The GiantSteps Project: A Second-Year Intermediate Report

Peter Knees,[1] Kristina Andersen,[2] Sergi Jordà,[3] Michael Hlatky,[4] Andrés Bucci,[5] Wulf Gaebele,[6] Roman Kaurson[7]
[1] Dept. of Computational Perception, Johannes Kepler University Linz, Austria
[2] Studio for Electro-Instrumental Music (STEIM), Amsterdam, Netherlands
[3] Music Technology Group, Universitat Pompeu Fabra, Barcelona, Spain
[4] Native Instruments GmbH, Berlin, Germany [5] Reactable Systems S.L., Barcelona, Spain
[6] Yadastar GmbH, Cologne, Germany [7] JCP-Connect SAS, Rennes, France
info@giantsteps-project.eu

ABSTRACT

We report on the progress of GiantSteps, an EU-funded project involving institutions from academia, practitioners, and industrial partners with the goal of developing new concepts for intelligent and collaborative interfaces for music production and performance. At the core of the project is an iterative, user-centric research approach to music information retrieval (MIR) and human computer interaction (HCI) that is designed to allow us to accomplish three main targets, namely (1) the development of intelligent musical expert agents to support and inspire music makers, (2) more intuitive and collaborative interfaces, and (3) low-complexity methods addressing low-cost devices to enable affordable and accessible production tools and apps. In this paper, we report on the main findings and achievements of the project's first two years.

1. INTRODUCTION

The stated goal of the GiantSteps project is to create the so-called "seven-league boots" for future music production.[1] Built upon an iterative and user-centric research approach to music information retrieval (MIR) and human computer interaction (HCI), the project is developing digital musical tools and music analysis components that provide more intuitive and meaningful interfaces to musical data and knowledge in order to empower music practitioners to use their creative potential.[2] In particular, we want to achieve this by targeting three directions:

1. Developing **musical expert agents**, i.e., supportive systems for melody, harmony, rhythm, or style to guide users when they lack inspiration or technical or musical knowledge.

2. Developing **improved user interfaces** and paradigms for (collaborative) musical human-computer interaction that are easily graspable by novices and lead to unbroken workflows for professionals.

3. Developing low-complexity algorithms for music analysis addressing **low-cost devices** to enable affordable and accessible production tools and apps.

In order to meet these goals, GiantSteps is set up as a transdisciplinary research project, carried out by a strong and balanced consortium including leading music information research institutions *(UPF-MTG, CP-JKU)*, leading industry partners *(Native Instruments, Reactable, JCP-Connect)* and leading music practitioners *(STEIM, Red Bull Music Academy/Yadastar)*.[3]

With this consortium, the project aims at combining techniques and technologies in new and unprecedented ways, all driven by users' practical needs. This includes the combination of state-of-the-art interfaces and interface design techniques with advanced methods in music information retrieval (MIR) research that have not yet been applied in a real-time interaction context or with creativity objectives. In addition to this, the industrial partners ensure alignment of the developments with existing market requirements, allowing for a smooth integration of outcomes into real-world systems.

In this report, we describe the findings and achievements of the project within the first two years. Section 2 outlines the user-centric design approach that generates the requirements for the technical developments. Section 3 deals with the advances in MIR that are necessary in order to enable music expert agents, novel visualizations, and new interfaces as discussed in section 4, as well as more effective algorithms for low-resource devices. Section 5 describes the outcomes of the project in terms of market-released products. Section 6 concludes with an outlook to the project's final year and beyond.

2. USER-CENTRIC DESIGN APPROACH

The project's constellation as a collaboration between music research institutions, manufacturers of software and hardware for music production, R&D companies and music practitioners allows us to engage in an ongoing conversation with the professional makers of Electronic Dance

[1] *http://www.giantsteps-project.eu*
[2] Note that parts of this paper have already been published in [1].

[3] *http://redbullmusicacademy.com*; From the webpage: "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."

