



HAL
open science

Being the Artificial Player: Good Practices in Collective Human-Machine Music Improvisation

Marco Fiorini

► **To cite this version:**

Marco Fiorini. Being the Artificial Player: Good Practices in Collective Human-Machine Music Improvisation. EAI ArtsIT 2024, Nov 2024, Abu Dhabi, United Arab Emirates. hal-04760154

HAL Id: hal-04760154

<https://hal.science/hal-04760154v1>

Submitted on 30 Oct 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.



Distributed under a Creative Commons Attribution 4.0 International License

Being the Artificial Player: Good Practices in Collective Human-Machine Music Improvisation

Marco Fiorini^[0009-0004-0523-2481]

IRCAM - STMS
Sorbonne Université, CNRS UMR 9912
1, place Igor Stravinsky, 75004 Paris, France
marco.fiorini@ircam.fr

Abstract. This essay explores the use of generative AI systems in co-creativity within musical improvisation, offering best practices for this paradigm. Using practice-based research in the interdisciplinary fields of scientific and artistic exploration, this study documents the author’s artistic collaboration with an experimental music ensemble. This partnership resulted in a studio album and will continue with international concerts and workshops. The research is supported by observations and surveys on computational creativity and co-creativity in Human-Computer Interaction (HCI), along with reflections on musical improvisation from scientific, cognitive, and musicological perspectives. These insights form the foundation of the article, which later specifically examines the co-creative system for improvisation, Somax2, developed within the European project REACH (Raising Co-Creativity in Cyber-Human Musician-ship). The analysis situates Somax2 in a new context, detailing generative agent models, the taxonomy of interactive systems between player and instrument paradigms, and cognitive behaviors in improvised music. This research confers on Somax2 a novel multifaceted identity, offering artists new perspectives on its use and positioning it as an optimal case for defining the new proposed framework of Environmental or Ecological Referent in HCI contexts of music improvisation.

Keywords: Human-Computer Interaction · Co-Creativity · Music Improvisation · AI · Generative Agents.

1 Introduction

The analysis in this paper combines aspects of academic research in the domain of Human-Computer Interaction (HCI), with elements of computational creativity and artistic research practice in the specific area of music improvisation. The attempt of this study is to interrogate the role of generative AI systems in multi-modal contexts of musical co-creativity. Foundational studies linked to the concept of co-creativity in improvised human-machine interaction situations can be found in [6] and are fully theorized in [7,8]. Therefore, in this paper we will use the term *co-creativity* in the same sense than these works, that is of new forms emerging from the complex dynamics of interaction.

The methodology of this study is based on the combination of scientific and artistic research, following the model of practice-based research. In this way, not only can an artistic process serve as the purpose of scientific research, but it can itself become one of the vehicles of such research. This methodological paradigm has already been explored in different contexts, notably in the ones of mixed multi-modal performances involving music and dance [23,33,35]. The uniqueness of this type of approach lies in the fact that AI is thought of not as a generative process in its own right, divorced from human interaction and its potential understanding, but in its agential role: as one of the contributors to the performative aspect, in this case in the context of improvised music practices.

This kind of meeting of scientific research and artistic experience is now the basis of many interesting projects, and has led to the creation of international journals [46] and collection databases of multidisciplinary projects [64]. At the same time, it may prove to be an extremely important process in understanding how to integrate new Machine Learning systems and various generative AI tools within creative and artistic applications, shedding light on potential aspects of explainability, and fostering the importance of a responsible human-machine co-interaction. In essence, are we in control of the AI or are we controlled by it? Or perhaps there is an alternative trajectory to this duality?

The article is structured as follows: first it gives an overview of human-machine co-creativity, starting from more general concepts of creativity and zooming out to specific cases of computational creativity. Next, the specific area to this analysis, that of music improvisation, and its definition in the musical, cognitive, and computational domains is framed. Then, the case of interaction with the tellKujira¹ ensemble will be discussed, exposing the tools and strategies used to carry out a concrete sonic creation and recording project. Finally, new HCI ideas will be introduced to frame the results of this multidisciplinary artistic/scientific research.

2 On Human-Machine Co-Creativity

Although various definitions of creativity have been attempted over the years [74,42,71], there is no universally accepted definition of it, nor does this research intend to seek to establish one. Instead, in the following section, an attempt will be made to give special emphasis to the concepts of computational creativity and creativity in music. With stress on the processes of human-machine interaction in music systems, we will explore the notion of co-creativity resulting from this intersection [6,7,8].

Recent applications of artistic co-creativity tools range from songwriting [25] to music harmonization [80], addressing also pedagogical purposes [75]. Focusing on design and evaluation of co-creative artificial intelligence systems [47] and their implication with users, these studies reveal interesting findings on the different perspectives based on the specific background of the users involved:

¹ <https://tellkujira.bandcamp.com>

More advanced users value co-creative systems that serve in an adaptive “follower” role, while novice users typically benefit from systems that provide support using levels of knowledge which the users themselves do not have. (Truesdell et al., [75])

Similarly, recent studies have shown how in live performance contexts involving dance and music, artificial intelligence systems can function as “tools” for musicians or as “actors” capable of intervening in the performance on a collaborative level [23]. This taxonomy, emerging in this specific case from mixed artistic practices of music and dance, cannot help but be reminded of Rowe’s seminal idea of a duality between “player” and “instrument” paradigms as the basis of interactive music systems [66]. In more recent declinations, these paradigms see their natural development in interactive systems based on co-creative agents [59,9,48], developed over a long and prolific genealogy and type of applications, including various generative AI and machine learning techniques [73].

Although there are now considerable examples of different types of systems of agent architectures, in the following article we will focus on the analysis of systems dedicated to music improvisation. In his seminal composition and research work, *Voyager* [50], which still turns out to be a fundamental milestone in the literature of agent-based music systems, Lewis argues that music computer programs are not objective, but rather represent the particular ideas of their creators. In particular, through the notion of “emotional transduction”, Lewis introduces the notion of a bidirectional transfer of intentionality through sound, encompassing emotional and mental intention through embodied meaning in the performance of the human and the artificial agents. This idea is highly rooted in the spirit of “one’s own sound”, typical of African-American music-making [31], becoming a carrier for identity.

