MUSIC IN THE DIGITAL AGE

INTERNATIONAL CONFERENCE | DIGITAL TECHNOLOGIES AND ARTISTS' RIGHTS

ATHENS, OCTOBER 22-24, STAVROS NIARCHOS FOUNDATION CULTURAL CENTER

PRESENTATIONS

Music in the Digital Age: Streaming & Artificial Intelligence has been a three-day international forum organized by APOLLON (Greek CMO for musicians' neighboring rights) and FIM (International Federation of Musicians). Against the backdrop of **AI-generated content and the dominance of streaming platforms**, the conference examined how **revenue models, legal frameworks and artistic labor are being reshaped** in a digital economy that prioritizes scale over sustainability.

Bringing together artists, journalists, industry professionals, legal experts, academics, policy makers and technologists, the event focused on three core questions:

- How to build **sustainable and equitable compensation** models in a saturated streaming market?
- How to protect creators' rights when **AI is trained on and competes with their work?**
- What role should legislation play in **securing ethical AI and fairer digital markets** without stifling innovation?

Pavlos Antoniadis (IRCAM/University of Ioannina) challenges the industrial standards of AI music generation. Moving beyond the "black box" of Big Data models like Suno or Udio, he explores the mechanics of AI through the lens of "tacit knowledge"–the embodied, subjective experience of music-making.

PRESENTATIONS | PAVLOS ANTONIADIS

The Mechanics of Artificial Intelligence

Athens, October 23, 2025
Stavros Niarchos Foundation Cultural Center

Good morning, everyone

Thanks very much to the organizers, and particularly to Dinnos Georgountzos, for their kind invitation. I would have loved to share the same physical space with you, but I'm currently in Beijing, China for the 10th edition of the TENOR conference, by the title Notational Intelligence. Hence, my current video intervention.

I would like to contribute to panel A, the mechanics of AI, by providing an alternative to the current industry standards of AI generation, as represented by tools like Suno or Udio. My vision entails the conception of “musicking” as process rather than product, a narrative of human-machine synergy rather than human replacement, and a technique which does not necessarily capitalize on big data and deep learning architectures.

Instead, the alternative I present is based on tacit knowledge. Tacit knowledge is defined as a personal, subjective, and experiential, difficult to document in the usual data forms, the know-how of experiencing music-making with others. This form of knowledge helps us identify the deep epistemic biases running through the current industrial standards.

I define these biases as every form of explicit or implicit, theoretical or technical constraint in knowledge representation with unreasonable, indeterminant, and uncontrollable results in the model's interaction with the world. At IRCAM, we recently signed a manifesto for responsible artificial intelligence serving inventive and diverse musical creation which identifies the main problems, ecological, societal, economical, and artistic of the current industrial standard.

My own claim is that all these problems refer to the definition of epistemic biases above. In tracing the philosophical and technical criteria of epistemic biases, I will attempt a rapid introduction to the basics of machine learning and artificial intelligence in music, then to the basic models of commercial music generation, and subsequently to associated biases and alternatives.

Why does a machine need to learn from the world rather than execute hardcoded rules as in this first example based on the father of computer music, Max Mathews? The answer is that sometimes rules might be hard to implement due to the complex, big, and noisy data as the following example by Yannis Xenakis demonstrates.

In these cases, we might choose to learn the rules from the data ending, up to predict and generate new data. If we compare definitions by engineers with definitions by musicians, as shown in this slide, we notice that musicians might use the algorithms in unexpected ways which are tapping into their creative potential.

Classic types of supervised or unsupervised machine learning, but also the state-of-the-art deep architecture types such as RNNs, VAEs, GANs, transformers, or diffusion models give rise to the most important systems out there. They share basic datasets such as MIDI, audio, and spectrograms, but often disregard aspects of musicking such as embodiment and interactivity. Their capitalization on big data is associated not only with societal problems such as the micro work happening in refugee camps, whereby refugees label data for ridiculous compensations, but also to a deeply problematic approach to time and change since these models deal with abstractions of time rather than dynamic change.

Whether on linear or non-linear approaches to temporality, time as an open-ended horizon is absent from these models. In his book "Logiciel", philosopher AA Cavia describes deep neural networks as topological structures with three functions: encoding, embedding, and decoding.

Encoding of the world into data is crucial for the data types discussed above. In the embedding, the data are transformed through hidden layers which allow for a detection of patterns in different hierarchical scales or abstraction levels of the latent space. The decoding of the structure in an output is probabilistic distribution which can be directly mapped to the input layer. Epistemic violence then occurs in all three stages, in how the world is encoded, embedded, and decoded in DNNs.

But there is an even deeper problem if you excuse the pun. The semantic blind spot of meaning creation in these models that work based on correlations and not on causations. Judea Pearl has analyzed the theoretical limits of the models. They can reason only at the level of association, but not of intervention or counterfactuals.

What if I do X or what would have happened if my input data were not X but Y? This is what only humans can do, imagining different worlds and intervening to make them happen. Models, current models can't do that. Marie Shanahan underlines another problem, the latent anthropomorphism of LLMs which is integrated through the uncritical use of epistemic terms like "thinks", "knows", "believes".

In the same vein, music diffusion problems are not composing or performing music but rather recombine, no matter how convincingly these bits of information to generate statistical simulacra.In my work since 2013, including the development of the GestCOM system, I have enthusiastically adopted probabilistic designs, machine learning-based modeling, and more recently, AI systems such as SOMAX II.

Beyond the very technical differences of such tools from industrial standards at all stages of the data treatment, encoding, embedding, and decoding, the fact that they are often based on one-shot gestures, rather than Big MIDI, audio and spectrogram Data, or the fact that they employ machine listening and feedback circles of human-machine interaction as a prerequisite to music generation, and they are for the greater part explainable as opposed to the black box effect of deep learning, such attempts capitalize on a multiplicity of representation and interaction, attempting to capture the fleeting social and interactive nature of music-making that lives equally in scores, recordings, videos, and above all, live performance.

With Mikhail Malt, we create dialogues between human and machine based on mutual training and output processes, which, although generated by well-defined datasets, reclaims the intuitiveness of a real improv albeit based on something as complex as "Evryali" by Iannis Xenakis.

This was rather my second example, whereby I create new flexible mappings between music recording and the system of automatic improvisation. And here's Mikhail Malt and me playing Xenakis.

In yet other cases of multimedia composition, "Nǐmen shì chóngzi", (Mandarin for "You're Bugs"), is a work that reflects on problems of authorship, authenticity, and authority in musical AI, namely the disturbing potential of human inferiority - the fact that we are all potentially bugs with its multiple connotations, and when we might have to accept a lower status in a new or not-so-new hierarchy.

Concluding the technical details of musical AI are not merely means that can be used for positive or negative ends, as the usual platitude goes. They rather carry deeper meanings about control and time. They encode, manipulate, and decode not only data, but the world itself. A different musical AI is possible, an AI restoring the possibility of disturbing interventions and of singular imagining of multiple futures. Thank you.

Pavlos Antoniadis,
IRCAM/University of Ioannina