

Institution: 10007140 Birmingham City University		
Unit of Assessment: UoA11. Computer Science and Informatics		
Title of case study: Intelligent Music Production: saving time for experienced engineers, lowering barriers to entry for hobbyists		
Period when the underpinning research was undertaken: October 2014 to Dec 2020		
Details of staff conducting the underpinning research from the submitting unit:		
Name(s):	Role(s) (e.g. job title):	Period(s) employed by submitting HEI:
Dr Ryan Stables	Associate Professor in Audio Engineering	2012 – Present
Dr Jason Hockman	Associate Professor in Audio Engineering	2015 – Present
Period when the claimed impact occurred: Dec 2016 to Dec 2020		
Is this case study continued from a case study submitted in 2014? N		
1. Summary of the impact (indicative maximum 100 words) <p>Dr Ryan Stables and colleagues' research into how computers can be trained to understand the human vocabulary of sound has led to a level of intelligent automation that is unprecedented in commercial music production. Since 2014, the BCU team has introduced a suite of adaptive software plugins (SAFE) which have been used by well-known music studios to create commercial recordings, as well as downloaded by 15,000 music-creators in 20 countries. The software makes the technical and creative process easier for professionals to perform complex and tedious mixing tasks, and for amateurs and early careerists to bypass many years' costly learning of how to process sound. The team have developed a next-generation Digital Audio Workstation (DAW), facilitating real-time collaborative music production and embedding machine learning capabilities into the production workflow; this is being commercialised by a university spin-out, Semantic Audio Labs (SAL), which raised an initial GBP500,000 of seed-funding, employs 6 full-time staff, and has a number of pending patents.</p>		
2. Underpinning research (indicative maximum 500 words) <p>Audio production and engineering are technical disciplines that inherently require extensive experience or investment to achieve desired results. In order to lower the barriers to entry for everyday users such as amateur musicians and podcasters, Dr Ryan Stables, Dr Jason Hockman and a team of 3 PhD students have developed systems that combine heuristic knowledge of audio processing with machine learning and intelligent systems. This includes systems to mix music automatically, to route audio signals based on their metadata, and to allow for abstract control of audio effects, reducing the complexity of a number of arduous tasks in sound engineering.</p> <p>Semantic Audio Feature Extraction (SAFE): In 2014 Stables led the development of SAFE, a suite of audio effects plugins which used a wide range of in-built semantic feature extraction algorithms to allow users to control them with natural language [R1]. The research demonstrated that real-time information gathered from the signal flow of a typical audio processing chain could be used to derive abstract parameter spaces, which were confluent with human language. To support this, a widely used dataset of sound transformations with accompanying descriptive metadata was constructed. A number of statistical models were then used to explore relationships between features of the dataset and concepts in natural language [R2]. The researchers found that there are strong commonalities in the way users describe effects</p>		

processing, and this allows the development of recommendation engines and abstract interfaces to help apply audio effects more effectively based on synonyms and antonyms.

An example of an abstract interface, which was developed using the findings of the SAFE project, was the low-dimensional equaliser [R3]. One of the key issues with sound equalisation is the large number of complex parameters. Using the data collected in [R1] and [R2], Stables and colleagues were able to develop a 2-dimensional interface, with assignable descriptions. This enables users to easily transition between terms such as warm and bright, without needing to understand the underlying high-dimensional parameter space.

Faders: In 2017, Dr Stables' team developed an intelligent Digital Audio Workstation (DAW), which allows users across the globe to produce music collaboratively in real-time, with in-built intelligent systems. This is a culmination of several studies on automatic mixing, intelligent routing algorithms, novel interface design, low-latency processing, and music information retrieval. The blueprint for Faders [R4] summarises the challenges of building a system that supports this form of intelligent multi-channel adaptive processing, and suggests ways that the audio engine can distribute data in order to be both portable and efficient. It outlines BCU's plugin framework JSAP, which is now being used by a number of third-party developers, and is becoming a de facto standard for audio processing modules on the web.

