



HAL
open science

A posthumous improvisation by Toots Thielemans

Marc Chemillier, Shlomo Dubnov, Ke Chen, Mikhail Malt, Gérard Assayag

► **To cite this version:**

Marc Chemillier, Shlomo Dubnov, Ke Chen, Mikhail Malt, Gérard Assayag. A posthumous improvisation by Toots Thielemans. Toots Thielemans (1922-2016). A Century of Music across Europe and America, May 2022, Bruxelles, Belgium. hal-03565439v2

HAL Id: hal-03565439

<https://hal.science/hal-03565439v2>

Submitted on 12 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A posthumous improvisation by Toots Thielemans

Marc Chemillier (EHESS), Shlomo Dubnov (USCD), Ke Chen (UCSD), Mikhail Malt (IRCAM), Gérard Assayag (IRCAM)

In jazz, the presence of the musician is of such central importance that the recordings considered to be the best were often live recordings, preferably made in small intimate clubs. The evolution of live performance technologies tends to blur this notion, first with the gigantism of concerts, and more recently, with the appearance of simulacra in rap and pop music in the form of holograms (Tupac), virtual stars (Hatsune Miku), video game avatars (Travis Scott), and the most recent example based on voice synthesis (Drake and The Weeknd). Toots Thielemans was one of those jazzmen whose presence left an indelible mark on a concert or a recording session, so strong was his musical personality and his harmonica playing immediately recognizable. We propose here to try an experiment consisting in creating a musical avatar of Toots Thielemans.

We use a music improvisation software developed at IRCAM in collaboration with the CAMS (EHESS) in Paris (Assayag et al. 2006). This machine learning system allows to capture the phrases played by an instrumentalist and to extend them by a virtual improvisation system restoring the sound of the musician, his phrasing and his accents, but playing something different. The system is able to synchronize with the pulse of a human orchestra and can also follow a given chord progression.

The overall software architecture is based on two main components : the corpus preprocessing and the machine improvisation. Corpus preprocessing deals with extraction of Toots solos from mixed recordings by learning the sound characteristics of Toots harmonica from a short solo example, as will be explained below. The machine Improvisation component consists of three main modules. A first one built in Max, the body, containing the listening machine, the sound memory unity, the tempo following functions and allowing to interface external MIDI controllers. The mind, an external set of algorithms, made in Common Lisp and embedded in the OpenMusic computed aided composition software (Assayag et al. 1999), in charge of calculating new improvisation sequences (according to harmonic and other constraints) sent to the main interface. And the third piece, the Antescofo Max object (Cont 2010; Echeveste et al. 2013), that is the clock and the bridge allowing to synchronize the body and mind of our software, by messages. Currently, this last module is also able to infer tempo, adapting past recorded sound slices to actual improvisation sequences calculated by our OpenMusic mind taking into consideration swing and tempo changes.

To realize the experiment and to make the machine learn Toots' playing, it was necessary to have Toots' solos in separate tracks. We used a system of separation of the audio sources based on state of the art machine learning that was developed by the authors (Chen et al. 2022), which comprises of a three-component pipeline including a sound event detector, an embedding processor, and a query-based source separator. The system was able to extract both the leading instrument and the remaining accompaniment of the jazz music for the

machine's improvisation. The advantage of this model is that it only requires a small audio clip of the source to extract it from the audio mixture without being pre-trained for solely separating this source. We used the only solo part, a two-sec harmonica clip around 5:56-5:58 in the jazz song "Body & Soul" (from the album *Affinity* with Toots and Bill Evans, 1979), to extract the whole harmonica piece of this song. And we used an existing accompaniment part, about an 80-sec clip around 2:44-4:09, to extract the whole accompaniment. The separation results are correct enough to be used for the following machine improvisation.