Another important aspect brought up by Lewis is the one of structure formation in emergent music, dealing with continuous awareness of the performers, and the role of improvisation in composition practices. Here, through different axis of intervention, processes of interactivity, negotiation and different perspective enter the game, opening an important door in the creativity domain, where personalities and identities are formed through sound:

Rather than asking if computers can be creative and intelligent, *Voyager* asks us where our own creativity and intelligence might lie. (Lewis, [50])

As a particular case, using Rowe’s taxonomy between “player” and “instrument” paradigms [66], *Voyager* can be thought of as an extreme case of a “player”. Being able to function independently without being controlled by a musician, and integrating into its very essence and nature notions of musical behavior and de-instrumentation, makes it an independent improviser. Building on top of Rowe’s taxonomy, Emmerson extends the description of the “player” paradigm, detailing the generation of another performer by the computer:

This added musician may progressively be gaining almost human characteristics such that the equivalent of a Turing test might be some measure of successful performance. (Emmerson, [34])

Trying to evaluate the results of a human-machine improvisation system based on analytical abstraction and re-interpretation of Voyager, called *Favola*, Mogensen [54] argues that, in this specific scenario, what can be perceived as “convincing” in an improvised duo piece, might lie on the convincing nature of the human performer. This doubt actually opens up a new possibility in the taxonomy analyzed by Rowe, in which a system can be evaluated, as continuously oscillating between the player and instrument paradigms. Indeed, if one views the interactive system as an extension of the improvised actions of the human performer, such a system can quickly lean toward the instrument axis rather than remain in the player position:

In other words, the evaluation of *Favola*’s output is only possible in the context of specific performances that includes human improvisers. (Mogensen, [54])

This consideration debates what and how to consider compelling in the output of a human-machine interaction. At the same time it opens an avenue in which the interaction itself is necessary not only for the evaluation of the system’s output, but also for the evaluation of its creative and emergent phase [29]. This idea positions itself against the general trend of generative AI, by giving considerable importance to the role of the human-in-the-loop [56].

The paradigms described here obviously fall within the context of computational creativity. Introduced decades ago by Boden [14], this has only recently considered and studied seriously in light of the latest rise of AI systems. Among the attempts to define computational creativity, one of the most suitable turns out to be that expressed by Wiggins:

The study and support, through computational means and methods, of behavior exhibited by natural and artificial systems, which would be deemed creative if exhibited by humans. (Wiggins, [79])

From here, it is possible to extrapolate once again the importance of creative agents intervening and modeling behaviors that can produce creative interactions and outcomes, later explored in systems such as the already discussed Voyager. Therefore, in addition to the concepts of novelty, value and surprise already expressed by Boden [15], it is crucial to look for expressions of co-creativity in the very moment of the development of the interaction between human and artificial agents. As formalized by Assayag et al. in the notion of human-machine co-creativity [6,7,8], this can lead to the manifestation of two new components that, as we shall see, are closely related to musical improvisation, namely emergence [22] and non-linear dynamics [57].

Although it is now clear that computers might produce results assessed as creative, and thus could be called creative according to certain criteria, it is still

very difficult to assess an artificial system as absolutely (or generally) creative. Velardo, using the concept of Recursive Ontology [77], tried to address this issue by assuming that:

There is no difference between what *appears* to be creative and what it is *really* creative. In other words, if a system produces creative artefacts, it can be deemed to be creative (Velardo, [78])

However, such thinking is problematic for two reasons: (i), it introduces the notion of a creative artefact as a system output that is independent of the system itself; (ii), it separates the system output from the process of creation, and thus interaction with human interventions. Regarding the issue of creative artefacts, several authors have proposed a number of properties that can characterize them, but an interesting viewpoint is that of Colton and Wiggins [28], who introduce a few good maxims to follow when evaluating and pursuing the path of computational creativity:

When we celebrate an artefact such as a musical composition, a painting, a theorem or a poem, we are also celebrating the creative act which brought it into being. The artefact resulting from a creative act should be seen as an invitation to engage in a dialogue with the artefact and/or the creator and/or the culture and/or yourself. (Colton and Wiggins, [28])

This invitation to celebrate the creative act and to value the dialogue between human and machine is not only a modern take on the fundamental work of Benjamin on the concept of “aura” in the mechanical reproduction of art [11], but is also the basis of a number of systems, termed genre-agnostic, developed in recent years in the field of musical improvisation [32]. Among these, as will be later justified in the following sections, the choice of the following study therefore fell on the practical use of Somax2 [9], because of its particular predisposition for use in collective free improvisation contexts. One of the main reasons for this choice, in addition to the predisposition of the specific model for the context of improvisation, lies precisely in its ability to adapt to agnostic situations, and in its emphasis on human-machine interaction.

In fact, the qualities of novelty, value and surprise [53], together with the role of the agent itself [34], can be taken into account as a kind of Turing test [76] for musical creativity. But the thesis of this research is precisely that this would address only the evaluation of the output phase, without focusing on the aspects of true co-creation. As noted also by Ariza, this could be potentially misleading when considered in the context of generative music systems:

Use of the Turing test in the evaluation of generative music systems is superfluous and potentially misleading; its evocation is an appeal to a measure of some form of artificial thought, yet, in the context of music, it provides no more than a listener survey. (Ariza, [3])

3 On Music Improvisation

Music improvisation is undoubtedly a very layered process that can be analyzed and studied from different perspectives. Moreover, there are traces of improvisation in music from the most diverse social, cultural, geographical and historical backgrounds. Specifically, the type of improvisation discussed in this essay is that defined by Bailey as non-idiomatic, or free improvisation [10]. This definition is related to a phenomenon particularly in vogue in the European and American scene since the 1960s, drawing from elements of the jazz tradition, expanding them with influences from the contemporary compositional avantgardes of the 20th century, and integrating within it identity concepts in which different levels of creativity can be seen:

Diversity is its most consistent characteristic. It has no stylistic or idiomatic commitment. It has no prescribed idiomatic sound. The characteristics of freely improvised music are established only by the sonic-musical identity of the person or persons playing it. (Bailey, [10])

Although free improvisation can thus be seen as an expression of highly individual characteristics, such as the concept of identity, the most interesting aspect of this practice emerges in the contexts of collective improvisation. Here, in fact, not only do the concepts of identity and intention become parallels of social and anthropological patterns [50,31], but complex dynamics of interaction are introduced, leading to developments applicable to human-machine interaction processes as well.

