Springer Proceedings in Humanities and Social Sciences

Alessandra Micalizzi *Editor*

Artificial Creativity

Looking at the Future of Digital Culture

Springer

Springer Proceedings in Humanities and Social Sciences

Aims and Scope

Springer Proceedings in Humanities and Social Sciences collates the most current, selected research presented at international and regional conferences, seminars, and workshops to bring it to a global readership. The series features volumes (in electronic and print formats) of selected contributions from conferences in all areas of both the humanities and social sciences, including interdisciplinary work, covering philosophy, religious studies, history, language and literature, communications and cultural studies, anthropology, sociology, developmental studies and political science and policy. In addition to an overall evaluation by the publisher of the topical interest, research quality, and timeliness of each volume, each contribution is refereed to standards comparable to those of leading journals, resulting in authoritative contributions to the respective fields. Springer's production and distribution infrastructure ensures rapid publication and wide circulation of the latest developments in the most compelling and promising areas of research in HSS today. The series publishes both Open Access and non-Open Access works.

This series promotes the inclusion of 'Selected Proceedings', a middle ground between edited volumes and full conference proceedings. The series is specifically looking for submissions from conferences with a defined, focused theme, or a regional focus in a specific discipline.

The series offers the option for conferences to use Springer Nature's conference management platform, Equinocs, free of charge.

For more information, contact the listed publishing editor.

Alessandra Micalizzi
Editor

Artificial Creativity

Looking at the Future of Digital Culture

Springer

Editor
Alessandra Micalizzi
Sociology, Communication and
Cultural Processes
Pegaso University
Naples, Italy

ISSN 2948-2526 ISSN 2948-2534 (electronic)
Springer Proceedings in Humanities and Social Sciences
ISBN 978-3-031-88333-0 ISBN 978-3-031-88334-7 (eBook)
https://doi.org/10.1007/978-3-031-88334-7

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

If disposing of this product, please recycle the paper.

Foreword

Notes on Generative AI

"AI" as a Cultural Perception

There is not one specific technology or a single research project called "AI." However, we can follow how our cultural perception of this concept evolved over time and what it was referring to in each period. In the last 50 years, when an allegedly uniquely human ability or skill is being automated by means of computer technology, we refer to it as "AI." Yet, as soon as this automation is seamlessly and fully successful, we tend to stop referring to it as an "AI case." In other words, "AI" refers to technologies and methodologies that automate human cognitive abilities and are starting to function but are n't quite there yet. "AI" was already present in the earliest computer media tools. The first interactive drawing and design system, Ivan Sutherland's *Sketchpad* (1961–1962)[1], had a feature that would automatically finish any rectangles or circles you started drawing. In other words, it knew what you were trying to make. In the very broad understanding just given, this was undoubtedly "AI" already.

My first experience with a desktop paint program running on an Apple II was in 1984, and it was truly amazing to move your mouse and see simulated paint brushstrokes appear on the screen. However, today we no longer consider this to be "AI." Another example would be the Photoshop function that automatically selects an outline of an object. This function was added many years ago—this, too, is "AI" in the broad sense, yet nobody would refer to it as such today. The history of digital media systems and tools is full of such "AI moments"—amazing at first, then taken for granted and forgotten as "AI" after a while. (In academic studies of AI history, this phenomenon is referred to as the *AI effect*.) Thus, today *creative AI* refers only to recently developed methods where computers transform some inputs into new media outputs (e.g., text-to-image models) and specific techniques (e.g., certain types of AI models). However, we must remember that these methods are neither the first nor

[1] Ivan Sutherland, "Sketchpad: A Man-Machine Graphical Communication System," AFIPS '63: Proceedings of the Spring 1963 Joint Computer Conference (May 21, 1963), https://dl.acm.org/doi/10.1145/1461551.1461591.

the last in the long history and future of simulating human art abilities or assisting humans in media creation. I expect that after a certain period, GenAI technology will be taken for granted, becoming ubiquitous and thus invisible—and some other cultural use of computers will come to be seen as "AI."

From Representation to Prediction

Historically, humans created images of existing or imagined scenes using a number of methods, from manual drawing to 3D CG (see below for explanation of the methods). With AI generative media, a fundamentally new method emerges. Computers analyze patterns in large datasets of existing media. Using these learned patterns, they can then create new, previously unseen still and moving images that exhibit similar characteristics. This process forms the core of generative AI technology.

One can certainly propose different historical paths leading to visual generative media today or divide one historical timeline into different stages. Here is one such possible trajectory:

1. Creating representations manually (e.g., drawing with variety of instruments, carving, etc.). More mechanical stages and parts were sometimes carried out by human assistants typically training in their teacher's studio—so there is already some delegation of functions.
2. Creating manually but using assistive devices (e.g., perspective machines, camera lucida). From *hands* to *hands + device*. Now some functions are delegated to mechanical and optical devices.
3. Photography, X-ray, video, volumetric capture, remote sensing, and photogrammetry. From *using hands* to *recording information using machines*. From *human assistants* to *machine assistants*.
4. 3D CG: You define a 3D model in a computer and use algorithms that simulate effects of light sources, shadows, fog, transparency, translucency, natural textures, depth of field, motion blur, etc. From r*ecording* to *simulation*.
5. Generative AI: Using media datasets to predict still and moving images. From *simulation* to *prediction.*

"Prediction" is the actual term often used by AI researchers in their publications describing visual generative media methods. So, while this term can be used figuratively and evocatively, this is also what happens scientifically when you use image generative tools. When working with a text-to-image AI model, the artificial neural network attempts to predict the images that correspond best to your text input. I am certainly not suggesting that using all other already accepted terms such as "generative media" is inappropriate. But if we want to better understand the difference between AI visual media synthesis methods and other representational methods developed in human history, employing the concept of "prediction" and thus referring to these AI systems as "predictive media" captures this difference well.

Visual AI and Media Accumulation

I will use the term "visual AI" to refer to computational methods that use machine learning for generating and editing visual content, trained on vast amounts of images and videos found across the web. In other words, this is my shortcut for saying "generative AI used to make and edit images, video and animation."

Visual AI is the fourth significant *data* effect of the web—a global accumulation of networked, hyperlinked cultural content that began to grow quickly after 1993. Although people have been sharing texts and images on the internet since the 1970s, this process picked up speed after 1993, when the first visual browser, Mosaic, was introduced on January 23 of that year.

I have observed several repercussions of the growth of information on the web over the next 30 years. If we wish to situate the development of visual AI in the early 2020s in this timeline, here are four such effects. Certainly, others can be also named, so this is only one list of techno-cultural developments technologies enabled by the web I am particularly interested in:

1. The first effect is the switch from categorical, hierarchical, and structured organization of information (exemplified by library catalogs and early web directories) to search engines in the late 1990s. There was so much content that organizing it in conventional ways was no longer practical, and search became the new default. Note that *web search is based on a prediction of what will be most relevant to the user* as opposed to giving you a precise and definite answer. Note that generative AI is also predictive—it predicts possible text, images, animation, or music in response to your question or prompt. The regime of absolute certainty, i.e., a truth versus a lie typical for human civilization is replaced by predictions, as statistics becomes foundation of human sciences in the twentieth century and data science and AI in recent decades.
2. The second major effect is the rise in popularity of data visualization during the 2000s. The field comes into its own around 2005. As a part of this development, the new field "artistic data visualization" develops in the same decade, along with other new cultural fields: data art and data design. (In our lab, we created Phototrails, Selfiecity, and On Broadway in 2012–2014. These were first interactive visualizations of millions of Instagram images.)[2] If search attempts to find the most relevant items in the giant data universe, visualization tries to show parts of this universe in one image, revealing patterns and connections.
3. The third effect is the emergence of "data science" as the master discipline of the new big data era at the end of the 2000s. While many techniques employed in data science have already been available for decades, the rapid increase in unstructured data in the 2000s motivated the development of a separate data science field—the key new profession of the data society. My own version of

[2] See "Projects," Cultural Analytics Lab, accessed September 27, 2024, https://lab.culturalanalytics.info/p/projects.html.

> this stage was "cultural analytics," an idea I introduced in 2005 and developed over the following 15 years in our lab. Cultural analytics applies the paradigm of data science to cultural content, using computational techniques to analyze and visually represent large collections of digital media, enabling the exploration of patterns and trends across entire cultural datasets.[3]

The next, but certainly not the last, effect of the growth of online visual digital content is visual AI which becomes popular in early 2020s. DALL-E was released in 2020, Midjourney in 2022, and Adobe Firefly and Runway Gen-1 in 2023. Today (2025), hundreds of other AI image, video, and animation tools exist, and image generation is also available in all popular AI text bots. (A bit earlier around 2017, a particular AI method for media generation called GAN became already popular with digital artists.)

(It is relevant to mention that visual AI and generative AI in general build on 20 years of research. The key breakthrough was the idea to use web content universe as a source of data for machine learning, without labeling it. This idea was already articulated in the research papers published around 2001.)

Let's see what kind of pattern is established by these four effects. Search is the first method to deal with the new scale of content on the web. Data science focuses on finding patterns, relations, clusters, and outliers in big data and also predicting future data. Data visualization tries to summarize datasets visually. And now generative AI explores "big content" in yet another way, generating new content which combines many patterns from existing media.

To put this differently, generative AI synthesizes new content that has statistical properties similar to existing content. But it's not a copy of what already exists. AI generates new content (texts, images, animation, 3D models, music, singing, etc.) by interpolating between existing points in the latent space. This space contains numerous patterns and structures extracted by artificial networks from billions of image-text pairs, trillions of text pages, and other large collections of existing human cultural artifacts. AI predicts what could exist between these points in space of patterns. For example, it can predict a "painting" made by artists A, B, and C, using techniques D and E, with content F, G, and E, with mood, colors M and N, proportion W, composition K, etc.

Note that the three earlier developments all approach big data by summarizing it. Web search reduces billions of web pages to the top results. Data visualization reduces it to a diagram. Data science reduces it by using summary statistics, cluster analysis, regression, or latent space projection. But visual AI is doing something new. It also first reduces big data during learning and then generates new data points.

One way to sum up all this is to say that we moved from probabilistic search (1999) to probabilistic media generation (2022). But certainly, generative AI and its subset visual AI are not the last effect of the existence of web data; others will likely emerge in the future.

[3] See Lev Manovich, *Cultural Analytics* (The MIT Press, 2020).

AI and Modernism

After training on trillions of text pages or billions of images taken from the web, AI models can generate new texts and visuals on the level of highly competent professional writers, artists, photographers, or illustrators. These capacities of the AI models are distributed over trillions of connections between billions of artificial neurons rather than determined by standard algorithms. In other words, we developed a technology that, in terms of complexity, is extremely similar to the human brain. We don't fully grasp how our AI technology works, just as we don't fully comprehend human intellect and creativity.

On the surface, *the logic of modernism appears to be diametrically opposed to the process of training generative AI systems*. Modern artists desired to depart from classical art and its defining characteristics such as visual symmetry, hierarchical compositions, and narrative content. In other words, their art was founded on a fundamental rejection of everything that had come before it (at least in theory, as expressed in their manifestos). AI models are trained in the opposite manner, by learning from historical culture and art created up to now. AI model is analogous to a very conservative artist studying in the "meta" *museum without walls* that houses historical art.

But we all know that art theory and art practice are not the same thing. Modern artists did not completely reject the past and everything that came before them. Instead, *modern art developed by reinterpreting and copying images and forms from old art traditions*, such as Japanese prints (van Gogh), African sculpture (Picasso), and Russian icons (Malevich). Thus, the artists only rejected the dominant high art paradigms of the time, realistic and salon art, but not the rest of human art history. In other words, it was deeply historicist: rather than inventing everything from scratch, it innovated by adapting certain older aesthetics to contemporary art contexts. In the case of geometric abstract art created in 1910s, these artists used images that were already widely used in experimental psychology to study human visual sensation and perception.[4]

When it comes to artistic AI, we should not be blinded by how these systems are trained. Yes, AI models are trained on previously created human art and culture artifacts. However, their newly generated outputs are not mechanical replicas or simulations of what has already been created. In my opinion, these are frequently *genuinely new* cultural artifacts with *previously unseen content, aesthetics, or styles*.

Of course, simply being novel does not automatically make something culturally or socially interesting or significant. Indeed, many definitions of "creativity" agree on this point: it is the creation of something that is both original and worthwhile or useful.

