LUISS

Corso di laurea in Giurisprudenza

Cattedra

di Intellectual Property

The insurmountable rise of Generative AI in the music industry and its challenges for copyright law

Prof. Christophe Geiger

RELATORE

Prof. Paolo Marzano

CORRELATORE

Jazmín Chloé Greco – Matr. 154123

CANDIDATO

Anno Accademico 2023/2024

하늘에서 떨어진 별인 줄 알았어요

소원을 들어주는 작은 별

TABLE OF CONTENTS

INTRODUCTION

It is the year 2022, and a new phenomenon is launched into the market, taking the world by storm. The LLM model ChatGPT, launched by OpenAI, has democratised and advertised to the public the power and capabilities of Generative AI in the present era. Every digital user questions whether it is a reliable service to use and how much its answers can be trusted.

The strong impact ChatGPT had on the global market, however, cannot compare to the immense efforts and research done in the field long before the explosive phenomenon. The world of LLMs is just the tip of the iceberg when talking about artificial intelligence, and a deeper insight is necessary to understand the roots, psychology and functioning of such an interesting, yet risky, invention.

The focus of this discussion, however, will shift to a specific market and area of the broader subject of Generative AI. It will cover the connection of music and artificial intelligence, an argument that may sound very new and unexplored – though that is not the case – and it will discuss what kind of impact it has on copyright frameworks or, seen in another point of view, how many different copyright frameworks can cope with this new and (not so) unexpected phenomenon.

People worry about AI taking their jobs, but musicians could face something more worrying: AI taking their voice. As the debate around AI regulations rages on, creative industries are taking notice, and music is one of the biggest. The music industry saw the impact of AI-powered voice cloning after Tiktok user @ghostwriter77 released an AI-generated song from Drake and The Weeknd that went viral. It had all the imprints of a Drake song and inspired a sternly worded statement from his record label, Universal Music Group. Then, YouTube took the video down for featuring an unauthorized sample — something that has nothing to do with its AI aspects.

"*At the core of music is math, and every mathematical combination has already occurred in some way, shape, or form. It's the performance of that math that changes depending on the singer or the song style*" Justin Blau, co-founder of Royal and a DJ under the name 3LAU, told *The Verge*. "*Saying something is derivative is a pretty hard argument for copyright owners to make because we all borrow ideas from things that we've heard before. AI just does it at a way faster speed.*"

This example of voice cloning is just one of the many that raises questions and doubts on the balance between authors' copyright and the goal of scientific and cultural progress that needs to be made in the context of artificial intelligence and its impact in many fields.

Before diving deep into the examination of the subject of interest, a brief summary will be made on the structure and contents of this thesis.

The first chapter introduces the discussion by giving a thorough and detailed outlook of artificial intelligence, its history and psychology and its current structure and methods of training and existing models. Then it dives deeper into the world of Generative music, by giving definitions, describing its birth and explaining its impact on the music market. An explanation of generative music techniques is further explored, while also mentioning the type of outputs available in generative music models. Last, but not least, a chronological history of generative music is described with examples and cases to better understand its evolution and focus during the years, from being a research study subject in academic settings, to being exploited and marketed thanks to the rise of startups and big tech companies.

The second chapter will start the main discussion of the thesis that is carried out on two fronts. The first, which will be Chapter 2's main focus, is the debate of the legality of AI training; while the second, discussed in Chapter 3, will analyse the copyrightability of AI-generated output, after it has been trained and tested. The training process of an artificial intelligence model is a fundamental, if not the most important thing, step for the design of a particular AI model. The training will influence and direct the output that will finally be generated through inputs and prompts by developers and users. The concept of AI training, however, finds its roots in fundamental and human rights stated in international conventions such as the Berne Convention, the Universal Declaration of Human Rights, the WIPO Treaty and the International Covenant on Economic, Social and Cultural Rights. The possibility and right to train an AI model find their existence and justification in right such as the right to scientific research and cultural progress or the right to freedom of expression and information that, however, need to be balanced with other important and "clashing" fundamental rights such as the right to intellectual property and copyright. The first part of the second chapter will discuss the roots and philosophy behind the drafting of such important principles, and in the second part there will be a global overview of AI and copyright regulations from western and eastern outlooks on AI training and text and data mining, which will indicate very different approaches on developers' freedom when designing an AI model.

The third and last chapter will have an emphasis on the second half of the debate mentioned previously. After discussing the legality of AI training, the first thought that comes to mind is to examine whether copyright applies to AI-generated outputs. On this still fresh and emerging matter, the US Copyright Office has weighed in. With a fundamental and preparatory analysis of copyright principles – which are largely similar worldwide, with slight variations – the US Copyright Office registration refusal of an AI-generated image piece gives a cue on how copyrightability can be definitely assessed when talking about outputs generated by an artificial intelligence model. On the other hand, a few judicial cases, including *Li v. Liu*, will be examined to support or refute this perspective and highlight differing approaches across jurisdictions. Moreover, a paragraph will be dedicated to the analysis – even if at its primordial stage – to the first lawsuit brought by music labels against AI music generator models to defend the rights of the artists protected by such music companies. Finally, there will be a discussion on the proposal of a new right, a limitation-based remuneration right, to finally establish a balance between the creation of AI-generated outputs that take big inspiration from already existing protected works and the rights of authors and musicians that create works that are used for the training of such AI models.

CHAPTER 1

THE GENERATIVE AI PHENOMENON: HISTORY AND CASES OF GENERATIVE MUSIC

1.1. History and Psychology of Artificial Intelligence

Before the AI phenomenon could take the world by storm, making every digital user concerned about its safety and correct functioning, it was thoroughly studied and discussed for a very long time. Contrary to everyone's beliefs, AI was not born in 1950 with the famous question "*Can machines think?*"[1] posed by the well-known mathematician Alan Turing; instead, the first ever example of mathematical calculations made by a machine was presented by William Schickard in 1623, though it was not automatic. Automation in calculations was then reached by Blaise Pascal – with his *pascaline* – a few years later, and Gottfried Wilhelm von Leibniz through the creation of the *stepped reckoner* in 1673, a calculator that could perform all four basic arithmetic operations. Nonetheless, the most famous and recognised character in the world of AI is the already mentioned mathematician Alan Turing. The journey before reaching his most famous seminal paper, "*Computing Machinery and Intelligence*", was preceded by various studies and introductions to the concepts of calculability and computability[2], which are still of extreme importance in our present days. On the other side of the planet, a few years before the publishing of the Turing test, the first ever contribution regarding artificial intelligence was presented by the mathematicians W. McCulloch and W. Pitts through a study which introduced the first ever mathematical model of a neural network[3]: the unit of this model, a simple formalized neuron, is still the standard of reference in the field of neural

---

[1] In an attempt to explain the "imitation game", Turing stated: "*I propose to consider the question, "Can machines think?" This should begin with definitions of the meaning of the terms "machine" and "think." The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words "machine" and "think" are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, "Can machines think?" is to be sought in a statistical survey such as a Gallup poll. But this is absurd*."

