CONVERSATIONS WITH EXPERT USERS IN MUSIC RETRIEVAL AND RESEARCH CHALLENGES FOR CREATIVE MIR

Kristina Andersen
Studio for Electro Instrumental Music (STEIM)
Amsterdam, the Netherlands
kristina@steim.nl

Peter Knees
Dept. of Computational Perception
Johannes Kepler University Linz, Austria
peter.knees@jku.at

ABSTRACT

Sample retrieval remains a central problem in the creative process of making electronic dance music. This paper describes the findings from a series of interview sessions involving users working creatively with electronic music. We conducted in-depth interviews with expert users on location at the Red Bull Music Academies in 2014 and 2015. When asked about their wishes and expectations for future technological developments in interfaces, most participants mentioned very practical requirements of storing and retrieving files. A central aspect of the desired systems is the need to provide increased flow and unbroken periods of concentration and creativity.

From the interviews, it becomes clear that for Creative MIR, and in particular, for music interfaces for creative expression, traditional requirements and paradigms for music and audio retrieval differ to those from consumer-centered MIR tasks such as playlist generation and recommendation and that new paradigms need to be considered. Despite all technical aspects being controllable by the experts themselves, searching for sounds to use in composition remains a largely semantic process. From the outcomes of the interviews, we outline a series of possible conclusions and areas and pose two research challenges for future developments of sample retrieval interfaces in the creative domain.

1. MOTIVATION AND CONTEXT

Considerable effort has been put into analysing user behaviour in the context of music retrieval in the past two decades [35]. This includes studies on music information seeking behaviour [14,17], organisation strategies [15], usage of commercial listening services [36], the needs or motivations of particular users, such as kids [28], adolescents [34], or musicologists [29], and behaviour analysis for specific tasks, e.g., playlist and mix generation [13], or in specific settings, e.g., riding together in a car [16] or in music lessons in secondary schools [49].

© Kristina Andersen, Peter Knees. Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). **Attribution:** Kristina Andersen, Peter Knees. "Conversations with Expert Users in Music Retrieval and Research Challenges for Creative MIR", 17th International Society for Music Information Retrieval Conference, 2016.

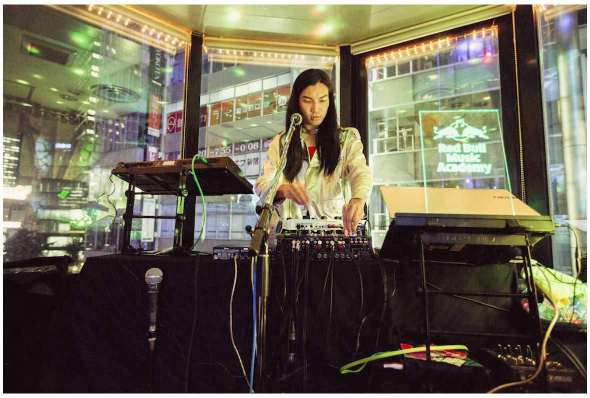

Figure 1. Live electronic music performance at the Red Bull Music Academy 2014

In this paper, we want to address music retrieval from the perspective of music producers, thus investigate the user behaviour of a group that deals with audio retrieval professionally on a daily basis, but has received comparatively less attention in MIR research so far—as have other questions from the area of Creative MIR [27].

The majority of today's electronic music is created from pre-recorded or live-generated sound material. This process often combines sound loops and samples with synthesized and processed elements using a so-called digital audio workstation (DAW), an electronic device or computer application for recording, editing and producing audio files. In these systems, Music Information Retrieval (MIR) methods, e.g., for content analysis, gain importance. In essence, future tools and applications need to be aware of the nature and content of the music material, in order to effectively support the musician in the creative process.

However, user studies on retrieval for musicians and producers are scarce. Cartwright et al. [6] investigate potential alternatives to existing audio production user interfaces in a study with 24 participants. In another example, Bainbridge et al. [4] explore and test a personal digital library environment for musicians, where based on a spatial paradigm musicians should be able to capture, annotate, and retrieve their ideas, e.g., using query-by-humming. In this paper, our approach is not to test an existing system, but to gain an understanding of the processes involved for music producers, who are used to working with existing music software suites.