[4] For the detailed analysis of these relations between modern art and experimental psychology, see Paul Vitz and Arnold Glimcher, *Modern art and Modern Science: The Parallel Analysis of Vision*, 1983.

However, estimating what percentage of all novel artifacts produced by generative AI are also useful and/or meaningful for a larger culture is not a feasible project at this time. For one thing, I am not aware of any systematic effort to use such systems to "fill in," so to speak, a massive matrix of all content and aesthetic possibilities by providing millions of specifically designed prompts. Instead, it is likely that, as in every other area of popular culture, only a small number of possibilities are realized over and over by millions of users, leaving a long tail of other possibilities unrealized. So, if only a tiny fraction of the vast universe of potential AI artifacts is being realized in practice, we can't make broad statements about the originality or utility of the rest of the universe.

A Letter to a Young Artist

> "I have completed the construction of my burrow and it seems to be successful." ... [T]he most beautiful thing about my burrow is the stillness. Of course, that is deceptive. At any moment it may be shattered and then all will be over. For the time being, however, the silence is with me." (Franz Kafka, *The Burrow,* 1924)

The key difference between me, a human, and generative AI: I am limited, but AI is unlimited. Yes, of course: it has significant limits now, in practice. But it advances fast, and what it can already do today is beyond what we could have imagined a year or two ago. Instead of dwelling on what AI can't do at this particular moment, it is safer to assume that what it "can" will only multiply.

Because of how human skills, learning, and memory work, I have limitations. I can't draw in hundreds of styles of other artists or effortlessly combine them together. I don't have knowledge of the immense *museum without walls* distributed over the web and museum databases. But AI can. And it will only get better.

I can't simply sit down and start writing summaries of numerous topics in the history of culture. AI can. I can't instantly make hours of music that mixes the languages of different composers and map them into new instruments. AI can.

"I can't... but AI can." (Endless other examples can be added.)

So why make art now? And what art will still be meaningful to make?

What is interesting about human art now is our limits—and obsessions. Our inability to instantly think and paint exactly like any one of the millions of artists who lived. Our inability to quickly change. The way I walk, talk, my habits. My constraints. This is what makes me human as opposed to an AI. The latter will continue to evolve. But human evolution does not work on the same scale.

Note that this is not about simulating my idiosyncrasies and thus making AI "more human." Yes, we can do it, but that's not interesting. It is like taking a Boing 777 around the block to get groceries. Its forcing superhumans to act like humans, and this is a banal and weak strategy.

And there is another crucial point to make. *What makes art "human" is not our intentions, plans, ideas, or meanings. For over 100 years, modern artists did their* *best to remove all this from their art making.* If you give AI a direction, it can perfectly simulate ideas, plans, and meanings. So, this is not relevant.

The only relevant thing is our limitations. Our inability to compete with the superhuman. With the web, with search engines, with recommendation engines, with huge databases, with machine learning algorithms, with generative AI—and other superhuman computer technologies to come.

Therefore, "human artists making art with AI tools" is a meaningless idea. You want to collaborate with Gods? A mortal "collaborating" with Apollo, Athena, Hephaestus, Hermes, Zeus?

Instead, nurture your limitations. Be extremely limited—not unlimited. Don't be "creative." Forget the meaningless idea that AI will help us, "expand our creativity."

Work within constraints—the ones you already—or the ones you can make on purpose. White on white. Black on black. This is the right direction. Instead of a vast surface of "endless possibilities," concentrate in a single spot and go as deeply as possible.

(Think like Morandi rather than like Picasso.)

Make a tiny hole in the vast surface of everything that was already created and everything that is still possible and keep digging. When you get completely tired digging meters of wrong underground paths, get lost again and again, and want to give up, it means you are finally close to something. Keep going.

Because AI is so vast and endless in its knowledge and skills, you needed to work on the micro-scale. Very narrow. So narrow that AI can't quite get there. Through the needle eye. Only in this way can you compete with superhuman generative AI.

The artist needs to become a mole. And you need to be constantly stressed and worried because AI can discover your hole at any time and, in an instant, destroy all the underground pathways you have spent years making. But perhaps this stress, this endless anxiety is the right motivation for making something original and authentic in the end. Making your art in secret, knowing that you can be discovered and erased tomorrow by AI progress.

References

Manovich, L. (2020). *Cultural analytics*. The MIT Press.

Sutherland, Ivan. "Sketchpad: A Man-Machine Graphical Communication System." In *AFIPS '63: Proceedings of the Spring 1963 Joint Computer Conference*, 21 May 1963, https://dl.acm.org/doi/10.1145/1461551.1461591.

"Projects." *Cultural Analytics Lab*, accessed 27 Sept. 2024, https://lab.culturalanalytics.info/p/projects.html.

Vitz, P., & Glimcher, A. (1983). *Modern art and modern science: The parallel analysis of vision*.

Acknowledgement

This book represents the last step of a wider project proposed by the Centre of Research in Digital Humanities of Pegaso University and funded by the same university.

A board of scholars with interdisciplinary background worked together to discuss about the relation between generative engine and the production of creative outputs in music, audiovisual narratives, texts, and images.

Thanks to the reciprocal collaboration and the peculiar perspective each of them offered to the topic, it was possible to complete the project titled ACAI—*Artificial Creativity: Perceptions and Practices in Applying AI to Artistic and Cultural Productions* (PRA2023004).

I would also like to thank the publisher, who believed in this project without hesitation. A special thanks to all the authors for their efforts in writing and editing the papers. The richness of perspectives is evident by the table of contents, where we tried to identify four different threads about the impact and the application of AI: language, narratives, music production, and the multimedia outputs.

I intend to express my gratitude to all the members of the board of Artificial Creativity Projects that are not part of the publication but have done an important work behind the scene: Anna Rinaldin, Gilda Policastro, Lenoardo Galteri, Lara Balleri, Sara Selmi, Gaia Turconi, Federico Ambruosi, Francesco Epifani, and Fabio Pelagalli. We hope to share other academic adventures all together.

Editorial Board

Alessandra Micalizzi (Pegaso University)—Coordinator
Giuliano Cenati (Pegaso University)
Gilda Policastro (Pegaso University)
Anna Rinaldin (Pegaso University)
Fabio Pelagalli (NABA—University of Foggia)
Corrado Santoro (SAE Institute)
Teresa Savoia (Pegaso University)
Mirko Tavosanis (University of Pisa)
Massimiliano Zanon (Politecnico of Milan)

Contents

the US-based Anthropic,[3] the French Mistral,[4] the fertile open source research, and others). These models fall under the category of generative AI (GenAI), which refers to deep-learning models that, starting from "raw" data, learn to generate statistically probable outputs on demand (Martineau 2023). GenAI models generally encode a simplified representation of training data and use this encoded representation to generate output that is similar (but not identical) to the original. Generative AI can be opposed to Discriminative AI (e.g., classification and clustering models, such as facial recognition, or spam filtering), and presents unique characteristics and challenges to the user experience since its nature violates common Human-Computer Interaction (HCI) principle that states a system should respond consistently to user input (Weisz et al. 2023, 2024).

Nevertheless, GenAI is increasingly widespread and is set to change (or disrupt) various human practices. OpenAI's GPT-4o[5] and Google's Project Astra[6] promise to become overpowered and cross-modalities AI assistants for everyday activities, transforming, for example, the way we learn. Suno[7] and Udio[8] present a future "where anyone can make great music" and "anyone with a tune, some lyrics, or a funny idea can now express themselves in music", enabling the generation of songs from a text prompt. The same happens for image generation with the popular MidJourney,[9] OpenAI Dall-E,[10] and Adobe Firefly.[11]

There is a growing recognition that AI advancement should prioritize human needs to benefit humans first, without overshadowing human values, priorities, and lived experiences, and considering the numerous risks associated with this technology (Weidinger et al. 2022; Bender et al. 2021). This has invited researchers inside and outside HCI to research and design an AI that is Human-Centered (HCAI) (Xu 2019; Shneiderman 2022, 2020; Capel & Brereton 2023).

Capel and Brereton (2023) offer a review of HCAI research and give a comprehensive definition:

> Human-Centered Artificial Intelligence utilizes data to empower and enable its human users, while revealing its underlying values, biases, limitations, and the ethics of its data gathering and algorithms to foster ethical, interactive, and contestable use.

Whereas previous definitions of HCAI have left open who decides what is desirable for humans and who would benefit from AI, this definition focuses instead on ethical issues, bringing into consideration all those who might be impacted.

[3] https://www.anthropic.com.

[4] https://mistral.ai.

[5] https://openai.com/index/hello-gpt-4o.

[6] https://deepmind.google/technologies/gemini/project-astra/.

[7] https://suno.com.

[8] https://www.udio.com.

[9] https://www.midjourney.com.

[10] https://openai.com/index/dall-e-3.

[11] https://firefly.adobe.com.

In addition, the definition also emphasizes interaction, bringing considerations of actual use by end users to the forefront. Capel and Brereton in fact identify an emerging research area within HCAI involving interaction with AI, or Human-AI Interaction (HAII), that *"explicitly addresses the need to understand how people will interact with inferred models in embodied and situated contexts"*.

Following this approach, we will now explore the potential and challenges of GenAI within a specific domain and for a particular community of practice (Wenger 1999): music composition. This provides an illustrative example of the benefits and complexities of casting AI research and design in situated settings, but it is also particularly fitting since the social nature of music includes factors not taken into account by AI music models. These factors show the oversights of current AI research, which does not pay attention to the settings where the use of these systems occurs. Humans create music embedded in cultural contexts, motivated by social factors; AI systems, on the other hand, operate without these social dimensions, producing music only based on algorithms and predictive models, which risks limiting their impact by not fully reflecting human creative processes (Bown 2021).

In the next section, we present a critical review of some musical AI research and the use of generative systems in music composition. In particular, we focus on the role AI takes in human-AI interaction and the importance of context.

2 Making Music with AI

Around 1950, Hiller and Isaacson produced the first score composed with the creative input of algorithms: the Illiac Suite (Hiller & Isaacson 1958; Ji et al. 2023). According to Bown (2021), the approach used for the Illiac Suite has since marked the rest of the century: creative artists with advanced computer skills, programming computers for algorithmic composition tasks. Later, much academic research focused on transferring human tasks and decisions into the domain of the computer to make music creation more autonomous. The ultimate goal was to create fully autonomous creative machines (Bown 2021). Arguably, a significant portion of musical AI research can still be described in the same way.

A first major distinction that is worth making before proceeding is that between symbolic and audio music generation (Ji et al. 2023). Symbolic music generation has to do with representing, learning, and generating music as sequences of symbols. This representation usually consists of discrete sequences that contain musical elements, such as pitch and duration (e.g., MIDI-like events or ABC notation). In contrast, audio music generation models a continuous audio signal (examples now available to everyone are the previously mentioned Suno and Udio). We primarily focus on symbolic generation because it allows for fine-grained modifications compared to audio generation. Symbolic generation also affords more interactive uses, for example, involving Digital Audio Workstations and MIDI controllers, which may better align with our ultimate goal of facilitating human-AI interaction.

However, several considerations that we will make below may apply to audio generation.

Ji et al. (2023) offer a taxonomy of symbolic music generation through five categories:

1. *Generation from scratch*, or unconditioned generation. The main subtasks into which it is divided are melody generation, polyphonic generation, and multi-track generation.
2. *Conditional generation*, meaning generation of music conditioned by a specific input related to a musical context. It is subdivided into melody generation from chord progression, melody harmonization, accompaniment arrangement, and music inpainting.
3. *Controllable generation*, for example, by specifying the structure or attributes of the music. It is divided into style control, structure control, and sentiment control.
4. *Performance generation*, which is about incorporating expressive and dynamic elements typical of live performances. This can be the addition of performance characteristics to a given score, or the simultaneous creation of a musical score and its expressive performance characteristics.
5. *Interactive generation*, which is based on collaboration between humans and machines to create music together. It can be real-time cooperation between the human musician and the machine, where each plays in turn (call and response), or the generation of a musical accompaniment in response to the human performance.

However, we can envision how, depending on the user's intentions, each of these categories can take part in human-AI interaction systems for music composition: generation from scratch may be functional for a composer's initial process of free exploration; a guitarist may want to conditionally generate a rhythmic accompaniment out of a riff or may use inpainting (Hadjeres & Crestel 2021) to explore variations to a phrasing; controlling the sentiment of a chorus may help a composer explore variations based on their intentions—and so on.