[2] A mathematical problem is *computable* if it can be solved in principle by a computing device. Some common synonyms for "computable" are "solvable", "decidable", and "recursive". Hilbert believed that all mathematical problems were solvable, but in the 1930's Gödel, Turing, and Church showed that this is not the case. There is an extensive study and classification of which mathematical problems are computable, and which are not. In addition, there is an extensive classification of computable problems into computational complexity classes according to how much computation—as a function of the size of the problem instance—is needed to answer that instance.

[3] In the words of McCulloch and Pitts: "*Many years ago one of us, by considerations impertinent to this argument, was led to conceive of the response of any neuron as factually equivalent to a proposition which proposed its adequate stimulus. He therefore attempted to record the behaviour of complicated nets in the notation of the symbolic logic of propositions. The "all-or-none" law of nervous activity is sufficient to ensure that the activity of any neuron may be represented as a proposition. Physiological relations existing among nervous activities correspond, of course, to relations among the propositions; and the utility of the representation depends upon the identity of these relations with those of the logic of propositions. To each reaction of any neuron there is a corresponding assertion of a simple proposition. This, in turn, implies either some other simple proposition or the disjunction or the conjunction, with or without negation, of similar propositions, according to the configuration of the synapses upon and the threshold of the neuron in question*."

networks, taking the name of its creators, the McCulloch-Pitts neuron. Through their seminal paper, McCulloch and Pitts managed to demonstrate that every computable function can be represented by a network of neurons and logical connectives can be implemented through a more complex neural structure. Eight years later, in 1951, two students from Harvard University, Marvin Minsky and Dean Edmonds, inspired by the work of McCulloch and Pitts, created the first ever neural net-machine containing a number of synapses that – because they contain a certain amount of memory – hold the probability of receiving an output signal given an input signal: this machine, called SNARC (Stochastic Neural Analog Reinforcement Calculator), is considered one of the first pioneering attempts at the field of artificial intelligence. Following the great invention of SNARC, in the summer of 1956 a convention was held at Dartmouth College, the Dartmouth Summer Research Project on Artificial Intelligence, which is considered to be the founding event of artificial intelligence as a field. During the eight weeks of the brainstorming session, the experts discussed the possibility of simulating through a machine every aspect of human learning or thinking: the proposal of the project stated: "*An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer*", paving the way for discussion on computers, natural language processing, neural networks, theory of computation, abstraction and creativity.

Almost twenty years later, in 1971 two professors of the Carnegie Institute of Technology (now Carnegie Mellon University), H.A. Simon and A. Newell – the same creators of the program Logic Theorist[4] (LT) in 1955 – published a document titled "*Human Problem Solving: The State of Theory in 1970*"[5] in which they explained the theory of information elaboration psychology and their strategy to develop it in a time frame

[4] The *Logic Theorist*, developed by Allen Newell, Herbert A. Simon, and Cliff Shaw in the mid-1950s, is often regarded as one of the first artificial intelligence programs. This groundbreaking software was designed to replicate human problem-solving abilities by proving mathematical theorems, specifically targeting those in *Principia Mathematica* by Alfred North Whitehead and Bertrand Russell. The program utilized symbolic representation and heuristic methods to efficiently search for proofs, marking an early use of heuristic problem-solving in AI. Remarkably, it successfully proved 38 of the first 52 theorems in the text, sometimes offering more elegant solutions than the original authors. The Logic Theorist's ability to automate complex logical reasoning highlighted the potential of computers to perform tasks requiring human-like intelligence, laying the foundation for the emerging field of artificial intelligence.

[5] The of information elaboration psychology described in the publication examines how humans process and transform information to solve problems. According to this theory, problem-solving involves encoding problems into mental representations, which significantly influence the ease of finding solutions. Rather than exhaustively searching for answers, individuals employ heuristics—strategic shortcuts that simplify complex problems by focusing on the most promising paths. The process is sequential, with each step involving the elaboration of information, the generation of new insights, and the refinement of approaches based on feedback. Problem-solving is conceptualized as a search within a problem space, where each state represents a potential configuration of the problem, and the solution emerges through navigating this space. This theory bridges cognitive psychology and artificial intelligence, highlighting how computational models can mimic human cognitive processes to provide a deeper understanding of problem-solving behaviour.

of fifteen years. The main source of the paper is the creation of an artificial intelligence program, like LT, using the "think-aloud" protocol – a testing method through which experts talk and express their opinions, thoughts and problems in the interaction with a machine – with the help of experts capable of solving various problems of logic, arithmetics and so on. This method has then been the most efficient and most used in training artificial intelligence, especially to obtain information on human behaviour in the field of social sciences. The main question posed in the paper was "How can a human being learn how to solve problems?": attempting to answer the question, the authors developed the theory of production systems through which every human solves and answers questions and problems. In every production system, the routine has a bipartite structure consisting of a condition and an action. The condition defines a test: if the test is satisfied, the action can be executed; if not, the test is transferred to another production system. The study of this thesis continued in the following years and by 1978, with the help of the psychologist John R. Anderson, Simon and Newell unfolded the ACT[6] theory (Adaptive Control of Thought) on human cognition, later explained in the publication "*The Architecture of Cognition*", in which Anderson tries to understand the principles behind the control of thought in a way that exposes the adaptive function of these principles. In the following years, and especially with the beginning of the following decade, starting from 1980, the study and research in the field of AI has touched many other subjects, such as education and psychology, in which it has helped their improvement, giving researchers other points of view that they never thought to be feasible. Nowadays intelligent systems are present in every field, especially in daily activities, manifesting what researchers, many decades before, thought would happen in the future. There are programs, such as DeepBlue, capable of competing on par with chess champions; others are used in the field of astrophysics like the Remote Agent program, that was used by NASA in 1998 to manage the activities regarding space systems; the concept of automated cars that are capable of driving without the human input is not so strange today. Through the lens of human experience, the revolutionary impact of AI has covered every aspect of life, with little or even no notice at all: still, its journey is still long and filled with many uncertainties, in which its future is still mostly uncertain.

