ISSN: 2633-4828 Vol. 5 No. 4, December, 2023

International Journal of Applied Engineering & Technology

Artificial Intelligence Composition System Applying the Translation Process of Actor Network Theory

Yeongshin Choi[1] and Seungyon-Seny Lee[2*]

[1]*PhD. Department Music Technology, the Graduate School Sangmyung University, Republic of Korea*

[2]*Professor, Department Music Technology, The Graduate School Sangmyung University, Republic of Korea*

[1]*blissin12@gmail.com and* [2]*senylee01@gmail.com*

Date of Submission: 21st September 2023 Revised: 30th October 2023 Accepted: 26th November 2023

How to Cite: Yeongshin Choi and Seungyon-Seny Lee, (2023), Artificial Intelligence Composition System Applying the Translation Process of Actor Network Theory, International Journal of Applied Engineering and Technology 5(4), pp. 81-88.

Abstract -This study examined the relationship between humans and technology by employing an artificial intelligence composition system of popular music using cutting-edge technology to the actor network theory translation process. Utilization of artificial intelligence technology within the popular music industry has altered how humans produce and distribute music. Examined as an example of the algorithmic model's functions is an artificial intelligence configuration system with automation and generation functions. On the basis of the development process of the artificial intelligence configuration system, the translation process of actor network theory was evaluated. Throughout the system development process, human-centered actors construct networks to facilitate access to music creation. Successful system design increases the commercialization of AI configuration programs and fosters the formation of networks between various actors. As a rule, artificial intelligence configuration systems learn through deep learning learning models using midi files from existing sound sources or newly created midi files as data. Existing music principles are discovered and output the same results as the input data during the learning process. If the system is developed effectively, it will establish a network and be distributed to the public. Multiple networks of diverse technologies, machines, and human actors compose a network. Only when no one has the upper hand and everyone acts equally does the network solidify. Despite human superiority or mechanical optimism, this creates a new point of interest between the two. Actors are also continually switching roles within the network, receiving new roles and forming new relationships within those roles. It is necessary to pursue the coexistence of humans and technology in light of the evolving music market and the altered role of humans.

Keywords: Artificial Intelligence, AI Music Composition, Actor Network Theory, Generative Music

1. INTRODUCTION

The rapid rate of technological advancement is altering all aspects of society. Specifically, the development of cutting-edge technologies such as artificial intelligence and virtual reality is accelerating this transformation. Gradually, artificial intelligence has surpassed the concept of basic Technology and is now capable of exhibiting human-like abilities [1]. Recently released by OpenAI, the conversational AI ChatGPT generates more accurate and sophisticated responses than previous AIs. Evidently, technology is altering and developing at an alarming rate. Obviously, this results in a variety of issues requiring quick and adaptable human responses. Despite this, the emergence and development of advanced technology is being applied in all fields, and it is transforming the music industry.

At the center of this phenomenon, prominent K-pop agencies are investing in and creating metaverses and trading blockchain-based NFTs, enabling music consumption in a virtual world. Additionally, artificial intelligence technology is being applied to other types of content, such as voice restoration and personalized music recommendations, and the music creation system enables anyone to create music rapidly.

Initially, the AI composition system was designed to create music similar to existing works, but it is presently evolving into a system that enables users to create and share their own music based on their specific requirements. By demonstrating a creative process distinct from the traditional method of composing through professional composers, it offers consumers a new means of acceptance.

Technological advancements have increased our dependence on technology, requiring humans to be adaptable and receptive to change. To create high value, it is crucial to comprehend and embrace technological changes in the industrial environment. How are creators and consumers of

music adapting to the technological environment in order to create and consume music? How is their role evolving, and what should their direction be?

AI compositional systems can be viewed as an active manifestation of the agency of technology and machinery, since technology rather than humans produces musical results. As technology has altered the manner in which music is created, this study examines the relationship between humans and technology in each process and seeks directions for technology's application.