2.1 Context

Here, we argue for the need to broaden the perspective on AI for music generation, move beyond a purely technical focus, and consider music as a complex social activity shaped by context.

Despite progress, deep learning techniques for symbolic music generation face challenges such as lack of musical novelty and structure, limited emotional expression, limited user interaction, and lack of standards for music evaluation (Ji et al. 2023; Hernandez-Olivan & Beltrán 2023). Among future directions, Ji et al. call for improving and refining how artificial intelligence interacts with humans with more fine-grained controls, and researching real-life scenario applications to promote social progress and development. The latter appeal is supported by Huang et al. (2020) who suggest future AI system design to follow a study of the practices currently in use and commonly adopted by music composers, in order to adapt to them instead of demanding the opposite. However, doing this through a purely technical focus may be difficult as it neglects the social and cultural dimensions of music-making.

Bown (2021) shows how music is a culturally-rooted activity and suggests that there is some friction between engineering and sociological perspectives on musical AI systems. An engineering perspective, such as the one that is dominant, has no problem removing the social context from the equation to focus on technical and quantifiable aspects such as note choice, arrangement strategies, or imitation of a musical style. Instead, it is likely that a sociological perspective, which pays more attention to the various aspects that define human practices (such as identities, preferences, motivations, values, and cultural contexts), may be necessary to develop AI that is better able to support (or replicate) typical human creative processes.

Newman et al. (2023) highlight various contextual factors that influence what creators believe to be a good or bad placement of AI within their creative process. These factors are: *personal context*, which means the person's abilities, familiarity with the creative process and musical literacy; *social context*, that is, the community and culture of reference, as well as musical training and exposure; and the *creative goal*, or the purpose for which the work is created, (e.g., a commission or performance). These aspects influence musicians' views on the role AI can play in the creative process, and thus how it will be used.

However, as Dourish (2004) points out, the notion of context is often misrepresented in engineering approaches. These approaches, inherited from a positivist tradition, tend to think that context consists of a set of features of the environment, that can be coded and made available to software. For Dourish this reflects a misunderstanding of the nature and role of contextuality in everyday life. Instead, context is an outcome, something that is made, not a mere description of a setting. Context is produced, maintained, and enacted during an activity. It is also both a relational property, occurring between objects or activities, and an occasioned property, particular to each occasion of an activity or action. Context is dynamically defined and cannot be determined in advance. Dourish goes on to suggest that what constitutes context is a matter related to *practice*, following the notion of Wenger (1999). Contextual properties assume meaning and relevance through forms of practice, and it is not possible to separate the meaning of technology from how people use it (and how they appropriate it (Dix 2007)) because when they do so, new meanings of that technology are created and communicated by people while their practice evolves (Dourish 2004).

Furthermore, as Small (1998) suggests, music is not merely a collection of artifacts, like records and songs, but a social activity through which participants explore their relationships and identities in relation to others. The focus should be not so much on musical artifacts like songs but on the actions of creating, perceiving, and responding to them, what Small calls *Musicking*. Viewing music as a social activity emphasizes the importance of understanding the social context in which music is created and experienced. This leads us to ask how designed musical AI systems can become meaningful within the situated practices of those involved in musicking processes, such as, in our case, composers.

2.2 AI-Roles

We have shown that, to design the interaction between humans and musical AI, we may need to consider not only technical and quantifiable aspects, but also user intentions, socio-cultural factors, the role of context (which cannot be reduced to a set of codifiable features), and to expand the idea of music from an outcome to a complex and deeply rooted practice. But what is the role of AI in this interaction?

In mapping Human-centered AI research, Capel and Brereton (2023) discuss what they name Human-AI Teaming, a strand of research that posits that by working together both AI and people can perform better and increase their capabilities. Unlike Human-in-the-loop approaches or level-3 automation of autonomous vehicles,[12] the human is not just there to do what the machine cannot do on its own, but there is a synergistic collaboration between the two. This idea of collaborative teams of humans and AI is getting significant traction in and out of HCI research (e.g., Wang et al. 2020; Seeber et al. 2020; McComb et al. 2023; Wu et al. 2021; Capel and Brereton 2023) and can be traced back as far as Licklider's idea of symbiosis between humans and computers (Licklider 1960). Also, some argue for a paradigm shift in HCI, moving from designing human-computer interactions to human-computer integration and teaming (Xu 2019; Farooq & Grudin 2016; Rapp 2023).

This formulation often stands in contrast to the idea of AI as a tool. Seeber et al. (2020) develop a research agenda around the idea of machine teammates collaborating with humans and, while addressing several critical issues, their intent is clearly to move beyond the idea of tool use in favor of collaboration with autonomous machines: *"We propose that AI will not (just) be the functionality of a tool but rather a machine teammate characterized by a high level of autonomy, based on superior knowledge processing capabilities, sensing capabilities, and natural language interaction with humans"* (Seeber et al. 2020, p. 9). McComb et al. (2023) present a 2x2 matrix to isolate "different AI capabilities in teaming", which crosses two key dimensions: mode (reactive vs. proactive) and focus (problem orientation vs. process orientation). This matrix describes four quadrants of possible AI systems: AI-as-Analytics, AI-as-Tool, AI-as-Partner, and AI-as-Guide. McComb et al. explain how *"AI agents will work not just as tools but also as members of a team"*, while showing a clear preference for the last two "archetypes".

On the opposite side, Sarkar (2023) speaks of an "agentistic turn" to denote this attribution of agency to AI systems and warns that it could obscure the large

[12] https://www.sae.org/blog/sae-j3016-update.

amount of labor that fuels AI, often conducted in the Global South at low wages by, for example, data labelers. Sarkar calls for abandoning the idea of partnership and collaboration with AI in favor of a "more equitable" position of AI as a tool. However, critiques of human-AI teaming go beyond just the choice of terminology and metaphors (which, however, leaning on the idea of generative metaphors by Agre 1997, pp. 33–48, can embody the questions that are relevant to a field of research and the methods useful for answering them, bringing some phenomena to the center of inquiry and marginalizing others). According to Cabitza et al. (2021), the dominant HAII approach, which they term "agential AI", views AI systems as autonomous agents interacting with humans in a dyad. In this paradigm, intelligence and agency are attributed to both humans and machines as separate entities that collaborate or compete. Cabitza et al. discuss how there are intrinsic issues in the dyadic HAII paradigm, arising from neglecting the relational, collaborative, and contextual aspects typical of real-world decision-making processes, and from a cognitivist view that conceives of AI systems as autonomous agents, which can trigger negative bias such as *automation bias*, *automation complacency* or conversely *"prejudices against the machine"*. These biases can in turn have consequences of deskilling, avoidance of responsibility, and lack of vigilance due to excessive complacency. For Cabitza et al., this suggests the need to treat AI differently and they propose an alternative approach based on the conception of AI as a component of a Knowledge Artifact, i.e., an ecosystem of tools aimed at supporting a collective of competent decision-makers working collaboratively (which, with some simplification, we might call "community of practice" Wenger 1999).

The work of Muller and Weisz (2022) addresses some of the concerns raised by Cabitza et al. and Sarkar by extending the framework of Shneiderman (2022), which already uncoupled human initiative from AI initiative by representing them as independent dimensions. Muller and Weisz provide a way to better analyze the interactions between humans and AI in organizational contexts. First, whereas Shneiderman analyzed AI applications as static points in the two-dimensional space of human and AI initiative, they propose instead to analyze them as dynamic sequences of steps, in which initiative can shift within the same application. It thus emerges that initiative and control change dynamically during use. This analysis at the level of individual steps makes it possible to identify points where there is an imbalance or discrepancy between desired and actual initiative, whether on the human or AI side. Second, they observe that many AI applications in organizational contexts involve not just a single user, but a plurality of human stakeholders in different roles, and show how, by making them explicit, the power relations, interests, and values of these different stakeholders can be analyzed, beyond just the end user interacting directly with the application.

Returning to music AI systems, another perspective is that of Gioti (2021), who draws on Latour and Gell's ideas on agency to propose a vision in which AI is not a substitute for human creativity but is an agent that contributes to "distributed human-computer co-creativity." Gioti takes interactive music systems as an example of systems in which distributed agency between human and non-human actors is most evident and in which both the creative process and authorship are distributed among them. More specifically, Gioti notes how there is a negotiation between human creative intentions and the specific way in which technological artifacts are designed, their *technological directionality*, i.e. their scope, and suggests that the relationship between human intentionality and technological directionality can be improved with AI.

However, the balance between automation, agency distribution, and control is delicate. Newman et al. (2023) show how critical it is for music creators to maintain control, agency, intention, and choice and that this affects the roles granted or allowed to AI by the users. Louie et al. (2020) study the effect of steering tools for AI music models and observe how the need for control can change dynamically depending on the user's mental state and creative goal. During exploratory phases users seek inspiration, even the most unexpected, and may be willing to cede control. However, during other phases, such as when the user wants to focus on details, more control is needed. We can link this to what Muller and Weisz (2022) said about the *"dynamic shifts and exchanges of human and machine initiative"*, and with the need to better understand how context plays a role, as we argued above.

Finally, Suh et al. (2021) investigate the role of AI in human-human collaboration for music creation and its impact on interpersonal dynamics. Their study concludes that AI can help ease the underlying tensions often present in creative collaborations by facilitating the flow of ideas and group cohesion. However, the researchers also observed a shift in human creative and collaborative roles: participants reported feeling more like curators or co-producers rather than co-composers, as they primarily focused on evaluating AI-generated output rather than directly developing ideas. This resulted in a patchwork of AI-produced creative ideas, leading to lower creative involvement.

2.3 Researching Music Composition as a Situated Practice

Some new perspectives suggest innovative research directions that challenge the current AI paradigm. For instance, Bown (2021) hypothesizes the creation of machine learning algorithms that embody human-like musical behaviors. This would involve training AI models not only on musical content but also on the context, associations, and cultural meanings of the musical corpus used, and using search-based and evaluative models along with predictive ones. Rohrmeier (2022), on the other hand, highlights how musical creativity actually requires general artificial intelligence, going far beyond mere replication or generation of musical outcomes within limited domains. This is considered an AI-complete problem, meaning we would need models for musical cognition, for the external world, a model of bodies, instruments, musical interaction and performance, and a model of meta-creativity. As one might guess, these paradigm shifts would not only be about the architectures used, but also about a profound epistemological and philosophical movement in the field of AI research.

However, one can wonder what can be done today, with available technologies, to design a human-AI interaction more in line with the needs of music makers and the ways in which they work. Human practices and socio-cultural activities such as musicking are already established and are proven to hold values for human learning and identity creation. We, therefore, take on a different approach: instead of overriding them in favor of autonomous musical agents, we as designers, HCI practitioners, and AI researchers should aim to assist and empower these activities, whether through AI-imbued tools, AI assistants, or AI as collaborators. Yet, ignoring the importance of human intentionality, the need for control and agency, and the importance of context can be detrimental to the design of human-AI interactive music systems that are truly human-centered. In the next section, we propose and discuss a study that might respond to these concerns, and present some preliminary results.

3 An Ethnographic Study

In this section, we propose an ethnographic approach that foregrounds the situated nature of music composition. Ethnography is a well-known method within HCI research (e.g., Rapp 2021) and has also been applied to AI research (Marda & Narayan 2021; Christin 2020; Blackwell 2021; Van Voorst & Ahlin 2024; Seaver 2017), feeding into a broader conversation on the lack of, and thus the need to integrate the social sciences into AI research to mitigate the exclusive use of quantitative methods. In fact, the uncritical and positivist use of quantitative techniques can lead to ignoring the context and causes of certain outcomes and how they occur (Sloane & Moss 2019; Marda & Narayan 2021; Dahlin 2021).

Existing ethnographic and autoethnographic studies have also contributed to the intersection of music and AI. For example, McGarry et al. (2021) conducted an ethnographic study focusing on the workflow of two professional music producers as they worked on a Digital Audio Workstation. The study highlighted the importance of metadata and the organization of audio resources in coordinating the creative process. The authors suggested the potential use of AI for automating process tracking and documenting song data provenance. Nicholls et al. (2018) take an autoethnographic approach to examine how two musicians interact during songwriting and create a model of the collaborative process. The data they gather would serve to inform a collaborative AI system that works with musicians. Noel-Hirst and Bryan-Kinns (2023) provided an autoethnographic exploration of explainable AI (XAI) models in music composition. They show how the use of XAI models can influence the musical creative process through unexpected uses.