This work is organized a follows. In section 2, we identify and briefly discuss existing MIR approaches in the context of music production. In section 3, we describe our motivation for engagement with expert users, our approach of conducting semi-structured interviews, and information on the interview background. The main part of the paper is presented in sections 4 and 5, where we highlight interesting outcomes of the interview sessions and distill central topics. Corresponding to this, we conclude this work by posing two research challenges for Creative MIR (section 6).

2. MIR RESEARCH IN MUSIC MAKING

Existing MIR research targeted at composition support and music production deals with browsing interfaces to facilitate access to large collections of potentially homogeneous material, such as drum samples [41]. Exploration of sample libraries, e.g., [8, 50], is often supported or driven by methods to automatically extract music loops from music files. Given the prevalent techniques of sampling and remixing in today's music production practise, such methods are useful to identify reusable materials and can lead to inspirations for new compositions [40, 51] and new interfaces for remixing [23] In terms of retrieval in the creative domain, and in contrast to consumer-based information systems, the query-by-example paradigm, implemented as query-by-humming or through other vocal inputs [5, 25, 31], is still an active research field.

Other MIR systems facilitate musical creation through automatic composition systems [10] or mosaicing systems that "reconstruct" the sound of a target piece by concatenating slices of other recordings [38, 45, 54]. This principle of concatenative synthesis can also be found in interactive systems for automatic accompaniment or improvisation such as the OMax system by Assayag et al. [3], Audio Oracle by Dubnov et al. [19], or Cont's Antescofo [11].

Other MIR systems emphasize the embodiment of creativity expression. For instance, Schnell et al. [44] propose a system that combines real-time audio processing, retrieval, and playback with gestural control for re-embodiment of recorded sound and music. The Wekinator [22] by Fiebrink is a real-time, interactive machine learning toolkit that can be used in the processes of music composition and performance, as well as to build new musical interfaces and has also shown to support the musical expression of people with disabilities [32].

3. WORKING WITH EXPERT USERS

Standards for user involvement in the field of Human Computer Interaction (HCI) have evolved from a traditional approach of metric user-testing of already designed systems, to understanding of users and their context through ethnographic methods and scenarios, towards an emerging focus on developing empathy with the user's experience of life. Wright and McCarthy state that "'knowing the user' in their lived and felt life involves understanding what it feels like to be that person, what their situation is like from their own perspective." [52]

This is especially important in the case of the creative expert users, who are not just looking to complete a series of tasks, but rather are engaging **with** the technology in order to express themselves **through** it. As such they can be seen to be not only using the technology, but rather collaborating with it as described by Tom Jenkinson (aka Squarepusher): "Through his work, a human operator brings as much about the machine to light as he does about himself ... The machine has begun to participate." [30]

This paper describes some of our efforts at building such an understanding. Eventually, our work will aim to create musical tools that provide new interfaces to the selection of sounds and musical data through music analysis algorithms. The underlying concern will be to not just improve existing user interfaces for the creation of electronic music through increases in efficiency, but facilitate increased flow and unbroken periods of concentration and creativity. To do so, we are engaging with expert users throughout the entire project, allowing them a strong peer position in the conceptualisation and evaluation of any ideas.

Our main users are the participants at the Red Bull Music Academy (RBMA), cf. fig. 1, an event held yearly with a carefully selected group of professional electronic dance music makers on the point of breaking though. [1]

Our sustained involvement with this group of expert users is key to our strategy of building detailed understandings of current forms of electronic music making, cf. [1]. We hope that this will allow us to go beyond user testing, and instead aim for a coherent impression of how an interface may benefit the real-life creative process of the users. To this end, we are committed to conducting interviews in a fashion that fits within the work-flow and interpersonal communication style of these music professionals, we ultimately aim to support creatively with the outcomes of the project. What we need to understand is: How do they organise their work, what are their needs, and ultimately what are their mental models of their music?

We conducted 33 in-depth interviews with expert users on location at the Red Bull Music Academy in Tokyo (2014) and Paris (2015). The aim of the 2014 sessions was to establish understandings of existing work practices among users, and the 2015 sessions were set up to investigate a number of emergent themes in more detail. Our interviews were executed in an open conversational structure, engaging the interviewees directly as peers, while aiming to support them to go beyond evaluation of current interfaces and into the imagination of new and unknown interfaces for their own creative practice.