Actor-Network Theory, advocated by science and technology scholar Bruno Latour, moves away from anthropocentric thinking that ascribes agency and dignity only to humans and argues that nonhumans, such as machines and matter, have agency. His argument demonstrates that both humans and non-humans have agency and allows for an examination of the interests of both humans and technology in relation to the changes brought about by technology. Consequently, the purpose of this study is to investigate the relationship between humans and technology via the system design process of an artificial intelligence composition system based on the four steps of Callon's translation in ANT theory.

2. Actor-Network Theory

Bruno Latour, Michel Callon, and John Law created Actor-Network Theory (ANT) to provide a comprehensive and in-depth perspective on the complex relationship between humans and technology. Instead of viewing humans and non-humans as subjects and objects, respectively, it views them as equal actors. In this context, nonhumans encompass everything other than humans, including technology, objects, nature, and insects. In the 1980s, ANT was proposed as a framework for the sociological comprehension of science and technology and as a method for examining the relationships among science, technology, and humans[2]. It contends that modern dichotomies should be abandoned in favor of equal symmetries between human and nonhuman, society and science, subject and object, and so on. It is centered on the notion that science and technology, previously perceived as separate, are in fact intertwined and inseparable[3].

People and technology, humans and non-humans, influence each other in an infinite number of permutations that form society. It considers humans and non-humans as individual actors and attempts to view the outcomes or phenomena as the result of the relationships between human and non-human actors. In other words, ANT is suited for observing and analyzing events and technologies that evolve or change over time.

Human and non-human actors collaborate with other actors to form innumerable heterogeneous networks. "Black boxes" are used to refer to networks that are exceptionally resilient. We do not examine the network components in their robust state. For instance, when an electric car is introduced, we do not examine the technology that went into it or any technical coupling or other difficulties. Nevertheless, it is crucial to accurately identify human and non-human actors in a robust network and disclose what they have been doing to one another. By tracing the intricate relationships between humans and nonhumans, we can gain a deeper understanding of the networks through which science and technology have produced technological products. In addition, by analyzing the successes and failings of networks between actors, we can gain a deeper understanding of one another and generate new perspectives and inventive outcomes.

During the formation of a network, a human actor is typically the primary actor. The primary actor combines with other actors based on their comprehension or intention of the situation or environment to form a network, and it is crucial to comprehend and coordinate actors with varying interests and objectives. The procedure is extremely intricate and difficult. Existing network objectives are adjusted, modified, bypassed, and reconfigured, frequently redefining and transforming actor roles[4].

An actor network is a significantly larger network of actors with diverse interests who collaborate for a common purpose. Translation is the examination of the process of expansion and contraction in the formation of a network. There are numerous methods to examine network formation, but Callon's translation process is an effective method for analyzing network actors. This is the most frequently cited ANT procedure, which consists of four steps: problematization - intersegment - enrollment - mobilization[5].

In a network, actors establish "Obligatory Passage Points" (OPPs) that must be traversed. An OPP is a crucial component in establishing a network and can be thought of as the actors' shared purpose or objective. Actors associate and disassociate for this purpose.

The stages of Callon's translation procedure are as follows: In problematization, the central actor attempts to disrupt the existing network of actors and connect to a new network based on their particular intentions and goals. In Intersegment, the central actor negotiates to separate other actors from the existing network and connect them to his network. In Enrollment, other actors are connected to the network through interest and given new roles.

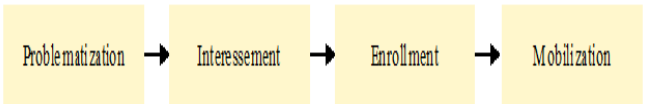

Figure 1. Carlong's ANT Translation Step 4

The translation process does not always adhere to these four stages, and the process of establishing a new network may manifest itself differently. Nonetheless, since the translation process focuses on the roles and processes of multiple actors, it is evident that the expression of agency in a variety of ways is an essential observational subject. For the formation and expansion of robust networks, effective translation of network formation processes is crucial.