Previous studies have primarily focused on specific aspects of music production or the adoption of AI tools by composers. However, as we portrayed in broad strokes in Sect. 2, there is a need to investigate music composition as a situated practice, considering composers' personal motivations, and artistic sensibilities, and foregrounding the broader socio-cultural context that influences their work. Moreover, studying how composers' intentions and needs influence the compositional process, and how control is distributed in human-human collaboration, might illuminate which role is best suited for the AI in music composition practice (and when).

We propose an ethnographic study, which will involve 18 interviews and approximately 60 hours of participant observation. Our goal is to use the findings to better inform the design of human-AI interactive systems. These systems should support and empower music composers to explore new musical possibilities, rather than supplant their creative abilities.

Combining interviews with participant observation will offer an in-depth look at the compositional process and the socio-cultural setting in which it is embedded, providing insights into how AI can support and enhance composers' practice. Interviews will help us understand the goals and motivations of composers, as well as the socio-cultural context that influences them, and make them reflect on their creative process. Participant observation will allow us to see the creative process in action in real-life situations, peek into tacit knowledge, and check whether what emerged during the interviews is reflected in daily practice. The study is not yet completed, but we can present three preliminary findings from the first round of interviews.

3.1 Preliminary Findings

3.1.1 Intentionality Guides Composition

Intentionality, which can be described as the "about-ness" of our actions (Dourish 1999), is central to music composition, guiding both specific compositional strategies and the overall process of search and evaluation of creative outcomes. Whether composers start with a melody, harmony, or specific timbre often depends on what they intend to convey. Moreover, these intentions are often negotiated directly or indirectly in relation to the intended use (e.g., a commercial, a soundtrack, an album track) and with other stakeholders involved in the process.

As we discussed before, the existing literature acknowledges the importance of maintaining control and creative agency and the dynamic nature of initiative in HAII. However, in our study, we observed a specific connection between intentions and a process of meaning-making. Composers seek to construct a "coherent discourse" in the musical piece through the process of composition. Moreover, as one participant pointed out, music composition is not about striving for a perfect form, but rather about "moving closer" to what you create, and "making your own what you create". The use of GenAI may risk inhibiting this process if the negotiation of control and initiative is not well managed, and if the system lacks adaptivity toward the composer's evolving intentions.

The ability of these systems to flexibly accommodate composers' intentionality and their process of meaning-making can be essential to effectively support their creative practice. We might also consider how to support processes of appropriation

(Dix 2007), by designing for customization and configurability of musical AI systems, making their intent clear (or in the words of Gioti 2021: their technological directionality), and not forcing the user into overly narrow workflows, providing instead freedom of interaction by *"offering the user a myriad of ways to achieve a product's functionality"* (Wensveen et al. 2004).

3.1.2 Music Creation Is a Collaborative Process

Music creation is often a collaborative process, involving bandmates, arrangers, clients, and various stakeholders who contribute to and have an interest in the final product. For example, we observed how some compositional choices are extended to the recording studio, where they are negotiated, discussed, or delegated to the sound engineer.

This raises the question of whether AI systems should be designed to enhance these existing human-human interactions, or whether they should become an additional collaborator within a system where creative control is already dynamically distributed and negotiated.

As Cabitza et al. (2021) argued, the dominant "agential AI" approach might neglect the relational, collaborative, and contextual aspects typical of real-world decision-making, but the work of Muller and Weisz (2022) can help show how human-AI interaction can involve a plurality of stakeholders. Through the perspective of Value Sensitive Design (Friedman & Hendry 2019) the interests and values of these different stakeholders can be analyzed and reveal hidden roles and dynamics, such as indirect stakeholders who influence the use of the application (Muller & Weisz 2022).

In addition, we must also be wary of how human roles change as we introduce GenAI into compositional practice, as this can lead to a loss of control and ownership, and reduced creative engagement, as shown by Suh et al. (2021)

3.1.3 An Unstructured Creation Process

One participant described how initial musical ideas often emerge from "messing around", improvising, and even making mistakes, suggesting that creativity can often come from spontaneous and less structured activities. The creative process is not always linear or purposeful but can be serendipitous and emergent. The boundaries between rehearsal, practice, composition, recording, and leisure playing are often blurry.

Designing AI systems that sit at the boundary of these contexts and flexibly adapt to and support these emergent, blurred, and serendipitous aspects of the compositional process will be crucial for meaningful human-AI interactive systems for music creation.

4 Conclusion

In this paper, we discussed human-AI interaction through the lens of music composition with the help of GenAI. We highlighted that music is a contextual and socially situated activity and how this poses a limit in current AI research. We discussed different paradigms for the role AI could take in HAII, and that span from AI as a tool to AI as a teammate, and we tied all of this together in the recognition that we need to view music composition as a situated practice in order to design HAII systems that are human-centered and support human composers.

We then briefly presented an ongoing ethnographic study investigating music composition through interviews and participant observations, to inform the design of interactive AI systems that can meaningfully support composers. We pointed out the need to design AI systems that flexibly accommodate composers' evolving intentions rather than forcing rigid workflows, take into account that music composition is a collaborative process involving multiple stakeholders, and account for a creative process that can be unstructured, emergent, and serendipitous.

By foregrounding the situated nature of music composition, in this paper we emphasized the need to move beyond a purely technical focus on AI capabilities, and instead design human-AI interactive systems that are better aligned with human goals, practices, and socio-cultural contexts.

References

Agre, P. E. (1997). *Computation and human experience* (1st edn.). Cambridge University Press. https://doi.org/10.1017/CBO9780511571169

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). ACM, Virtual Event Canada. https://doi.org/10.1145/3442188.3445922

Blackwell, A. F. (2021). Ethnographic artificial intelligence. *Interdisciplinary Science Reviews, 46*(1–2), 198–211. https://doi.org/10.1080/03080188.2020.1840226

Bown, O. (2021). Sociocultural and design perspectives on AI-based music production: Why do we make music and what changes if AI makes it for us? In E. R. Miranda (Ed.), *Handbook of artificial Intelligence for music* (pp. 1–20). Springer. https://doi.org/10.1007/978-3-030-72116-9_1

Cabitza, F., Campagner, A., & Simone, C. (2021). The need to move away from agential-AI: Empirical investigations, useful concepts and open issues. *International Journal of Human-Computer Studies, 155*, 102696. https://doi.org/10.1016/j.ijhcs.2021.102696

Capel, T., & Brereton, M. (2023). What is human-centered about human-centered AI? A map of the research landscape. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Hamburg, Germany, CHI'23 (p. 23). https://doi.org/10.1145/3544548.3580959

Christin, A. (2020). The ethnographer and the algorithm: Beyond the black box. *Theory and Society, 49*(5–6), 897–918. https://doi.org/10.1007/s11186-020-09411-3

Dahlin, E. (2021). Mind the gap! On the future of AI research. *Humanities and Social Sciences Communications, 8*(1), 1–4. https://doi.org/10.1057/s41599-021-00750-9

Dix, A. (2007). Designing for appropriation. In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI...But Not as We Know It - Volume 2*, BCS Learning & Development Ltd., Swindon, GBR, BCS-HCI'07 (pp. 27–30). https://doi.org/10.5555/1531407.1531415

Dourish, P. (1999). *Embodied Interaction: Exploring the Foundations of a New Approach to HCI*. https://www.dourish.com/publications/misc/embodied.pdf. Accessed 17 September 2024

Dourish, P. (2004). What we talk about when we talk about context. *Personal and Ubiquitous Computing, 8*(1), 19–30. https://doi.org/10.1007/s00779-003-0253-8

Farooq, U., & Grudin, J. (2016). Human-computer integration. *Interactions, 23*(6), 26–32. https://doi.org/10.1145/3001896

Friedman, B., & Hendry, D. F. (2019). *Value sensitive design: Shaping technology with moral imagination*. The MIT Press.

Gioti, A. M. (2021). Artificial intelligence for music composition. In E. R. Miranda (Ed.), *Handbook of artificial intelligence for music: Foundations, advanced approaches, and developments for creativity* (pp. 53–73). Springer. https://doi.org/10.1007/978-3-030-72116-9_3

Hadjeres, G., & Crestel, L. (2021). The piano inpainting application. https://doi.org/10.48550/ARXIV.2107.05944

Hernandez-Olivan, C., & Beltrán, J. R. (2023). Music composition with deep learning: A review. In A. Biswas, E. Wennekes, A. Wieczorkowska, R. H. Laskar (Eds.), *Advances in speech and music technology: Computational aspects and applications, signals and communication technology* (pp. 25–50). Springer. https://doi.org/10.1007/978-3-031-18444-4_2

Hiller, L. A., Jr., & Isaacson, L. M, (1958). Musical composition with a high-speed digital computer. *Journal of the Audio Engineering Society, 6*(3), 154–160.

Huang, C. Z. A., Koops, H. V., & Newton-Rex, E. (2020). AI song contest: Human-AI co-creation in songwriting. In *Proceedings of the 21st International Society for Music Information Retrieval Conference*. International Society for Music Information Retrieval (pp. 708–716).

Ji, S., Yang, X., & Luo, J. (2023). A survey on deep learning for symbolic music generation: Representations, algorithms, evaluations, and challenges. *ACM Computing Surveys, 56*(1), 1–39. https://doi.org/10.1145/3597493

Licklider, J. C. R. (1960). Man-computer symbiosis. *IRE Transactions on Human Factors in Electronics HFE-1*(1), 4–11. https://doi.org/10.1109/THFE2.1960.4503259

Louie, R., Coenen, A., Huang, C. Z., Terry, M., & Cai, C. J. (2020). Novice-AI music co-creation via AI-steering tools for deep generative models. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Honolulu, CHI'20 (pp. 1–13). https://doi.org/10.1145/3313831.3376739

Marda, V., & Narayan, S. (2021). On the importance of ethnographic methods in AI research. *Nature Machine Intelligence, 3*(3), 187–189. https://doi.org/10.1038/s42256-021-00323-0

Martineau, K. (2023). What is generative AI? https://research.ibm.com/blog/what-is-generative-AI

McComb, C., Boatwright, P., & Cagan, J. (2023). Focus and modality: Defining a roadmap to future AI-human teaming in design. *Proceedings of the Design Society, 3*, 1905–1914. https://doi.org/10.1017/pds.2023.191

McGarry, G., Chamberlain, A., Crabtree, A., & Greehalgh, C. (2021). The meaning in "the mix": Using ethnography to inform the design of intelligent tools in the context of music production. In *Audio mostly 2021* (pp. 40–47). ACM. https://doi.org/10.1145/3478384.3478406

Muller, M., & Weisz, J. (2022). Extending a human-AI collaboration framework with dynamism and sociality. In *2022 Symposium on Human-Computer Interaction for Work* (pp. 1–12). ACM. https://doi.org/10.1145/3533406.3533407

Newman, M., Morris, L., & Lee, J. H. (2023). Human-AI music creation: Understanding the perceptions and experiences of music creators for ethical and productive collaboration. In *Proceedings of the 24th International Society for Music Information Retrieval Conference* (pp. 80–88). International Society for Music Information Retrieval.

Nicholls, S., Cunningham, S., & Picking, R. (2018). Collaborative artificial intelligence in music production. In *Proceedings of the Audio Mostly 2018 on Sound in Immersion and Emotion* (pp. 1–4). ACM. https://doi.org/10.1145/3243274.3243311

Noel-Hirst, A., & Bryan-Kinns, N. (2023). An autoethnographic exploration of XAI in algorithmic composition. https://doi.org/10.48550/ARXIV.2308.06089

Rapp, A. (2021). In search for design elements: A new perspective for employing ethnography in human-computer interaction design research. *International Journal of Human–Computer Interaction, 37*(8), 783–802. https://doi.org/10.1080/10447318.2020.1843296

Rapp, A. (2023). Wearable technologies as extensions: A postphenomenological framework and its design implications. *Human–Computer Interaction, 38*(2), 79–117. https://doi.org/10.1080/07370024.2021.1927039

Rohrmeier, M. (2022). On creativity, music's AI completeness, and four challenges for artificial musical creativity. *Transactions of the International Society for Music Information Retrieval, 5*(1), 50–66. https://doi.org/10.5334/tismir.104

Sarkar, A. (2023). Enough with "human-AI collaboration". In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Hamburg, CHI EA'23 (p. 8). https://doi.org/10.1145/3544549.3582735

Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society, 4*(2), 1–12. https://doi.org/10.1177/2053951717738104

Seeber, I., Bittner, E., Briggs, R. O., De Vreede, T., De Vreede, G. J., Elkins, A., Maier, R., Merz, A. B., Oeste-Reiß, S., Randrup, N., Schwabe, G., & Söllner, M. (2020). Machines as teammates: A research agenda on AI in team collaboration. *Information & Management, 57*(2), 103174. https://doi.org/10.1016/j.im.2019.103174

Shneiderman, B. (2020). Bridging the gap between ethics and practice: Guidelines for reliable, safe, and trustworthy human-centered AI systems. *ACM Transactions on Interactive Intelligent Systems, 10*(4), 26:1–26:31. https://doi.org/10.1145/3419764

Shneiderman, B. (2022). *Human-centered AI*. Oxford University Press.