The unique interest of this experiment lies in the high barrier of acceptability raised by trying to simulate improvisations of a legendary human musician by the machine (Chemillier and Nika 2016). This goes beyond the "deep fake" challenge where the goal is to convert a sound signature of one voice to another. The task for Toots avatar combines both aspects of his unique harmonica instrument sound and his idiomatic style of improvisation in the specific interpretation of jazz on the harmonica. Will an amateur who is well familiar with the playing of Toots Thielmans be musically "taken in" by listening to this fake harmonica player? Moreover, it is possible to set the artificial improviser to produce solos that are more or less similar to the human it tries to model. At what threshold of departure from exact replication will the Toots fan consider that these improvisations are not real Toots? To address these challenges, we designed a test that involved listening to Toots' original solo on "Body & Soul" and our avatar, and trying to identify which one was the real one and which one was the fake one. A multimedia animation showing this test is available online (Improvisation and the computer 2023, <http://digitaljazz.fr/multimedia/improvisationordinateur/>). Notice that a similar test was presented in the exhibition devoted to Toots Thielmans at the KBR Royal Library in Brussels, but not involving a computer avatar. The curators set up an attractive interactive terminal where one could listen to audio clips of pieces with harmonica by Toots and other players, such as Stevie Wonder, Bob Dylan, or bluesmen, and try to identify whether they were Toots or not (fig. 1).

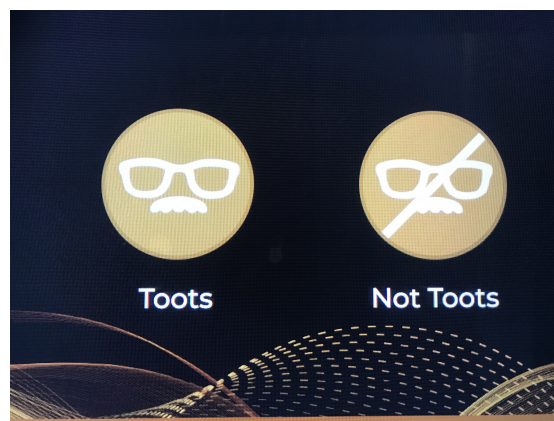


Figure 1. Screen display of the interactive terminal at KBR exhibition on Toots Thielmans.

Our experiment was designed in the following way. The harmonica part extracted from Toots was provided as input to the learning process of our improvisation software Djazz. We then improvised using Djazz trained on these data with the accompaniment of the piano-bass-drums part extracted from the recording. But the remaining harmonica notes that were not completely eliminated by the source separation algorithm conflicted with the harmonica notes generated by Djazz. So, to ensure clear accompaniment, we made by ear a MIDI transcription of the piano-bass-drums part of the original recording (piano by Bill Evans, bass

by Mark Johnson, and drums by Eliot Zigmund). Finally, our tests focused on an audio avatar played by Djazz with this MIDI accompaniment which was compared to the original version. The experiment was limited to an AA section of the song which, in its complete form, has an AABA structure.

The two examples (avatar and original recording) were tested during a seminar at EHESS in Paris, and then they were sent to Hugo Rodriguez in Brussels, curator of the Toots exhibition. He played them to various people some of them being well aware of Toots' style. They all distinguished the real Toots, but said the avatar was quite impressive. Hugo Rodriguez also made his own comments: "Overall, this is very good, and impressive! [...] In terms of style: between 00:38 and 00:48, in my opinion, Toots would not have repeated this chromatic motif so many times in a loop. I suppose that the software must have been inspired by what happens between 02:08 and 02:13 in the original version. There Toots does indeed repeat something, but the pattern is shorter, melodically unidirectional (he does this more often, rather than a loop with descending-ascending alternation, sometimes he also does repeated notes "turning" around a note), fits into a larger melodic phrase (so it is not the leading motif, nor an isolated pattern)" (Rodriguez 2022a).

This comment reveals some aspects of the Djazz process. The solo generated by Djazz takes motifs from the original solo and recombines them with respect to the pulsation and the harmonies. The notes played by Djazz are thus notes that were originally played by Toots (Hugo Rodriguez noticed that Djazz plays motifs that are "inspired by" Toots). In addition, the user can modify the output manually through an interface by adding some effects such as looping or accelerating (Chemillier 2023). In this interface, the yellow buttons on the right in the lower part are used for looping and those on the left are used for accelerating. The upper part is a visual representation of the structure of the song which in this case is two lines corresponding to AA, with the current bar lit in red (fig. 2).