The development of Faders allowed Stables and colleagues to continually collect and process anonymised music production data. To do this efficiently, the team developed a number of feature extraction tools, such as the JavaScript library, JS-Xtract [R5]. To extract semantic parameters from users' recordings, Faders uses a range of empirically developed music informatics techniques such as novel methods for automatic music transcription [R6]. Here, highly effective algorithms were developed, which approximate drum sequences being played by a musician via deep learning. BCU's systems out-perform all other drum transcription methods, confidently labelling and transcribing a user's musical signals. This work was presented in some of the leading conferences in our field including ISMIR, a highly ranked event in computer science. The work was part-funded through an EPSRC mini-project grant, listed below.

Other key researchers: Dr George Fazekas, Senior Lecturer, QMUL; Dr Thomas Wilmering, RA, QMUL; Professor Joshua Reiss, Professor of Audio Engineering, QMUL.

These researchers were co-authors on some of the publications associated with this case study. Their roles involved contributing to the data analysis, and proof-reading papers before submission. This project was lead and implemented primarily by BCU. None of the external collaborators plan to submit an impact case study relating to this project.

3. References to the research (indicative maximum of six references)

[R1] Stables, R., Enderby, S., De Man, B., Fazekas, G. and Reiss, J. D. (2014) "SAFE: A system for the extraction and retrieval of semantic audio descriptor developed as" 15th International Society for Music Information Retrieval Conference (ISMIR 2014), October 2014.

Available at: <http://www.open-access.bcu.ac.uk/3255/>

[R2] Stables, R., Enderby, S., De Man, B., Reiss, J. D., Fazekas, G., and Wilmering, T. (2016) "Semantic Description of Timbral Transformations in Music Production," Proceedings of the ACM International Conference on Multimedia, Pages 337-341. October 2016. Available at:

<http://www.open-access.bcu.ac.uk/4105/>

[R3] Stasis, S., Stables, R., and Hockman, J. (2016) "Semantically Controlled Adaptive Equalisation in Reduced Dimensionality Parameter Space," Applied Sciences, Vol. 6 (4), April 2016. Available at: <http://www.open-access.bcu.ac.uk/3258/>

[R4] Jillings, N. and Stables, R. (2017) "An Intelligent Audio Workstation in the Browser". In: 2nd Web Audio Conference (WAC 2017), August, 2017, London, United Kingdom. Available at: <http://www.open-access.bcu.ac.uk/4965/>

[R5] Jillings, N., Bullock, J. and Stables, R. (2016) "JS-Xtract: A Realtime Audio Feature Extraction Library for the Web" International Society for Music Information Retrieval Conference (ISMIR 2016), August 2016. Available at: <http://www.open-access.bcu.ac.uk/4099/>

[R6] Southall, C., Stables, R., and Hockman, J. (2016) "Automatic Drum Transcription using Bi-directional Recurrent Neural Networks." International Society for Music Information Retrieval Conference (ISMIR 2016), August 2016. Available at: <http://www.open-access.bcu.ac.uk/4101/>

ISMIR is the 8th ranked publication in Google Scholar's Multimedia subcategory of Engineering and Computer Science. It is one of the largest conferences in music processing, and has an acceptance rate of less than 50%. Applied Sciences is a popular journal in computer science and engineering. It has an impact factor of 1.689. [R3] was an invited publication: a pre-cursor to this won a best paper award at the 18th International Conference on Digital Audio Effects 2015.

Grants: R. Stables and C. Athwal, Large-scale Capture of Producer-defined Musical Semantics, ESPRC mini-project grant of £35k through the EPSRC Semantic Media Project:

<https://gtr.ukri.org/resources/outcome.html?type=collaboration&outcomeId=546114d5b9a6f7.63196086-1&grantRef=EP%2FJ010375%2F1>

4. Details of the impact (indicative maximum 750 words)

Since 2014 Dr Ryan Stables and colleagues' research has had significant creative and commercial impact on the field of music production through the provision of software tools that have made marked improvements on the ways in which professional audio producers and a wider community of music lovers and creators, work with sound. This has led to the creation of a spinout company, Semantic Audio Labs, that employs 6 full-time staff.