---

[6] Anderson, J.R., "The Architecture of Cognition.", 1983. This theory posits that human thought is organized into three types of memory systems: declarative, procedural, and working memory. Declarative memory stores factual knowledge, while procedural memory contains the knowledge of how to perform tasks. Working memory serves as a temporary workspace where information from both declarative and procedural memory is accessed and manipulated. ACT theory emphasizes the interaction between these memory systems in cognitive processes. Learning is viewed as the acquisition and refinement of both declarative and procedural knowledge, often transitioning from the former to the latter as skills become more automated. The theory also incorporates production rules—if-then statements that guide behaviour and decision-making based on the current context.

1.2. What does AI mean today?

1.2.1. Input and Training

In this brief paragraph the historical phases of development of AI training will be analysed, with a particular focus on the concept of the importance of the quality of the data given to feed the code. During the early stages of AI – specifically from 1950 until 1980 – the method of machine learning was described as "symbolic" or "old style", following the theories proposed and developed by Newell and Simon. In current days, however, the approach adopted by researchers and developers is in contrast with the previous one: in fact, the method is named "non-symbolic" or "sub-symbolic". This means that the approach is focused on automated learning through the method of Deep Learning (DL): such method is based on the theory of connectionism, which is the name of the approach to the study of mental processes and cognition that utilizes mathematical models, known as connectionist networks or artificial neural networks. Before focusing on the training process of AI it is fundamental to understanding how it works and what it takes to make it function in a way that generates outputs and, in an IP view analysis, operates.

1.2.1.a) Machine Learning

To give an extremely broad yet concise definition of artificial intelligence – on which there will be further discussion – such is considered the field of computer science that uses data to enable machines to be capable of having problem solving qualities. Let's imagine AI as a diagram: the term artificial intelligence itself is very vast and can indicate various things; it would be the bigger circle in the diagram that includes various other typologies of AI and even its training methods. Inside the AI diagram, the first sub-diagram that can be pictured is about Machine Learning: it is the broadest training input by which an AI could function. To be brief, a general machine powered by artificial intelligence is a machine in which the programmer gives a set of instructions by manually coding each decision. With ML the programmer "steps up the game" of the machine: basically, with this model, computer scientists train the machine by feeding it large amounts of data; then, the machine follows a set of rules – called algorithms – to analyse and draw inferences from such data[7]. The more data a machine parses, the better it can become at performing a task or taking a decision. Practically, Machine Learning models can learn and adapt automatically from experience without the need to be manually and explicitly programmed: these are used with the purpose of recognizing patterns

[7] Sarker, I.H. "Machine Learning: Algorithms, Real-World Applications and Research Directions", SN COMPUT. SCI. 2, 160, 2021. Iqbal H. Sarker emphasizes the pivotal role of ML in analysing vast datasets across various domains, including cybersecurity, healthcare, smart cities, and e-commerce, thereby enhancing the intelligence and automation of applications in these fields. He also addresses current challenges in ML, such as data privacy concerns, interpretability of models, and the need for real-time processing capabilities, suggesting potential research directions to tackle these issues. Overall, Sarker's work underscores the transformative impact of machine learning in the digital era and calls for continued exploration to address existing challenges and expand its applicability.

in a vast amount of data – known – and to make predictions on new or unknown data. In such a brief amount of time, Machine Learning has become one of the most popular[8] latest technologies in the fourth industrial revolution and is being utilized to transition to automation in various sectors such as conventional manufacturing and industrial practices. As a preface to further discussion, when discussing the training artificial intelligence algorithms, it is vital to take into consideration the importance of the quality of the data that is fed to the algorithm. In fact, the effectiveness and the efficiency of ML solutions depend on the nature and characteristics of data and the performance of learning algorithms.

ML models can be categorized in different approaches (or learning techniques):

- **Supervised Learning:** in this approach, the model is trained on a labelled dataset, where each input is associated with a desired output. The most common supervised tasks are *classification*, which separates the data into categories, and *regression*, that fits the data. A good example is found in email classification systems, when it categorizes spam/non-spam emails. The goal of this model is to learn a function that maps the inputs to the correct outputs. A specific type – one of the many existent – of Supervised Learning AI is, as already hinted in the history of artificial intelligence, Artificial Neural Networks (ANN) that are structured based on the human brain, and they are primarily used for image recognition and Natural Language Processing (NLP – such as chatbots, and translation services).
- **Unsupervised Learning:** in this model, the system works on unlabelled data and must identify hidden patterns or underlying structures without being guided by human interference and known outputs (i.e. *data-driven processes*). The most used application of this ML model is especially seen in anomaly detection, where the agent identifies data that do not follow the general pattern of the dataset given for training.
- **Semi-Supervised Learning:** as the name suggests, this model combines – commonly – a large set of unlabelled data and a small quantity of labelled data. It is a *hybridization* of the above mentioned supervised and unsupervised techniques. A perfect example is speech recognition, where only part of recordings are labelled. The ultimate goal of a semi-supervised learning model is to provide better outcome for prediction than that produced using labelled data alone from the model.
- **Reinforcement Learning:** this kind of approach requires the agent to learn behaviours in a dynamic environment by performing actions and receiving rewards or penalties. Here the machine has the

[8] [EC, Artificial Intelligence for Europe, COM (2018)237 final, p.10]. "Machine learning is the dominant AI technique disclosed in patents and is included in more than one-third of all identified inventions (134,777 patent documents). Filings of machine learning-related patent have grown annually...", WIPO, 'WIPO Technology Trends 2019' (n 22) 14.

goal to maximise the cumulative reward over time through learning and using insights obtained in the environmental-driven approach. This model application is a powerful tool for increasing automation and it is commonly used in computer games like chess or Go and, of course, in robotics and autonomous driving.

As for the development process and setting of a ML model, there are various phases to follow in order to obtain the correct implementation of the task set to create such machine. Firstly, the data used for training need to be collected and pre-processed before training: the pre-processing consists of cleaning, transforming and splitting all the data into training, validation and test sets. Once the pre-processing is complete, the actual training process can begin: during this phase, the machine learns to map inputs into outputs based on the training instructions decided by the computer scientist. Of course, the training process need to be validated through a technique called "tuning", thanks to which the model is adjusted based on the parameters given, to avoid overfitting. Finally, the model is then tested again on a separate dataset to appraise its real-world performance (evaluation).

Machine Learning models are the ones that are most commonly used in digital services due to their simpler training and functionality. Since their tasks are considered rather straightforward, their functions can vary from image recognition to NLP, including one of the clearest examples for digital users that is personalised recommendations in music (e.g. Spotify), shopping (e.g. Amazon) and audiovisual content (e.g. Netflix). ML represents one of the first and most advanced frontiers in modern technology, with enormous potential across various sectors. Nonetheless, it presents flaws and ethical challenges that need to be addressed and regulated to a correct use of such services: ML models are already known to amplify biases in data fed for training, leading to unfair and discriminatory decisions; privacy is another concern that is extensively discussed by experts, especially in healthcare sector and, of course, social media. The biggest concern to be looking out for is the labelling of such models as "black boxes", since it's rather common that it is difficult to understand and explain the reason behind the decisions that are made.

