



Audio Engineering Society Convention e-Brief 165

Presented at the 137th Convention
2014 October 9–12 Los Angeles, USA

This Engineering Brief was selected on the basis of a submitted synopsis. The author is solely responsible for its presentation, and the AES takes no responsibility for the contents. All rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Audio Engineering Society.

The Open Multitrack Testbed

Brecht De Man¹, Mariano Mora-Mcginity², György Fazekas¹, Joshua D. Reiss¹

¹*Centre for Digital Music, School of Electronic Engineering and Computer Science, Queen Mary University of London, Mile End Road, London E1 4NS, United Kingdom*

²*Cognitive Science, School of Electronic Engineering and Computer Science, Queen Mary University of London, Mile End Road, London E1 4NS, United Kingdom*

Correspondence should be addressed to Brecht De Man (b.deman@qmul.ac.uk)

ABSTRACT

We introduce the Open Multitrack Testbed, an online repository of multitrack audio, mixes or processed versions thereof, and corresponding mix settings or process parameters such as DAW files. Multitrack audio is a much sought after resource for audio researchers, students and content producers, and while some online resources exist, few are large and reusable, and none allow querying audio fulfilling specific criteria. The test bed we present contains a semantic database of metadata corresponding with the songs and individual tracks, enabling users to retrieve all pop songs featuring an accordion, or all tracks recorded in reverberant spaces. The open character is made possible by requiring that the contributions, mainly from educational institutions and individuals, have a Creative Commons license.

1. INTRODUCTION

Many types of audio research rely on multitrack recorded or otherwise generated audio (and sometimes mixes thereof) for analysis or for demonstration of algorithms. In this context, a ‘mix’ denotes a summation of processed versions of (a subset of) these tracks, that can itself be processed as well. The availability of this type of data is of vital importance to many researchers, but also useful for budding mix engineers looking for practice material, audio educators, developers, as well as musicians or creative professionals in need of accompanying music or other audio where some tracks can be disabled.

Among the types of research that require or could benefit from a large number of audio tracks, mixes and/or processing parameters, are analysis of production prac-

tices [1, 2], source separation, automatic mixing [3–7], automatic multitrack segmentation [8], applications of masking and other auditory phenomena, and others that we haven’t thought of yet. The testbed was initially compiled with analysis of music production practices and automatic mixing applications in mind.

Existing online resources of multitrack audio content have a relatively low number of songs, show little variation in content, contain content of which the use is restricted due to copyright, provide little to no metadata, rarely have mixed versions including the parameter settings, and/or do not come with facilities to search the content for specific criteria.

[9] offers 24 different stem mixes for each of the three songs (4 stems per song). The Free Multitrack Download

Library corresponding with [10] includes multitracks for about 180 songs, where forum users can submit their mixed versions of the song in MP3 format. For the experiment in [3], stems from the Rock Band video game were used, readily extracted from the game but not shareable as audio only, since its use is restricted by copyright. The Structural Segmentation Multitrack Dataset [8] contains 104 songs including structural segmentation ground truth annotations. A recent dataset provides 122 royalty-free multitracks including melody annotation [11]. Multitrack recordings are also published together with [12]. Finally, various other websites publish multitrack recordings with or without mixes with various licenses, with little to no extra content or functionality.

To be useful for the wider research community, the content should be highly diverse in terms of genre, instrumentation, and quality, so that sufficient data is available for most applications. Where training on large datasets is needed, such as with machine learning applications, a large number of audio samples is especially critical.

Data that can be shared without limits, on account of a Creative Commons or similar license, facilitates collaboration, reproducibility and demonstration of research and even allows it to be used in commercial settings, making the testbed appealing to a larger audience.

Moreover, reliable metadata can serve as a ground truth that is necessary for applications such as instrument identification, where the algorithm's output needs to be compared to the 'actual' instrument. Providing this data makes the testbed an attractive resource for training or testing such algorithms as it obviates the need for manual annotation of the audio, which can be particularly tedious if the number of files becomes large.

Similarly, for the testbed to be highly usable it is mandatory that the desired type of data can be easily retrieved by filtering or searches pertaining to this metadata.

For this reason, we present a testbed that

- can host a large amount of data;
- supports a variety of data of varying type and quality, including raw tracks, stems, mixes (plural), and digital audio workstation (DAW) files;
- contains data under Creative Commons license or similar (including those allowing commercial use);
- offers the possibility to add a wide range of meaningful metadata;
- comes with a semantic database to easily browse, filter and search based on all metadata fields.

The testbed can currently be accessed via multitrack.eecs.qmul.ac.uk.

2. CONTENT

We publish a set of recorded and/or generated multitrack audio including stems and mixes thereof, without restrictions in terms of type (music, speech, movie soundtrack, game sound, ...), quality (professionally recorded as well as displaying interesting artefacts such as noise, distortion, reverberation or interference), or number of tracks (from a multi-microphone recording to a 96-track project with many takes). We do however require the audio to have a Creative Commons license or to be public domain, e.g. music of which the author has been deceased for over 70 years.