Borgo attempts to trace a comprehensive historical survey of the practice and study of improvisation and its use in different fields of research [18,19,20]. He notes how improvisation definitely involves different types of creativity, and how this definition is not only problematic, but in the field of music improvisation implies an important dualism between composition and improvisation:

If composition and performance are already interdependent, does all musical creativity therefore involve improvisation? (Borgo, [20])

This duality between composition and improvisation contrasts with what historically had always been the duality between composition and performance [12]. However Pressing, one of the first to study musical improvisation through cognitive models, and still a proponent of some of the most relevant studies on the subject, invites us to go beyond this duality, seeking to see improvisation and composition as the extremes of a continuum of musical practices:

Improvisation may be viewed as a special kind of aesthetically constrained motor performance that maintains a commitment to high levels of real-time decision making. (Pressing, [61])

Thus, starting with the focus now shifted to the very moment of interaction - similarly to what was introduced in the previous section regarding the split

between evaluation of interaction and evaluation of the output of a co-creative system - it is possible to summarize the characteristics of computational creativity and co-creativity within the specific framework of musical improvisation [19]:

While many issues persist about how to frame and explore improvisation, the field on the whole appears to be moving beyond a priori definitions of improvisation towards a more inclusive stance investigating in-the-moment creativity in all of its multilayered facets. There is broad agreement that improvisation involves novel output created in non-deterministic, real time situations by individuals and collectives involving certain affordances and constraints. (Borgo, [20])

This definition can be enriched by Pressing's important contribution regarding referent-based and referent-free improvisation [61,62]. Pressing, introduces the concept of referent as:

The referent is an underlying formal scheme or guiding image specific to a given piece, used by the improviser to facilitate the generation and editing of improvised behaviour on an intermediate time scale. The generation of behaviour on a fast time scale is primarily determined by previous training and is not very piece-specific. If no referent is present, or if it is devised in real-time, we speak of 'free' or 'absolute' improvisation. This is much rarer than referent-guided, or 'relative' improvisation. (Pressing, [61])

In practice, Pressing gives us a concise definition of real-time interaction, making a distinction between what can be developed on the spot, and what can be prepared, at the level of material and personal knowledge, in order to anticipate future developments. However, it is important to note that much collective improvised music does not fit under the umbrella of referent-free. At the same time this does not possess such a strong referent (such as a structure, or a harmonic grid) that it can be classified as referent-based. Here, then, it turns out that the ecological context of improvisation is crucial in collective interaction, which even leads to the creation of pseudo-referents [20].

This argument has been noticed not only in the cognitive field, but also in musicology and mathematics, leading to the creation of a model for Collective Free Improvisation (CFI) thanks to concepts from systems theory, which extends the concept of referent-based (or pseudo-referent-based) improvisation. In their work, Canonne and Garnier [22] divide collective interaction according to different time scales similarly to Pressing's theories [61,62], showing how different phenomena of imitation, intentionality and transition lead to the creation of individual and collective identities, even in the absence of a referent.

This model turned out to be fundamental to explore in depth concepts of joint action emerging from improvised interaction [70], similarly to what Corbett remarked in his listener's guide to free improvisation [30]. This can be applied in duo or in group contexts of various sizes [41], focusing on interaction and

structural parameters fundamental to make sense of the experience of musical collectivity [69], from which might rise the formation of real behaviors (i.e., with, against, without) [40]. Thanks to these studies, from a phenomenological point of view, improvisation can be seen as an act of “we-ness”, in which senses of agency, identity, reflexivity, dependence and integration merge with musical intention [69]. This gives rise to an ecological process that emphasizes the creation of both individual and collective [18,19,20] sound worlds, not only through the musical output itself, but especially through the interaction that results from collective co-creativity.

These concepts can only suggest the importance of society’s role in improvisation, and of framing each entity as an agent within this ecosystem. Obviously, in the most successful systems based on agency concepts, this social vision is not lacking at all. Lewis’ view of human-machine improvisational systems goes far beyond the mere concepts of implementation and design. Indeed, according to him, such systems allow us to explore how meaning is exchanged through sound [50], and can teach us how to live in this new world dominated by concepts of agency, indeterminacy, and difficulty of choice:

Interactions with these systems in musical performance produce a kind of virtual sociality that both draws from and challenges traditional notions of human interactivity and sociality, making common cause with a more general production of a hybrid, cyborg sociality that has forever altered both everyday sonic life and notions of subjectivity in high technological cultures. Being present at the creation of such a new mode of everyday life is simply too interesting to pass up, so that is one reason why I want my computers to improvise. [...] What we learn is not about machines, but about ourselves, and our environment. (Lewis, [51])

4 Agent-based Human-Machine Music Improvisation

The idea of combining the generative potential of AI with human creativity has opened up new paths for musical exploration. Recent studies by Herremans, Chuan and Chew [44], Carnovalini and Rodà [24], and Hernandez [43] indicate the increasing integration of AI tools into the creative process, highlighting the promise of AI-driven composition methods. At the intersection of scientific and artistic fields, the notion of the musical agent is essential. This concept has evolved into a taxonomy of varied applications [73,4]. One of the first significant examples of musical improvisation and co-creation was the aforementioned George E. Lewis’s Voyager system, which remains influential today [50]. Other notable human-machine systems for musical composition and improvisation based on reactive agents include Cypher [65], ListeningLearning [27], Virtualband [55], GenJam [13], The Continuator [60]. However, although conceived in another generation of AI systems, Voyager still remains the best example of a system that, according to Rowe’s taxonomy, personifies the player paradigm [66].

Seeking to find an example that can simultaneously prove optimal in improvised music contexts and embody the instrument paradigm of the said taxonomy, the interest of this research was directed toward the work conceived by the REACH (Raising Co-Creativity in Cyber-Human Musicianship) project at IRCAM [63]. Earlier investigations of the Music Representation team at IRCAM, such as Omax and its stylistic reinjection machine-listening paradigm [5], have further refined the idea of musical generative agents. As a subsequent result, the REACH project in particular is a major example of the musical co-creativity [6,7,8] of AI and of human-machine interaction in cyber-human contexts. In this project, Somax2 is a key piece of software developed as part of this initiative, together with Improtek [58], and Dicy2 [59]. Because of its particular propensity for use in free improvisation, and because of the characteristics that will be described in the next section, Somax2 was therefore chosen as a case study for this research.

4.1 Somax2

Somax2 is a co-creative system for human-machine music improvisation, developed in the Music Representation team at IRCAM - STMS lab under the ERC REACH (Raising Co-Creativity in Cyber-Human Musicianship) project, directed by Gérard Assayag [63]. The system has been extensively described, both practically and theoretically in [9,16,17] but, in this article, we would like to provide for the first time an analytical view of Somax2, in light of what has been introduced in the previous sections.

Somax2 is a reactive co-creative system that inherits from its ancestor Omax [5] its characteristic improvisation paradigm, composed of different stages of *listening*, *learning*, *modeling* and *generating*, extending it with the following peculiarities:

- *listen*: the system is based on the notion of exchange of *influences*. These consists of musical information regarding a specific event just played by an external musician (or an agent within the software), accompanied by annotated values for melodic height (pitch) and harmonic contour (chroma). These influences are thus the computer abstraction of cognitive processes and joint actions that underlie collective listening in improvised music [69,70]. With a fully customizable routing system, it is possible to create complex configurations of generative agents listening to multiple external musicians (or other generative agents) across multiple musical dimensions, thus providing a rich context for multi-modality and interaction.
- *learn* and *model*: the system, thanks to an initial round of learning, based on a SOM (Self-Organized Map) algorithm [49], creates its own model of musical structure. Then, each agent can be loaded with some music material which, after a segmentation and annotation phase, becomes the specific *corpus* assigned to the agent. This phase relies on a segmentation and annotation process with the same notion of multidimensional labels (onset, pitch, chroma) implied by the influences [9]. In this way, the system relies on

a statistical shallow-learning model in real time, which then generates the musical content by concatenative synthesis of the various output segments from the corpus [72].