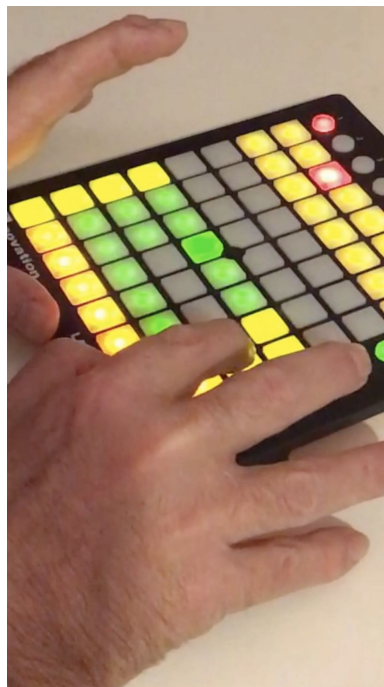


Figure 2. Interface of the Djazz software.

Hugo Rodriguez mentions a chromatic motif occurring at the beginning of the second A in Djazz solo which was extended by adding loops (loops are underlined below) :

Db-C-A#-B-Bb-A / Db-C-A#-B-Bb-A-Bb-A / Db-C-A#-B-Bb-A / Db-C-A#-B-Db-C-A#-B-Bb-A

The corresponding motif of the original solo that was picked by Djazz algorithm appeared at the beginning of the third A of the AABA form. It was repeated by Toots and extended by chromatic transpositions as follows (extensions are underlined below) :

Db-C-A#-B-A-Bb / Db-C-A#-B-A-Bb / Db-C-A#-B-A-Bb-G#-A / Db-C-A#-B-A-Bb-G#-A-G-Ab

In Toots' improvisation, the process is quite clear : notes are added progressively at the end of the motif A-Bb, then G#-A, then G-Ab and so on. In Djazz improvisation, the process is more obscure. Hugo Rodriguez concludes this comparison by pointing out “the absence of a feeling that the music was "saying something" in the Toots fake” (Rodriguez 2022b). This narrative aspect of music has been studied by him for other musical repertoires (Rodriguez 2021; see also for recent developments McAuley et al. 2021).

