

An AI-Based Design Framework to Support Musicians' Practices

Juan Martinez-Avila, Adrian Hazzard, Alan Chamberlain, Chris Greenhalgh and Steve Benford

The Mixed Reality Lab

University of Nottingham Nottingham, UK

[firstname.lastname]@nottingham.ac.uk

ABSTRACT

The practice of working musicians extends beyond the act of performing musical works at a concert. Rather, a significant degree of individual and collaborative preparation is necessitated prior to the moment of presentation to an audience. Increasingly, these musicians call upon a range of digital resources and tools to support this 'living' process. We present a speculative design paper in response to a set of ethnographies and interviews with working musicians to highlight the potential contemporary digital technologies and services can bring to bear in supporting, enhancing and guiding musicians' preparation and practice. We acknowledge the role that artificial intelligence and semantic technologies could play in the design of tools that interface with the traditional practice of musicians and their instruments.

CCS CONCEPTS

- **Human-centred computing**→**Scenario-based design**
- **Ethnographic studies** • **Applied computing**→**Sound and music computing**

KEYWORDS

Music, Practice, Rehearsal, Audio Technologies, AI, Artificial Intelligence, Semantic Media, Design, Ethnography, HCI, CSCW.

ACM Reference format:

Juan Martinez-Avila, Adrian Hazzard, Alan Chamberlain, Chris Greenhalgh and Steve Benford. 2018. An AI-Based Design Framework to Support Musicians Practice, preparation and Performance. In *Workshop Proceedings of AM '18, Wrexham, Wales, UK, September 12–14, 2018*.

1 Introduction

We define working musicians as those who regularly perform renditions of contemporary popular music, for example in function or tribute bands [6, 8] as well as the creation and performance of new, original music. They are typically experienced and instrumentally proficient, hired to perform a broad repertoire of popular music at social events (e.g. parties, weddings and music venues). They provide a service and often respond of the demands of any given performance situation. Their ecosystem is a complex and multivariant one that demands a continual refreshing of their individual instrumental capability; the learning and performance of new material; resource discovery [4]; and effective performance of a specific role in a complex and attuned collaborative act (i.e. group performance). The

environment of the working musician is also complex and physically loaded with instruments 'in hand', connected to various sound equipment and other tools and resources such as mobile devices and scores with 'living' annotations.

Drawing on a set of ethnographies and interviews with working musicians we illustrate a rich picture of their individual and collaborative preparation process, exposing some common themes in the process. We then present a speculative system design which draws on a range of contemporary technologies, such as audio feature analysis, conversational agents and semantic web technologies. We illustrate one view of a future system that supports the preparation practices of working musicians both at home and in the rehearsal studio. At the centre of our vision is a collaborative system that enables users to collate and semantically categorize, align and recall a range of media as controlled via networked devices and instrumental equipment. While this initial design sits on the boundary of artificial intelligence (AI), the semantic network at its core prepares the way for subsequent design that exploit further AI applications.

2 Related Work

The working musicians' preparation process draws on the use of multiple tools and resources, namely, reading and writing notation, and consulting online media resources [5, 10, 13]. Such recall of online resources evidences the need for these musicians to access multiple channels of information (audio, video and notation) and to be able to re-circulate, annotate, archive and appropriate media content to support their individual song-learning process [2, 12, 14]. Moreover, in some cases they may also use of social networks that to share and discuss ideas and resources or even carry out distributed collaborative work [5, 13]. There are many online sites and services that aim to support collaborative musical practice, but for the most part they are concerned with collaborative online recording and production (e.g. bandhub.com or www.soundbetter.com) or sourcing and connecting with new musicians (e.g. www.vampr.me). As yet, there is little in terms of musically oriented online collaborative platforms to support the practicing musician, such as those project management tools commonplace in the business world (e.g. basecamp.com). Musicians increasingly use digital tools in performance, such as displaying scores and 'charts' on digital tablets (e.g. www.padformusician.com) as they offer enhanced functionality in terms of storage and control via foot pedals (e.g. www.airturn.com). Prior research has seen investigations into

connecting human score annotations to digital music systems, such as [7]. Musical instruments can also act as controllers of other performance media or music related activities, such as controlling a digital score, visuals and MIDI piano in [11] or foot pedal control for a combination of effects and Digital Audio Workstation (DAW) control, such as the Pacer foot controller¹

3 ‘Picking’ Apart the Working Musician

In the following section we detail the findings from our ethnographies and interviews with working musicians by unpacking a series of vignettes on the key practices observed during their individual and collaborative preparation.