3. Artificial Intelligence Composition System

As artificial intelligence technology expands the capabilities of content creation and production, the music industry must adopt new communication strategies. Changes in the creation, consumption, and distribution of music have sparked new development and innovation through digital transformation, and are transforming the music industry's structure[6].

AI technology employs the recognition, prediction, automation, communication, and creation functions of algorithmic models to perform the five musical functions outlined in the table below [7].

Table 1. The five functions and musical functions of artificial intelligence

Function	Musical Function
Recognition	Recognize songs and melodies using voice recognition Recognize melodies using musical notation
Prediction	Predict music preferences based on user patterns Predict music results from existing data
Creation	Create new music
Communication	Perform, create, and communicate with AI
Automation	Automate musical tasks to improve output Explore distinct user requirements

The evolution and expansion of technology facilitated by these AI capabilities demonstrates the capacity to respect individual preferences, make creativity more accessible, recognize and combine diverse data, and accomplish faster processing through automation. Combined, AI-enabled music features provide the public with a broader and more diverse range of musical enjoyment options. It is essential to note that each feature is applied interactively, not independently.

AI compositional systems are the result of the intersection of two very different disciplines of science and art: artificial intelligence research, which seeks to determine whether machines can behave similarly to humans, and music composition, which employs computer algorithms to create music. This research began in 1989, when cognitive scientist Peter M. Todd set out to compose a few simple melodies, and garnered momentum in the 2000s[8].

Modern AI composition systems primarily use machine learning methods. They aim to generate music that is similar to or identical to the existing musical data they are fed, with the primary design goal being how much like existing music the system can create. Different developers use different algorithmic approaches for various purposes. They can be categorized according to various perspectives or methods, but in this study, they are divided into two categories based on their function and application. First, there are two categories of compositions: those that imitate the style of established composers and those that use the established system to compose freely as the user desires.

Composition systems based on the musical style of existing composers include the American-developed Emily Howell and Kullita and the Spanish-developed Iamus. This method employs AI technology to analyze the compositions of existing composers, identify their characteristics, and compose by imitating and evolving their style. This method emphasizes how similar the AI composition system is to human music as a composing entity on par with humans.

Conversely, systems that allow users to compose their own music include Aiva in the UK, Amper music in the US, and Vio.Dio and Musia in Korea. This method involves programming automatic songwriting software developed through artificial intelligence technology and providing it to users on a platform basis. The user simply selects the desired genre and mood, and music is composed automatically to match. The composed music can be rearranged by the user and shared on the platform, enabling user collaboration and interaction. In particular, background music without lyrics that effectively conveys the mood of video or photo content is increasingly utilized in television entertainment programs. AI technology is accomplishing certain results in communicating with humans, and interest is growing despite the fact that it is not yet possible to completely replace human creativity. Both methods generate music using an AI-based automated composition system, but their objectives and implementations are distinct.

The focus of this study is artificial intelligence composition systems. This method enables users to create their own music by reducing the accessibility of creation, which was previously difficult due to professionalization. AIVA, Amper Music, Musia, and Vio.Dio are selected as

examples among the many systems, with a focus on systems that are presently providing services or in development.

Deep learning is a form of machine learning that generates rules by analyzing existing data and learning on its own. The objective is for the system to learn from the input data and output identical, similar, or existing data. This enables automatic composition by learning from a vast quantity of existing music data and identifying rules for creating new music.

As shown in Figure 2, AIVA is capable of providing genre-specific audio or creating music based on the user's preferences. When users submit a genre, composition, beat, duration, etc., music is generated based on the input conditions, and users can modify the melody, chords, and instrumentation to create music that suits their tastes.

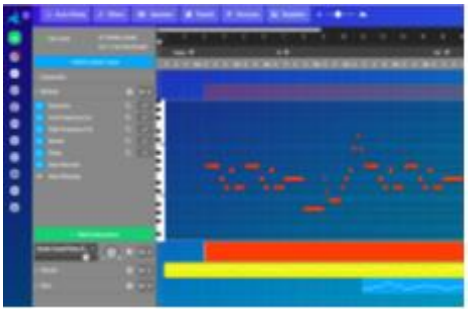

Figure 2. AIVA

Amper Music is comparable to AIVA in that users can select genres, genre-specific moods, and customize the composition and instrumentation. However, it has the potential to be used for other types of content by allowing users to input their own videos, generate music that matches the length of the video, and modify the length of the music in minute detail.