- *generate*: Somax2 agents, thus, do not simply generate musical outputs out of thin air, but rely on the notion of corpus as musical material from which to draw. The choice of the segment to be played is entrusted to a navigation model (see Figure 1) that searches for the best matches between the sequences received as input and converted into influences, and the position of equivalents within the loaded corpus in a certain agent (named *player*).

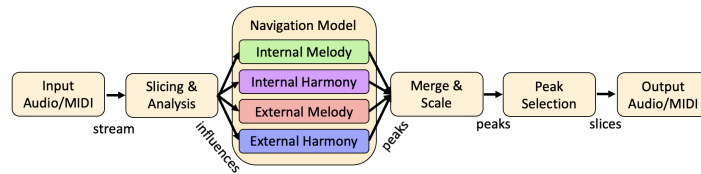


Fig. 1: The navigation model of Somax2, detailing how, at the occurring of every new influence, the system generates an output by a combination of dimensional weights and activation peaks on its own corpus memory [38].

This model is thus suited for the current study, and closely related to the concepts expressed in the previous sections, for two reasons:

- (i) The system is genre-agnostic, making itself capable of interacting regardless of musical genre or playing style. This element refers us back to what Colton and Wiggins expounded earlier [28], emphasizing the celebration of the musical artefact as something strictly bond to the act of creation itself, and incentivizing dialogue between the agents in play.
- (ii) The system generates the reception of influences and the selection of matches through a mechanism of activation *peaks* at certain points in the corpus corresponding to the matches found in the input. These peaks, however, are not simply immediate matches, as it is possible to define the length of the incoming music segment to be compared with what is stored in the individual agent’s memory. Therefore, the peaks are time-shifted and can be incremented by positive matches, decremented by negative matches, or maintain their own state of inertia for a given sequence of events.

The latter mechanism turns out to be a valid implementation of concepts of short-term memory [36] and cognitive resource management [61], which are fundamental in the development of improvised musical interaction [62]. Moreover, what is exhibited about peaks, as shown in Figure 2, is applied on each of the musical dimensions managed by the system (pitch/chroma, internal/external)

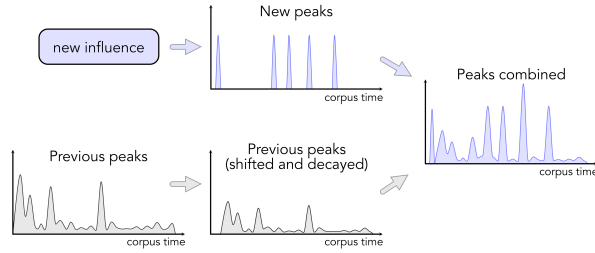
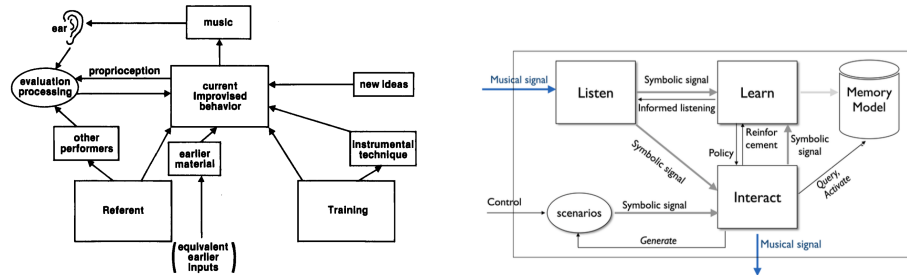


Fig. 2: The mechanism of memory peak profile update at the arrival of each new influence in Somax2 (Assayag et al., [9]).

and merged into a general overall peak profile, thus working equally across all dimensional levels.

Analyzing the model structure of Somax2, and the other agent-based HCI systems developed within the REACH project, like Improtek [58] and Dicy2 [59], it is practically immediate to draw conceptual parallels between the architecture of one of these generative agents [32] and the cognitive model developed by Pressing [61].



(a) Major factors influencing the production of an improvised music behaviour (Pressing, [61]).

(b) Feedback architecture of a generative REACH music agent, like the ones of Somax2 (Dubnov et al., [32]).

Fig. 3: Parallelism between the feedbacks of Pressing’s cognitive model and those of Somax2 generative agents architecture.

As can be seen from Figure 3, not only the feedback systems - defined by Pressing and necessary to maintain a high level of commitment and possibility for real-time decision making - are exactly carried over into the architecture of a Somax2 agent, but here we also find specific implementations of some defined theoretical models:

- The machine listening phase is obviously the implementation of the human listening system, via the ear. But the evaluation processing and awareness traits of other performers falls at the same time into this macro block, which is in a two-way informed listening link with the machine learning module.
- The musical training phase, can be seen as both related to current behavior and instrumental technique. Depending on the case, this concerns the machine learning phase, and the construction of a memory model, or the construction of the corpus itself.
- Previous material, equivalent to previous inputs, corresponds both to the incoming musical signal (thus in the case of Somax2 to the system of influences), and to a kind of global timestamp of the system as a whole.
- Output generation is obviously the musical signal, generated by human-machine interaction.
- Control via scenarios concerns the general configuration of listening and interaction with other musicians and the status of the current behavior, including in relation to the previous behavior.
- The referent, in the general case of a generative agent of this type, could correspond to the corpus present in the Somax2 player. However, as we will see later, in the specific case of Somax2 and its human-machine interaction, the system is placed in a hybrid position of the proposed taxonomy, thus creating an *Environmental Referent*, which is also constituted by the presence of the computer musician in control of the system.

Somax2 has already been highly employed in artistic and scientific research, being also at the center of publications on duo compositional re-improvisations [37] and philosophical studies on embodiment [68]. Thus, it has been chosen in this research, to be the main practical environment to operate with, and to extend previous existing research to the domain of CFI, with the aid of some sort of referent [61]. In the context of the case study described in this article, this consists of a specific environment composed by the corpus of choice, and the intervention of the human computer musician. The choice hopes to prove context and substance to support the implications tackled in this research, as well as to give some solid framework for good practices in human-machine music interaction.