Figure 1. Impressions from the user sessions: participatory workshop, mock-up prototype for music event live manipulation, visual sound query interface, and practitioners at the Red Bull Music Academy in Tokyo (from left to right).

Music (EDM), that we consider our key users, and to tailor musical tools to their needs. This takes the form of generating user requirements and testing prototypes with end-users in an iterative process throughout the project.

The overall goals of this user involvement are to establish a range of current creative practices for musical expression in EDM, explore mental models of musical qualities, produce user-generated ideas through explorative making, and inspire design and non-design related tasks within the project, cf. [2, 3]. To this end, we conduct a series of different workshop and interview sessions involving expert users (cf. fig. 1). The user sessions comprise interface-specific and work-practice-related interviews and cognitive walkthroughs, e.g., to identify breaks in workflows, as well as ad-hoc, open-ended interviews, carried out on location at STEIM, Native Instruments, Music Hack Days, and the Red Bull Music Academy, resulting in interactions with over 200 individual users so far.

To ensure traceability of the identified goals and requirements throughout the process of developing prototypes, we have set up a system for managing prototypes, generating keyword requirements, and exploring ideas in functional and non-functional prototypes. Fig. 2 illustrates the overall flow of user involvement in the project. From user conversations, we extract the most pertinent ideas as keywords that are either addressed in a concrete technical implementation or — if not yet at that point — a conceptual prototype. Either can be exposed to the user or, in case of a technical prototype, quantitatively evaluated (particularly low-level MIR algorithms). To close the circle, results are evaluated with users, leading to a new round of user conversations informing the project's next iteration.

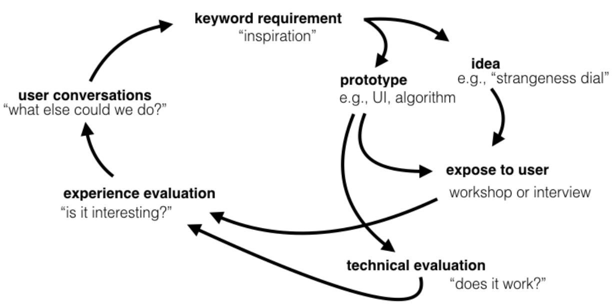

Figure 2. User involvement flow as a circular process.

As concrete examples, from our open-ended interview sessions, a number of ideas and requirements have emerged addressing the areas of structuring audio repositories and describing and finding sounds, embodiment and physical devices, and the role of the (collaborating) machine in music creation. In the following sections we will lead with example statements from our expert users to demonstrate how these findings are informing the technical research in the project and to illustrate the circular flow of the development process depicted in fig. 2. More details on the studies and general considerations of the methodological approach can be found in [4, 5, 6, 7].

3. MIR FOR MUSIC PRODUCTION

Music information retrieval plays a central role in the project. The goal of the research is to develop high-performance and low-complexity methods for music and audio analysis that allow for extraction of musical knowledge in order to drive intelligent composition agents and visualizations (cf. section 4). The following user quotes demonstrate a need for accurate music analysis methods.

> *"Onset detection, beat detection, tempo detection and harmony detection is pretty much all we need. [...] Being able to pick out a small musical phrase of any kind in a big bunch of noises could be extremely helpful for people like me. Instead of spending nights equalizing something to get out a small musical idea."* [Tok003]

> *"...if you had technology that could tag all your drum samples, that this one is like dirty or distorted, 43Hz is the dominant frequency..."* [Tok007]

Inspired by these and other user statements, we developed new and improved MIR algorithms for onset detection, beat and downbeat tracking [8, 9, 10, 11], tempo estimation [12, 13], key detection [14], chord extraction, melody extraction, style description and classification, instrument classification, drum sample description and recommendation, and drum transcription [15] for electronic dance music (EDM).