1.2.1.b) Deep Learning

Deep Learning is a subset of a wider family of artificial neural networks-based machine learning that focuses on neural networks with many layers, known as deep neural networks. These networks are designed to mimic the structure and function of the human brain[9], allowing computers to learn from vast amounts of

[9] At the heart of deep learning is the concept of artificial neurons, which are modelled after the neurons in the human brain. Each artificial neuron receives input, processes it through an activation function, and passes the output to the next layer of neurons. The strength of the connection between neurons is represented by "weights," which are adjusted during training to improve the model's accuracy. Just as the brain strengthens or weakens synaptic connections through experience, deep learning algorithms

data in a way that can recognize complex patterns, such as images, sounds, and text. It is possible to distinguish neural networks from deep neural networks: the first ones are layers of interconnected nodes through which input data is processed and passed onto the next layer; the training happens in the so called “hidden layer”, that applies a nonlinear function to the input data and passes it to the next layer creating an output. On the other hand, deep neural networks are neural networks with more than one hidden layer: the increased depth of the nodes allows a hierarchical analysis of the data between the different layers (e.g. in image recognition, a deep neural network the different layers – lower, middle and upper – can identify different parts of the picture such as edges, shapes, and complex structures like objects and faces). The training process of such approach is evidently different from the one applied to ML models, since it requires additional steps such as backpropagation and gradient descent. At first, input data is fed to the machine to analyse it and create an output: such output (actual), thanks to the additional steps mentioned, is compared to the predicted output, with the computing of the errors; these errors are then propagated back into the system in order to allow its upgrade. The difference between ML and DL models is pretty clear after this description: while ML models are made of a simpler structure, they require smaller datasets and computational powers (a simple CPU is enough for ML to work) and are preferred to complete simpler tasks, DL models are more complex, require vast datasets and computational powers (GPU or TPU to manage the larger number of calculations) and are programmed to handle more complex tasks like computer visions and autonomous systems. In fact, the most notorious examples of Deep Learning machines are convolutional neural networks (CNNs) – used for tasks such as image classification, object detection and facial recognition – and automated machines like self-driving vehicles.

Deep learning represents the cutting edge of machine learning, enabling breakthroughs in various fields that were previously impossible to reach. While it shares common ground with traditional machine learning, it distinguishes itself through its ability to automatically learn complex features from large datasets, its reliance on high computational power, and its exceptional performance in tasks involving unstructured data like images, text, and speech. However, the power of deep learning comes with challenges, including its need for large datasets, high computational costs, and difficulties in interpretability. Despite these challenges, deep learning continues to be a driving force in the advancement of AI.

Footnote 9 continued
adjust these weights using a process called backpropagation, which helps the network learn from errors and optimize its predictions over time.

1.2.1.c) Natural Language Processing and Large Language Models

Another sub diagram in the field of Artificial Intelligence is certainly Natural Language Processing, which constitutes an extremely important characteristic of current AI software. This subfield focuses on the interaction between computers and human (natural) language, with the goal of enabling machines to read, understand, interpret and generate human language in a way that is meaningful, understandable and useful. This can be tricky, especially when considering natural language, because natural communication often appears ambiguous, context-dependent and extremely complex for rigid and structured language that only computers understand. Nevertheless, the scope of closing the gap between computers and humans is helpful for generating and supporting the growth and progress of human knowledge, thanks to the contribution of principles from linguistics, computer science, artificial intelligence and cognitive psychology. The importance of NLP, as stated before, is crucial in AI systems, as it allows the analysis and understanding of textual and spoken data, broadening even more the system's computational power and data availability for training purposes.

To better understand its potentialities, it is important to understand how NLP is composed and its process on analysing data in natural language. First and foremost, NLP involves several key components, each addressing a specific challenge related to language understanding and generation. There is syntax, that refers to the grammatical structure of the language, focusing on the arrangement of words and phrases to create correctly formed sentences: this quality requires the analysis of sentence structure to extract the meaning and identify the relationship between words. The second component of NLP systems is semantics, and it focuses on understanding the meaning of the words and text: a first challenge appears as this skill is necessary to handle ambiguities in natural language, such as polysemy and homonymy. Thirdly, there is pragmatics, which addresses the context in which language is used, including understanding intentions, social background and situational factors of the speaker: this proves essential in performing tasks like question-and-answer dialogue systems, where context and prior knowledge are necessary to interpret user queries effectively. Discourse allows the processing and understanding of larger bodies of text, like paragraphs or entire documents, to extract the relations of sentences between one another. At last, there is phonology and morphology, fundamental in speech recognition and synthesis, though it is less directly involved in textual NLP.

A discussion on NLP cannot be made ignoring its latest evolution, Large Language Models, which constitutes a subfield of NLP, representing a significant advancement in the field. LLMs, as evolutionary as it is currently presented, sets the foundations for Generative Artificial Intelligence thanks to the union of NLP and DL, which allows the understanding and generation of human-like text. These models, such as

OpenAI's *Chat GPT* series and Google's *Bard*, rely on transformer architectures for processing sequential data and capturing intricate linguistic patterns, enabling nuanced language generation. Trained on diverse datasets encompassing books, articles, and web content, LLMs excel in applications ranging from content creation to conversational agents. However, their operation raises critical questions concerning ethical considerations, computational resource demands, and intellectual property rights, especially as their outputs increasingly intersect with human creativity[10]. The implications of LLMs extend beyond technical innovation, presenting challenges in regulating their use and addressing biases inherent in training data[11]. As generative AI continues to evolve, LLMs remain central to discussions on the balance between technological progress and societal impacts. Moreover, since it is the closest technological discovery to imitate human creativeness, it raises serious philosophical and ethical questions that intersect human and algorithmic creativity, threatening – apparently – the displacement of human creators by AI-generated works. In essence, LLMs can be seen as a transformative application of NLP principles, pushing the boundaries of what NLP systems can achieve while also highlighting areas requiring ongoing research and ethical scrutiny, thanks to its massive predictive power[12].