Sloane, M., & Moss, E. (2019). AI's social sciences deficit. *Nature Machine Intelligence, 1*(8), 330–331. https://doi.org/10.1038/s42256-019-0084-6

Small, C. (1998). *Musicking: The meanings of performing and listening*. Music/Culture, University Press of New England.

Suh, M. M., Youngblom, E., Terry, M., & Cai, C. J. (2021). AI as social glue: Uncovering the roles of deep generative AI during social music composition. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama, Japan, CHI'21 (p. 11). https://doi.org/10.1145/3411764.3445219

Van Voorst, R., & Ahlin, T. (2024). Key points for an ethnography of AI: An approach towards crucial data. *Humanities and Social Sciences Communications, 11*(1), 337. https://doi.org/10.1057/s41599-024-02854-4

Wang, D., Churchill, E., Maes, P., Fan, X., Shneiderman, B., Shi, Y., & Wang, Q. (2020). From human-human collaboration to human-AI collaboration: Designing AI systems that can work together with people. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–6). ACM. https://doi.org/10.1145/3334480.3381069

Weidinger, L., Uesato, J., Rauh, M., Griffin, C., Huang, P. S., Mellor, J., Glaese, A., Cheng, M., Balle, B., Kasirzadeh, A., Biles, C., Brown, S., Kenton, Z., Hawkins, W., Stepleton, T., Birhane, A., Hendricks, L. A., Rimell, L., Isaac, W., Haas, J., Legassick, S., Irving, G., & Gabriel, I. (2022). Taxonomy of risks posed by language models. In *2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 214–229). ACM. https://doi.org/10.1145/3531146.3533088

Weisz, J. D., Muller, M., He, J., & Houde, S. (2023). Toward general design principles for generative AI applications. https://doi.org/10.48550/arXiv.2301.05578

Weisz, J. D., He, J., Muller, M., Hoefer, G., Miles, R., & Geyer, W. (2024). Design principles for generative AI applications. https://doi.org/10.1145/3613904.3642466

Wenger, E. (1999). *Communities of practice: Learning, meaning, and identity* (first paperback edn.). Cambridge University Press.

Wensveen, S. A. G., Djajadiningrat, J. P., & Overbeeke, C. J. (2004). Interaction frogger: A design framework to couple action and function through feedback and feedforward. In *Proceedings of the 2004 Conference on Designing Interactive Systems Processes, Practices, Methods, and Techniques - DIS'04* (p. 177). ACM. https://doi.org/10.1145/1013115.1013140

Wu, Z., Ji, D., Yu, K., Zeng, X., Wu, D., & Shidujaman, M. (2021). AI creativity and the human-AI co-creation model. In M. Kurosu (Ed.), *Human-computer interaction. Theory, methods and tools*. Lecture Notes in Computer Science (vol. 12762, pp. 171–190,). Springer. https://doi.org/10.1007/978-3-030-78462-1_13

Xu, W. (2019). Toward human-centered AI: A perspective from human-computer interaction. *Interactions, 26*(4), 42–46. https://doi.org/10.1145/3328485

Recognition as 'Art' by Publics: How Generative AI Music Production Is Perceived

Alessandra Micalizzi

Abstract This study explores the intersection of artificial intelligence (AI) and human creativity by investigating how the perceived authorship of a musical piece influences its evaluation. Unlike previous studies that have focused on comparing the outputs of human and AI creators, this research delves into the perceptual aspects of AI-generated art. Through a carefully designed social experiment, we examined how participants rate the quality of music pieces without and with knowledge of their authorship. Our findings reveal a nuanced relationship between perceived authorship and aesthetic judgment, suggesting that while human biases can influence evaluations, the intrinsic qualities of a piece ultimately play a significant role. This research contributes to the growing body of work on AI and creativity by offering novel insights into the human perception of AI-generated art and the potential for human–AI collaboration in creative endeavours.

Keywords Computational creativity · Music production · AI · Artificial creativity · Computational artist

1 Computational Creativity and the Effect on Cultural Production: An Overview

The debate about the use of AI in creative production has been especially relevant in recent years, though it should be mentioned that artistic language has always been linked to the evolutions of techniques, tools, and technologies discovered and applied over time. In this context, the development of the cultural industries and their world of production has given an important push to the contamination between art and technologies.

A. Micalizzi (✉)
Centro Direzionale ISOLA, Pegaso University, Naples, Italy
e-mail: alessandra.micalizzi@unipegaso.it; a.micalizzi@sae.edu

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

A. Micalizzi (ed.), *Artificial Creativity*, Springer Proceedings in Humanities and Social Sciences, https://doi.org/10.1007/978-3-031-88334-7_12

As McCormack (2003) argued, innovation today generally ‘is not achieved within the precious bubble of fine art, but by those who work in the industries of popular culture–computer graphics, film, music videos, games, robotics and the Internet’ (McCormack, 2003, p. 5). This contamination started in the late 1950s due to the contributions of cybernetics and other informational sciences. In other words, artistic production has become increasingly based on communication and information processing technologies that have significantly expanded in power and diversity over the past 50 years (Boden & Edmonds, 2009). The terminology used to describe these emerging art forms is diverse and has yet to solidify into a widely accepted classification. Artists working in these areas tend to favour terms such as generative art, computer art, digital art, computational art, process-based art, electronic art, software art, technological art, and telematics.

In an interview, Catodo (Math in the Air, 2019) defines the computational artist as someone who uses ‘computational language to create his/her own artwork, that means creating an algorithm and thus the writing of a computer program’. In this sense, the word ‘generative’ refers to the use of a system, which may be automated, to produce a work of art. He specifies that ‘Computer programming is one of many possible systems that can be used in generative art, but it is not the only one’. For example, John Cage (1912–1992), the famous American composer, employed generative techniques in his works of aleatoric music by using proportions derived from the I Ching. Another renowned musician, Brian Eno, was the first to use the term ‘generative music’ to describe his compositions created by introducing delays in audio recording systems (ivi, n.d.).

Generative art is a realm of digital art practice that has experienced a boom since the start of the twenty-first century (Galanter, 2016); it is not to be confused with the wider concept of digital art (Paul, 2023). To offer a definition, we can say that computational arts are based on the contamination of languages, texts, and techniques, and they are deeply influenced by the contribution of digital computing.

Over the decades, here have always been supportive and enthusiastic voices opposing the (often large) groups of detractors (e.g., D’Isa, 2024) regarding experiments with particular technologies. The substantial difference between AI and its technological precursors lies in how its generative capacity operates: the new generation of AIs has the capacity to become more autonomous in the production process and can make decisions independently. We are talking about the so-called generative artificial intelligence that can be distinguished as generative adversarial networks (GANs), which are restricted to reproducing a specific artistic style, and creative adversarial networks (CANs), which can deviate from the learned style, thus facilitating the creation of new and potentially innovative works (Liu, 2023).

The increasing integration of AI tools in artistic production has pushed the debate to a crucial point: the recognition of these cultural products’ value (Mazzone & Elgammal, 2019). There are several critical aspects at work; part of the artistic world continues to interrogate the impact of AI on the concept, status, and role of art, as well as how these new methods of (re)producing reality relate to our collective imagery (Kalpokas, 2023).

Zeilinger (2021) argued that AI ‘has the potential to reshape the aesthetic, cultural, and socio-economic dimensions of creativity’, fundamentally destabilizing the traditional notion of the author. He suggested that AI-generated art is ‘based on Big Data, which is the most social thing we have’ and that generative artistic productions are the result of ‘an ongoing composition in which humans and non-humans participate’.

D’Isa (2024) agrees with this integrated vision and, above all, with the possibility of abandoning an old-fashioned definition of authorship, which he asserts is more relevant to market issues than to artist evolution and history. More specifically, his position can be synthetized in the following statement: ‘I rather think that the shock brought about by innovation has momentarily disrupted an interpretative habit, reminding us that it is not we who create, but the world itself’ (ivi, 2024 p. 145). The interpretative habit he refers to deals with the obsession of explicitly recognizing an author and thus ownership of the intellectual work, which undermines the Creativity 4.0 system (Csikszentmihalyi, 2014; Gruner & Csikszentmihalyi, 2018), especially when a technology like AI intervenes in systemic dynamics.

AI-generated art is never entirely detached from human experience; indeed, it is deeply intertwined with human expression and our perceptions of the world. Since AI creativity is data-driven, its primary function is to reorganize information in ways that are both novel and recognizable to human audiences. Although the execution of creativity is machinic, it is not derived from any intrinsic machine-based aesthetics (Oksanen et al., 2023). As a result, the aesthetic appeal of AI-generated art remains—at least for now—oriented around human sensitivities, both in terms of the learning process and the intended audience (e.g., Manovich, 2018; Manovich & Arielli, 2021, 2024).

An integration of human intentionality with technological practices might define a new landscape in which data, human beings, and AI collaborate. In this context, the incorporation of AI brings a distinctly posthuman dimension to the triangular interaction of humans, data, and AI. However, the debate is still open. Some scholars argue that ‘the notion that a machine learning system could produce art or be regarded as an artist is implausible, as an artist cannot be reduced to a mere machine with intent’. Other scholars extend Duchamp’s (1957) concept of art to include AI-generated outputs, to the extent that the action performed by machine learning can be seen as a process of selection. In this sense, we support Hertzmann’s (2018) view of art as a form of interaction, where AI serves as a medium—a co-protagonist in new forms of creation shaped by artists, technology, and society.

1.1 Artificial Music Production: Art or Consumer Products?

Several studies have focused on AI’s impact on artistic products, whether in terms of quality, production processes, or the roles of creators (Lee, 2024; Celis Bueno et al., 2024; Zhou & Lee, 2024). Other scholars have focused their attention on user perceptions of AI-generated outputs. The ongoing debate ranges from apocalyptic views portraying AI as a threat to creativity and artistic production to more optimistic positions that foresee a future of integration between human intent and AI capabilities (Tubadji et al., 2021; Millet et al., 2023).

In the domain of artistic production, the application of AI has expanded to include intangible outputs such as music. There has been substantial growth in research on both human and computational creativity in music in recent years, especially within the field of computer science. Consequently, experimental studies exploring public perceptions of AI-generated or AI-assisted music have proliferated, aiming to map different facets of this revolutionary and rapidly expanding phenomenon of computational creativity. Some studies have specifically examined biases related to authorship when evaluating musical works that were created or co-created by AI (Tigre Moura et al., 2023; Horton Jr et al., 2023).

Several studies have documented a growing acceptance of AI-generated or AI-assisted products. For instance, Elgammal et al. (2017) demonstrated that AI-generated artworks are often indistinguishable from those created by humans. Their findings were corroborated by Hitsuwari et al. (2023). Furthermore, Elgammal et al. (2017) stated that AI-generated works frequently outperformed human-made pieces in terms of perceived novelty, complexity, intentionality, and inspiration. An interesting example of the application of AI in music production and public perception comes from Hadjeres et al. (2017). They used the generative model DeepBach to produce new compositions in the style of Johann Sebastian Bach, and nearly half of the participants in their experiment believed the machine-generated compositions to be original Bach works. However, participants with more musical experience were less likely (around 40%) to attribute the compositions to Bach. These findings have been supported by studies including Hong and Curran (2019) and Tigre Moura and Maw (2021), which revealed that despite initial scepticism towards AI-generated music, awareness of AI's role had minimal effect on respondents' overall perceptions of the compositions.

This brief, inexhaustive overview highlights a peculiarity regarding public perceptions. Listeners often struggle to distinguish authorship in musical productions, particularly as AI becomes more adept at replicating melodies, rhythms, and sound quality. However, even when they recognize the high quality of AI-generated music, listeners may not necessarily prefer it due to the prevailing belief that 'true' creativity is an inherently human trait (Rohrmeier, 2022). Moreover, there is a pervasive instinct to 'defend' human creativity against the perceived threat of AI supplanting it (Tubadji et al., 2021; Millet et al., 2023).