Our methods for onset detection, beat tracking, and tempo estimation [4] have successfully competed in the scientific MIREX evaluation campaign and yielded the top ranks in their respective tasks in two consecutive years. [5] [6] Fur-

[4] made available via `https://github.com/CPJKU/madmom`

[5] `http://www.music-ir.org/mirex/wiki/2014:MIREX2014_Results`

[6] `http://www.music-ir.org/mirex/wiki/2015:MIREX2015_Results`

thermore, steps towards the optimization of algorithms for mobile platforms have been undertaken by establishing an audio analysis and benchmarking framework for the iOS mobile platform and a real-time-capable analysis library [7] for use in Pure Data and Max/MSP environments, both based on the Essentia audio analysis library. [8] The released libraries are not only of interest to researchers but also address music hackers and developers who often are music practitioners themselves. In addition to signal-based approaches to music analysis, we also investigate the potential of online resources to provide semantic information on music and music styles [17]. Developed software libraries and tools are made available via the GiantSteps GitHub account. [9]

To facilitate MIR research in these areas also outside the consortium, two test collections for tempo and key detection in EDM were created and released [18]. [10] The GiantSteps Key data set has already been adopted as an evaluation data set for the MIREX 2015 key detection task.

4. EXPERT AGENTS AND NEW INTERFACES

The musical knowledge extracted through MIR methods is used to inform supportive and inspirational music expert agents as well as enable new visualisations and interfaces. While users generally welcome the possibility of compositional support by intelligent systems, we found that this is a sensitive matter as it can not only disturb the creative process but also challenge the artistic identity of the user.

> *"It can turn to be pretty invasive pretty fast, and, like, annoying, if it doesn't work properly."* [Ber002]
>
> *"I am happy for it to make suggestions, especially if I can ignore them."* [Tok007]
>
> *"I'm sceptical about introducing, you know, stuff like melody into it, like, here's a suggested kind of thing which fits nicely the two or three patterns you already got in there, then you are really kind of like creating melodies for them, then it's like (laughs), then it's really like, you know, who is the composer of this?"* [Ber003]

Thus, we have to ensure that these systems are present when needed, but do not take over or inhibit the creative process. So far, the expert agents developed were encapsulated in designated UI modules that can be invoked when seeking inspiration but otherwise do not invade the existing workflow. Developed suggestion systems are concerned with rhythmic variations, tonality-aware scale restrictions, concatenative sound generation based on timbre, or arpeggiations, among others.

Due to the importance of rhythm and drum tracks in many contemporary electronic dance genres, quite some effort was devoted to rhythm pattern variation and generation. The goal of this research is to develop algorithms to recommend variations of a rhythm pattern to a user in an

[7] http://mtg.upf.edu/technologies/EssentiaRT~

[8] http://essentia.upf.edu

[9] https://github.com/GiantSteps

[10] The test collections together with benchmarking results comparing academic algorithms and commercial products can be found at http://www.cp.jku.at/datasets/giantsteps/.

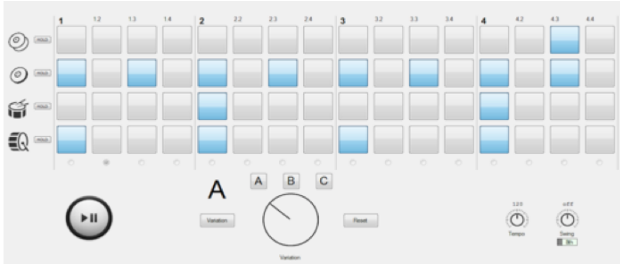

Figure 3. Intelligent user interface for rhythm variation, controllable via hardware interfaces and connected with digital audio workstation through MIDI in and out.

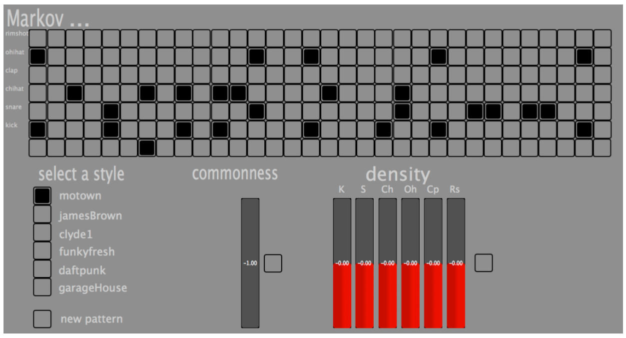

Figure 4. Drumming with style/Markov Drums, running as a Pure Data application inside a VST container.