3.1 Fieldwork

We chose to focus specifically on guitar and bass players so as to capture a corpus of rich observational and first-hand discussion with a comparable group of musicians. Nonetheless, observing these participants in group rehearsals has also permitted the observation of wider collaborative practice across a range of instrumentalists. Participants were recruited via online social networks and word of mouth with close acquaintances.

3.1.1 Recruitment and Data Capture. 22 participants were recruited to take part in semi-structured interviews, 5 observed during their individual practice and 9 bands observed in rehearsal. The interviews sought to promote discussion around the topics of instrument proficiency, experience, individual and collaborative practices, and the methods, tools and resources used. These were captured using an audio recorder for later transcription. Video and audio of the observation sessions were captured and reviewed. Emerging themes and accompanying vocalizations were subsequently identified.

3.3 Findings

We now illustrate three example cases that universally reflect the practice observed across our participants. Pseudonyms have been created for each participant.

3.3.1 Learning a Song. We present the example of Mike, which illustrates how musicians set about their individual preparation of new material. For an upcoming event one of Mike’s bands has been requested to play a specific song. At the time of observation Mike had previously worked on the main sections of this song but he still needed to learn the solo bass part. Mike was observed at his home, sitting at a desk with his laptop holding his bass guitar. He navigated to YouTube and searched for the official video of the song. Once found he began by playing along with the video, recapping the sections of the song he previously learned. At times, Mike was observed trying to simultaneously manage the controls of the online video with his right hand while playing the bass one-handed with his left. When he arrived at the solo part he stopped

playing his instrument to listen. He then scrubbed the media player back to the start point of the bass solo tried out playing the melody along with it as closely as he could. This prompted a discussion regarding accuracy of reproduction. Mike felt that those sections of a song that would be expected, or easily recognizable by the audience ought to be reproduced to a finer degree of accuracy, compared to other less prominent musical passages, which can just be more broadly representative of the original. This in turn prompted Mike to search for a video tutorial of someone playing the bass solo. He found a video of a bass player performing the song alongside a traditional notation and tablature notation visualization synchronized to the audio of the song. Using this resource, he continued to practice in real-time while watching this video, reproducing the notes with more fidelity.

Mike then shifted his attention to another band he performs with, to learn an original song written by one of the other band members. In this setting it was typical of Mike to create a chord chart (i.e. a notation of the chord changes of a song) as he listened along an MP3 recording that was sent to him via e-mail. Again, this process involved him holding the instrument, while operating the laptop’s media player and writing with pen and paper. He mentioned he keeps his ‘charts’ as a personal reference to use later on when the band works on the song collaboratively. Moreover, he kept multiple binders in which he archived his tablatures, lyrics and chord charts as a resource to refer to when performing material after extended periods.

3.3.2 Support Resources. Cindy describes the utility of paper-based resources to support her performance with the bass guitar. Cindy discussed her use of written ‘charts’, explaining how she creates multiple versions which contain differing degrees of information granularity. Cindy showed and discussed some examples of her charts. For instance, to support initial learning and orientation of a song Cindy often creates a chart with in-depth scaffolding of information, such as the lyrics, sectional descriptions, and chord progressions along with diagonal lines representing how many times the chord is repeated in each measure. As her familiarity with a song develops she typically creates a new abbreviated version of the above. At this stage she may compile multiple songs onto one sheet of paper, detailing the song titles alongside the sectional arrangement and their corresponding repetitions (e.g. ‘Verse Chorus x2’). The final version of the chart, reflecting the last stage of her scaffolding process, displays only the song titles, which she describes as the “joyous moment”, where no additional information about a song is required in order to perform it.