The interviews were audio recorded and fully transcribed. Three independent HCI researchers analysed the

[1] `http://redbullmusicacademy.com`; From the web page: "The Red Bull Music Academy is a world-travelling series of music workshops and festivals *[in which]* selected participants – producers, vocalists, DJs, instrumentalists and all-round musical mavericks from around the world – come together in a different city each year. For two weeks, each group will hear lectures by musical luminaries, work together on tracks and perform in the city's best clubs and music halls."

transcripts for content-rich quotes: short sentences or paragraphs describing a particular idea or concern. The keywords from these quotes were extracted and used for labelling the quote, e.g., "search", "finding", "colour", or "pace". Following this, keywords were manually clustered into bigger concepts or themes. From the material collected, we identified the themes of *Search*, *Categories*, *Visualisation*, *Colour*, *Organisation*, *Assistance*, *Workflow*, *Connections*, *Correction*, *Suggestions*, *Obstructions*, *Deliberate error*, *Tweaks*, *Interfaces*, and *Live*. The material we address in this paper belongs to the themes of *Search*, *Categories*, *Colour*, *Visualisation*, *Organisation*, *Suggestions*, and *Obstructions*.

4. INTERVIEW QUOTES AND FINDINGS

When asked about their wishes and expectations for future technological developments, most participants mentioned very practical requirements for storing and retrieving files. Sounds are usually stored as short audio files (samples) or presets, which can be loaded into playback devices in composition software.

> *"Because we usually have to browse really huge libraries [...] that most of the time are not really well organized."* (TOK003)

> *"If you have like a sample library with 500,000 different chords it can take a while to actually find one because there are so many possibilities."* (TOK015)

> *"Like, two hundred gigabytes of [samples]. I try to keep some kind of organisation."* (TOK006)

> *"I easily get lost... I always have to scroll back and forth and it ruins the flow when you're playing"* (PA011)

> *"...what takes me really long time is organising my music library for DJing. [...] Yes, it could be something like Google image search for example. You input a batch of noise, and you wait for it to return a sound."* (TOK011)

Even from this small selection of statements it becomes clear that organisation of audio libraries, indexing, and efficient retrieval plays a central role in the practice of music creators and producers. However, in the search tools provided by existing DAWs, this aspect seems addressed insufficiently. When asked directly: "How do you find what you are looking for?" answers indicated a number of personal strategies that either worked with, or sometimes in opposition to, the existing software design.

> *"You just click randomly and just scrolling, it takes for ever!"* (TOK009)

> *"Sometimes, when you don't know what you are looking for, and you're just going randomly through your samples, that might be helpful, but most of the time I have something in mind that I am looking for, and I am just going through all these sound files, and I am just waiting for the sound which I had in mind to suddenly appear. Or what comes the closest to what I had in mind. So I think that most of the time, I know what I am looking for, and then it is just a matter of time before I find it."* (TOK002)

> *"Part of making music is about being lost a little bit and accidentally stumbling upon stuff that you didn't think would work."* (TOK007)

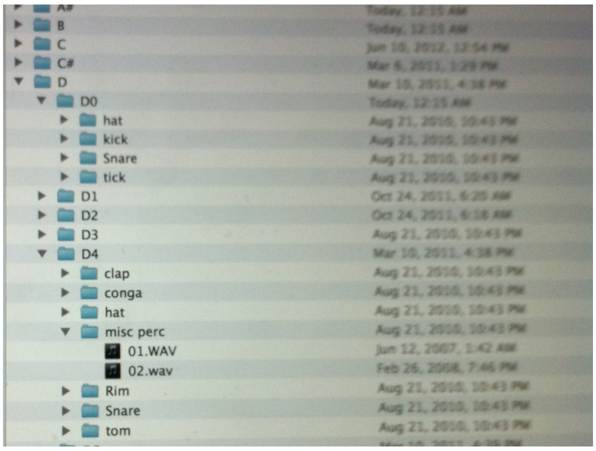

Figure 2. User sample file collection, photographed from laptop screen of expert music producer at RBMA 2014.