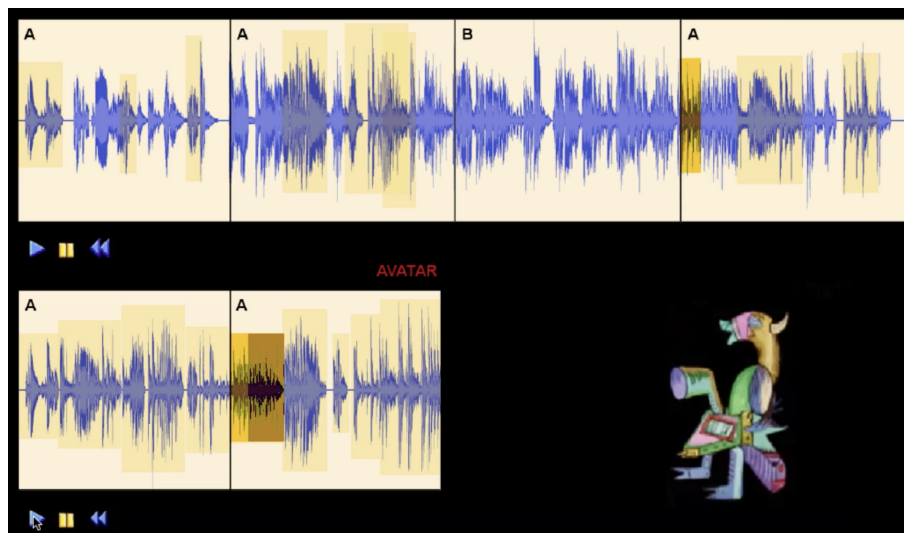


Figure 3. Comparison of the original AABA solo by Toots and the AA solo by Djazz

People who can distinguish the real Toots are generally convinced that Toot's *desire*, which includes multiple factors related to his creative intent, emotions, artistic imagination and interpretation of the jazz language, are not fully present in the avatar. This is what it means to say: “Toots wouldn't have done that”. It is the personality of Toots, that we try to emulate here in addition to his technical and musical skills. But in reality, the Djazz user also has his or her own desire when improvising the avatar. This makes a fundamental difference in terms of judgment of the result compared to common approaches, such as Turing test, where success (or failure) is defined in terms of producing a computer response that is indistinguishable (or noticeably artificial) from that of a human, but this could be any human and not a specific person, and even less a hybrid response that combines user's intent with that of a cloned musical virtuoso. Using the interface connected to the Djazz software, the user tries to find manual effects that might match Toots' phrases. These are not statistical fake effects added to the recombination process, but rather they express the user's own

desire. The idea behind Djazz is not just to provide clones of the original material, but to offer opportunities to improvise with it yourself. One of Hugo Rodriguez's interviewees put it this way: "In the AI solo, there are some pretty fast fragments (which made me think of the bebop period), which I think Toots wasn't going to do in a piece like "Body & Soul". But, I think that's just the point?" (Rodriguez 2022a, quotation from a musicologist who worked on Toots exhibition).

Furthermore, you can admit it's not the real Toots, but still love the avatar. Identifying the avatar doesn't mean you think it's bad. Another person interviewed by Hugo Rodriguez pointed to the avatar's personality: "Wow, this is great! After a non-analytical listening, it seems to me that he is a little more "nervous", that he makes more notes, that he takes less time, but I don't know enough the style of Toots to know! Maybe in his other solos Toots is like that. It is bluffing in any case" (Rodriguez 2022a, quotation from a musicologist who knows little about Toots).

During the conference dedicated to Toots Thielemans at KBR on May 10, 2022, we had the opportunity to test it out in front of a live audience. Unfortunately, as expected from this panel of Toots specialists, they were unanimous in finding the right Toots and we were disappointed that not a single person thought the real one was the avatar...! Jonathan De Souza from the University of Western Ontario in Canada who was invited speaker at the conference proposed a very interesting idea. Instead of using the original recording as reference, we should use the extracted harmonica part mixed with the MIDI transcription of the piano-bass-drums trio. In this way we could avoid bias in the comparison induced by the fact that the sound of the original recording is much better than the sound of our own mixing. Since then, we've carried out this new test on various occasions (and you can do it online as indicated above). Although there remains a majority of people finding the real Toots, there is a significant number of persons who think the real one is the avatar. Almost inadvertently, this experience leads one to think about the next research challenge of cloning Toots' persona, and not only his way of making and improvising music.

This also inspires us the perspective of a general frame for creative musical AI, linked to the emerging field of general intelligence. AI systems proceed by some form of imitations at the surface and formal level. This can be reasonably convincing as has been seen above, but will always lack the power of creative mind. It seems that what's missing now is not so much the technical depth of structural comprehension by the model than its hypothetical capacity of being a "desiring machine" according to the concept created by philosophers Deleuze and Guattari in their book *Anti-Œdipus* (Deleuze and Guattari 1971). For these authors, desire is not so much the expression of a lack, as Lacan puts it in his famous claim that "human desire is the desire of the other", where the other is understood as the locus of symbolic formation (Lacan 1953, 407; Cléro 2003), but a creative force rooted in reflexivity and interaction ("coupling machines", Deleuze and Guattari 1971, 5). Modern generative AI methods are particularly interesting considering the parallels of the imaginary and symbolic Lacanian concepts to problems of learning generative music representations and problems of modeling symbolic language-like aspects, respectively. Problems of finding optimal symbolization and identifying the symbolic units particular to a specific style or cultural expression seem to be emerging as a research topic important for general musical AI (Dubnov and Geer 2023; Dubnov, Huang, and Wang 2021). On this subject, it is noticeable that as soon as 1954, in the wake of cybernetics, Lacan initiated a reflexion on the machine

and its relation to the subject and the symbolic order (Lacan 1978, 63, 350). As pointed out by Patrick Juignet, for Lacan “the cybernetic machine is identified with a structure detached from the subject's activity, so that “the symbolic is the world of the machine”” (Juignet 2003).