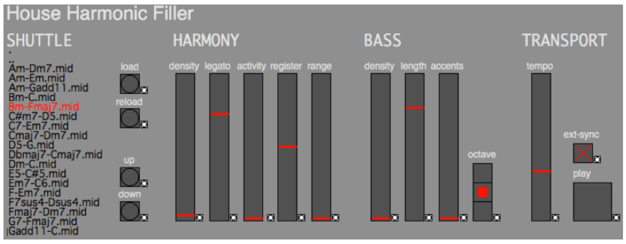

Figure 5. House Harmonic Filler, implemented as a Pure Data application with MIDI-learn control capabilities and master/slave synchronization with external MIDI clocks.

interactive way for usage in live situations and composition. So far, three different approaches were designed and compared, namely pattern variation based on database-retrieval, restricted Boltzmann machines [19], and genetic algorithms [20]. Fig. 3 shows the interface consisting of a simple drum pattern grid editor and a dial for effortless variation, which was received very positively due to its simplicity and creative output.

The prototype shown in fig. 4, is an interactive drum pattern generator based on Markov chains incorporating findings on rhythm similarity through user studies [21, 22, 23]. It addresses rhythm variation from a performance perspective, allowing continuous variations to be controlled by the performer on the basis of high-level musical parameters such as density, syncopation, commonness, amount and rate of variation, while maintaining the drumming style loaded or predefined by the user.

Other prototypes, aim at chord variation in live performance contexts of, currently, House music (see fig. 5), visual browsing interfaces for concatenative synthesis drum generation (see fig. 6), or integrate multiple prototypes to control several facets simultaneously [24].

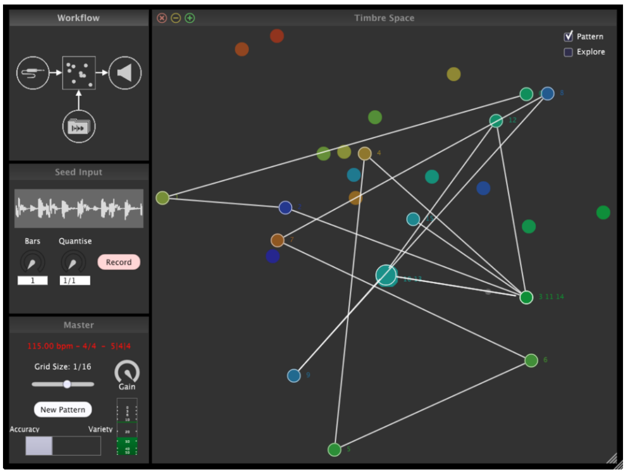

Figure 6. RhythmCAT, a VST-based software instrument that generates new patterns through granular synthesis of sounds clustered in 2D based on timbral similarity.

The integration of these agents in the workflow of music creators is inherently tied to the development of suitable interfaces for either existing desktop-based production and performance suites (i.e., digital audio workstations, such as Apple's Logic, Ableton's Live, Avid's ProTools, Steinberg's Cubase, or NI's Maschine), tangible and/or tabletop user interfaces like the Reactable [25], or smaller multi-touch interfaces of affordable portable devices such as tablets and smartphones. For instance, a developed *automatic tonalizer* expert agent integrates with the Reactable by displaying a virtual keyboard that is restricted to notes that match the scale of sample objects positioned on the table. The impact of the *intelligent arpeggiator, scaler, and chorder* agent can be controlled by hardware dials on a new hardware interface (cf. section 5). Other interface developments relate to the collaborative control of multi-dimensional parameter spaces, leading to intuitive, expressive and tangible input modalities [26, 27].