This highlights a key element of much creative work, the element of the accidental, sometimes caused by the positive and negative effects of malfunctioning of sound-editing software. Our *finding 1* is that serendipity is highly important to support creative work and that when existing software is not providing this desired functionality, workarounds will be created.

Often, users are constructing their own personal systems for searching, sometimes working with the structures available to them, but often developing idiosyncratic and personal strategies of misuse or even randomness. We also see users painstakingly creating hierarchical categorisation schemes manually in order to stay in control of their own sound collections as seen in Figure 2.

Searching for sounds to use in composition remains a broadly semantic process, where the user has a certain structure of meaning in mind when querying the collection. Current solutions in this direction rely on databases with sets of meta-data (tags) for the available sounds. However, all available solutions that come with pre-tagged sounds use a static semantic structure, which cannot adapt to the users individual understanding of the sounds. This is especially problematic when the user has a specific target sound in mind, but does not know how this would be described in the pre-tagged semantic structure of the database. In short, as our *finding 2* we see that our users have mental images of sound that they translate to verbal expressions.

> *"So it would be really useful to for example have some kind of sorting system for drums, for example, where I could for example choose: 'bass drum', and here it is: 'bass' and 'bright', and I would like it to have maybe bass drum 'round' and 'dry', and you can choose both, the more I choose, of course, the less results I will have [...] So it is filtering it down, that is really helpful, if it works well of course."* (TOK002)

> *"It would be even more useful to be able to search for a particular snare, but I can't really imagine how, I need something short, low in pitch, dark or bright in tone and then it finds it..."* (TOK003)

> *"There are a lot of adjectives for sound, but for me, if you want a 'bright' sound for example it actually means a sound with a lot of treble, if you say you want a 'warm' sound, you put a round bass, well, round is another adjective."* (TOK009)

We also see that the semantic descriptions used in these descriptions are very individually connoted and often stemming from non-auditory domains, such as haptic, or most prominently, the visual domain (e.g., round, dark, bright). Thus, the mental images expressed verbally are often actually rooted in another domain.

In fact, one solution to the problem of indexing suggested by our participants is to organize sounds by sources, by projects in which they have been used, or to colour-code them.

> *"If I have hundreds of tracks, I have to colour code everything, and name properly everything. It's a kind of system, and also kind of I feel the colours with the sounds, or maybe a rose, if kind of more orange, and brownish and maybe... I use that kind of colour coding."* (TOK006)

Finding 3 is that we see a need for semantic representations of sounds, but it's not only a matter of just tags and words, but rather an ability to stay much closer to the vocabulary and mental representations of sound of each user.

Additionally, the question arises whether the interviewees really want a direct representation of their mental map in the data structure, or if indeed they rather expect something more akin to a machine collaborator, that could come up with its own recommendations and suggest a structure based on the personal requirements of the individual user.

> *"I'd like it to do the opposite actually, because the point is to get a possibility, I mean I can already make it sound like me, it's easy."* (TOK001)

> *"What I would probably rather want it would do is make it complex in a way that I appreciate, like I would be more interested in something that made me sound like the opposite of me, but within the boundaries of what I like, because that's useful. Cause I can't do that on my own, it's like having a band mate basically."* (TOK007)

Again we see the computer as a potential collaborator, one that might even be granted some level of autonomy:

> *"Yeah, yeah, well I like to be completely in charge myself, but I like to... I don't like other humans sitting the chair, but I would like the machine to sit in the chair, as long as I get to decide when it gets out."* (TOK014)

Finding 4 is that instead of computer-generated and similarity-based recommendations, desired features are surprise, opposition, individuality, and control over the process (with the possibility to give up control when needed).

5. INTERVIEW CONCLUSIONS

From the interviews outlined in the previous section, we see that central concerns in everyday work in music production are core topics of MIR: indexing, retrieval, browsing, recommendation, and intuitive (visual) interfaces. More than in music retrieval systems built for consumer or entertainment needs, the expert user in a music production environment will actually evaluate a large part of the returned items to find the best—a process that is integral to some as a way to get inspired.