1.2.2. Output

In the context of artificial intelligence, outputs refer to the results produced by an AI system after processing input data. These outputs can take various forms depending on the type of AI model, the nature of the input, and the specific task or application for which the AI is designed. Understanding the concept of AI-generated outputs is crucial for evaluating the effectiveness, reliability, and ethical implications of AI systems in general. There are different types of outputs that an AI can produce through its training, such as:

- **Predictive outputs:** are forecasts or estimations generated by AI models based on patterns identified in historical data. These outputs are widely used in various fields, including finance and

[10] These models, due to their large scale and complexity, can unintentionally reinforce biases, spread misinformation, or be misused in harmful ways. Furthermore, the high computational costs associated with training these models raise concerns about environmental sustainability and accessibility. The authors emphasize the need for careful consideration of these risks and suggest strategies for mitigating them, such as improving model interpretability, ensuring fairness, and developing frameworks for responsible AI usage.

[11] One key point discussed is the concept of *stochastic parrots*, where these models essentially mimic and regurgitate patterns found in the data they are trained on, without truly understanding the content, leading to concerns about the reinforcement of biases and the propagation of harmful stereotypes.

[12] *See* Chen, Y. et al. “Expected Returns and Large Language Models” (2022), SSRN, pp. 1-62. A study conducted in 2022 by Yifen Chen, Brian Kelly and Dacheng Xiu on the application of LLMs in financial forecasting demonstrated the astonishing capabilities of these systems through the extraction of contextualized representations from news articles to predict stock returns, aiming to enhance the accuracy of financial predictions by leveraging the advanced capabilities of LLMs in understanding and processing complex textual information.

marketing (i.e. prediction of stock prices based on previous market data) or healthcare (i.e. prediction of a patient's disease progression).

- **Classification outputs:** it involves categorizing input data into predefined classes or labels. AI models that produce classification outputs are commonly used in tasks such as image recognition, spam detection, and medical diagnosis.
- **Generative outputs:** create new data that resembles the input data they were trained on. These outputs can include text, images, audio, or even more complex data like 3D models. The most famous and well-known example of Generative AI is *ChatGPT* (General Pre-trained Transformer) a chatbot and virtual assistant developed by OpenAI. It is based on large language models (LLMs) and enables users to refine and steer conversations towards desired length, format, style, level of detail, and language. ChatGPT uses deep learning to generate human-like text based on user prompts. Following the discussion, there will be a more in-depth analysis on Generative AI and its connection with Intellectual Property.
- **Decision-making outputs:** AI systems designed for decision-making produce outputs that guide or determine actions. These outputs are typically used in autonomous systems, such as self-driving cars or automated trading systems. Moreover, it is a thoroughly discussed topic in the legal environment, especially in the sentencing and parole decisions. Nonetheless, it has various applications that could range from a simpler task (i.e. legal research and case analysis), to a more complex one (i.e. contract drafting and litigation prediction). It will be discussed further in detail on such applications and their consequences in the juridical world.
- **Regression outputs:** are continuous values predicted by AI models, often used in quantitative tasks where the goal is to predict a numerical value. A clear example could be the prediction of house prices based on location, size and number of bedrooms.

After briefly describing the typologies of outputs that an AI can produce, these need to manifest certain characteristics[13] in order to be considered acceptable. First of all, AI-generated outputs' quality needs to be

[13] *See* Floridi, L. & Cowls, J. "A Unified Framework of Five Principles for AI in Society.", Harvard Data Science Review (1), pp. 1-14. Luciano Floridi and Josh Cowls propose a comprehensive framework to guide the ethical development and deployment of artificial intelligence (AI) in society. The authors argue that AI systems should be designed with an emphasis on fairness, accountability, transparency, and inclusivity, to ensure they benefit all of humanity. They outline five core principles: (1) the principle of *beneficence*, which advocates for AI to be developed for the public good; (2) the principle of *non-maleficence*, ensuring that AI does not cause harm or exacerbate existing inequalities; (3) the principle of *autonomy*, supporting individuals' rights to control and make decisions about how AI is used; (4) the principle of *justice*, which emphasizes fairness and equal access to AI benefits, and (5) the principle of *explicability*, ensuring that AI systems are understandable and transparent in their functioning; Floridi, L. "The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities", Oxford, 2022.

evaluated in terms of accuracy – how close the outputs are to true values – and precision – how consistent those outputs are; also, interpretability is a key factor in analysing the nature of the output. The interpretability refers to the extent to which humas can understand the rationale behind the decisions or results shown: some AI outputs, like decision trees, have a very linear structure and are inherently interpretable; on the other hand, more complex AIs like DNNs are considered "black boxes". In fields such as finance, medicine and especially law, the interpretability factor is essential for trust and accountability of the machine. Another very important aspect that an AI should present is generalization. This concept embodies the capability of an AI to produce accurate outputs for new and unseen data, meaning that it is crucial for the model to produce accurate data in a real-world scenario. Poor generalization is considered "overfitting", which occurs when a model performs excellently on the trained data, but poorly on new information. Lastly, the production of AI-generated outputs raises concerns on its ethical and legal implications, especially regarding accountability – whether such output causes harm in particularly important fields (i.e. healthcare or criminal justice) – transparency and misuse (regarding the generation of deepfakes or misinformation, posing important concerns on privacy, security and public trust).

As a final consideration, it is necessary to understand that outputs are a crucial and important part of an AI model, just like input materials, and can potentially offer benefits in the legal environment, from increasing the efficiency and accuracy in legal processes to providing valuable insights and provisions thanks to its enormous computing power. Nonetheless, the integration of AI models in this field must be approached with caution, due to its enormous impact in matters like transparency, accountability, fairness and data privacy, but it can develop into an extremely powerful system for decision-making[14], contributing to a more efficient and just legal system.

1.3. Generative Music: a definition

After a brief introduction on the technical aspects of artificial intelligence systems and their function, it is now time to analyse the topic that is currently discussed in every aspect of our everyday lives, that is generative AI. Generative Artificial Intelligence represents one of the most transformative domains in the field of computer science, offering capabilities that go beyond traditional computational paradigms. At its

[14] *Ibidem,* Floridi argues that AI represents a novel form of agency, distinct from human intelligence, which necessitates a re-evaluation of traditional ethical frameworks. In the context of decision-making, the author highlights concerns such as bias, accountability, and transparency, discussing the potential for AI systems to perpetuate existing biases present in their training data, leading to unfair or discriminatory outcomes. He emphasizes the importance of explicability, advocating for AI systems whose decisions can be understood and scrutinized by humans to ensure accountability, focusing on the necessity of interdisciplinary collaboration in creatin an unitary and ethical framework for AI systems.

core, generative AI involves the use of advanced algorithms and machine learning models to produce original content, including text, images, music, and more, based on learned patterns and structures from data. Unlike conventional AI systems that primarily analyse and interpret existing data, generative AI focuses on the creation of novel outputs, mimicking human-like creativity in its process. This distinction underscores its potential to revolutionize industries such as art, entertainment, healthcare, and education, where creativity and innovation play pivotal roles. The technical foundation of generative AI lies in its ability to model complex probability distributions and generate data samples that align with those distributions.