5. PRODUCT INTEGRATION

Apart from inclusion into publicly accessible developer libraries (cf. section 3), the maturity of the technical developments in the project have allowed us to integrate some of the project's outcomes into market-ready commercial products already. For instance, the Intelligent Arpeggiator, Scale, and Chord Engine has been integrated and released by NI into the Komplete Kontrol Plugin, a plugin shell that is shipped with the Komplete Keyboard (fig. 7) for seamlessly browsing through Komplete instruments, and the iMaschine 2 app for iOS (fig. 8) which was the no.1 app on the US iTunes store for several weeks end of 2015.

The same features were also released as a free update to existing Maschine customers with the free Maschine 2.2 "Melody" update in Nov. 2014, reaching +100k users. The developed Automatic Tonalizer has been integrated by Reactable Systems and will be contained in a future release (see fig. 9). This integration effort will intensify in the third and final year of the project, as more ideas and prototypes mature.

Figure 7. The Native Instruments Kontrol keyboard released in 2015 containing the Intelligent Arpeggiator, Scale, and Chord Engine developed within GiantSteps.

Figure 8. The Native Instruments iMaschine 2 app released in 2015 containing GiantSteps technology.

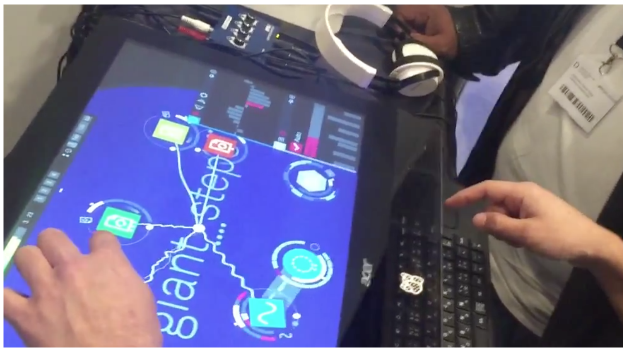

Figure 9. The Reactable Automatic Tonalizer being showcased at Musikmesse 2015.

6. CONCLUSIONS AND OUTLOOK

The consortium and orientation of GiantSteps allow for a genuinely target-user-focused MIR and HCI research approach. This unusual combination of disciplines makes it possible for user's requests and desires to be present in the earliest stages of the MIR algorithm design, a process users are otherwise often excluded from.

While the first two years have primarily been concerned with the *extraction* of musical knowledge by means of MIR technology and the *application* of this knowledge in music expert agents, the third year will have a focus on *interacting* with this knowledge, thus stressing the need for intuitive interfaces as well as for possibilities for collaboration — both, with other musicians and intelligent machines.

Through this process the underlying question for the user remains: Which technology would you want, if you could have anything at all? In practice, as the process moves on, this is refined to questions like: When is an algorithm good enough? How can you annotate or "mark" musical fragments, so that they remain available to you (see also [28])? Can you imagine a system that is able to make valuable suggestions in real time? Could these suggestions also serve as push-back and creative obstructions? and finally: What will it mean for your music, if it works?

Throughout our conversations with users, there are strong desires for improved retrieval mechanisms and inspirational systems that help exploring "the other," e.g., through non-obvious, serendipitous recommendations.

> *"...what takes me really long time is organizing my music library for DJing. [...] it could be something like Google image search for example."* [Tok011]

> *"Because we usually have to browse really huge libraries [...] that most of the time are not really well organized."* [Tok003]

In relation to supportive and recommendation systems, i.e., to the question "how do we want the computer to 'help' us in our creative work process?", beside issues of artistic control and the fear of making predictable sounds, it becomes apparent that the desired features of recommenders in a creative context go beyond the query-by-example-centered paradigm of finding similar items and even also beyond the goal of serendipitous suggestions, cf. [29].

> *"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]

So the desired functionality of the machine is to provide an alter-ego of sorts, which provides the artist with opposite suggestions, that still reside within the artist's idea of his own personal style. This can be related to the artistic strategy of "obstruction" to assess the quality of a piece in the making, by changing the perception of the freshly edited music through changes in acoustics and hardware to render the piece "strange" [30].