It is thus evident that the role of the public in the recognition of artwork is crucial. In our opinion, it is important to investigate which psycho-social and socio-cultural mechanisms influence the evaluation of an artistic production—in our case, a sound production—to fully understand the role that AI plays in user recognition of 'artistic value' in creative output and identify the main drivers that can orient this process.

2 Methodology

The results presented in this chapter are part of a wider study aimed at reconstructing the prejudice level of people—experts and everyday users alike—on creative AI-generated products (CAIGP). More specifically, we wanted to understand whether and to what extent users were influenced by the authorship of the creative product when evaluating it. In detail, our goals entail the following: (a) to reconstruct the level of knowledge around the general theme of computational creativity; (b) to identify which aspects affect the definition of creative and artistic products; and (c) to reconstruct components of public prejudice toward CAIGP with a focus on image production and soundtracks.

The research plan is based on mixed methods (Micalizzi & Lelicanin, 2023) and it is organized in three main stages:

- First, we analysed the imagery of AI described and contained in a sample of early-twentieth-century literary production.
- Second, we conducted a social experiment comparing AI-produced soundtracks with one produced by a young composer.
- Third, we collected 2500 CAWI questionnaires (Boreham & Wijnant, 2013) from a representative sample of the Italian population with a specific section addressing the evaluation of visual stimuli (most of which were produced by AI and only one by a human creator).

In this paper, we present the main results of the second stage—the part of the research that combined an experiment with in-depth interviews to verify whether there was a real bias towards the use of AI as a creative tool, as shown by previous studies (e.g., Latikka et al., 2023; Magni et al., 2024), and what factors could influence it. We tried to pay specific attention to variables such as the respondents' gender, artistic skills, and, even more specifically, audio skills.

2.1 The Experiment Design

The experiment was based on the participant listening to two tracks—one produced by AI and one composed by a young artist. The participants were divided into three groups:

- Those who were correctly informed about the authorship of the tracks (well-informed)
- Those who were poorly informed about the authorship of the tracks (misinformed)
- Those who were not informed at all (uninformed)

The participants from the first two groups were first subjected to an in-depth interview on the topic of computational creativity to ascertain their level of knowledge

Fig. 1 Evaluation scale of the listening section

on the topic. The third group (uninformed) did not receive any introduction before listening. Then, all three groups participated in an in-depth talk about the evaluation.

The participants evaluated the songs according to five specific criteria (emotional impact, melody, originality, tone, and general evaluation) on a 5-point scale (Fig. 1).

2.2 Sample

The experiment was carried out in two waves, in 2022 and at the beginning of 2024 (Micalizzi, 2024). The first wave included 30 individuals, 10 for each group (well-informed, misinformed, and uninformed). The second wave involved 65 individuals. In this case, it was not possible to divide them into three equal groups. The well-informed group was the most represented, and the misinformed were less represented. In both waves, the sample was distributed equally in terms of gender and role (i.e., whether the individual works in the creative industries or not). This difference means that we cannot consider gender as a significant variable, although some interesting reflections emerged at a qualitative level. Finally, the sampling was based on a snowballing method (Parker et al., 2019).

3 Results: Creativity Is Not Rocket Science

The presentation of the results is organized in three sections. The first one discusses the interviews carried out before the listening experiment. The second one focuses on the main results of the listening experiment. The last one discusses the results of the final interviews.

3.1 Before Listening

When asked to give a definition of ‘computational creativity’, the interviewees contextualized the concept within the general framework of using technologies in art productions. However, it remained a vague topic that was less explored than others, and the respondents showed a low level of awareness about it. They were only able to identify some products made by AI that they enjoyed online after having read the correct definition.

Could these products be considered ‘artistic’ and/or ‘creative’? To answer this question, we explored the participants’ opinions about the components that characterize a creative product. It was clear that they considered a creative product to be the result of three main elements: (a) the content it expresses; (b) the aesthetics, which depend on the recognition of others (Csikszentmihalyi, 1996, 2014; Gruner & Csikszentmihalyi, 2018); and (c) the author’s intentions, which represent a pivotal rule in defining a product as ‘creative’.

Moreover, not all respondents agreed in identifying a creative product as a work of art. They always considered the outputs of generative AI to be cultural products, but they did not necessarily consider them to be creative. This position is aligned with that of Smith and Cook (2023), who argued that AI applied to creative productions could open a new era of the readymade. The respondents in this study highlighted some aspects that can contribute to identifying artistic products:

- Author’s intention: The work of art is not produced for commercial purposes but arises from the artist’s urgent desire to communicate something. His intention makes it art regardless of the type of output. It is close to the definition offered by Merriam-Webster (2024), which considers art to involve the conscious use of human skills.
- Imperfection: Works of art carry with them the distinctive human trait of imperfection, which is also an element of their originality. AI is able to produce imperfections that are called hallucinations, but these are considered a mistakes—something unexpected that needs to be corrected (Maleki et al., 2024).
- Recognition: According to the interviewees, the work of art is the outcome of an ‘interpretative agreement’ of publics and experts, not only of the artist. This confirms the positions of Simonton (2009) and Boden (1998), who emphasize the central role of being recognized as valuable.
- Non-reproducibility: While recalling a now famous phrase by Walter Benjamin (1963), the respondents wanted to stress a different point—that art and its new forms of expression (e.g., multimedia installations, AR showrooms) are the result of the connection among users, artworks, and contexts. The focus of the new forms of art is the experience and not simply the artistic object. For this reason, they are not reproducible.

In this framework, some respondents highlighted how the introduction of the use of AI is a ‘need’ related to the distortion of the artwork concept associated with commissioned, accelerated, and ephemeral production, which requires a time-based creativity that cannot respect an artist’s timeframe. From this perspective, creative production with AI becomes a performed task, without a specific intention of the author. It is conditioned by the skills of the prompter (that is, the one who writes prompts); some authors argue that it is precisely for this reason that the use of AI is inseparable from human input (Manovich, 2018; Kalpokas, 2023). Creative production with AI also responds to the needs of the market: it is fast, versatile, and adaptable to the tastes of the audience, with greater computing power.

However, some respondents, especially those who have experienced it, emphasized the supportive role of AI in creative productions, noting that it is only possible to delegate marginal and mechanical steps of the process. Young interviewees who were also creative workers were more sensitive to the topic, stressing the fact that nowadays we need to distinguish consumer creative productions from proper artworks. Consumer creative productions are characterized by purely entertainment purposes, conceived for the market and designed according to its indications (Smith & Cook, 2023). This makes them reflect homogeneous trends and be capable of producing an economic return. Proper artworks, in contrast, are based on the intention of the artist who produces them, without regard of the market return but rather in accordance with his/her expressive need. They have no economic value except after being placed in the art market circuit—and not always successfully (Hertzmann, 2018). The positions of our interviewees seem to ignore the third possibility of abandoning the concept of authorship and art and instead observing AI productions as new languages and new forms of aesthetics (Manovich & Arielli, 2021 p. 2024).

Our study confirms the gender bias in attitudes toward technologies (Verbick, 2002): the female respondents felt uncomfortable expressing an opinion on the use of AI in creative productions due to a declared 'personal incompetence' on the topic. The idea that art is something aesthetic and negotiated, not a science, was found in women. Finally, they are more inclined toward the apocalyptic scenario linked to an excessive use of the machine for a skill that is considered purely human, creativity.

3.2 The Listening Experiment

All the respondents were invited to evaluate two soundtracks: the first one was always the human-produced track (track 1), and the second one was produced by AI (track 2). They were two piano compositions based on the same references.

After listening to track 1, the participants expressed their evaluation through keywords associated with the very slow rhythm, such as peace, relaxation, sadness, and nostalgia. At the same time, some respondents stressed the technique they perceived in the execution, with terms such as mastery and competence. Moreover, it was considered as something already listened to, famous, or 'classic' in its structure.

The second track was associated with other words, such as tension and surprise, for its fast pace. In fact, it stimulated emotions that were not always positive: some interviewees considered it to be disturbing, desperate, anxiety-inducing, and 'dark' in its rhythmic traits. At the same time, the repetitiveness of the scale construction caused boredom, one of the words most repeated by interviewees in their evaluation of the piece.

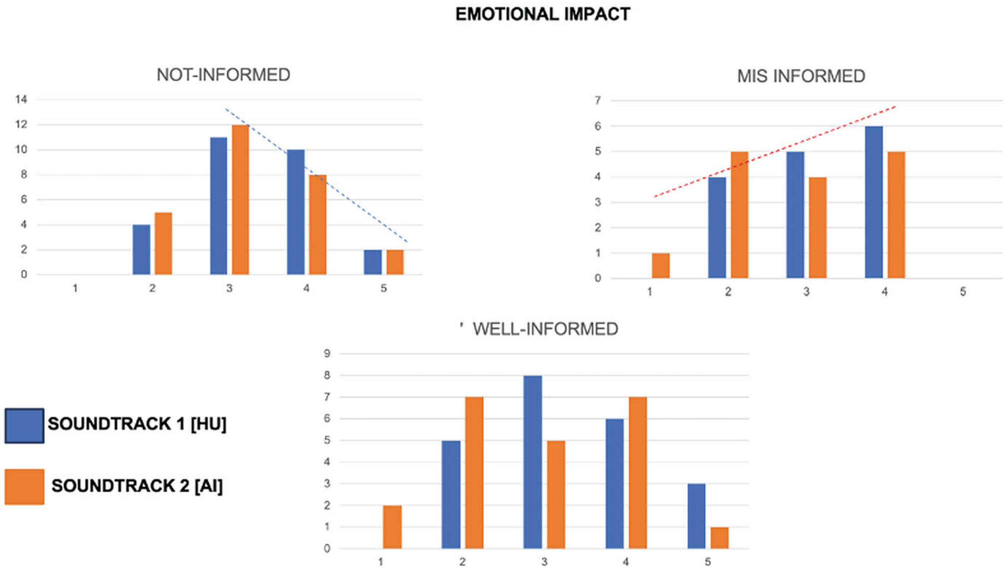

Fig. 2 Scores for emotional impact

The respondents expressed their opinion using five criteria—emotional impact, melody, rhythm, originality, and tone—on a 5-point Likert scale. Then, we asked for a general evaluation.

Figure 2 shows that in the well-informed group, there was not a clear preference for one of the two soundtracks, even though the first one seemed to be more appreciated. The clear differences were among the uninformed and misinformed groups, where there was a more positive trend among the misinformed respondents. While the uninformed group shows an inverted peak with a decrease in the highest scores, the misinformed group seemed to appreciate more the track they considered produced by AI.

Considering the second criterion of melody, we did not register significant differences within and among groups. In any case, the first soundtrack was the most appreciated. If we compare the evaluation of the second track between the well-informed and misinformed groups, we can see that, even though it was less appreciated than the first track, the well-informed group tended to be less harsh, expressing higher scores (4 and 5) than the misinformed ones for the second track (which they thought was produced by a human).

If we look at the evaluation of melody, we can see the uninformed group is more generous with the first soundtrack, and this evaluation trend is confirmed by the two other groups. Interestingly, the misinformed group gave better scores for the second track, which they considered to be produced by a young composer (Fig. 3).

Rhythm is the dimension that is least relevant for our experiment, since the results are coherent with the characteristics of the track, regardless of the authorship (whether real or attributed) (Fig. 4).