While the promise of generative AI is immense, its popularity also presents critical challenges, particularly in areas such as ethical responsibility, intellectual property rights, and societal impact. The ability of AI systems to produce content that closely mirrors human creativity raises questions about authorship and ownership, necessitating a re-evaluation of existing legal frameworks. Moreover, the reliance on training data introduces potential biases, which can perpetuate inequities or skew the diversity of generated outputs. Addressing these issues requires an interdisciplinary approach, integrating technical innovation with ethical oversight and policy development. As an integral subset of this broader field, generative music epitomizes the convergence of technological sophistication and artistic endeavour. By leveraging the principles of generative AI, it challenges traditional notions of creativity and expands the horizons of musical expression, as explored in the following discussion.

Generative music represents a groundbreaking approach to musical creation, wherein compositions emerge through algorithmic processes rather than traditional human-driven methods. This field combines principles from computer science, music theory, and artificial intelligence to enable the autonomous production of complex and adaptive musical structures. By leveraging computational systems to create, analyse, and modify musical works, generative music provides a framework that transcends the boundaries of conventional composition. It involves the synthesis of sound and form through automated processes, ranging from simple rule-based algorithms to advanced machine learning systems, allowing for the creation of dynamic and evolving auditory experiences. The conceptual foundation of generative music can be traced back to early experiments in algorithmic composition, which utilized mathematical rules and formulas to guide musical creativity. Pioneers such as Iannis Xenakis[15] and Brian Eno[16] laid the groundwork for this interdisciplinary field by integrating technological innovation with artistic vision. These early efforts explored the potential of computer systems to automate aspects of composition, challenging traditional notions of authorship and creativity. In contemporary practice, generative music has expanded significantly, driven by advancements in computational power and AI, offering unprecedented opportunities to redefine musical creativity. Central to the development of generative music is the use of algorithms, which serve as the foundation for generating patterns and structures. These algorithms may range from deterministic systems, where outputs are predictable based on predefined rules, to stochastic models that introduce elements of randomness and variability. For instance, early implementations of generative music employed probabilistic techniques, such as Markov chains, to dictate transitions between musical elements. These models enabled the creation of compositions that, while algorithmically driven, retained a sense of human-like fluidity and coherence[17]. The integration of artificial intelligence, particularly machine learning, has further transformed the landscape of generative music. Modern AI techniques enable systems to analyse vast datasets of musical works, extracting patterns and stylistic nuances that inform the generative process. Neural networks, including recurrent neural networks (RNNs) and transformer models, have been instrumental in advancing the capabilities of generative systems. These architectures allow for the synthesis of intricate and contextually rich compositions, pushing the boundaries of what automated systems can achieve[18] [19].Generative music extends beyond the realm of static

---

[15] Iannis Xenakis was a groundbreaking composer and theorist who played a pivotal role in the development of generative music. He pioneered stochastic music, using probability and statistical models to create complex, dynamic compositions. Xenakis was an early adopter of computer-aided composition, most notably through his UPIC system, which allowed users to graphically design sound. His innovative blending of mathematics, architecture, and music expanded the possibilities of musical creation, influencing the evolution of generative techniques and inspiring future generations of composers and technologists. *See* Xenakis, I. "Formalized Music: Thought and Mathematics in Composition", Pendragon Press, 1992.

[16] Brian Eno is a renowned musician, composer, and producer, widely credited as a pioneer in the field of generative music. He introduced the concept of ambient music, aiming to create immersive soundscapes that evolve over time. Eno utilized algorithmic processes and chance operations to generate music that could develop continuously with minimal human intervention. His groundbreaking works, such as *Music for Airports*, emphasized mood and atmosphere, influencing the broader understanding of music as an evolving, generative art form. Eno's innovations have left a lasting impact on electronic music and the development of generative techniques. *See* Eno, B. "A Year of Swollen Appendices", Faber & Faber, 1996.

[17] By integrating probability and rule-based systems, this showcases the potential for machines to produce music that resonates with the organic flow and complexity of human-made compositions, paving the way for future advancements in generative music.

[18]*See* Huang, C. et al. "Music Transformer: Generating Music with Long-Term Structure" (2018), International Conference on Learning Representations, pp. 1-14. The *Music Transformer*, utilizing a self-attention mechanism, captures long-term dependencies, allowing for the creation of compositions that maintain structural integrity over time. The model effectively captures long-term patterns, enabling the production of compositions that mirror the continuity and expressiveness of human-created music. This approach significantly improves the ability to generate pieces with sustained structure and natural progression, contributing to the evolution of AI-driven musical creativity.

[19] *Ibidem*, The *Transformer* architecture revolutionizes generative models by eliminating the need for recurrence and convolution. This allows for more efficient parallelization and better handling of long-range dependencies. The model's versatility and efficiency have broad applications, including significant advancements in generating structured, coherent musical compositions.

compositions, encompassing adaptive and interactive systems that respond to real-time inputs. These systems leverage environmental data or user interactions to shape musical outputs dynamically, creating personalized and immersive auditory experiences. For example, generative techniques are employed in video game soundtracks to produce music that evolves in response to gameplay, enhancing the player's emotional engagement[20]. While the technical advancements underpinning generative music are remarkable, they also raise important philosophical and ethical questions. The shift from human-centric to machine-driven creativity challenges traditional concepts of authorship and originality, necessitating a re-evaluation of intellectual property frameworks. Additionally, the reliance on computational systems introduces concerns about bias and the potential homogenization of musical diversity, as generative models are often trained on datasets that reflect existing cultural and stylistic norms. Generative music epitomizes the convergence of technology and artistry, offering a paradigm that redefines the processes and possibilities of musical creation. As the field continues to evolve, it holds the potential to not only expand our understanding of music as a form of expression but also to influence broader cultural and technological landscapes. By examining the methodologies, applications, and implications of generative music, we gain insights into a transformative domain that bridges the gap between human creativity and computational innovation.

From a legal perspective, generative music can be defined as music that is created, modified, or arranged by an algorithm or system that operates autonomously or semi-autonomously, without direct human intervention in the composition of the music itself. The key feature distinguishing generative music from traditional music composition is the role of the algorithm, machine learning model, or artificial intelligence in producing or influencing the output.