This must of course be done with a strong consideration of how each musician's notion of "strange" depends on personality, emotions, preferences, and style of music, cf. [31].

> *"No, it should be strange in that way, and then continue on in a different direction. That's the thing about strange, that there's so many variations of strange. There's the small, there's the big, there's the left, there's the right, up and down."* [Strb006]

In addition to the more concrete steps of elaborating on interaction with musical knowledge, we will keep exploring these open questions. Throughout this process we are determined to not only investigate if these ideas work, but maybe more importantly, if they are interesting and productive as interfaces for creative expression in digital sound.

Acknowledgments

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).

7. REFERENCES

[1] P. Knees, K. Andersen, S. Jordà, M. Hlatky, G. Geiger, W. Gaebele, and R. Kaurson, "GiantSteps – Progress Towards Developing Intelligent and Collaborative Interfaces for Music Production and Performance," in *2015 IEEE International Conference on Multimedia & Expo (ICME) Workshop Proceedings*, 2015.

[2] K. Andersen and D. Gibson, "The Instrument as the Source of new in new Music," in *Proceedings of the 2nd Biennial Research Through Design Conference (RTD)*, Cambridge, UK, 2015.

[3] K. Andersen, "Using Props to Explore Design Futures: Making New Instruments," in *Proceedings of the ACM CHI 2014 workshop on Alternate Endings: Using Fiction to Explore Design Futures*, Toronto, Canada, 2014.

[4] K. Andersen and F. Grote, "GiantSteps: Semi-Structured Conversations with Musicians," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA)*, Seoul, Republic of Korea, 2015.

[5] F. Grote, K. Andersen, and P. Knees, "Collaborating with Intelligent Machines: Interfaces for Creative Sound," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA)*, Seoul, Republic of Korea, 2015.

[6] F. Grote, "Jamming with Machines: Social Technologies in Musical Creativity," in *Proceedings of the 8th midterm Conference of the European Research Network Sociology of the Arts*, Cluj, Romania, 2014.

[7] ——, "The Music of Machines – Investigating Culture and Technology in Musical Creativity," in *Proceedings of the XIII. Conference on Culture and Computer Science (KuI)*, Berlin, Germany, 2015.

[8] S. Böck and G. Widmer, "A Multi-Model Approach to Beat Tracking Considering Heterogeneous Music Styles," in *Proceedings of the 15th International Society for Music Information Retrieval Conference (ISMIR)*, Taipei, Taiwan, 2014.

[9] M. Davies and S. Böck, "Evaluating the Evaluation Measures for Beat Tracking," in *Proceedings of the 15th International Society for Music Information Retrieval Conference (ISMIR)*, Taipei, Taiwan, 2014.

[10] F. Korzeniowski, S. Böck, and G. Widmer, "Probabilistic Extraction of Beat Positions From Neural Network Activations," in *Proceedings of the 15th International Society for Music Information Retrieval Conference (ISMIR)*, Taipei, Taiwan, 2014.

[11] F. Krebs, S. Böck, and G. Widmer, "An efficient state space model for joint tempo and meter tracking," in *Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR)*, Málaga, Spain, 2015.

[12] S. Böck, F. Krebs, and G. Widmer, "Accurate Tempo Estimation based on Recurrent Neural Networks and Resonating Comb Filters," in *Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR)*, Málaga, Spain, 2015.

[13] F. Hörschläger, R. Vogl, S. Böck, and P. Knees, "Addressing Tempo Estimation Octave Errors in Electronic Music by Incorporating Style Information Extracted from Wikipedia," in *Proceedings of the 12th Sound and Music Conference (SMC)*, Maynooth, Ireland, 2015.

[14] Á. Faraldo, E. Gómez, S. Jordà, and P. Herrera, "Key Estimation In Electronic Dance Music," in *Proceedings of the 38th European Conference on Information Retrieval (ECIR)*, Padua, Italy, 2016.