Indeed, reflexivity, self motivation, creative intent, interaction are major challenges on the horizon of musical AI and creative agentivity. A synthesis between a notion of autonomous agency rooted in the “ego” as inaugurated by Freud and later developed by Lacan into his theory of the subject on one hand, and the idea of co-creative emergence where the subject dissolves into the mere interaction of production machines as in Deleuze and Guattari is still missing, and beyond the scope of this article. We might just say here as a possible research route that the imaginary and symbolic orders are on the side of the subject, thus on the side of desire as motivation, whereas the real order (that complements Lacan’s real-imaginary-symbolic trilogy, Lacan 1953) would be a host to co-creative interaction of machines, thus on the side of desire as creative force, breaking free from the transcendence induced by the language/symbolisation order. This dichotomy is at the very heart of the REACH project that underlies the work presented here (Reach 2021).

Now to back to Toots, where is his personality in all this? Obviously we shouldn’t pretend that our AI models in any way represents Toots’s personality as a creative source. If we feel a subject here, it is as Deleuze and Guattari rightly say, a “residuum” in the interplay of machines subjected to real production (Deleuze and Guattari 1971, 19). This is particularly true in the case of the complex experience we described, where we actually intervened as a musician ourselves! However we can fugitively feel the presence of Toots, due the process of using his real sound, and reproducing patterns statistically.

What we can say is that whatever remains of Toots' musical personality is not present as a cause, but as an effect (not as a deep structure, but as a surface structure produced by the latter). It is just a recording, an image, that we reenact in a smart way, and of course there are traces of his unique way of playing that persist. So the question is: “Is there more to a musical personality, than just its shallow structure as it can be recorded, imitated, and even transformed?” The answer could be : for a dead musician no, as they can be predictable in any aspects of their production, since they will not surprise us anymore. A dead musician's musical personality is an archive, and manipulable as such. In the interaction process (s.a. when somebody “plays Toots” with the pad) we introduce a novel causal thread (our own desire), and reenact the archive thus creating a strange mixture of our acting personality and images of a past personality. Archive personalities are animated (like voodoo) to create an effect without cause, their personality is an image of personality like an actor seen in the cinema. So it's literally “fiction.” This is consistent with Auslander's view that musicians on stage appear before us under a dual identity: as real people playing music (like actors playing their roles in the theater), and as musicians involved in narratives, and, therefore, implicitly dramatic (Auslander 2013, 56). In their position against psychoanalysis Deleuze and Guattari have indeed vowed to avoid transcendence and interiority, and are as such a real precursor of machine interaction. An old debate is thus receiving new light in the development of AI.

Fundings: This research is supported by the European Research Council (ERC) REACH project,

under Horizon 2020 programme (Grant 883313) and by Agence Nationale de la Recherche (ANR) project MERCI (Grant ANR-19-CE33-0010).

Assayag, Gérard, Georges Bloch, Marc Chemillier, Arshia Cont and Shlomo Dubnov. 2006. "OMax brothers: a dynamic topology of agents for improvisation learning." *AMCMM '06: Proceedings of the 1st ACM workshop on Audio and music computing multimedia*. (October 2006): 125-132.

Assayag, Gérard, Camilo Rueda, Mikael Laurson and Olivier Delerue. 1999. "Computer-assisted composition at IRCAM: From PatchWork to OpenMusic." *Computer music journal* 23, no. 3: 59-72.