In order to facilitate this process and to enable creative work to make better use of MIR tools, we identified four crucial findings:

1. Surprise and serendipity in recommendation and retrieval are important to support creative work
2. Users have personal mental images of sound
3. There is a need for semantic representations of sounds for retrieval, which are not just tags and words but rather reflect those mental images (which can be visual or haptic)
4. Instead of "more of the same" recommendations, desired features are surprise, opposition, individuality, and control over the recommendation process

The desires for systems that respond to personal vocabularies and individual mental images of sound alongside the desire to have a controllable element of otherness and difference, constitute both a challenge and an opportunity. However, this also goes someway towards illustrating how individualized the creative expert user needs may turn out to be. While we can try to find common concerns, it is clear that no system will fit the requirements of all users. In the creative domain even more than in consumer-oriented

MIR, allowing personalisation of the systems is a central requirement, cf. [4]. This is reflected in the two following areas, which take the findings into a bigger context and outline two conceptual ideas for future Creative MIR research, namely "The Collaborative Machine," building upon findings 1 and 4 to go beyond the idea of a traditional recommender system, and "Synaesthetic Sound Retrieval," based upon findings 2 and 3, as an approach beyond tag-based "semantic retrieval".

5.1 The Collaborative Machine

The collaborative machine can be imagined as a virtual bandmate who assesses, critiques, takes over, and occasionally opposes.

It appears that in creative work as well as in the consumer world, successful imitation is not enough for the machine to be recognized as "intelligent" anymore. While this is a first and necessary step in creative and intelligent behavior, a machine requires more multi-faceted and complex behavior in order to be considered a useful advice-giver or even collaborator. However, no matter how well grounded or wise they can be, artificial knowledge and expert agent-based advice might be completely useless or, even worse, annoying and even odious. Aspects of such human behavior, as well as of surprise, opposition, and obstruction, should contribute to making the interaction with the machine more interesting and engaging.

Can we imagine an intelligent machine providing the user with creative obstructions in the place of helpful suggestions?

A creative obstruction is based on the artistic technique of "defamiliarisation" as defined by Shklovsky [47]—a basic artistic strategy central to both Surrealism and Dada. It is based on the idea that the act of experiencing something occurs inside the moment of perceiving it and that the further you confuse or otherwise prolong the moment of arriving at an understanding, the deeper or more detailed that understanding will be. This technique and the findings from the interviews can be directly translated into new requirements for recommendation engines in music making.

This need for opposition goes far beyond the commonly known and often addressed *needs for diversity, novelty, and serendipity* in recommendation system research, which has identified purely similarity-based recommendation as a shortcoming that leads to decreased user satisfaction and monotony [7, 48, 53]. This phenomenon spans multiple domains: from news articles [37] to photos [42] to movies [18]. One idea proposed to increase diversity is to subvert the basic idea of collaborative filtering systems of recommending what people with similar interests found interesting ("people with similar interests also like...") by recommending the opposite of what the least similar users (the k-furthest neighbors) want [43]. Indeed it could be shown that this technique allows to increase diversity among relevant suggestions.

In the context of experimental music creation, Collins has addressed the question of opposition in the Contrary Motion system [9] using a low-dimensional representation of rhythm. The system opposes a piano player's rhythm in real time by constructing a structure located in the space of actions "where the human is likely not to be" [9]. The hypothesis underlying the system is that being confronted with an oppositional music style can be stimulating for a musician. Experiments where the opposing structure is sonified using a different instrument have indeed shown that musicians start to experiment and play with the opposing agent. For future work, it would be interesting to see whether computer-system-created music (or a system that suggests fragments) will be accepted by experts or declined, cf. [39].

5.2 Synaesthetic Sound Retrieval

Multiple search avenues allow the user to use many different ways to describe the searched-for sound. This includes acoustic sketching, e.g., [5, 25, 31], as well as graphical representations.

In a number of quotes in section 4, sounds are described by shapes (round), brightness (bright, dark) and textures (soft, dry). While these might be regarded as unusual descriptors of sound, there is some evidence that many humans make to some degree use of *synaesthetic* connections between visual perceptions and sound. In the Creative MIR scenario, we make use of a weak definition of synaesthesia as cross-modal associations, cf. [20, 26], and, in the context of computer science, "the more general fact that digital technologies offer, if not a union of the senses, then something akin: the inter-translatability of media, the ability to render sound as image, and vice versa." [12]

Focusing on the visual domain, through the interviews, a number of ideas and notions came up in addition to the importance of brightness, shape, and texture for sound retrieval. More precisely, colour plays a central role: *"I see the music sometimes as more aesthetic and something that I can see more than something that I can hear"* (PA013), *"When I listen to music I see colours. [...] I remember colours."* (PA011), *"Different sounds to me have specific colours. ... [For finding files,] I don't have the actual ability to use images [now], so I just use colour."* (PA009).