The variety of Creative Commons licenses available also allows artists or institutions to share material with different restrictions, including whether or not the material can be used commercially. In any case, however, the owner of the content is required to be properly attributed with every use of their work. With the exception of content creators who are under a contract that prohibits them from releasing their intellectual property, any institution or artist can benefit from sharing their work on this testbed with a wide community of researchers, students, educators, developers and creative professionals. Any audio shared in this way increases the exposure of the artist and all personnel involved in the production of the music and their affiliation, as this is all included in the metadata corresponding to every song. Furthermore, through dissemination of their work, artists can expect it to be reworked and used in creative applications. In case the owner thinks sharing a song would damage record sales, one can share tracks and stems through this platform while not releasing the final mix.

We add extensive metadata to every song, track, stem and mix, such as title, artist, composer, and recording artefacts of a song; instrument, microphone and preamp of a track; DAW format, mixing engineer and sampling rate of a mix; and many more. This allows searching for content that meets a set of specific criteria, which is essential for a body of audio of this size, as browsing through everything manually is highly impractical.

We welcome contributions by any institution or individual, all of which will be clearly acknowledged on the website (unless an anonymous contribution is desired, in which case we are happy to license the content in our

name), be it multitrack audio, stems, mixes, DAW files, or any combination of the aforementioned. Pending a built-in upload facility, the authors can be contacted for any contribution. Content is being uploaded and annotated at the time of writing.

3. INFRASTRUCTURE

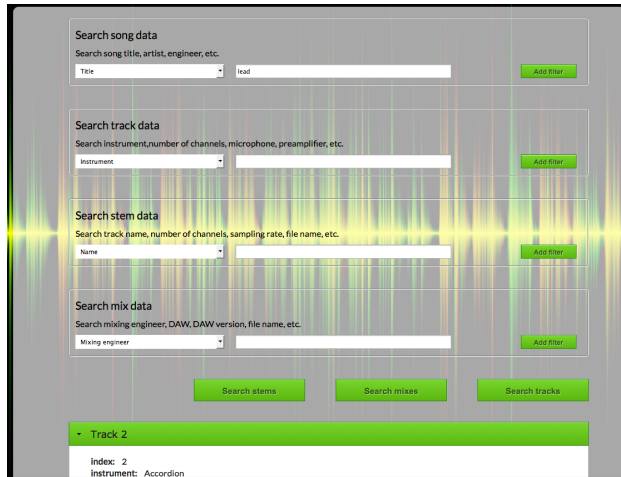


Fig. 1: Current interface of the testbed, allowing filtering/searching on various metadata fields

We have chosen a triplestore database to store statements containing metadata related to the songs, raw tracks, stems and mixes¹. Semantic databases allow the linking of data by storing subject-predicate-object structured triples [13]. One can then navigate the network formed by linked statements and for instance find more songs of the same artist, engineer or contributing institution. Our implementation features:

- A database which offers a SPARQL endpoint to query and insert data through HTTP requests.
- A REST web service, which receives JSON objects, parses them and stores the different elements in RDF format. These linked elements are then stored in the database.
- A web application offering two functionalities:
 - An interface to insert data. Access to this interface is restricted to authorised users. Data input is laid out in four different sets of elements: song elements, track data, stem data and data related to mixes.

¹<http://franz.com/agraph/allegrograph/>

- An interface to search for data according to a number of criteria, shown in Figure 1. The web application points at the SPARQL endpoint directly, dynamically building SPARQL queries without using a web service. Access to this interface is not restricted, although the data is.

A depiction of a scaled-down network formed by the linked data is shown in Figure 2.

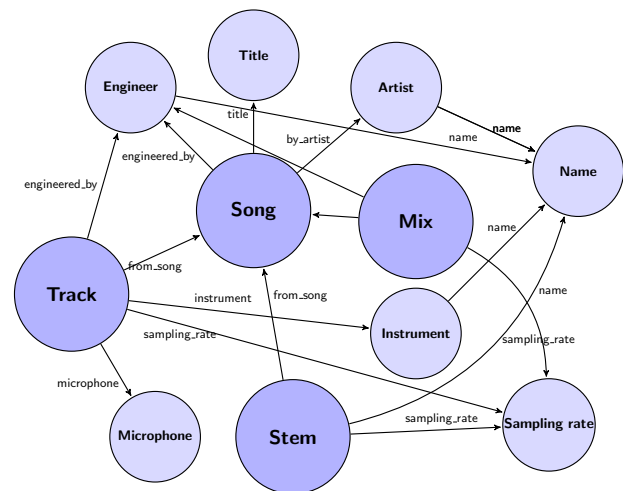


Fig. 2: Example of linked data network, showing only a subset of the features, with class elements (larger nodes), other elements (smaller nodes) and connections through properties (edge labels).