Figure 3. Amper Music

Musia, which was created in Korea, has a limited duration because it can only generate eight segments, but users can make a variety of adjustments. If users input a chord, a corresponding melody can be created, and if they input a melody, a corresponding chord can be created. The function of domestic corporations' non-audio content differs from the aforementioned instances. It provides the actual music, which is composed by its own AI system. Users can select completed music based on their desired mood, theme, or musical genre. Users cannot modify the melody or chords, but they can remove a specific instrument to create a solo instrumental rendition from a multi-instrumental composition. It appears as shown in Figure 4.

Figure 4. Musia

As shown in Figure 5, Vio.Dio serves a purpose distinct from the preceding three instances. It provides the music itself, which is composed using artificial intelligence. Users can select the completed music based on their desired mood, theme, or genre. In contrast to previous programs, users cannot modify the melody or chords, but they can remove specific instruments from a multi-instrumental composition to create a solo instrumental piece.

Figure 5. Vio.Dio

Table 2 below summarizes the functions of the four programs above.

Table 2. An artificial intelligence composition program that allows users to participate

	AIVA	Amper Music	Musia	Vio.Di
Target for use	User			
Scope of Use	Melody Chord	Melody Chord Video	Melody Chord	reated music
Music Genre	Various genres			Background music
User input status	Possible			Impossible

Thus, AI composition systems can generate a variety of music based on the user's selections, modifications, and preferences. Generally speaking, users can select the genre, mood, beat, speed, composition, instruments, chords, and melody, but each program has limitations on the instruments that can be conveyed, the length, and the scope of creation, so the user's involvement varies. However, regardless of a person's musical proficiency, anyone can easily express their artistic creativity, and with assistance, they can express more original music.

Evidently, there are still debates and issues regarding whether human creativity is possible with machines and copyright issues of existing music used for data analysis, but because AI technology has produced some achievements and demands, a new artistic paradigm and institutional direction are required.

4. Translation of the Artificial Intelligence Composition System Design Process

This study's AI composition system employs deep learning algorithms to enable machines to compose music independently. Through deep learning technology and non-human actors such as music big data, composing, once believed to be the exclusive domain of humans, is now possible via non-human technology. It is an example of the active manifestation of the agency of non-human actors, according to the ANT theory. This disrupts the conventional method of human creation and necessitates a shift away from an instrumental perspective on technology.

However, technology's agency as a non-human actor is also unachievable without human intervention. Artificial intelligence does not independently produce artworks. It still necessitates human involvement at some point. This study aims to move away from the dichotomous dominance of AI systems or humans and instead examine AI composition systems and humans in an equal relationship.

This chapter describes AI compositional systems from an ANT standpoint. It analyzes widely used machine learning methods in AI composition systems using the four-step translation procedure of Callon, a common network analysis method in ANT theory.

Using techniques from artificial intelligence, system developers create automated composition systems in order to remove obstacles to music creation. David Cope, a longtime developer of automated songwriting systems, began automating his songwriting when he ran out of creative ideas and required assistance composing [9].