[15] M. Leimeister, "Feature Learning for Classifying Drum Components from Nonnegative Matrix Factorization," in *Proceedings of the 138th International Audio Engineering Society Convention (AES)*, Warsaw, Poland, 2015.

[16] B. Lehner, G. Widmer, and S. Böck, "A low-latency, real-time-capable singing voice detection method with LSTM recurrent neural networks," in *Proceedings of the 23rd European Signal Processing Conference (EUSIPCO)*, 2015.

[17] P. Knees, "The Use of Social Media for Music Analysis and Creation Within the GiantSteps Project," in *Proceedings of the First International Workshop on Social Media Retrieval and Analysis (SoMeRA)*, Gold Coast, Queensland, Australia, 2014.

[18] P. Knees, Á. Faraldo, P. Herrera, R. Vogl, S. Böck, F. Hörschläger, and M. Le Goff, "Two Data Sets for Tempo Estimation and Key Detection in Electronic Dance Music Annotated from User Corrections," in *Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR)*, Málaga, Spain, 2015.

[19] R. Vogl and P. Knees, "An Intelligent Musical Rhythm Variation Interface," in *Proceedings of the 21st International Conference on Intelligent User Interfaces (IUI)*, Sonoma, CA, USA, 2016.

[20] C. Ó Nuanáin, P. Herrera, and S. Jordà, "Target-Based Rhythmic Pattern Generation and Variation with Genetic Algorithms," in *Proceedings of the 12th Sound and Music Conference (SMC)*, Maynooth, Ireland, 2015.

[21] D. Gómez-Marín, S. Jordà, and P. Herrera, "Strictly Rhythm: Exploring the effects of identical regions and meter induction in rhythmic similarity perception," in *Proceedings of the 11th International Symposium on Computer Music Multidisciplinary Research (CMMR)*, Plymouth, UK, 2015.

[22] ——, "Evaluating rhythm similarity distances: the effect of inducing the beat," in *Proceedings of the 15th Rhythm Production and Perception Workshop (RPPW)*, Amsterdam, the Netherlands, 2015.

[23] ——, "PAD and SAD: Two Awareness-Weighted Rhythmic Similarity Distances," in *Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR)*, Málaga, Spain, 2015.

[24] Á. Faraldo, C. Ó Nuanáin, D. Gómez-Marín, P. Herrera, and S. Jordà, "Making Electronic Music with Expert Musical Agents," in *ISMIR 2015: 16th International Society for Music Information Retrieval Conference Late Breaking/Demo Session*, Málaga, Spain, 2015.

[25] S. Jordà, M. Kaltenbrunner, G. Geiger, and R. Bencina, "The reacTable," in *Proceedings of the International Computer Music Conference (ICMC)*, Barcelona, Spain, 2005.

[26] K. Gohlke, M. Hlatky, and B. de Jong, "Physical Construction Toys for Rapid Sketching of Tangible User Interfaces," in *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction (TEI)*, Stanford, CA, USA, 2015.

[27] C. Ó Nuanáin and L. Ó Sullivan, "Real-time Algorithmic Composition with a Tabletop Musical Interface: A First Prototype and Performance," in *Proceedings of the 9th Audio Mostly: A Conference on Interaction With Sound (AM)*, Aalborg, Denmark, 2014.

[28] 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 *Proceedings of the 6th ACM International Conference on Multimedia Retrieval (ICMR)*, New York, NY, USA, 2016.

[29] P. Knees, K. Andersen, and M. Tkalčič, "'I'd like it to do the opposite': Music-Making Between Recommendation and Obstruction," in *Proceedings of the 2nd International Workshop on Decision Making and Recommender Systems (DMRS)*, Bolzano, Italy, 2015.

[30] K. Andersen and P. Knees, "The Dial: Exploring Computational Strangeness," in *Proceedings of the 34th Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA)*, San Jose, CA, USA, 2016.

[31] B. Ferwerda, "The Soundtrack of My Life: Adjusting the Emotion of Music," in *CHI Workshop on Collaborating with Intelligent Machines: Interfaces for Creative Sound*, Seoul, Republic of Korea, 2015.