4.2 The case of tellKujira

The artistic project in which the means of scientific research mentioned in this article were applied began in late 2023 and continues at the time of writing this essay, with future projections. After two phases of artistic residencies², which lasted first three weeks and then one week, a studio recording at Duna Studio (Italy) in June 2024 consecrated the work done for the subsequent vinyl LP release, scheduled for 2025. Public returns of this collaboration are expected

² <https://reach.ircam.fr/index.php/2023/11/20/tellkujira-and-marco-fiorini-in-residence-at-ircam/>

through concerts³ and international teaching workshops hosted by leading institutions of higher music education, such as the one that is taking place throughout the year 2024 at the Alfredo Casella Conservatory in L’Aquila, Italy⁴. This is because, not only do we believe it is important to disseminate such research work, but also because the collaborative and collective aspect is even more amplified by this kind of pedagogical collective experience.

The ensemble tellKujira was born in autumn 2020. Since then it has self-produced a series of artistic residencies, focusing on improvisation and collective writing, instant composition, timbre, prepared instruments, alternative tunings. It is in this process that a flexible musical path was born, in a continuous search for structure, open to contamination and which aims to discover a unique blend composed of abstract soundscapes, at times ambient, polyrhythms, industrial timbres. A versatile blend, which does not hesitate to deconstruct any type of language:

We imagined a kind of rock band in the form of a classic quartet, with two electric guitars instead of two violins. By virtue of this chamber choice, every individualism has been set aside in favor of a “We”, a focal center that was also the center of our music. Collective research from a perspective that is not necessarily aimed at performance immediately has become the main vector for the emancipation of the quartet from individual individuals and for the search for a common language. (author’s interview with tellKujira)

The following testimony immediately brings to mind the concept of we-ness explored in various studies on improvisation [69,30]. Indeed, it was clear early on that integrating into such an ensemble would not be the same as considering oneself a mere “+1,” but that collective personal and artistic growth work would have to accompany the work of aesthetic research and scientific application. In the following case, we will consider the intervention of a computer musician (the author) by giving him the role-name of RIM (Réalisateur en Informatique Musicale) as conceived at IRCAM by Pierre Boulez [81].

4.3 On Interaction

The repertoire we worked on consisted of three pieces averaging 10 minutes in length each, that had initially originated from collective improvisations and extemporaneous composition sessions, and later re-opened and reworked again with the same idea, but this time as a quintet, with the addition of a RIM (the author) operating Somax2. The kind of work done, could be defined as a particular case of collective practice of free improvisation, from which semi-free forms of spontaneous composition spring from improvisational practice. This kind of methodology, from which extemporaneous structures, evocations, notations made of gestures and planes of interaction originate, has been widely employed by renowned improvisers, such as Anthony Braxton, Stefano Battaglia,

³ <https://reach.ircam.fr/index.php/2024/09/28/tellkujira-and-marco-fiorini-live-at-forli-open-music-festival/>

⁴ <https://reach.ircam.fr/index.php/2024/02/28/sound-thinkin-workshop-on-music-improvisation-with-co-creative-agents/>

Jacob Anderskov [1,2] through music productions and pedagogical workshops, and also reported in musicological and cognitive studies [21].

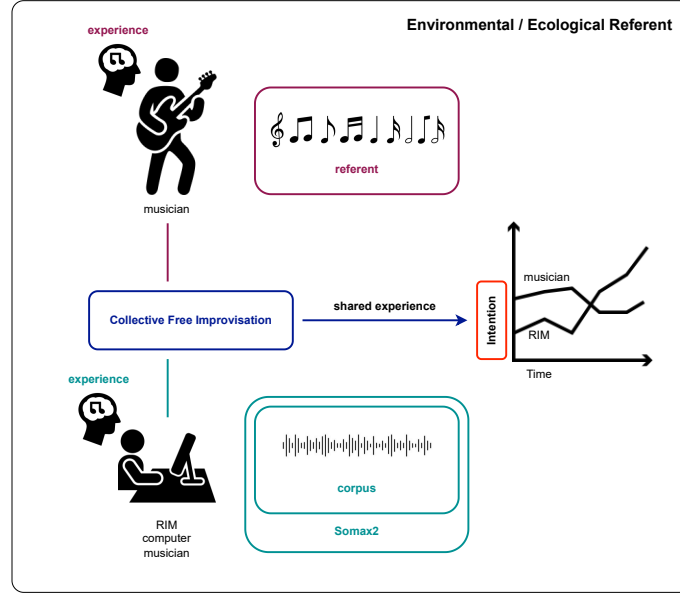


Fig. 4: The proposed definition of Environmental / Ecological Referent, composed of its various characteristic modules, in the case of a duo interaction. If more musicians are involved in the framework, their contribution to the CFI has to be taken in consideration, equally as the ones depicted here.

An important aspect in considering this kind of interaction not only between musicians but with the addition of generative agents is that, as mentioned in the previous sections, there is the impression that a new concept of referent comes into play. In parallel with similar ecological theories in other fields, like to one of Lindblom [52] on embodied social cognition, and Gibson [39] on ecological psychology, we might call this new framework an *Environmental Referent* or *Ecological Referent*. As proposed in Figure 4 and described in Table 1, this is composed of the musician or musicians interactive in the environment, and their instruments, the presence of the RIM, the corpus (i.e., the musical material loaded into the generative agents, with which one decides to interact), the generative agents themselves, the context of CFI collectivity in which one finds oneself (expanded by one's personal and shared musical experience, and thus by instrumental and human practice), and the intentionality of the interaction gestures [22].

In Table 1, each element and agent of the framework has been assigned a specific affordance. Inspired by Gibson’s theory of affordances [39] and by Borgo’s reflections on music improvisation [20], we refer to this as perceived action possibilities in the specific proposed framework, within the domain of human-machine music improvisation. It is worth noting that this framework can be applied to both types of non-idiomatic improvisation formulated by Pressing (referent-based and referent-free) [61], as, by expanding on some of the concepts peculiar to the field of HCI, it is also valid in cases of referent-free interaction.

Table 1: Elements of the Environmental / Ecological Referent, with relation to their corresponding agents and identified affordances in the framework.