Auslander, Philipp. 2013. "Jazz Improvisation as a Social Arrangement." In *Taking It to the Bridge: Music As Performance*, edited by Nicholas Cook and Richard Pettengill. Ann Arbor: University of Michigan Press.

Chen, Ke, Xingjian Du, Bilei Zhu, Zejun Ma, Taylor Berg-Kirkpatrick and Shlomo Dubnov. 2022. "Zeros-Shot Audio Source Separation through Query-based Learning from Weakly Labeled Data." *Thirty-Sixth Association for the Advancement of Artificial Intelligence Conference AAAI 2022*.

Chemillier, Marc and Jérôme Nika. 2016. "« Étrangement musical » : les jugements de goût de Bernard Lubat à propos du logiciel d'improvisation ImproteK." *Cahiers d'ethnomusicologie* 28: 61-80. <https://journals.openedition.org/ethnomusicologie/2496>

Chemillier, Marc. 2023. "The smile of Velonjoro. Bi-musicality and the use of artificial intelligence in the analysis of Madagascar zither music." submitted to *Analytic Approach to African Music*.

Cléro, Jean-Pierre. 2003. "Concepts lacaniens." *Cités* 4, no. 16: 145-158. <https://www.cairn.info/revue-cites-2003-4-page-145.htm>

Cont, Arshia. 2010. "A coupled duration-focused architecture for realtime music to score alignment." *IEEE Trans Pattern Anal Mach Intell* 32, no. 6: 974-987.

Deleuze, Gilles and Félix Guattari. 1971. *Anti-Œdipus. Capitalism and schizophrenia*. Translated by Robert Hurley, Mark Seem, and Helen R. Lane. University of Minnesota Press, [1983].

Dubnov, Shlomo and Ross Greer. 2023. *Deep and Shallow Machine Learning in Music and Audio*. Chapman & Hall. ISBN 9781032133911.

Dubnov, Shlomo, Kevin Huang, Cheng-i Wang. 2021. "Towards Cross-Cultural Analysis using Music Information Dynamics." *CoRR*. <https://arxiv.org/abs/2111.12588>

Echeveste, José, Arshia Cont, Jean-Louis Giavitto and Florent Jaquemard. 2013. "Operational semantics of a domain specific language for real time musician-computer interaction." *J. of Discrete Event Dynamic Systems* 23, no. 4: 343-383.

Improvisation and the computer. 2023. Toots Thielemans. Multimedia animation.
<http://digitaljazz.fr/multimedia/improvisationordinateur/>

Juignet, Patrick. 2003. "Lacan, le symbolique et le signifiant." *Cliniques méditerranéennes* 68, no. 2: 131-144. <https://www.cairn.info/revue-cliniques-mediterraneennes-2003-2-page-131.htm>

Lacan, Jacques. 1953. "Symbolique, imaginaire, réel." *Bulletin interne de l'Association française de psychanalyse*, Paris.

Lacan, Jacques. 1978. *Séminaire II. Le moi dans la théorie de Freud et dans la technique psychanalytique*. Paris: Le Seuil.

Lacan, Jacques. 2013. *Le Séminaire, livre VI, Le désir et son interprétation*. Texte établi par J.-A. Miller. Paris: La Martinière/Le Champ Freudien.

McAuley, J. Devin, Patrick C. M. Wong, Lucas Bellaiche and Elizabeth Hellmuth Margulis. 2021. "What Drives Narrative Engagement With Music?." *Music Perception* 38, no. 2: 509-521. <https://doi.org/10.1525/mp.2021.38.5.509>

REACH. 2021. Raising CoCreativity in Cyber-Human Musicianship. Project granted by the European Research Council. <http://www.repmus.ircam.fr/reach>

Rodriguez, Hugo. 2021. "Sémantique et pragmatique de la musique. Une approche cognitive basée sur les travaux de Philippe Schlenker et sur les œuvres de Franz Liszt." thèse, Université libre de Bruxelles.

Rodriguez, Hugo. 2022a. Personal email. January 19, 2022.