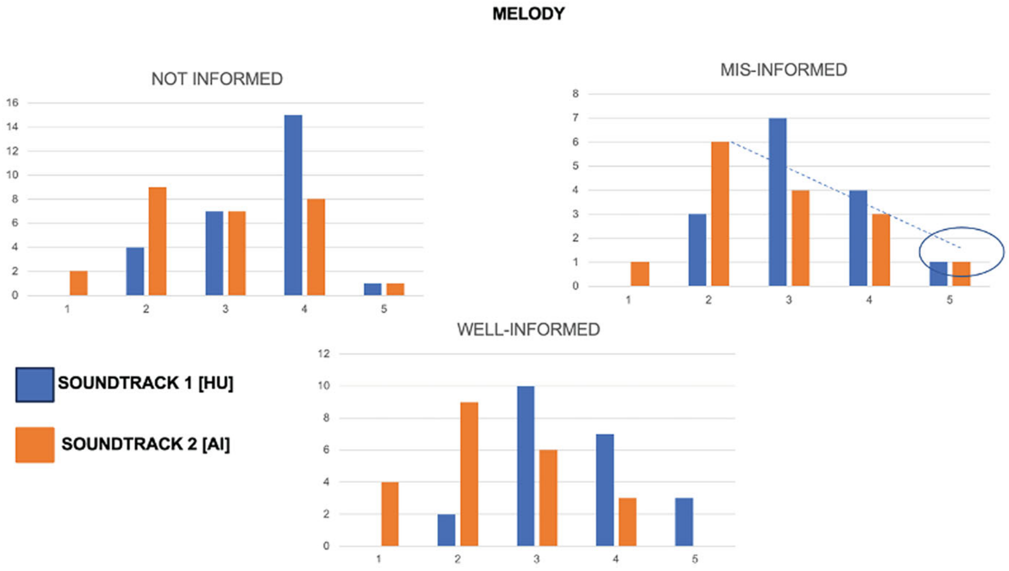

Fig. 3 Scores for melody

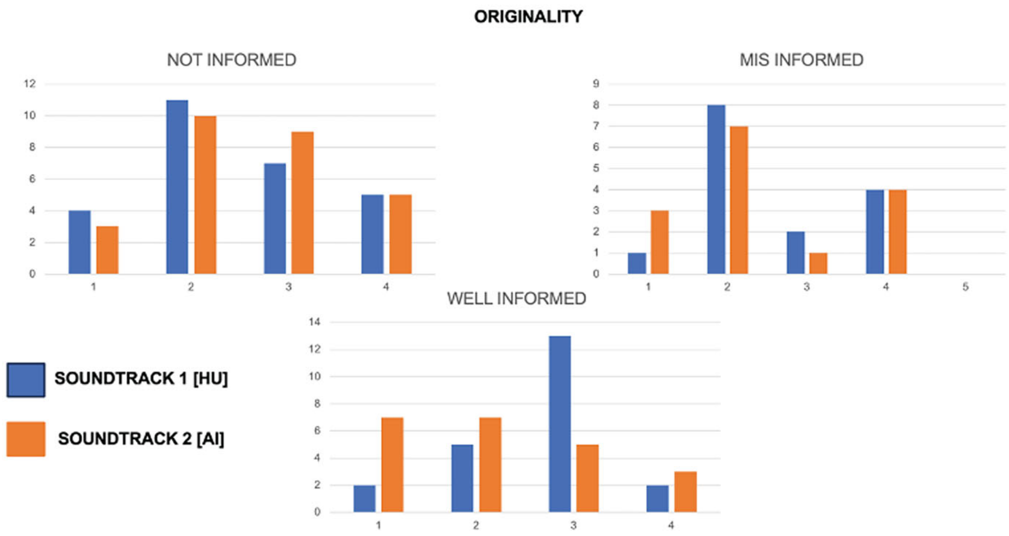

Fig. 4 Scores for originality

All the groups expressed a sense of familiarity for both the soundtracks (criterion: originality) in terms of knowing or simply recognizing the references. However, if we compare the scores between the uninformed and well-informed groups, we can see how authorship affects the evaluation of the first track, which registers a higher score in the second group (if we add up the 3 and 4 evaluations).

Finally, tone is the only category in which we can see a clear difference among the three groups. More specifically, the misinformed group had more appreciation for the track that they believed to be produced by a real composer (Fig. 5).

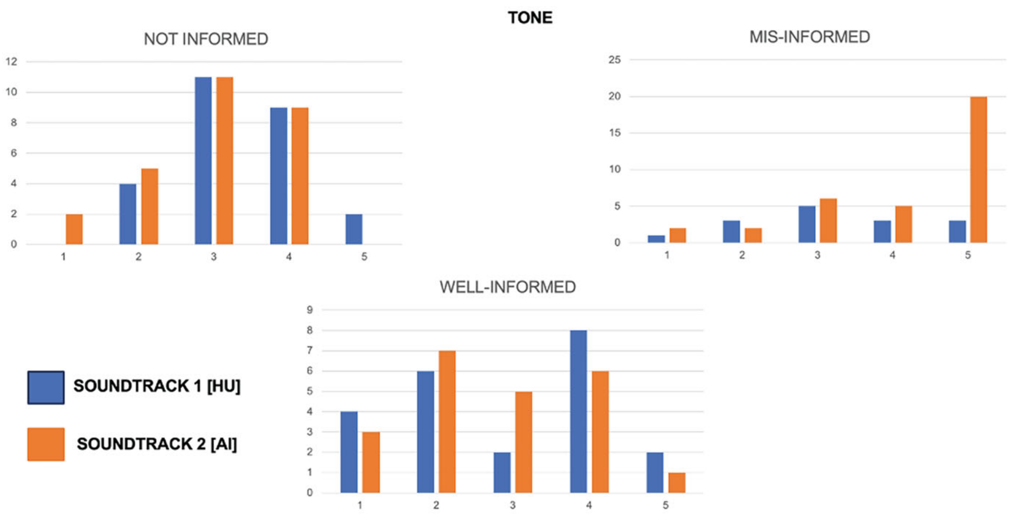

Fig. 5 Scores for tone

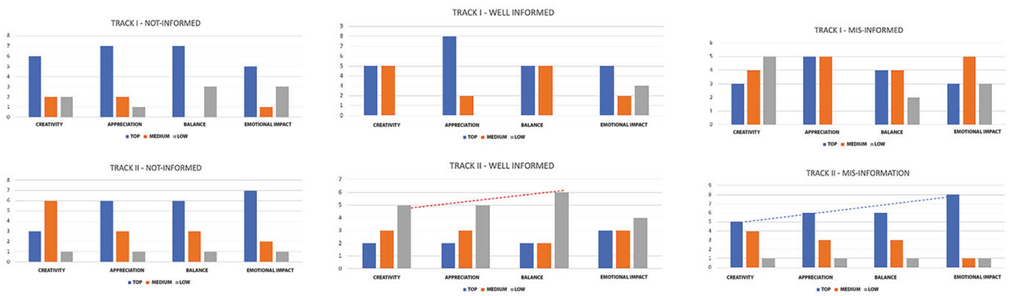

Fig. 6 Comparison of the evaluations given during the first wave (3-score scale)

If we consider the score of the general evaluation parameter, both the tracks were not considered appealing. However, comparing the results among the groups, we can see a better score for the first one; it is more evident with the misinformed respondents who considered the track as produced by AI.

This is an extremely interesting finding, and it confirmed the effect of procedural reactivity to the test, which was probably caused by the purpose of the study. The participants felt as if they could be deceived in the listening test and tended to give answers that did not meet the expectations. This was even more evident among the ‘experts’. Figure 6 presents the results of the first wave of the experiment, where the difference among the three groups is neat and evident.

To simplify the reading, we merged the results into 3 points on the scale (from 5 points). Focusing on track 2 and the difference between the well-informed and misinformed groups, the pattern of judgements (blue column and grey column) is clearly opposite to each other. This highlights how the authorial factor leads to a higher appreciation for the track that is considered to have been produced by a human being. The data from the first wave seem to confirm a possible authorship bias, which was only partially highlighted by the current wave. However, it should be specified that the results cannot be compared completely: the first survey took place at the end of 2022, a period in which information and debates on AI were certainly less present and therefore less consciously or unconsciously assimilated by respondents. Moreover, the size of the sample was more limited (30 participants), and the Likert scale included 4 dimensions that were similar but not identical to the ones used for the second wave—creativity, balance, emotional impact, and general evaluation.

3.3 After Listening

The final section aimed to investigate the participants' level of awareness about their prejudices against AI in the process of evaluating the two soundtracks. All the respondents seemed to be aware of the role of the information received (group 2 and group 3) in their process of attributing value to the track. However, we identified three main rationalizations. First, the most cited reason for recognizing the role of authorship was the fact that it is part of the artistic product, indivisible from the quality of the final output. A smaller but more conscious part of the sample emphasized that their prejudices are anchored in the idea that AI tools cannot achieve the same quality and expertise in creative products since creativity is a typical human skill.

In addition, there was a numerically less relevant part of the sample that declared that they were not influenced in any way by the product's authorship—both in this listening experiment and in the evaluation of any other creative work—due to the fact that even an AI-based production has an important human component, that is, writing the prompt and conducting the work of selection (in some cases even manipulation) on the AI output.

4 Conclusion

The findings in this research are consistent with those of previous studies (e.g., Hong, 2021) while offering a more integrated view of the connection between creativity and the use of artificial intelligence. The matter of AI's visibility in daily operations underscores a perception shaped by the lack of transparency regarding the role that technology plays in routine activities (Micalizzi, 2024). This is particularly relevant in the creative field, where the role of AI initially seems negligible, especially in relation to the evaluation of the final production.

This observation challenges the widespread belief that AI is neither 'creative' nor 'intelligent' but rather designed to perform specific tasks, which makes it 'effective' and 'efficient'. The participants stressed that although creative outputs generated by AI could be considered cultural products, they do not meet the criteria to be labelled as works of art.

Despite its limitations, our experiment provides important insights into the biases and perceptions surrounding AI's contribution to cultural production. On the one hand, interviewees—especially non-experts—were impressed by the high quality of AI-generated outputs. On the other hand, these outputs were seen as familiar and somewhat repetitive and characterized by a sense of predictability, which made them less innovative but more appealing to the market due to their consumable nature.

A key distinction that continues to differentiate humans is the intentionality behind communication. As some scholars argue (e.g., Hertzmann, 2018, 2020), art involves an interactive negotiation, and the uniquely human component lies in the communicative intention driving the creative process (Esposito, 2022). Moreover, while generative AI systems are becoming increasingly autonomous, artists retain control over each phase of production, including selecting inputs, making choices, and refining outputs. Additionally, the respondents pointed out that humans make unintentional mistakes, which reflect a level of imperfection and distinctiveness intrinsic to human nature.

Furthermore, novelty—a hallmark of artistic creativity—seems absent in AI-generated works, as they are often closely tied to their source material. To create genuinely original and innovative combinations, human intervention remains essential. Therefore, there is potential for positive integration, where machines act as tools that augment human creativity, but the human remains the true originator of the creative process.

The central concern surrounding the use of generative AI, in creative fields and beyond, is rooted in ethical issues, particularly regarding how much control humans can exert over this technology. Unlike earlier technologies, generative AI can 'learn', enabling it to function independently from human decision making. Thus, it is essential to establish shared ethical frameworks that safeguard human agency while promoting the responsible adoption of new technologies.

Finally, the participants expressed concern about the potential homogenization of public taste, driven by increasingly repetitive and short-lived content. This flattening of expectations in creative production is already happening, fuelled by the dominance of algorithms and market-driven strategies that favour only what has already proven to be successful.

The issue of AI education—including its potential applications in everyday life and professional settings—and its ethical implications remains critical to ensuring that this tool is used responsibly and that individuals engage with the socio-technological context in a manner that upholds shared ethical principles.

References

Benjamin, W. (1963). The work of art in the age of its technological reproducibility. In *Philosophers on film from Bergson to Badiou: A critical reader* (pp. 44–79). Columbia University Press.

Boden, M. A. (1998). Creativity and artificial intelligence. *Artificial Intelligence, 103*(1–2), 347–356. https://doi.org/10.1016/S0004-3702(98)00055-1

Boden, M. A., & Edmonds, E. A. (2009). What is generative art? *Digital Creativity, 20*(1–2), 21–46. https://doi.org/10.1080/14626260902867915

Boreham, R., & Wijnant, A. (2013). Developing a web-smartphone-telephone questionnaire. In *IBUC 2013 15th international Blaise users conference. Washington, DC*. Available at: http://www.blaiseusers.org/2012/papers/04b.pdf

Celis Bueno, C., Chow, P. S., & Popowicz, A. (2024). Not "what", but "where is creativity?": Towards a relational-materialist approach to generative AI. *AI & Society*. https://doi.org/10.1007/s00146-024-01921-3

Csikszentmihalyi, M. (1996). *Creativity: Flow and the psychology of discovery and invention*. HarperCollins.

Csikszentmihalyi, M. (2014). Society, culture, and person: A systems view of creativity. In *The systems model of creativity* (pp. 47–61). Springer.

D'Isa, F. (2024). *La rivoluzione algoritimica delle immagini. Arte e Intelligenza artificiale*. Luca Sossella Editore.

Duchamp, M. (1957). The creative act. *Art News, 56*(4).