Such a colour-coding, for instance, takes the role of "semantic tagging". The fact that a system needs time to learn a user's associations first, i.e., that it might not work perfectly out of the box but learn their associations over time (personalisation), is understood and accepted:

> *"You could imagine that your computer gets used to you, it learns what you mean by grainy, because it could be different from what that guy means by grainy."* (PA008)

The models learned from personal categories and (visual) tagging could then be applied on new collections for personal indexing.

For browsing sound, the idea of tapping into the visual domain is well-established. Most proposed sound browsing systems are based on 2D arrangements of sounds [8, 21, 41, 46]—even including personalised adaptation of the arrangement [24]. In these systems, the visual aspect is the spatial arrangement of sounds, however, this does not reflect the mental models but rather requires the user to learn the mapping provided by the system. Grill and Flexer [26] get closer to a synaesthetic representation by visualizing perceptual qualities of sound textures through symbols on a grid. [2] To this end, they map bipolar qualities of sound that describe spectral and temporal aspects of sound to visual properties. The spectral qualities of pitch (high vs. low) and tonality (tonal vs. noisy) are mapped to brightness and hue, and saturation, respectively. The temporal (or structural) qualities of smoothness vs. coarseness, order vs. chaos, and homogeneity vs. heterogeneity are associated with the jaggedness of an element's outline, the regularity of elements on the grid, and a variation in colour parameters, respectively.

To the best of our knowledge, there is no system that matches a visual query representing a mental image of sound to a sound collection for retrieval. Developing such a system would, however, pose an interesting challenge.

6. POSED CHALLENGES

Results that enable the two conceptual ideas discussed above can not be trivially achieved. Therefore, to conclude this paper, we want to pose two topics as challenges for future work in Creative MIR to the wider community. Both of these topics should allow for alternative retrieval paradigms particularly relevant in creative work. As discussed before, they require high levels of personalisation in order to facilitate "semantic" retrieval.

Challenge 1: developing models for exploring dissimilarity in search
To arrive at an artificial collaborator capable of inspiring by opposing, the concept of opposition needs to be explored first, cf. [2]. Music similarity is a multi-dimensional concept and while proximity can be easily, "semantically" defined through minimizing distance measures, the concept of dissimilarity is by far more difficult to capture as it "spreads out" to different directions and dimensions of sound. Finding dissimilar sounding audio from a given query is therefore more challenging and requires individual user models of music perception as well as a solid understanding of usage context in order to derive an understanding of sounding "different".

Challenge 2: developing retrieval methods for visual queries
This challenge is to develop a software interface for sound search based on queries consisting of sketches of mental images, cf. [33]. A central requirement for such an interface is that it needs to be able to deal with different sound properties and different types of sounds, such as effects, samples, ambient, tonal, or textured recordings, and therefore comprise different simultaneous representational models for indexing. For instance, while tonal aspects might be best represented using symbolic music notation, noise sounds should be modeled primarily via their textural properties. It is expected that modeling and indexing will heavily draw from audio content processing and analysis methods—again—in order to cover a wide range of sound property dimensions.

We hope that these challenges will drive the discussion on Creative MIR and its applications in music production and help reflecting upon and advancing the field of music retrieval also beyond the specific area of study of this work.

[2] http://grrrr.org/test/texvis/map.html

7. ACKNOWLEDGEMENTS

This work is supported by the European Union's seventh Framework Programme FP7 / 2007–2013 for research, technological development and demonstration under grant agreement no. 610591 (GiantSteps). Figure 1 courtesy of Red Bull Music Academy.

8. REFERENCES

[1] K. Andersen and F. Grote. "GiantSteps: Semi-structured conversations with musicians". In *Proc CHI EA*, 2015.

[2] K. Andersen and P. Knees. "The dial: Exploring computational strangeness". In *Proc CHI EA*, 2016.