Our classes are taken from existing ontologies²³, or extend classes from these ontologies. A “Track”, for example, is an instance of the <http://purl.org/ontology/studio/multitrack#AudioTrack> class defined in the Multitrack Ontology [14], from which we also used the *Instrument* class; a “Song” is an instance of <http://purl.org/ontology/mo/Composition> from the Music Ontology [15], etc. We have extended these by creating the classes “Stem”, “Mix” and “Engineer” as well as numerous properties, such as *engineered_by*, *from_song*, *bit_depth*, *number_of_channels*, etc. Hence, a track X is *from_song* Y, which has a *name* Z and is *by_artist* A, which is a MusicGroup with a list of members. The track is one of a number of tracks from that song, was *engineered_by* B, features an *instrument*, was recorded using a *bit_depth* and a *sampling_rate*, etc.

²<http://musicontology.com/>

³<http://motools.sourceforge.net/studio/multitrack>

4. FUTURE WORK

First and foremost, we will continue to look for contributors to expand the testbed, and promote its use and support community-based evaluation of research ideas and algorithms using this as the testbed [16, 17]. As in [2], we will organise mixing experiments to provide different mixes of source material found on the testbed, for analysis of music production practices.

We would also like to solicit feedback from users in terms of functionality, metadata, and any other features that could be useful to the research or end user community.

5. ACKNOWLEDGMENTS

This work was supported in part by EPSRC Grant EP/K007491/1, Multisource audio-visual production from user-generated content. We would like to thank the members of the Sound Software project (<http://soundsoftware.ac.uk>) for fruitful discussions concerning the nature and structure of the testbed, as well as everyone who has contributed content to the testbed so far.

6. REFERENCES

- [1] P. Pestana and J. Reiss, “Intelligent audio production strategies informed by best practices,” in *53rd Conference of the Audio Engineering Society*, 2014.
- [2] B. De Man, B. Leonard, R. King, and J. D. Reiss, “An analysis and evaluation of audio features for multitrack music mixtures,” to appear in *15th International Society for Music Information Retrieval Conference (ISMIR 2014)*, October 2014.
- [3] J. Scott and Y. E. Kim, “Analysis of acoustic features for automated multi-track mixing,” in *12th International Society for Music Information Retrieval Conference (ISMIR 2011)*, 2011.
- [4] G. Bocko et al., “Automatic music production system employing probabilistic expert systems,” *129th Convention of the Audio Engineering Society*, 2010.
- [5] J. D. Reiss, “Intelligent systems for mixing multichannel audio,” in *17th International Conference on Digital Signal Processing (DSP)*, 2011.
- [6] B. De Man and J. D. Reiss, “A knowledge-engineered autonomous mixing system,” in *135th Convention of the Audio Engineering Society*, 2013.
- [7] E. Perez-Gonzalez and J. D. Reiss, “Automatic mixing,” in *DAFX: Digital Audio Effects, Second Edition*, pp. 523–549, John Wiley, 2011.
- [8] S. Hargreaves, A. Klapuri, and M. B. Sandler, “Structural segmentation of multitrack audio,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 10, pp. 2637–2647, 2012.
- [9] M. Cartwright, B. Pardo, and J. D. Reiss, “MIX-PLORATION: Rethinking the audio mixer interface,” in *International Conference on Intelligent User Interfaces*, 2014.
- [10] M. Senior, *Mixing Secrets*. Taylor & Francis, 2012.
- [11] R. Bittner et al., “MedleyDB: a multitrack dataset for annotation-intensive MIR research,” to appear in *15th International Society for Music Information Retrieval Conference (ISMIR 2014)*, October 2014.
- [12] J. Fritsch and M. D. Plumbley, “Score informed audio source separation using constrained nonnegative matrix factorization and score synthesis,” in *IEEE ICASSP’13*, pp. 888–891, 2013.
- [13] G. Fazekas and T. Wilmering, “Semantic web and semantic audio technologies,” tutorial presented at the *132nd Convention of the Audio Engineering Society*, 2012.
- [14] G. Fazekas and M. B. Sandler, “The Studio Ontology framework,” in *12th International Society for Music Information Retrieval Conference (ISMIR 2011)*, 2011.
- [15] Y. Raimond, S. A. Abdallah, M. B. Sandler, and F. Giasson, “The Music Ontology,” in *8th International Society for Music Information Retrieval Conference (ISMIR 2007)*, pp. 417–422, 2007.
- [16] J. S. Downie, “The music information retrieval evaluation exchange (2005–2007): A window into music information retrieval research,” *Acoustical Science and Technology*, vol. 29, no. 4, pp. 247–255, 2008.
- [17] E. Vincent et al., “The signal separation evaluation campaign (2007–2010): Achievements and remaining challenges,” *Signal Processing*, vol. 92, no. 8, pp. 1928–1936, 2012.