Another, and probably, most suitable definition for jurists is offered by AI music engineer and researcher Valerio Velardo, creator of *The Sound of AI*[21], which describes generative music in a vaguer and more general approach. In fact, Velardo says that generative music is "*the art and science of developing computer*

[20] *See* Collins, K. "Game Sound: An Introduction to the History, Theory, and Practice of Video Game Music and Sound Design", MIT Press, 2018. Karen Collins traces the history of game audio, from the limited, iconic chiptunes of early games to the complex, adaptive soundtracks of modern titles and explores the theory behind video game music, emphasizing how sound can shape the player's experience and emotional engagement with the game world. He also looks at sound design practices, including how game music is integrated with gameplay mechanics, often responding dynamically to the player's actions. Its study published in a book (2008) highlights the interdisciplinary nature of game audio, drawing connections between music, sound design, technology, and game theory, ultimately showcasing how video game music has become a key element of the medium's storytelling and immersive experience.

[21] The Sound of AI is a project founded by musical engineer Valerio Velardo, an AI music ecosystem which includes various services such as a YouTube channel, Academy, Community where experts from various fields (i.e. engineers, researchers, students, entrepreneurs) can benefit by the music AI knowledge shared through the various activities followed and also thanks to the Open Source Research. For more information consult: https://thesoundofai.com/

programmes that create music with a varying degree of autonomy" considering that *art and science* means the attempt to replicate what humans can do in a creative setting through the use of an AI software – including the extremely ambiguous meaning of art mixed with technology – and *varying degree of autonomy* describing the difference between AI-generated and AI-assisted works, about which will be further discussed in Chapter 3.

In brief, generative music rises the discussion about intelligent and creative tasks[22]. While the firsts consist of two qualities, which correspondingly are *objective success metrics* – that are quantifiable standards used to evaluate the effectiveness of strategies and actions – and a *well-defined problem description*, building tasks that can be easily achieved by machines (i.e. object detection, translation), creative tasks don't present the same qualities, since they are characterized by variability and ambiguity. In the artificial intelligence field, the combination of these qualities is defined as *computational creativity*, which is harder for machines to produce, and has taken the name of Generative AI.

At last, generative music is only one of the many names that describe the same field, which has been named differently by various important figures during the years. The term generative music is known to have been coined by the famous musician Brian Eno[23]; before the diffusion of machine learning, it was commonly used to refer music created by algorithms as *algorithmic composition*; the term *music metacreation*[24] (MuMe) is introduced by Philippe Pasquier[25]. Other definitions are more general and related to specific fields like *procedural music*, which refers to generation of music in video games, and *music AI*, which focuses solely on the creative impact required by artificial intelligence systems in the creation of an art, such as music.

In the next two paragraph a brief outline will be made on both training and output of generative music and, secondly, on generative music techniques used in current AI systems. Finally, the last paragraph will focus

---

[22] It is important, however, to analyse the philosophical interplay between intelligence and creativity, emphasizing that the two are not merely separate traits but deeply intertwined in complex ways. Actually, creativity often requires intelligence, but intelligence alone does not guarantee creative outcomes. In fact, cognitive processes involved in creative tasks challenge traditional notions of intelligence, suggesting that both must be considered in tandem to fully understand human potential. This perspective highlights the importance of integrating both creativity and intelligence in educational and developmental contexts.

[23] Eno, B. "A Year of Swollen Appendices", cit. The term intended by Eno refers to a description of music that is *ever-different and changing* and which is created by a machine system

[24] It defines the science of machines addressing creative tasks, consisting in a subfield of computational creativity that addresses specifically musical tasks like improvisation, composition and performance. More on the argument *see* Pasquier, P. et al. "An Introduction to Musical Metacreation" (2017), Computer Entertainment, Volume 14, Issue 2, pp.1-14..

[25] Philippe Pasquier is a professor in the School of Interactive Arts and Technology of Simon Fraser University's Faculty of Communication, Arts and Technology since January 2008. There, he conducts both a scientific and artistic research agenda. His research focuses on building deeper theories for endowing machines with autonomous behaviours, with a focus on creative and artistic applications.

on the explanation of the history and evolution of generative music providing examples and cases that will make further discussion more clear.

1.4. Training and generative music output (symbolic v. audio)

It is important to consider that the field of generative music, as recent as it seems – just like generative AI in general – has actually deep roots way before the birth of computers. Nonetheless, through the development of technology and computational power, it has reached the popularity we are aware of today. It is clear – but it should not be taken for granted – that the generative music field necessarily requires the collaboration and integration of different subjects and disciplines, such as: AI, digital signal processing, software engineering, music cognition, music theory, musicology and music information retrieval. This revolutionary field, however, also poses some very distinctive challenges in music representation and output evaluation. With regards to music representation, the use of generative systems does not help in defining the complexity of formalising music: in fact, there is not a perfect methodology to grant a perfect representation, even though various models are available to do so. The second, output evaluation in the context of music seems to be a very dangerous topic to discuss: more specifically, the concept of music output evaluation does not recall the quality of objective success metric due to the fact that it is not something that is objectively measurable in terms of good or bad performance. Another question arising from this particular challenge is: who should be the person or subject in charge of evaluating such output? And what are the characteristics that make one choose one subject or the other (listeners, authors, technologists, musicologists…)? Answering this pivotal question seems exceedingly difficult and also very daring, since music is not an objective matter to be evaluated. Nonetheless, it is crucial to have knowledge of the quality and typologies of music representations available today.

Music representation is the conceptual framework used to encode and convey musical information, allowing for the communication, analysis, and reproduction of music. It serves as a bridge between the abstract ideas of music and their tangible expression, providing a structured way to document and interpret musical elements such as melody, rhythm, harmony, and dynamics. At its core, music representation is about translating the ephemeral nature of sound into a form that can be preserved, studied, and shared. This can take various forms, from traditional notations and digital codes to auditory recordings and graphical visualizations. Each form of representation captures different aspects of music, whether it's the symbolic structure, the acoustic properties, or the performance nuances.

By formalizing music into representational systems, it becomes possible to analyse its structure, reproduce performances accurately, and innovate within musical creation. Music representation is thus an essential concept in both musicology and music technology, underpinning the ways we understand, interact with, and advance the art of music. In a practical point of view, the author of the musical piece appoints all the information necessary to perform such piece. Then, the same author will represent what is written on paper into pressure air waves (sound). Finally – and it is considered a third level of representation – the listener, who plays a pivotal part, contributes to music representation through brain waves[26].

The music representation goal for generative music models follows the objective of encoding music in a digestible format for machines, consisting of objectivity and quantifiability, easy manipulation, acquisition of all musical details and compactness – meaning that it does not have to be described with a lot of parameters. In fact, if each, or most, of the qualities are satisfied, it is possible to qualify that as a good music representation and surely can solve most of the problems that generative models face without a good input.

After a general description of what music representation is and how it can be described as "good" for generative models, it is now time to discuss what typologies of representation can be offered through generative music. First, and also the oldest up to date, is the symbolic representation. In this type of musical depiction, symbols – or "tokens" – are used to represent music into notes or instruments. This kind of representation is very much similar to a traditional music score[27], but in a digital format.