[3] G. Assayag, G. Bloch, M. Chemillier, A. Cont, and S. Dubnov. "OMAX Brothers: A dynamic topology of agents for improvisation learning". In *Proc ACM MM Workshop on Audio and Music Computing for Multimedia*, 2006.

[4] D. Bainbridge, B.J. Novak, and S.J. Cunningham. "A user-centered design of a personal digital library for music exploration". In *Proc JCDL*, 2010.

[5] M. Cartwright and B. Pardo. "Synthassist: Querying an audio synthesizer by vocal imitation". In *Proc NIME*, 2014.

[6] M. Cartwright, B. Pardo, and J. Reiss. "Mixploration: Rethinking the audio mixer interface". In *Proc IUI*, 2014.

[7] Ò. Celma and P. Herrera. "A new approach to evaluating novel recommendations". In *Proc RecSys*, 2008.

[8] G. Coleman. "Mused: Navigating the personal sample library". In *Proc ICMC*, 2007.

[9] N. Collins. "Contrary motion : An oppositional interactive music system". In *Proc NIME*, 2010.

[10] N. Collins. "Automatic composition of electroacoustic art music utilizing machine listening". *CMJ*, 36(3):8–23, 2012.

[11] A. Cont. ""Antescofo: Anticipatory synchronization and control of interactive parameters in computer music". In *Proc ICMC*, 2008.

[12] C. Cox. "Lost in translation: Sound in the discourse of synaesthesia". *Artforum International*, 44(2):236–241, 2005.

[13] S.J. Cunningham, D. Bainbridge, and A. Falconer. "'More of an art than a science': Supporting the creation of playlists and mixes". In *Proc ISMIR*, 2006.

[14] S.J. Cunningham, J.S. Downie, and D. Bainbridge. "'The pain, the pain': Modelling music information behavior and the songs we hate". In *Proc ISMIR*, 2005.

[15] S.J. Cunningham, M. Jones, and S. Jones. "Organizing digital music for use: An examination of personal music collections". In *Proc ISMIR*, 2004.

[16] S.J. Cunningham, D.M. Nichols, D. Bainbridge, and H. Ali. "Social music in cars". In *Proc ISMIR*, 2014.

[17] S.J. Cunningham, N. Reeves, and M. Britland. "An ethnographic study of music information seeking: Implications for the design of a music digital library". In *Proc JCDL*, 2003.

[18] T. Di Noia, V.C. Ostuni, J. Rosati, P. Tomeo, and E. Di Sciascio. "An analysis of users' propensity toward diversity in recommendations". In *Proc RecSys*, 2014.

[19] S. Dubnov, G. Assayag, and A. Cont. "Audio oracle: A new algorithm for fast learning of audio structures". In *Proc ICMC*, 2007.

[20] K.K. Evans and A. Treisman. "Natural cross-modal mappings between visual and auditory features". *JOV*, 10(1):6, 2010.

[21] M. Fernström and E. Brazil. "Sonic browsing: An auditory tool for multimedia asset management". In *Proc ICAD*, 2001.

[22] R. Fiebrink. *Real-time Human Interaction with Supervised Learning Algorithms for Music Composition and Performance*. PhD thesis, Princeton University, NJ, USA, Jan 2011.

[23] J. Forsyth, A. Glennon, and J.P. Bello. "Random access remixing on the iPad". In *Proc NIME*, 2011.

[24] O. Fried, Z. Jin, R. Oda, and A. Finkelstein. "AudioQuilt: 2D arrangements of audio samples using metric learning and kernelized sorting". In *Proc NIME*, 2014.

[25] O. Gillet and G. Richard. "Drum loops retrieval from spoken queries". *JIIS*, 24(2-3):159–177, 2005.

[26] T. Grill and A. Flexer. "Visualization of perceptual qualities in textural sounds". In *Proc ICMC*, 2012.

[27] E.J. Humphrey, D. Turnbull, and T. Collins. "A brief review of Creative MIR". In *ISMIR Late-Breaking News and Demos*, 2013.

[28] M. Hutter and S.J. Cunningham. "Towards the design of a kids' music organizer". In *Proc CHINZ*, 2008.