Element	Agent	Affordance
Instrument	Musician	Control
Generative AI Agent	Somax2 [9]	Player/Instrument [67]
Somax2 UI	RIM	Control
Referent [61]	Musician	Past + present material
Corpus [72]	Somax2	Present material
CFI [22]	Musician + RIM	Shared environment
Individual Experience	Musician, RIM	Past context
Shared Experience	Musician + RIM	Present context
Intention [22]	Musician + RIM	Shared goal and outcome

4.4 On Control

The definition proposed in the previous section is not only meant to be viewed from the specific point of view of interaction, but also proves to be fundamental in expressing reflections on the notion of control in HCI with AI agents in creative contexts. Here, in fact, the terms *ecological* or *environmental*, takes on an empowering meaning of responsibility. We shall in fact consider the mindset with which the systems in question are approached, and in the possible philosophical, musicological and ethical implications that may arise, as well as in those of concrete future implementations [26]. Indeed, the position of this study is to donate an empirical framework for the use of AI in creative domains, from which humans can concretely project fruits of their imagination, but without abandoning the concept of co-creative interaction [6,7,8], and thus working through bidirectional feedback within a shared ecosystem:

Designing computer programs that will recognize and reason about human musical concepts enable the creation of applications for performance, education, and production that resonate with and reinforce the basic nature of human musicianship. (Rowe, [67])

Therefore, considering the proposed definition of Environmental or Ecological Referent, Rowe’s taxonomy can be extended, and by placing Voyager on the

extreme axis of the player paradigm, Somax2 can assume a hybrid position (see Figure 5).

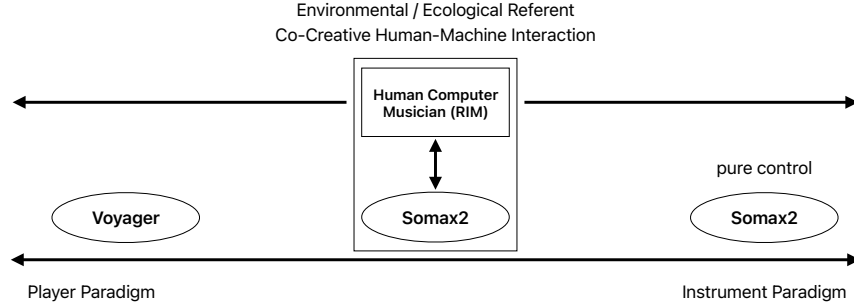


Fig. 5: The positioning of Voyager and Somax2 according to the model proposed in this study over Rowe's extended taxonomy.

Although, by its own conception, Somax2 can probably be found at the exact midpoint of the axis between the player and instrument paradigms, total control by a RIM could move it completely toward the instrument extreme - thus weakening its co-creative AI intent, but repositioning it as a Digital Musical Instrument (DMI) [45]. In light of the above, however, thanks to the intervention of the RIM in the context of the Environmental/Ecological Referent, the new "Somax2 + RIM" system is able to move freely along the entire axis of interaction proposed by Rowe. This implication opens a new track in the normally exposed duality of controlling AI or being controlled by it, consistent with what Caramiaux and Donnarumma expressed in their studies on AI, music and dance, making thus us reflect on how we can "Being the Artificial Player":

The use of ML and AI in music-related research does not have to be constrained by a paradigm of control of humans over algorithms, of humans over musical forms, or of humans over other humans. (Caramiaux and Donnarumma, [23])

Another key aspect, not yet addressed by this research, as a full public restitution of this work has yet to take place, is precisely that of the audience perception of a co-creative AI system being employed in an artistic creation following the practices proposed by this study. In the absence of qualitative and quantitative measures, one can only speculate on what might be happening, but we think this is a very fascinating track to follow in order to reevaluate the role of the artist in contemporary instances of interdisciplinary production, and at the same time demystify certain controversial positions regarding the role of AI in the creative field.

5 Conclusions

This paper sought to provide an overview of the applications of generative agent-based AI systems in the specific context of an improvised music interaction. Perspectives were offered on various definitions of creativity, also addressing computational creativity to co-creativity in the context of interactive systems. An excursus on music improvisation and its insights in musicological, philosophical and cognitive fields is at the same time included to highlight the specific particularities of the research in question, which was carried out through a practice-based method, thanks to a continuous intersection of scientific and artistic research. Concretely, the study led to an artistic realization by the author as RIM added to the tellKujira ensemble for a long-term collaboration. The choice of the particular AI software Somax2 was justified as the optimal candidate for the specific context, and its use in this particular study led to the formulation and proposition of a new framework of collective co-creation within HCI improvised music: the Ecological or Environmental Referent. The Somax2 system was also analyzed according to the proposed hybrid methodology, and repositioned on a specific taxonomic axis of agent-based systems paradigms within HCI applications.

Unfortunately, during the research documented in this article, it was not possible to collect quantitative data regarding user utilization of Somax2. However, this, along with data collected from musicians interacting with Somax2, was covered in a parallel study conducted by the author regarding embodiment factors in generative AI algorithms as applied to duo interactions [68]. Nevertheless, it is the author's intention to collect these kinds of measures in future research regarding collective improvisation, to evaluate the framework proposed in this research through objective evaluation metrics. In addition, future studies will focus on the audience perspective in perceived co-creative HCI music improvisation performances. Moreover, at the time of writing, a study is being conducted on the choices made by a RIM in co-creative musical interactions, according to the framework proposed here.

The above study is intended to stand as a basis for best practices in the use and interaction of generative AI in co-creative art application domains. By insisting on the ecological importance of this practice, the responsibility of employing certain systems, and the enhancement of the artist's creative thinking, this study hopes to establish a good foundation for future developments of fundamental evolutions in the area of human-computer interaction in co-creative fields.

6 Acknowledgments

This research is supported by the European Research Council (ERC) as part of the Raising Co-creativity in Cyber-Human Musicianship (REACH) Project directed by Gérard Assayag, under the European Union's Horizon 2020 research and innovation program (GA #883313).