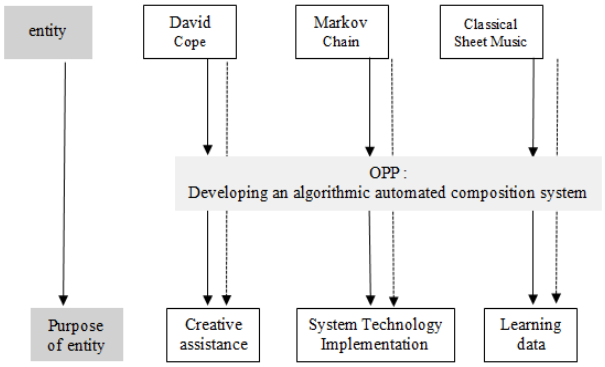

Figure 6. Problematization in the system design procedure of David Cope

David Cope initiated the problematization of extant AI composition systems by establishing the development of an algorithm-based automatic composition system as an Obligatory Passage Point. He began his research by concentrating on the creation of an imitation system that autonomously generates works in his style using artificial intelligence that has learned his style. After achieving success in composing music in his own style by identifying patterns in his own works, he widened his focus to include the works of renowned composers, identifying patterns in their existing works and composing new works in their style. Below is depicted Cope's design procedure.

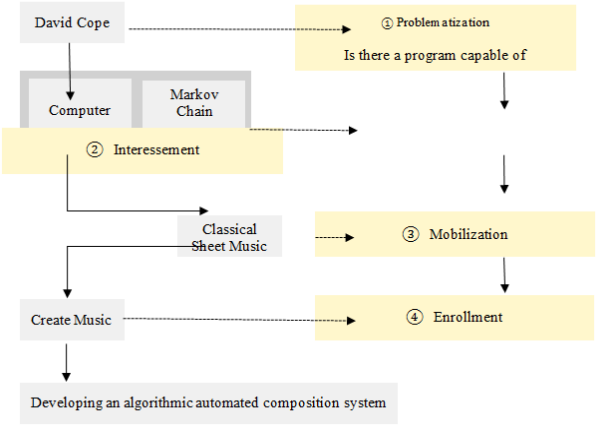

Figure 7. The Translation Process of the System Design Procedure of David Cope

David Cope recruits probability-based Markov chains as actors to construct an algorithmic automated composition system. Classical music, which is free of copyright issues, is used as input data for the algorithmic system in order to generate music for his purposes that resembles extant musical styles. Using Markov chain technology and classical music data, they are able to successfully create an AI composition system. However, the public viewed it as music that merely imitated the composer and lacked originality and human emotion. There is a betrayal in this procedure.

A new network is formed for the purpose of developing technology to create music that is original, as opposed to

imitating the music of established composers. Various systems are created so that anyone can compose music in a variety of methods. There is a creation of new problematization and networks.

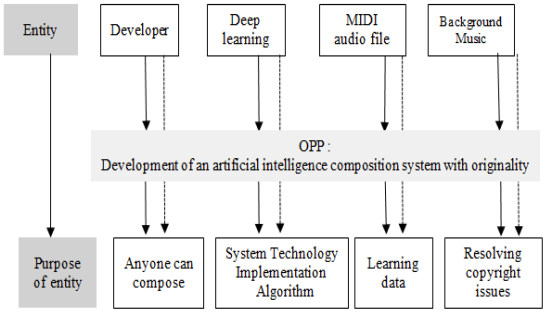

Figure 8. Problematization in Deep learning method AI system Design Stage during Translation Process

Rather than creating a copy of an existing composer's composition, the network is rebuilt by making the development of an original system an obligatory passage point for developers. As observed earlier in the system design process, artificial intelligence produces results by aggregating data about humans and identifying rules and patterns in the data. It is crucial to repeatedly learn from significant amounts of data[10]. Increased computer capacity and large amounts of big data have resulted in faster and more potent computational resources, allowing for the implementation and study of larger neural networks. The Internet and the proliferation of computing power have enabled deeper deep learning structures.

Because it can extract accurate data, it saves time and effort during the system design process and connects deep learning algorithms to actors. Outside of deep learning, developers are also working on algorithms that use musical evolutionary mechanisms to precisely simulate the way humans think. Therefore, deep learning techniques are not black boxes. In this research, the translation process of the user-created system can be implemented as follows.

4.1 Problematization

Once considered to be the exclusive domain of humans, advanced AI technology now makes creative work possible. It is necessary to develop an AI composition system to lower the high barrier to entry for music creation, not just for a few professionals. As a result, the network will be constructed through the creation of a composition system that anyone can use creatively as a required passage point. To this end, two problems are proposed.