Rodriguez, Hugo. 2022b. Personal email. January 26, 2022.

Biographies

Marc Chemillier

Musician, computer scientist and anthropologist, Marc Chemillier studied jazz piano (Schola Cantorum, CIM). He entered the ENS de Fontenay-aux-roses in mathematics in 1981 and studied harmony-counterpoint at the CNSM in Paris. He made a PhD thesis in collaboration with IRCAM. In ethnomusicology, he worked on the harp of the Nzakara of the Central African Republic (CD *Musiques des anciennes cours Bandia* in 1995), then on the zither of Madagascar. In 2000, he created with the OMax Brothers (Gørard Assayag, Marc Chemillier, Shlomo Dubnov, Georges Bloch) the OMax improvisation software. Director of studies at the EHESS in Paris, he published *Les Mathématiques naturelles* in 2008 (Odile Jacob) and continues his research on computer-assisted improvisation and its anthropological and social issues. In 2021, he published the book-CD *Artisticiel* with Bernard Lubat and Gørard Assayag.

Ke Chen

Audio and Music Researcher. Ke Chen now is a Ph.D student in Computer Music at UC San Diego, California, USA. He is co-advised by Prof. Shlomo Dubnov and Prof. Taylor Berg-Kirkpatrick. His research interest lies in the interdisciplinary between music and computer science, focusing on: Music Generative System and Music Information Retrieval (including singing melody extraction, music source separation, and music recommender system). He proposed a music generative framework, Music SketchNet, to allow the user to specify their own ideas in automatic music generation. He also proposed a zero-shot audio separator to allow a more efficient separation pipeline to separate all possible sources. Ke Chen obtained his Bachelor's degree in Software Engineering at Fudan University, China. He is also the website maintainer of China Conference on Sound and Music Technology (CSMT) and New Interfaces for Musical Expression (NIME).

Mikhail Malt, having a twofold training, scientific and musical (Engineer, composer and musical conductor) started out his musical career in Brazil as both flutist and orchestral conductor. He has a PHD grade with a thesis at the "Ecole des Hautes Etudes en Sciences Sociales" dedicated to the use of Mathematical models in Computer Assisted Composition, and a HDR (Habilitation à diriger des recherches) degree with the dissertation : "Representation in Computer Aided Composition and Computational Musicology". Nowadays he is researcher in the Music Representations Team - UMR 9912 STMS (IRCAM, CNRS, UPMC), associated researcher at IReMus - Sorbonne Universités and Computer Music Designer Teacher in the Educational Department at Ircam, Paris-France. He is currently pursuing his research and composition activities in the fields of musical modeling and discovery, computer assisted performance, and musical representation epistemology.

Shlomo Dubnov is a computer music researcher and composer. He is a Professor in the Music Department and Affiliate Professor in Computer Science and Engineering and a founding faculty of the Halicioлbu Data Science Institute in the University of California, San Diego, where he has been since 2003. He is the Director of the Center for Research in Entertainment and Learning (CREL) at UC San Diego's Qualcomm Institute. He is best known for his research on machine improvisation in computational creativity and stylometry of music. He is also known for his contributions to the field of computer audition by inventing the method of information dynamics and use of spectral flatness.

Gérard Assayag has been the director of the STMS research lab (Sciences et Technologies de la Musique et du Son, a joint unit by IRCAM / Ministry of Culture, CNRS and Sorbonne University) from 2011 to 2017. He currently heads the IRCAM Musical Representations team he founded in 1992. His research interests are centered on music representation issues, involving programming languages, creative artificial intelligence, and computational music modeling. Among other international events, Gérard Assayag has founded with Marc Chemillier the ImproTech festival-workshop series on improvisation and digital intelligence (IRCAM Paris 2004, New York 2012, Philadelphia 2017, Athens 2019, Uzeste 2023). Gérard Assayag has written or co-authored several seminal works on computer music research. He has been awarded in 2020 the European Research Council (ERC) Advanced Grant for a project entitled REACH (Raising co-creativity in cyber-human musicianship).