Impact on creativity: intelligent automation for professional audio producers and hobbyists

Digital Audio Workstations (DAWs) are the means by which all music producers and other sound designers record, mix, compose and edit audio; they are used both in professional sound studios and by millions of hobbyists on their personal computers. The software systems, SAFE plug-ins and Faders, developed by Stables and his team promoted a change in audio engineering practice by providing enhanced functionality to audio content creators with a wide range of backgrounds and levels of experience. The systems provide methods for automating and abstracting parametric complexity, using a similar philosophy to functions such as auto-focus and red-eye removal in photography, or content-aware-fill in image-editing.

The SAFE plugins provide extensions to a DAW which allow musicians and producers to apply processing using natural language, and since 2014 have been downloaded from BCU's servers over 15,000 times by users in over 20 countries. Faders, a next-generation DAW, saves 50% in production time according to a senior podcaster and consultant at Don't Skip: *"By providing AI-automation and simplified interfaces, this has saved over 50% of the time it takes to get from a simple idea, to having a professional release."* [S01]. Faders provides extended functionality such as novel interfaces for editing, mixing and mastering audio, and it is able to continuously improve the effectiveness of its algorithms through automatically collected usage data and user analytics.

The SAFE plugins are installed in at least 10 professional recording studios in at least 3 countries, including Wave (London, Amsterdam, New York), a prolific studio in the advertising sector; and Abbey Road Studios (London) who have produced a wide range of high-grossing commercial recordings over the past 90 years, including almost all of the Beatles' hits between 1962 and 1970. For audio producers and sound engineers who work at such studios, "the reduction in complexity of audio processing tools is a necessary advancement in our industry." [S02] Wave's senior sound designer has worked with large international marketing campaigns for clients such as BMW, Adidas, Bose, Compare the Market, and Honda. He says: *"The tools that have been developed through the SAFE program, and now through Faders have caused a*

change in the way I mix audio. At Wave, we are now able to spend more time focussing on the creative aspects of the production process, and we can trust the way machines recommend processing parameters.” [S02] The plugins have been used on commercial projects that will be seen by potentially millions of TV viewers. [S02]

A Grammy Award-winning recording engineer with 30 years’ experience says that, in his view, this is the first system of its kind: *“Music production software in general has a long history of being nostalgic, either trying to emulate old analogue hardware, or traditional mixing interfaces. The system that Semantic Audio Labs is developing ... provides users with a means to make rapid adjustments to a set of recordings without the need for extensive training. It is a positive step towards providing automation for arduous unnecessary tasks, [saving] time for experienced engineers like me, allowing us to produce content more effectively.” [S03]*

The system also lowers the barriers to entry for aspiring and early-career producers, hobbyists and music teachers or lecturers who are able to produce music and podcasts with their students collaboratively over the web. From a lecturer’s perspective, *“Music production is often a solitary process, whereby the traditional digital audio workstations work in isolation. As Faders is a web-based platform, multiple people are able to produce a piece of music simultaneously. I am able to demonstrate concepts to my students whilst they follow along in real time and make their own edits... this will transform the industry, and contribute significantly to music pedagogy.” [S03]*

Film composer Rael Jones (The Salisbury Poisonings, My Cousin Rachel, Suite Française) has tried the new software. *“These are quite simple effects and would be very intuitive for the amateur musician. There are similar commercially available technologies but they don’t take a semantic input into account as this does... for some amateur musicians this could be an interesting educational tool to use as a starting point for exploring sound,”* he told the BBC. [S04]