First, is it possible to use deep learning technology from any of the algorithms employed in the system's development? Second, is it possible to create non-classical pop, jazz, and background music using data from existing songs? Third, can non-experts utilize it?

Human brain neural networks were implemented on computers as part of artificial intelligence algorithms, offering the possibility of replacing human intelligence; consequently, machine learning methods were frequently used in creative development. Due to the limitations of accurate classification and rule generation, it was challenging to construct a flawless system. Nevertheless, with the advent of deep learning technology, it is possible to accurately classify existing data and generate rules, in contrast to the current machine learning methods.

As a consequence, numerous programmers are attempting to incorporate deep learning methods into non-human actors in order to create more comprehensive systems. In addition, the previous system design attempted to construct actors based on pop or jazz, background music such as movie music or game music, etc., based on the high utilization and commerciality of users, as opposed to the manner of famous composers centered on classical music.

4.2 Interessement

Interessement is the process of negotiating to bring the actors invoked in the problematization process into one's network for mutual benefit. As a prerequisite for constructing an AI composition system, the deep learning method will be able to establish a computer algorithm-based entity. As a result of the successful application of deep learning methods in other fields, the likelihood of employing them to construct a composition system network has increased. However, since deep learning is based on large data sets, it links existing music data to new actors. It connects audio files, spectrograms, and midi files, of which midi files can reduce the time and effort required to learn existing music data, making them essential training data and crucial network actors.

4.3 Enrollment

As actors in the network are given new roles, they are enrolled as actors in the developer's network. As an actor in his network, the developer enrolls one of the deep learning algorithms (CNN, RNN, GAN, etc.) and a midi file of pop, jazz, or background music. The issue with existing music is copyright, however. At this point, the developer revisits the problematization stage in order to construct a new network to solve the copyright issue. Using the same procedure as before, the developer decides to generate their own training data as a solution to the copyright issue during the problematization and interessement stages. Instead of the existing music data, the production midi files for each genre are constructed as a new network actor.

Deep learning algorithms and genre-specific midi files are assigned important roles in the construction of an AI composition system, which is successfully constructed. Since the network was created by human and non-human actors, the mobilization process has been completed. However, there are still some developers who pursue and

develop AI algorithms other than deep learning, or who concentrate on less commercial and popular classical music. In these situations, the actor's refusal to construct the network is considered a betrayal. The developer's network is a loose network that can be altered based on the results at any time.

4.4 Mobilization

The actors in the network constructed by the developers - deep learning methods, midi files of various genres, etc. - mobilize successfully when they fulfill their duties, achieving the Obligatory Passage Point of constructing an AI composition system. This network is not yet completely blackboxed and is still under construction. However, this network of developers subtly disrupts the traditional human compositional network's black box.

Table 3. Translation Contents of the Design Process of the Automatic Composition System of Artificial Intelligence

Problematization (Identifying the Problem When Constructing a New Network)	Interessement (Attracting Actors into the Network)	Enrollment (Assigning roles to Attracted Actors in the Network)	Mobilization (Actors Fulfilling Their Roles)
Is artificial intelligence capable of producing original music? Can originality be achieved?	Machine Learning Technology → Deep Learning Technology	CNN RNN GAN	Establishment of an AI compositional system
	Traditional classical music→ jazz, film music, background music	Created Music	
	Midi file	Sound Source	

The four stages of translation in the process of designing an AI composition system can be schematized as follows.

