Web Audio embedded DAW and effect chain for recording and playback.**

## 10. REFERENCES


[1] A. Allik, F. Thalmann, and M. Sandler. MusicLynx: Exploring music through artist similarity graphs. In *Companion Proc. (Dev. Track) The Web Conf. (WWW 2018)*, 2018.

[2] J. Anden and S. Mallat. Deep scattering spectrum. *IEEE Trans. Sig. Proc.*, 62(16), 2014.

[3] T. Bertin-Mahieux, D. Ellis, B. Whitman, and P. Lamere. The Million Song Dataset. In *Proc. 12th Int. Soc. Music Information Retrieval (ISMIR 2011)*, 2011.

[4] M. Buffa et al. WASABI: a two million song database project with audio and cultural metadata plus WebAudio enhanced client applications. In *Proc. 3rd Web Audio Conf. (WAC 2017)*, 2017.

[5] M. Buffa, A. Hallili, and P. Renevier. MT5: a HTML5 multitrack player for musicians. In *Proc. 1st Web Audio Conf. (WAC2015)*, 2015.

[6] M. Buffa and J. Lebrun. Real time tube guitar amplifier simulation using webaudio. In *Proc. 3rd Web Audio Conference (WAC 2017)*, 2017.

[7] M. Buffa and J. Lebrun. Web audio guitar tube amplifier vs native simulations. In *Proc. 3rd Web Audio Conf. (WAC 2017)*, 2017.

[8] M. Buffa, J. Lebrun, J. Kleimola, O. Larkin, and S. Letz. Towards an open web audio plugin standard. In *Companion Proc. (Dev. Track) The Web Conf. (WWW 2018)*, 2018.

[9] M. Buffa, J. Lebrun, J. Kleimola, O. Larkin, S. Letz, and G. Pellerin. WAP: Ideas for a Web Audio plug-in standard. In *Proc. 4th Web Audio Conf. (WAC 2018)*, 2018.

[10] M. Fell, Y. Nechaev, E. Cabrio, and F. Gandon. Lyrics segmentation: Textual macrostructure detection using convolutions. In *Proc. 27th Int. Conf. on Computational Linguistics (COLING2018)*, 2018.

[11] S. Letz, Y. Orlarey, and D. Fober. Compiling Faust audio DSP code to WebAssembly. In *Proc. 3rd Web Audio Conf. (WAC 2017)*, 2017.

[12] G. Meseguer-Brocal, A. Cohen-Hadria, and G. Peeters. DALI: A large dataset of synchronized audio, lyrics and notes, automatically created using teacher-student machine learning paradigm. In *Proc. 19th Int. Soc. of Music Information Retrieval (ISMIR 2018)*, 2018.

[13] Y. Orlarey, D. Fober, and S. Letz. Syntactical and semantical aspects of Faust. *Soft Computing*, 8(9), 2004.

[14] J. Pauwels, K. O'Hanlon, G. Fazekas, and M. Sandler. Confidence measures and their applications in music labelling systems based on hidden Markov models. In *Proc. 18th Int. Soc. Music Information Retrieval (ISMIR 2017)*, 2017.

[15] J. Pauwels and M. Sandler. A web-based system for suggesting new practice material to music learners based on chord content. In *Joint Proc. 24th ACM IUI Workshops (IUI2019)*, 2019.

[16] J. Pauwels, A. Xambó, G. Roma, M. Barthet, and G. Fazekas. Exploring real-time visualisations to support chord learning with a large music collection. In *Proc. 4th Web Audio Conf. (WAC 2018)*, 2018.