To support amateur musicians as well as researchers, hardware and software manufacturers, and developers, Stables co-authored (with Brecht De Man, BCU; Joshua Reiss, QMUL) a book on automated audio engineering techniques, Intelligent Music Production, published in 2019 by Taylor & Francis (print run of 10,000 copies). A comprehensive guide to the emerging field of automation in music mixing and mastering, it provides an introductory read for beginners, as well as a crucial reference point for experienced researchers, producers, engineers and developers. [S05] Similarly, an annual Workshop on Intelligent Music Production founded in 2015 by Stables allows the cross-fertilisation of academia and industry, drawing over 100 delegates from leading audio companies (Apple, Sony, BBC, Native Instruments, Focusrite and Music Tribe) and universities around the world [S06].

BCU’s JavaScript library of feature extraction tools, JS-Xtract, has been used by at least 20 third-party developers for projects ranging from mobile app development to generative birdsong synthesis. The JSAP plug-in framework, which allows developers to build, integrate and commercialise their audio plugins, has been used by at least 5 third parties under an open-source license to develop web-audio enabled software such as FXive (<https://fxive.com>), and Piper (<http://github.com/piper-audio>). This has not only created additional revenue streams for developers but will also standardise the development of web audio plugins. The co-founder of real-time online sound effect generation service Nemisindo says: *“In order for our models to be deployed on the web, we heavily rely on the technologies developed by Ryan and his team at BCU... we are able to build modules which are transferable, scalable, and can be easily integrated into other systems. Sounds that Nemisindo have created using JSAP have been featured in major film and TV releases, including Counterpart (Starz), Will (TNT) and Clash of Futures (LooksFilm).” [S07]*

Impact on commerce: creating a spinout company and 6 full-time jobs

The web-based intelligent DAW was commercialised as Faders (<https://faders.io>) through the formation of a spinout company Semantic Audio Labs Ltd in 2018, just a year after receiving GBP500,000 seed funding from a consortium of investors, for 2 years’ development [S08]. The company has employed 6 full-time members of staff since December 2018, and over 10

freelance developers sporadically. As of December 2020, Faders is in the beta testing pre-release phase, with a community of more than 1,000 beta users/testers from over 10 countries, who are a mix of professionals and hobbyists, including award-winning music producers. The platform is due for public release in 2021 (slightly delayed due to COVID), and within the year Semantic Audio Labs anticipates 100,000 subscribers (based on estimates provided by our out-sourced marketing team), and an additional Series A funding round in the region of GBP5,000,000 to employ a further 25 staff.

Stables has two pending patents with Semantic Audio Labs (UK patent applications GB1914571.3 and GB 2007751.7 [S09]), which were submitted in December 2019 and May 2020. These cover some of Faders' novel workflow interactions, which emulate audio objects on a stage, and methods for real-time collaborative audio production. An IP holding company has been founded to commercialise the IP generated through Semantic Audio Labs. This company, Semantic Audio IP Ltd, has an exclusive license with Semantic Audio Labs [S10].

5. Sources to corroborate the impact (indicative maximum of 10 references)

[S01] Testimonial: Senior podcaster and consultant, Don't Skip.

[S02] Testimonial: Senior Sound Designer, Wave Studio.

[S03] Testimonial: Grammy Award-winning recording engineer

[S04] Training computers to understand the language of music:

<https://www.bbc.co.uk/news/science-environment-29146655>

[S05] Intelligent Music Production, A book published by Taylor and Francis (2019):

<https://www.amazon.co.uk/Intelligent-Music-Production-Theoretical-Overview/dp/1138055182>

[S06] Workshop on Intelligent Music Production: <https://intelligent-music-production.github.io/>

[S07] Testimonial: Co-founder and Executive Director of Nemisindo.

[S08] Account of Semantic Audio Labs Ltd

[S09] Patent applications (GB1914571.3 and GB 2007751.7): <https://www.ipo.gov.uk/p->

<https://www.ipo.gov.uk/p->

<https://www.ipo.gov.uk/p->

[S10] Account of Semantic Audio IP Ltd