3.3.2 A Band Rehearsal. This last example describes a band rehearsal. After setting up their equipment the band discussed and agreed to start with “the new stuff”. In response, members of the band then turned to written charts and notes created prior to the rehearsal. For example, Stuart, one of the guitarists, took out a chord chart to play along with, and Kit, the bass player and lead singer, both referred to the same sheet of paper which contained

¹ <https://www.nektartech.com/pacer-midi-daw-footswitch-controller.html>

the lyrics he had written down. During rehearsal the band frequently stopped whenever mistakes occurred, or they fell out of sync. During these breaks they would typically discuss and seek to clarify aspects of the song, such as its structure or tempo. On several occasions, there was a collective confusion on the song’s structure. In reply, one of the musician’s, using a mobile device, would search for the original song recording online and play it, amplifying it through a microphone so everybody could hear it.

Kit took the leading role of directing the actions of the band. For example, he brought pre-prepared set lists for each one of the members. He would also signal song arrangement while performing by calling out instructions through the microphone (e.g. “Middle 8”) or would also cue them in whenever they stopped and had to start over (e.g. ‘let’s go from the top’). Furthermore, he would provide information about songs, ranging from general aspects such as its key or structure to the extent of teaching the other band members how a section ought to be played. For example, during one song the second guitarist, Quentin, had created a chord chart with a few mistakes. Kit then borrowed a guitar from the Stuart and demonstrated the correct progression to Quentin, who then requested that he dictate the chords, so he could amend his chord chart.

3.3 Summary Findings

Our data set illustrates the complex environment where a number of key themes emerge. There is a cyclic process of individual and collaborative preparatory work, where each member sources and prepares new material in order to contribute effectively when the group convene at rehearsal. The sourcing of materials often involves an auditioning process where multiple versions (e.g. YouTube videos) are previewed to find the one that best address their needs. When the musicians convene collectively at rehearsal, resources are often revisited, shared, integrated and updated. Band members assume differing roles, sharing information through a variety of means, such as written resources or through demonstration. Throughout all these processes the musician balances the instrument in hand, operating computers, writing notes and annotations, a physically demanding task. At the core of this practice is the assimilation and recall of information across multiple channels.

4 An Initial Design Framework

We now present an initial design framework of a fictitious new system termed ‘Musicians Support Central’ (MSC). MSC takes a holistic view on the themes arising from our data set, presenting a snapshot of how contemporary technologies could complement and interface within this setting. For this we draw on: online media and cloud services; conversational agents [3]; networked musical artefacts (IoT) [1]; and feature recognition and analysis technologies [9]. Figure 1 highlights an overview of relationships between musicians’ preparation practice and these technologies as implemented into MSC. The following design fiction outlines a descriptive scenario of MSC in use.

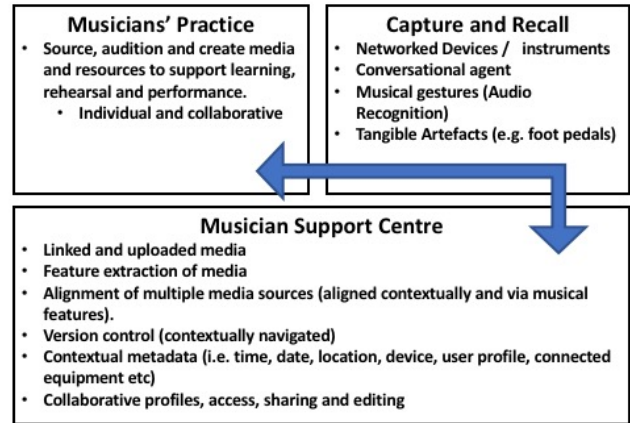


Figure 1: ‘Musicians Support Central’ Design Overview

Olivia is a guitarist in an established function band. They have an extensive and established set list but continue to add new material in response to new chart hits and requests that arise from specific bookings. The band have an upcoming wedding booking where the bride and groom have requested *Breakin’ Up is Hard to Do*, by Neil Sedaka. Olivia, and the rest of her band begin their preparation process of learning and integrating this song into their set. At home, Olivia sits down with her guitar and laptop. On her laptop Olivia logs into to ‘Musicians Support Central’ (MSC) a new online system with accompanying hardware tools to support the working musician. It contains resources for Olivia’s band. Each band member also has their own profile shared across the band environment where they can link and upload individual and shared resources.