[29] C. Inskip and F. Wiering. "In their own words: Using text analysis to identify musicologists' attitudes towards technology". In *Proc ISMIR'15*, 2015.

[30] T. Jenkinson. "Collaborating with machines". *Flux Magazin (www.fluxmagazine.com)*, March 2004.

[31] A. Kapur, M. Benning, and G. Tzanetakis. "Query-by-beat-boxing: Music retrieval for the DJ". In *Proc ISMIR*, 2004.

[32] S. Katan, M. Grierson, and R. Fiebrink. "Using interactive machine learning to support interface development through workshops with disabled people". In *Proc CHI*, 2015.

[33] P. Knees and K. Andersen. "Searching for audio by sketching mental images of sound – A brave new idea for audio retrieval in creative music production". In *Proc ICMR*, 2016.

[34] A. Laplante and J.S. Downie. "Everyday life music information-seeking behaviour of young adults". In *Proc ISMIR*, 2006.

[35] J.H. Lee and S.J. Cunningham. "Toward an understanding of the history and impact of user studies in music information retrieval". *JIIS*, 41(3):499–521, 2013.

[36] J.H. Lee and R. Price. "Understanding users of commercial music services through personas: Design implications". In *Proc ISMIR*, 2015.

[37] L. Li, D. Wang, T. Li, D. Knox, and B. Padmanabhan. "Scene: A scalable two-stage personalized news recommendation system". In *Proc SIGIR*, 2011.

[38] E. Maestre, R. Ramírez, S. Kersten, and X. Serra. "Expressive concatenative synthesis by reusing samples from real performance recordings". *CMJ*, 33(4):23–42, Dec 2009.

[39] D.C. Moffat and M. Kelly. "An investigation into people's bias against computational creativity in music composition". *Assessment*, 13(11), 2006.

[40] B.S. Ong and S. Streich. "Music loop extraction from digital audio signals". In *Proc ICME*, 2008.

[41] E. Pampalk, P. Hlavac, and P. Herrera. "Hierarchical organization and visualization of drum sample libraries". In *Proc DAFx*, 2004.

[42] A.-L. Radu, B. Ionescu, M. Menéndez, J. Stöttinger, F .Giunchiglia, and A. De Angeli. "A hybrid machine-crowd approach to photo retrieval result diversification". *Proc MMM* 2014.

[43] A. Said, B. Fields, B.J. Jain, and S. Albayrak. "User-centric evaluation of a k-furthest neighbor collaborative filtering recommender algorithm". In *Proc CSCW*, 2013.

[44] N. Schnell, F. Bevilacqua, N. Rasamimanana, J. Blois, F. Guédy, and E. Fléty. "Playing the 'MO' – Gestural Control and Re-Embodiment of Recorded Sound and Music". In *Proc NIME*, 2011.

[45] D. Schwarz. "Current research in concatenative sound synthesis". In *Proc ICMC*, 2005.

[46] D. Schwarz and N. Schnell. "Sound search by content-based navigation in large databases". In *Proc SMC*, 2009.

[47] V. Shklovsky. *Art as Technique. In: Russian Formalist Criticism: Four Essays, 1965. Trans. L.T. Lemon and M.J. Reis.* University of Nebraska Press, Lincoln, USA, 1917.

[48] B. Smyth and P. McClave. "Similarity vs. diversity". In *Proc ICCBR*, 2001.

[49] D. Stowell and S. Dixon. "MIR in school? Lessons from ethnographic observation of secondary school music classes". In *Proc ISMIR*, 2011.

[50] S. Streich and B.S. Ong. "A music loop explorer system". In *Proc ICMC*, 2008.

[51] H.L. Tan, Y. Zhu, S. Rahardja, and L. Chaisorn. "Rhythm analysis for personal and social music applications using drum loop patterns". In *Proc ICME*, 2009.

[52] P. Wright and J. McCarthy. "Empathy and experience in HCI". In *Proc CHI*, 2008.

[53] Y.C. Zhang, D. Ó Séaghdha, D. Quercia, and T. Jambor. "Auralist: Introducing serendipity into music recommendation". In *Proc WSDM*, 2012.

[54] A. Zils and F. Pachet. "Musical mosaicing". In *Proc DAFx*, 2001.