References

1. Anderskov, J.: Action vs. Reaction, <https://www.researchcatalogue.net/view/481737/481738>. Artistic Research Project at RMC, (2016).
2. Anderskov, J.: Sonic Complexion, <https://www.researchcatalogue.net/view/740085/1373319>. Artistic Research Project at RMC, financed by the Danish Cultural Ministry, (2020).
3. Ariza, C.: The Interrogator as Critic: The Turing Test and the Evaluation of Generative Music Systems. In: *Computer Music Journal*, 33-2, pp. 48–70, MIT press (2009)
4. Armitage, J., Magnusson, T.: Agential Scores: Artificial Life for Emergent, Self-Organising and Entangled Music Notation. In: *Proceedings of the International Conference on Technologies for Music Notation and Representation – TENOR (2023)*
5. Assayag, G., Bloch, G., Chemillier, M., Cont, A., Dubnov, S.: Omax brothers: a dynamic topology of agents for improvisation learning. In: *Proceedings of the 1st ACM workshop on Audio and music computing multimedia*, Santa Barbara, United States (2006)
6. Assayag, G.: Creative Symbolic Interaction: Key Note Talk. 40th Intl. Comp. Mus. Conf. and 11th Sound and Music Comp. Conf. (ICMC / SMC joint conf.), Athenes, Greece. pp. 1–6 (2014)
7. Assayag, G.: Co-créativité humains-machines. Une réflexion sur les indisciplines symboliques. In: *Revue Francophone d’Informatique et Musique*, 8-7, Culture du code (2020)
8. Assayag, G.: Human-Machine Co-Creativity. In Bernard Lubat, Gérard Assayag, Marc Chemillier. *Artisticiel / Cyber-Improvisations, Phonofaune, Dialogiques d’Uzeste* (2021)
9. Assayag, G., Bonnasse-Gahot, L., Borg, J.: Cocreative Interaction: Somax2 and the REACH Project. In: *Computer Music Journal*, 1-19, MIT press (2024)
10. Bailey, D.: *Musical improvisation: Its nature and practice*. Da Capo Press (1992)
11. Benjamin, W.: The Work of Art in the Age of Mechanical Reproduction. In: *Illuminations*, edited by Hannah Arendt, pp. 217–251. New York: Schocken (1936-1969)
12. Benson, B.: *The Improvisation of Musical Dialog: A Phenomenology of Music*. Cambridge University Press, New York NY (2003)
13. Biles, J.: GenJam: A genetic algorithm for generating jazz solos. In: *Proceedings of the international computer music conference, ICMC (1994)*
14. Boden, M. A.: *Artificial Intelligence & Natural Man*, Harvester Press (1977)
15. Boden, M. A.: Creativity and Artificial Intelligence. In: *Artificial Intelligence*, 103, pp. 347–356 (1998)
16. Borg, J.: The Somax 2 Theoretical Model. STMS-IRCAM, Technical Report (2021)
17. Borg, J.: The Somax 2 Software Architecture Rev. 0.2.0. STMS-IRCAM, Technical Report (2021)
18. Borgo, D.: Sync or Swarm: Musical Improvisation and the Complex Dynamics of Group Creativity. In: *Lecture Notes in Computer Science 4060*, pp. 1–24 (2006)
19. Borgo, D.: The Art of Improvisation in the Age of Computational Participation. In: *The Routledge Companion to Jazz Studies*, pp.411–421 (2018)
20. Borgo, D.: The Complex Dynamics of Improvisation. In: Bader, R. (eds) *Springer Handbook of Systematic Musicology*. Springer Handbooks. Springer, Berlin, Heidelberg (2018)
21. Canonne, C.: Rehearsing free improvisation? An ethnographic study of free improvisers at work. In: *Music Theory Online*, 24(4), (2018)

22. Canonne, C., Garnier, N. B.: A Model for Collective Free Improvisation. In: Conference: Mathematics and Computation in Music - Third International Conference, MCM 2011, Paris, France, June 15-17, (2011)
23. Caramiaux, B., Donnarumma, M.: Artificial Intelligence in Music and Performance: A Subjective Art-Research Inquiry. In: Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity, inPress. (hal-03046229) (2020)
24. Carnovalini, F., Rodà, A.: Computational Creativity and Music Generation Systems: An Introduction to the State of the Art. In: Frontiers in Artificial Intelligence, 14-3 (2020)
25. Cheatley, L., Ackerman, M., Pease, A., and Moncur, W.: Co-creative songwriting for bereavement support. In: Proceedings of the 11th International Conference on Computational Creativity, ICCO'20 (2020)
26. Ciuffo, T.: Design Concepts and Control Strategies for Interactive Improvisational Music Systems (2003)
27. Collins, N.: LL: Listening and learning in an interactive improvisation system. Technical Report, University of Sussex (2011)
28. Colton, S., Wiggins, G. A.: Computational creativity: The final frontier? In: Frontiers in Artificial Intelligence and Applications 242, pp. 21–26 (2012)
29. Comajuncosas, J. M.: Assessing creativity in computer music ensembles: a computational approach. PhD Thesis, Universitat Pompeu Fabra (2015)
30. Corbett, J.: A listener's guide to free improvisation. University of Chicago Press (2016)
31. Douglas, R. L.: Formalizing an African-American Aesthetic. In: New Art Examiner, pp. 18–24 (1991)
32. Dubnov, S., Assayag, G., Gokul, V.: Creative Improvised Interaction with Generative Musical Systems. In: IEEE 5th International Conference on Multimedia Information Processing and Retrieval (MIPR), pp. 121–126, (hal-04010750) (2022)
33. Edmonds, E., Candy, L.: Relating theory, practice and evaluation in practitioner research. *Leonardo*, 43(5):470–476 (2010)
34. Emmerson, S.: Combining the acoustic and the digital: Music for instruments and computers or prerecorded sound. In: R.T. Dean ed., *The Oxford Handbook of Computer Music*, Oxford University Press, New York, ch 9, pp. 167–188 (2009)
35. Fdili Alaoui, S.: Making an interactive dance piece: Tensions in integrating technology in art. In Proceedings of the 2019 on Designing Interactive Systems Conference, pp 1195–1208 (2019)
36. Fetsch, C., et al.: Neural Correlates of Reliability-based Cue Weighting during Multisensory Integration. *Nature Neuroscience*, 15(1), pp. 146–154 (2012)
37. Fiorini, M., Colombo, L.: A.I. Komboï: Re-Imagining Xenakis' Komboï through Cyber-Human Co-Creative Improvisation Practice. In: Journées d'Informatique Musicale 2024, May 2024, Marseille, France. (hal-04572304) (2024)
38. Fiorini, M.: Somax2 User's Guide. (2024). STMS-IRCAM, Technical Report, <https://github.com/DYCI2/Somax2/blob/76d152fb526412211a5646ad0dc3492498fc0555/Somax2%20User's%20Guide.pdf>
39. Gibson J. J.: *The Theory of Affordances. The Ecological Approach to Visual Perception*. Houghton Mifflin Harcourt (HMH), Boston (1979)
40. Golvet, A., Goupil, L., Saint-Germier, P., Matuszewski, B., Assayag, G., Nika, J., Canonne, C.: With, against, or without? Familiarity and copresence increase interactional dissensus and relational plasticity in freely improvising duos. In: *Psychology of Aesthetics, Creativity, and the Arts*, American Psychological Association (2021)