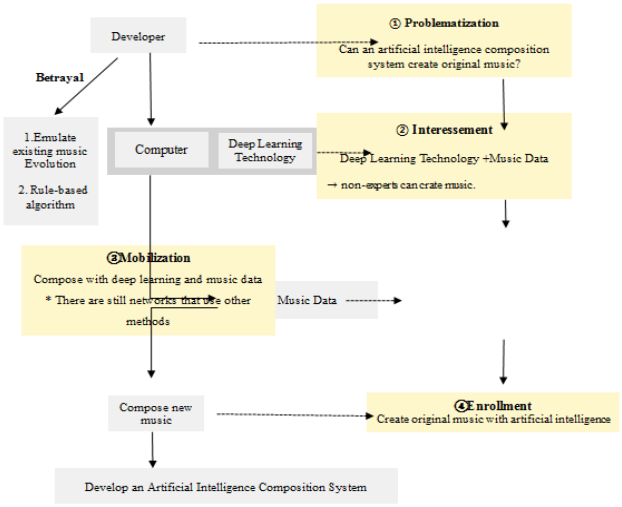

Figure 8. Four Steps to Translation in Designing an AI Composition System

5. CONCLUSION

This study employs ANT to compare humans and machines on equal terms and examines the network of artificial intelligence composition systems created by the shifting roles of human and non-human actors at various levels. In the evolving sociocultural trend of digital development, non-human actors such as technology and machines are interconnected with human actors in a network. In the world, humans and technology are intertwined and interdependent, perpetually producing new production and existence.

AI compositional systems, which have emerged as a result of the convergence of humans and technology, have the advantage of enabling anyone, regardless of musical expertise, to create music swiftly and easily. Users can express their latent artistic creativity within the system by modifying and augmenting it according to their requirements, and they can express more original music with the assistance of a portion of the composition process.

Technology can be regarded as a tool, however, because music creation tasks are performed by machines rather than humans. A dichotomy and hierarchical relationship exists between humans and technology. Nonetheless, if one examines the network of systems, it becomes apparent that neither human nor non-human actors can accomplish anything without the other. This is due to the fact that the creation of new concepts or alterations by machines transcends human consciousness. Creativity in the process of designing a system and creativity in the process of making adjustments under given circumstances can be viewed as heterogeneous effects induced by non-human actors such as machines and technology.

It is indisputable that machine-generated music is the result of the interaction between humans and non-humans and

represents an extend of the creative process that was previously exclusive to humans. Due to the development of ever-evolving technologies, current AI compositional systems are not yet stable and robust black boxes. They are still betrayed by their resistance to new creative methods and the influx of machine-generated music. By constructing a network of large and small AI compositional systems, the network is not wholly lost.

Human actors perform a crucial role in network construction. Human actors are also responsible for connecting non-human actors to the network. In order to build and expand a solid network for a common purpose, it will be necessary to understand the interests of each actor and to continuously monitor them. Therefore, it is necessary to continually contemplate how humans, machines, and technology can coexist.

ACKNOWLEDGEMENT

This research was supported by a 2022 Research Grant from Sangmyung University.

REFERENCES

[1] ChatGPT. (2013). https://openai.com/blog/chatgpt.
[2] Latour, B., (1987). Science in action : how to follow scientists and engineers through society. Harvard University Press, ISBN-10 0674792912 ; ISBN-13: 978-0674792913, 288, 133-145.
[3] Kim, H. S., (2009). Two Cultures, Technology, and Relational Ontology. Moonhwagwahak, 392, 40-60.
[4] Kim, H.S., (2011). Technoscience and democracy from the perspective of ANT. Journal of Korean Social Trend and Perspective, 83, 11-46.
[5] Hong, S.W., (2010). ANT with 7 themes. Eum, ISBN: 978-89-93166-22-4, 336, 15-36.
[6] Alpern, A., (1995). Techniques for algorithmic composition of music. On the web: http://hamp. hampshire. edu/adaF92/algocomp/algocomp, 120, 2-3.
[7] Jung, D. H., (2020). AI Business Model Ends in One Volume. Cheongram, ISBN; 9788935213177, 332, 47-61.
[8] Choi, Y. S., (2023). A Study on Generative Music UsingActor Network Theory. Sanmyung University, 180, 89.
[9] Choi, Y. S., (2023). Ibid., 151.
[10] Kim, J.H.,(2020). AI's Strongest Class. Maekyung Publishing, ISBN; 9791164841950, 374, 190-191.