Olivia selects ‘make a new song’ which opens up a template song page. In a separate browser page, she then searches for YouTube videos of the song in question, discarding some videos due to poor sound quality and information until she finds a version considered appropriate to use. She links this video into ‘Musicians Support Central’ (MSC) and the video now appears hosted in the ‘make new song’ template. This linking process automatically initiates the system’s feature extraction tool that analyses the audio of the video, subsequently capturing its harmonic progression, melody extraction, tempo, and segment (i.e. structure) which can be represented in a number of customizable ways. Olivia chooses to create a ‘chord chart’, which produces visual representation to display on her tablet or print off. When she plays back the YouTube video within the MSC the chart scrolls in alignment underneath the video view. She can add, or personalize this ‘chart’ in many ways, creating multiple linked versions with differing content and detail. Finally, Olivia shares these resources to the other band members’ profiles. These resources are contextually tagged with the user profile, date, time, location, and networked device. Olivia then sets about learning the song using these resources, guitar in hand. To support this process Olivia uses the system’s conversational agent to navigate through the resources, and control playback of the media via vocal commands, so she can keep both hands on her instrument. Furthermore, her guitar is

connected to her laptop. Using the MSC feature extraction and analysis tool Olivia can record her performance and align it with other MSC media for subsequent review or use her guitar as a controller for the MCS. For instance, Olivia wants to practice the chorus section of the song. She says “*MSC navigate video one*” and the conversational agent seeks the video and configures the system to listen to her guitar. She then plays the chorus riff on her guitar and the system analysis and matches her musical gesture (i.e. sequence of notes and rhythms) against those found in the YouTube video, and then starts playback from that point. The combination of hands-free vocalizations and performed musical gestures complements the ‘instrument in hand’ set-up.

We now move on to the band’s next group rehearsal where they practice *Breakin’ Up is Hard to Do*. Similarly, the other band members have been working with MSC, sourcing, linking, and analyzing resources to support their preparation. The band members have set-up their equipment in the rehearsal room and respective tablets are connected to the local WIFI network which are logged in to their MSC profiles. Other items of the bands equipment are also connected to the system. For instance, Olivia’s MIDI effects foot pedal controller is connected to the MSC, where the MIDI messages (in addition to controlling her effects) are mapped to the transport controls of the MSC recorder. The MSC system is aware of patch change instances within Mainstage. Drawing on a number of contextual information (i.e. network devices, location, users logged into MSC), the MSC identifies the ‘this particular band’ have convened to rehearse, subsequently loading the band’s song resources as a result. One tablet within the rehearsal room is designated as a ‘master’ where the systems conversation agent and audio analysis tools are subsequently active. Olivia calls out “*MSC, display group charts for ‘Breakin’ Up is Hard to Do*” and the personalized ‘charts’ for each band member are displayed on their respective tablets. Alex—the singer—then says “*MSC Record, Breakin Up is Hard to Do, take 1*” which arms the system to record. Olivia selects the guitar effect required for the intro of the song which also starts the MSC recorder. Whilst recording, metadata such as date, time, users, and configured equipment will be captured and subsequently linked to the recorded audio to support subsequent analysis and recall.

6 Final Thoughts

The initial design framework we have provided outlines some of the many possible interactions and uses a system such as this could have. As earlier work has proposed and shown [3] there are challenges that we can face when dealing with AI systems, particularly in regard to music performance. However intelligent systems could prove invaluable for the musician of the future and this research aims to inform research relating to this. Furthermore, we also need to consider further steps and opportunities to make the system more responsive, ‘intelligent’ and smart (not just connected). By taking an approach that first examined the practices of actual musicians, which as ethnographically informed we were able to start to appreciate the plethora of computer-based interactions, materials and channels used by musicians, and the

socio-technical nature of this interaction. Our future work aims to further unpack some of these interactional features in order to better appreciate what is involved in the practices of musicianship and to appropriately understand the implications that this has for the design and development of future systems.