41. Goupil, L., Saint-Germier, P., Rouvier, G., Schwarz, D., Canonne, C.: Musical coordination in a large group without plans nor leaders. *Scientific Reports* 10, 20377 (2020)
42. Hennessey, B. A., Amabile, T. M.: Creativity. In: *Annual review of psychology*, 61 pp. 569–598 (2010)
43. Hernandez-Olivan, C., Beltran, J. R.: Music Composition with Deep Learning: A Review. (2021)
44. Herremans, D., Chuan, C., Chew, E.: A Functional Taxonomy of Music Generation Systems. In: *ACM Comput. Surv.* 50-5, Association for Computing Machinery (2017)
45. Hunter, T. G.: Designing Digital Musical Instruments: Adopting an ecological lens to identify design considerations. In *Companion Publication of the 2021 ACM Designing Interactive Systems Conference (DIS '21 Companion)*. Association for Computing Machinery, New York, NY, USA, 33–35 (2021). <https://doi.org/10.1145/3468002.3468238>
46. *Journal of Artistic Research*, <https://www.jar-online.net/en>
47. Kim, J., Maher, M. L.: Evaluating the Effect of Co-Creative Systems on Design Ideation. In: *Proceedings of the Second Workshop on the Future of Co-Creative Systems* (2021)
48. Koch, J., Ravikumar, P. T., Calegario, F.: Agency in Co-Creativity: Towards a Structured Analysis of a Concept. In: *Proceedings of the Second Workshop on the Future of Co-Creative Systems* (2021)
49. Kohonen, T.: The Self-Organized Map. In: *Proceedings of the IEEE* 78(9), pp. 1464–1480 (1990)
50. Lewis, G. E.: Too Many Notes: Computers, Complexity and Culture in 'Voyager'. In: *Leonardo Music Journal*, 10, pp. 33–39 (2000)
51. Lewis, G. E.: Why Do We Want Our Computers to Improvise? In: *The Oxford Handbook of Algorithmic Music* (2018)
52. Lindblom, J.: *Embodied Social Cognition*. Springer (2015)
53. Maher, M.: Evaluating creativity in humans, computers, and collectively intelligent systems. In: *Proceedings of the 1st DESIRE Network Conference on Creativity and Innovation in Design*, Aarhus, Denmark, pp 22–28 (2010)
54. Mogensen, R.: Evaluating a swarm algorithm as an improvising accompanist: an experiment in using transformed analysis of George E. Lewis's *Voyager*. In: *Conference on Computer Simulation of Musical Creativity CSMC16* (2016)
55. Moreira, J., Roy, P., Patchet, F.: Virtualband: Interacting with stylistically consistent agents. In: *Proceedings of the 14th international society for music information retrieval conference, ISMIR* (2013)
56. Mosqueira-Rey, E., Hernández-Pereira, E., Alonso-Ríos, D. et al.: Human-in-the-loop machine learning: a state of the art. In: *Artificial Intelligence Review*, 56 (2023)
57. Mouawad, P., Dubnov, S.: On Modeling Affect in Audio with Non-Linear Symbolic Dynamics. In: *Advances in Science, Technology and Engineering Systems Journal*, Vol. 2, No. 3, pp. 1727–1740 (2017)
58. Nika, J., Chemillier, M., and Assayag, G.: ImproteK: introducing scenarios into human-computer music improvisation. *ACM Computers in Entertainment* (2017)
59. Nika, J., Déguernel, K., Chemla-Romeu-Santos, A., Vincent, E., and Assayag, G.: DYCI2 agents: merging the "free", "reactive", and "scenario-based" music generation paradigms. *International Computer Music Conference* (2017)
60. Pachet, F.: The continuator: Musical interaction with style. In: *Journal of New Music Research*, 32-3 (2003)

61. Pressing, J.: Cognitive Processes in Improvisation. In: Cognitive Processes in the Perception of Art. Ed. W.R.Crozier and A.J. Chapman, Elsevier Science Publishers B.V., North-Holland (1984)
62. Pressing, J.: Improvisation: Methods and Models. In: Generative processes in music. Ed. J. Sloboda, Oxford University Press (1987)
63. REACH (Raising Co-Creativity in Cyber-Human Musicianship) Project, <https://reach.ircam.fr>
64. Research Catalogue - an international database for artistic research, <https://www.researchcatalogue.net>
65. Rowe, R.: Machine listening and composing with cypher. In: Computer Music Journal, 16-1, 43, MIT press (1992)
66. Rowe, R.: Interactive Music Systems. MIT Press, Cambridge, MA (1993)
67. Rowe, R.: Machine Musicianship. MIT Press, Cambridge, MA (2001)
68. Saint-Germier, P., Canonne, C., Fiorini, M.: The corpus' body. Embodied Interaction from Machine-Learning in Human-Machine Improvisation. In: Proceedings of the International Conference on AI and Musical Creativity, AIMC (2024)
69. Saint-Germier, P., Goupil, L., Rouvier, G., Schwarz, D., Canonne, C.: What it is like to improvise together? Investigating the phenomenology of joint action through improvised musical performance. In: Phenomenology and the Cognitive Sciences (2021)
70. Saint-Germier, P., Paternotte, C., Canonne, C.: Joint Improvisation, Minimalism and Pluralism about Joint action. Journal of Social Ontology, 7(1), pp. 97–118 (2021)
71. Sarkar, P., Chakrabarti, A.: Assessing design creativity. In: Design Studies, pp. 1–36 (2011)
72. Schwarz, D., Beller, G., Verbrugge, B., Britton, S.: Real-Time Corpus-Based Concatenative Synthesis with CataRT. In: Proceedings of the 9th International Conference on Digital Audio Effects (DAFx) (2006)
73. Tatar, K., and Pasquier, P.: Musical agents: A typology and state of the art towards Musical Metacreation. In: Journal of New Music Research, 48-1, Routledge (2019)
74. Taylor, C. W.: Various approaches to and definitions of creativity. In: The nature of creativity, pp. 99–121 (1998)
75. Truesdell, E. J. K., Smith, J. B., Mathew, S., Katuka, G. A., Griffith, A., McKlin, T., Magerko, B., Freeman, J., Boyer, K. E.: Supporting Computational Music Remixing with a Co-Creative Learning Companion. In: Proceedings of the 12th International Conference on Computational Creativity, ICCO'21 (2021)
76. Turing, A. M.: Computing Machinery and Intelligence. In: Mind 49, pp. 433–460 (1950)
77. Velardo, V.: Recursive ontology: A systemic theory of reality. In: Axiomathes 26.1, pp. 89–114 (2016)
78. Velardo, V.: Towards a Music Systems Theory: Theoretical and Computational Modelling of Creative Music Agents. PhD Thesis, University of Huddersfield (2017)
79. Wiggins, G. A.: A preliminary framework for description, analysis and comparison of creative systems. In: Knowledge-Based Systems, 19(7) pp. 449–458 (2006)
80. Zacharakis, A., Kaliakatsos-Papakostas, M., Kalaitzidou, S., Cambouropoulos, E.: Evaluating human-computer co-creative processes in music: A case study on the CHAMELEON melodic harmonizer. In: Frontiers in Psychology, 12 (2021)
81. Zattra, L.: Les origines du nom de RIM (Réalisateur en informatique musicale). In: Journées d'Informatique Musicale, France (2013)