Elgammal, A., Liu, B. Elhoseiny, M. & Mazzone, M. (2017). CAN: Creative adversarial networks, generating "art" by learning about styles and deviating from style norms. In *8th International conference on computational creativity (ICCC)*. https://doi.org/10.48550/arXiv.1706.07068.

Esposito, E. (2022). *Comunicazione artificiale. Come gli algoritmi producono l'intelligenza sociale*. Bocconi University Press.

Galanter, P. (2016). Generative art theory. In *A companion to digital art* (pp. 146–180). Wiley.

Gruner, D. T., & Csikszentmihalyi, M. (2018). Engineering creativity in an age of artificial intelligence. In I. Lebuda (Ed.), *The Palgrave handbook of social creativity research* (pp. 447–462). Palgrave Macmillan.

Hadjeres, G., Pachet, F., & Nielsen, F. (2017, July). Deepbach: A steerable model for bach chorales generation. In *International conference on machine learning* (pp. 1362–1371). PMLR.

Hertzmann, A. (2018). Can computers create art? *Arts, 7*(2), 18. https://doi.org/10.3390/arts7020018

Hertzmann, A. (2020). Computers do not make art, people do. *Communications of the ACM, 63*(5), 45–48. https://doi.org/10.1145/3347092

Hitsuwari, J., Ueda, Y., Yun, W., & Nomura, M. (2023). Does human–AI collaboration lead to more creative art? Aesthetic evaluation of human-made and AI-generated haiku poetry. *Computers in Human Behavior, 139*, 107502. https://doi.org/10.1016/j.chb.2022.107502

Hong, J. H. (2021). Artificial intelligence (AI), Don't surprise me and stay in your lane: An experimental testing of perceiving humanlike performances of AI. *Human Behavior and Emerging Technologies, 3*(5), 1023–1032.

Hong, J. W., & Curran, N. M. (2019). Artificial intelligence, artists, and art: Attitudes toward artwork produced by humans vs. artificial intelligence. *ACM Transactions on Multimedia Computing, Communications, and Applications, 15*(2s), 1–16. https://doi.org/10.1145/3326337

Horton, C. B., Jr., White, M. W., & Iyengar, S. S. (2023). Bias against AI art can enhance perceptions of human creativity. *Scientific Reports, 13*, 19001. https://doi.org/10.1038/s41598-023-45202-3

Kalpokas, I. (2023). Work of art in the age of its AI reproduction. *Philosophy & Social Criticism*. https://doi.org/10.1177/01914537231184490

Latikka, R., Bergdahl, J., Savela, N., & Oksanen, A. (2023). AI as an artist? A two-wave survey study on attitudes toward using artificial intelligence in art. *Poetics, 101*(2), 101839. https://doi.org/10.1016/j.poetic.2023.101839

Lee, H. K. (2024). Reflecting on cultural labour in the time of AI. *Media, Culture & Society, 46*(6), 1312–1323. https://doi.org/10.1177/01634437241254320

Liu, B. (2023). Arguments for the rise of artificial intelligence art: Does AI art have creativity, motivation, self-awareness and emotion? *Arte, Individuo y Sociedad., 35*(3), 811.

Magni, F., Park, J., & Chao, M. M. (2024). Humans as creativity gatekeepers: Are we biased against AI creativity? *Journal of Business and Psychology, 39*, 643–656. https://doi.org/10.1007/s10869-023-09910-x

Maleki, N., Padmanabhan, B., & Dutta, K. (2024). AI hallucinations: a misnomer worth clarifying. In *2024 IEEE conference on artificial intelligence (CAI)* (pp. 133–138). IEEE. https://doi.org/10.1109/CAI59869.2024.00033

Manovich, L. (2018). *AI aesthetics*. Strelka Press.

Manovich, L. & Arielli, E. (2021). *Artificial aesthetics: A critical guide to AI, media and design*, 4–27, available at https://manovich.net/content/04-projects/175-artificial-aesthetics/artificial_aesthetics.chapter_1.pdf [accessed 09/09/2024].

Manovich, L. & Arielli, E. (2024). *Artificial aesthetics: Generative AI, art and visual media*. Available at https://manovich.net/index.php/projects/artificial-aesthetics [accessed 09/07/2024].

Math in the Air. (2019). *Intervista all'artista computazionale Catodo: algoritmi, coding e arte generativa*. Available at https://www.mathisintheair.org/wp/2019/03/intervista-allartista-computazionale-catodo-algoritmi-coding-e-arte-generativa/ [accessed 22/09/2024].

Mazzone, M., & Elgammal, A. (2019). Art, creativity, and the potential of artificial intelligence. *Arts, 8*(1), 26. https://doi.org/10.3390/arts8010026

McCormack, J. (2003). Art and the mirror of nature. *Digital Creativity, 14*(1), 3–22.

Merriam-Webster. (2024). *Art*. Available at https://www.merriam-webster.com/dictionary/art [accessed 22/09/2024].

Micalizzi, A. (2024). Artificial creativity. Perceptions and prejudices on AI music production. In X. S. Yang, S. Sherratt, N. Dey, & A. Joshi (Eds.), *Proceedings of Ninth International Congress on Information and Communication Technology. ICICT 2024. Lecture notes in networks and systems* (p. 1004). Springer. https://doi.org/10.1007/978-981-97-3305-7_39

Micalizzi, A., & Lelicanin, M. (2023). *Studiare i creative media. Ricerca e analisi dei processi di creazione, condivisione e appropriazione culturale*. Utet Università.

Millet, K., Buehler, F., Du, G., & Kokkoris, M. D. (2023). Defending humankind: Anthropocentric bias in the appreciation of AI art. *Computers in Human Behavior, 143*, 1–9. https://doi.org/10.1016/j.chb.2023.107707

Oksanen, A., Cvetkovic, A., Akin, N., Latikka, R., Bergdahl, J., Chen, Y., & Savela, N. (2023). Artificial intelligence in fine arts: A systematic review of empirical research computers in human behavior. *Artificial Humans, 1*(2). https://doi.org/10.1016/j.chbah.2023.100004

Parker, C., Scott, S., & Geddes, A. (2019). *Snowball sampling*. SAGE research methods foundations.

Paul, C. (2023). *Digital art*. Thames & Hudson.

Rohrmeier, M. (2022). On creativity, music's AI completeness, and four challenges for artificial musical creativity. *Transactions of the International Society for Music Information Retrieval, 5*(1), 50–66. https://doi.org/10.5334/tismir.104

Simonton, D. K. (2009). Varieties of (scientific) creativity: A hierarchical model of domain-specific disposition, development, and achievement. *Perspectives on Psychological Science, 4*, 441–452. https://doi.org/10.1111/j.1745-6924.2009.01152.x

Smith, A., & Cook, M. (2023). AI-generated imagery: A new era for the 'readymade'. *SA '23: SIGGRAPH Asia 2023 Art Papers, 5*, 1–4. https://doi.org/10.1145/3610591.3616432

Tigre Moura, F., & Maw, C. (2021). Artificial intelligence became Beethoven: How do listeners and music professionals perceive artificially composed music? *Journal of Consumer Marketing, 38*(2), 137–146. https://doi.org/10.1108/JCM-02-2020-3

Tigre Moura, F., Castrucci, C., & Hindley, C. (2023). Artificial intelligence creates art? An experimental investigation of value and creativity perceptions. *The Journal of Creative Behavior, 57*(4), 534–549. https://doi.org/10.1002/jocb.600

Tubadji, A., Huang, H., & Webber, D. J. (2021). Cultural proximity bias in AI-acceptability: The importance of being human. *Technological Forecasting and Social Change, 173*. https://doi.org/10.1016/j.techfore.2021.121100

Verbick, T. (2002). Women, technology, and gender bias. *Journal of Computing Sciences in Colleges, 17*(3), 240–250.

Zeilinger, M. (2021). Generative adversarial copy machines. *Culture Machine, 20*, 1–23

Zhou, E., & Lee, D. (2024). Generative artificial intelligence, human creativity, and art. *PNAS Nexus, 3*(3), 52. https://doi.org/10.1093/pnasnexus/pgae052

Creativity and Digital Music Education in the AI Era

Antonella Coppi

Abstract The twenty-first century has brought a wave of global technological advancement to make information more accessible than ever before in human history. Countless industries have benefitted from this newfound availability of information, including the education sector. The internet provides students with a wealth of information from which to draw, while teachers can access a breadth and depth of resources that dwarf what was available to their twentieth-century counterparts. These technological advances have fundamentally changed education and will continue to do so as long as technological evolution persists at such a rapid pace (Rivoltella, Studi Sulla Formazione/Open J Edu, 26(2), 63–67. https://doi.org/10.36253/ssf-14975, 2023). The latest breakthrough in technology, artificial intelligence (AI), has the potential to upend the education system as we know it. This contribution approaches using artificial intelligence in music education as part of the creative cognitive process. AI in education is a very diverse field, dating from about 1970 (Carbonell, IEEE Trans MMS, 11(4), 190–202, 1970), and it has its own developed methodologies, techniques, and traditions. The field is highly interdisciplinary, involving substantial contributions from the fields of music, education, AI, cognitive psychology, the psychology of music, social psychology, anthropology, philosophy, linguistics, human-computer interaction, and many other fields (Holland, Artificial intelligence in music education: a critical review. In E Miranda (Ed.), Readings in music and artificial intelligence, contemporary music studies, vol 20, Harwood Academic Publishers, 2000).

Keywords AI · Music education · Italian school music · Curriculum

Special thank goes to Sara Valente, for collecting data and sharing her useful reflections. Sara Valente is a primary school teacher and scholar of AI in music education at the State Conservatory "L. Marenzio", Brescia, Italy.

A. Coppi (✉)
Department of Human Science, IUL University, Florence, Italy
e-mail: a.coppi@iuline.it

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

A. Micalizzi (ed.), *Artificial Creativity*, Springer Proceedings in Humanities and Social Sciences, https://doi.org/10.1007/978-3-031-88334-7_13

1 Introduction

> John and I . . . were quite happy to nick things off people, because . . . you start off with the nicked piece and it gets into the song . . . and when you've put it all together . . . of course it does make something original (Paul McCartney quoted in Moore, 1992).

Artificial intelligence (AI) is rapidly transforming numerous industries, from healthcare to finance, and music education is no exception. Nevertheless, while the integration of AI technologies into music education offers new opportunities to learn, teach, and create music, what exactly is AI? The European Parliament has this to offer:

(AI) Is the Present and the Future of Technology. Artificial intelligence (AI) is the ability of a machine to exhibit human capabilities such as reasoning, learning, planning and creativity: it enables systems to understand their environment, relate to what they perceive and solve problems, and act toward a specific goal. AI systems are capable of adapting their own behaviour by analysing the effects of previous actions and working autonomously (European Parliament, last updated 06/28/2023).[1]

On 03.13.2024 the European Parliament approved the *Artificial Intelligence (AI) Law*, which guarantees security and respect for fundamental rights and promotes innovation. Members approved the regulation, the result of an agreement reached with member states in December 2023, with 523 votes in favour, 46 against, and 49 abstentions.[2] It aims to protect fundamental rights, democracy, the rule of law, and environmental sustainability from high-risk AI systems while promoting innovation and ensuring Europe's leadership in the field. The regulation establishes obligations for AI based on possible risks and level of impact.

The purposeful use of artificial intelligence (AI) in digital music education is crucial for enhancing rather than replacing creative action. This contribution explores the possibilities of integrating AI into music education to provide tools and resources that facilitate learning and student engagement (Panciroli & Rivoltella, 2020), offering new opportunities for exploring and experimentation. At the same time, it will highlight the importance of teaching and ethical education design in light of the recent approval of new EU regulations on the use of AI, the world's first, offering insights for future-oriented pedagogical reflection.

At this stage it would seem that AI can be an indispensable tool for implementing any kind of knowledge and developing work and productivity in any field, but flexibility in modality of application remains an issue. Thus, one of the main purposes of AI is to understand human intelligence, that is, the ability to reason, plan, solve problems, learn quickly and learn from experience (In 1967, Gottfredson, to bring this understanding to fruition and enable AI to interact with human intelligence, has attempted to devise solutions that would enable algorithms to

[1] https://www.europarl.europa.eu/news/it/press-room/20240308IPR19015/il-parlamento-europeo-approva-la-legge-sull-intelligenza-artificiale

[2] https://www.quotidianosanita.it/allegati/allegato1710335981.pdf