ACKNOWLEDGMENTS

This work was supported by the UK Engineering and Physical Sciences Research Council [grant numbers EP/L019981/1, EP/M000877/1] and the National Council of Science and Technology of Mexico (CONACyT).

REFERENCES

- [1] Steve Benford, Adrian Hazzard, Alan Chamberlain, Kevin Glover, Chris Greenhalgh, Liming Xu, Michaela Hoare, Dimitrios Darzentas (2016) “Accountable Artefacts: The Case of the Carolan Guitar”, Proceedings of the Conference on Computer Human Interaction, CHI’16, May 07 - 12, 2016, San Jose, CA, USA, 2016, ACM DOI: <http://dx.doi.org/10.1145/2858036.2858306>
- [2] Steve Benford, Peter Tolmie, Ahmed Y. Ahmed, Andy Crabtree, and Tom Rodden (2012) “Supporting traditional music-making: designing for situated discretion.” In Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work, pages 127–136. ACM.
- [3] Alan Chamberlain (2017) “Are the Robots Coming? Designing for Autonomy & Control in Musical Creativity & Performance.” Audio Mostly 2017: Augmented and Participatory Sound/Music Experiences, 23-26 August. Queen Mary University of London (London, UK) ACM - DOI <https://doi.org/10.1145/3123514.3123568>
- [4] Alan Chamberlain and Andy Crabtree (2016) “Searching for music: understanding the discovery, acquisition, processing and organization of music in a domestic setting for design”. in Personal and Ubiquitous Computing. pp. 1-13. Springer, ISSN 1617-4917, DOI: 10.1007/s00779-016-0911-2
- [5] Alan Chamberlain, Sean McGrath, and Steve Benford (2015) “Understanding social media and sound: music, meaning and membership, the case of Soundcloud”. DMRN+ 10: Digital Music Research Network, London
- [6] Robert R. Faulkner and Howard S. Becker. “Do You Know...?: The Jazz Repertoire in action”. University of Chicago Press, 2009.
- [7] Jérémie Garcia, Theophanis Tsandilas, Carlos Agon, Wendy Mackay. InkSplore: Exploring Musical Ideas on Paper and Computer. New Interfaces for Musical Expression (NIME 2011), May 2011, Oslo, Norway. 2011
- [8] Lucy Green. “How popular musicians learn: A Way Ahead for Music Education. Routledge, 2017.
- [9] Chris Greenhalgh, Steve Benford, Adrian Hazzard & Alan Chamberlain. (2017) “Playing fast and loose with music recognition.” In Proc of CHI ’17 Conference on Human Factors in Computing Systems, pages 4302–4313, New York, NY, USA, 2017. ACM. DOI 10.1145/3025453.3025900
- [10] Michaela Hoare, Steve Benford, Christopher Greenhalgh, and Alan Chamberlain. (2014) “Doing it for themselves: the practices of amateur musicians and DIY music networks in a digital age.” DMRN+9: Digital Music Research Network (EPSRC), December 2014, London.
- [11] Steve Benford, Chris Greenhalgh, Adrian Hazzard, Alan Chamberlain, Maria Kallionpää, David M. Weigl, Kevin R. Page and Mengdie Lin (2018) “Designing the Audience Journey through Repeated Experiences” In Proceedings of ACM CHI 2018 - Conference on Human Factors in Computing Systems. DOI - 10.1145/3173574.3174142.
- [12] Adnan Marquez-Borbon and Juan P. Martinez-Avila. 2018. The Problem of DMI Adoption and Longevity: Envisioning a NIME Performance Pedagogy. In Conference Proceedings of NIME’18, Blacksburg, VA, USA. June 3-6, 2018.
- [13] Sean McGrath, Alan Chamberlain & Steve Benford (2016) “Making Music Together: An Exploration of Amateur and Pro-Am Grime Music Production”, Audio Mostly 2016, ACM. DOI 10.1145/2986416.2986432
- [14] Janice Waldron. (2013) “User-generated content, YouTube and participatory culture on the web: Music learning and teaching in two contrasting online communities”. Music Education Research, 15(3):257–274