Symbolic representation generally describes different typologies of its concrete realisation. A perfect example of symbolic representation is the MIDI (Musical Instrument Digital Interface) format, which is a protocol that leverages hardware – it literally sends real-time input to instruments in order to receive the expected result – allowing software and hardware to communicate and synchronize with each other, leveraging both traditional and innovative methods to produce music.

[26] *See* Bellier, L. et al. "Music can be reconstructed from human auditory cortex activity using nonlinear decoding models." (2023), PLOS Biology 21(8), pp. 1-27. In this study, researchers explored how the human brain processes music by recording neural activity from 29 patients as they listened to a Pink Floyd song. Using advanced computational models, they successfully reconstructed the song from these brain recordings, demonstrating that it is possible to reproduce music based solely on neural data. The findings highlighted the significant role of the superior temporal gyrus, a brain region, in perceiving music. This research enhances our understanding of how the brain interprets complex sounds and could inform future developments in brain-computer interfaces.

[27]A musical score is a written or printed representation of music. It typically uses standardized symbols and notations to convey musical elements such as pitch, rhythm, dynamics, and articulation. A score serves as a detailed blueprint for musicians, providing all the necessary information to perform a piece of music accurately.

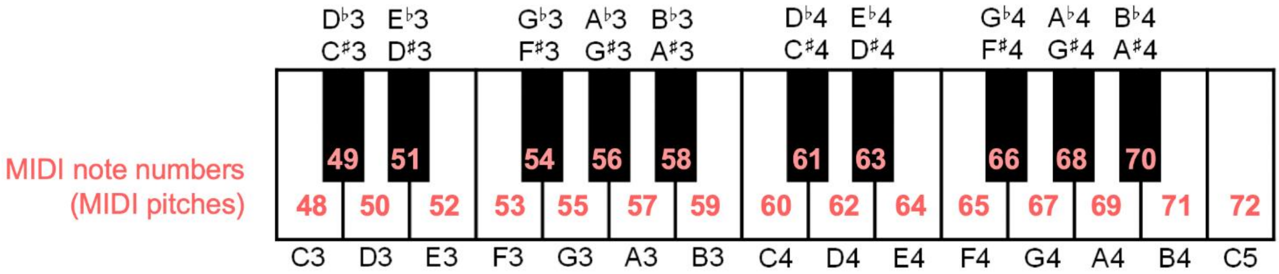


Fig. 1 Example of a MIDI representation of a piano keyboard (via audiolabs-enlargen.de)[28]

Another example is the Music XML format, which is a simple-structured xml file that is tailored to be used to represent music. It serves as a universal standard for sharing and distributing musical scores, enabling interoperability between different music software and hardware. Music XML captures detailed information about musical compositions, including pitch, rhythm, dynamics, and notational elements, making it a powerful tool for composers, performers, and music technologists.

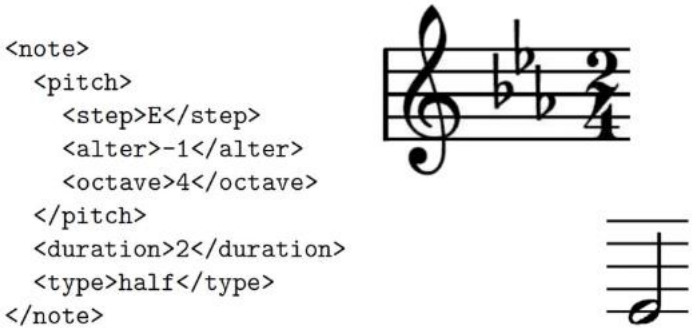


Fig. 2 Example of Music XML format (via Müller, Fundamentals of Music Processing, Springer, 2015)

As a final example, there is ABC notation, a simple and compact text-based music notation system that uses plain ASCII characters to represent musical scores. It is particularly popular for transcribing folk, traditional, and early music, but it can be used for any genre. ABC notation allows users to easily share and edit music using simple text files, making it accessible and convenient for musicians and music enthusiasts. It is particularly common for those who are not advanced in music production and representation, since it is easily readable and compact, utilizing simple and basic notation elements.

[28] The most important characteristics and qualities of MIDI protocol are timestamp, pitch (0-127) and velocity.

```
X:1
T:Twinkle, Twinkle, Little Star
M:4/4
K:C
C C G G | A A G2 | F F E E | D D C2 |
```

Fig. 3 Example of ABC notation of "Twinkle, Twinkle, Little Star" (via abcnotation.com)

At last, symbolic representation in generative music models is commonly used by authors and composers, with many systems available in the market, or with the option of creating self-made models. The training and generation of these systems is based on a symbolic input that generates an equal symbolic output, where the system is fed with symbolic materials and produces a symbolic piece that is the result of a mix-and-match process. In 2019 Open AI launched *MuseNet*[29], an AI music generator of 4-minute musical pieces with a maximum of ten instruments, ranging from Mozart to Beatles' style, through the analysis of large datasets of MIDI files, learning the relationship between notes, rhythms and instruments and with the ability to predict the next token. It presents a lot of advantages, since it is a very compact representation, easy to manipulate, clear and precise, it can capture long-term dependencies and it is usually in very small models; nonetheless, it also presents some disadvantages because it is oversimplified, with many musical and performance limitations and does not allow for the addition of production information – for example, the note expressed through a synthesiser it is difficult, if not impossible, to represent in a symbolic form.

The second typology of music representation that is commonly used in generative models is audio representation. Audio representation in the context of generative music refers to the various ways in which sound is captured, processed, and manipulated to enable the generation of new music using computational methods. It serves as a foundational element in the field of generative music, providing the means by which algorithms can interpret and produce sound. These representations are crucial for both understanding existing music and creating new compositions, as they translate the intangible nature of sound into data that can be analysed and synthesized by machines.

One of the primary manifestations of audio representation is the ability to model and replicate musical elements such as melody, harmony, rhythm, and timbre. By converting sound into a digital format, generative music systems can learn patterns, generate variations, and even create entirely new pieces. Audio

---

[29] MuseNet is an advanced deep learning model developed by OpenAI designed to generate music. It uses a transformer architecture similar to that employed in natural language processing models, allowing it to compose complex musical pieces in a variety of styles and instrumentations. The model is capable of generating music across multiple genres, from classical to jazz to pop, and can blend styles or continue a given musical piece in a coherent way.

representation can take several forms, each with distinct characteristics and uses, which are essential for different stages of the generative process.

The first typology is the waveform representation. This is the raw, unprocessed form of audio data, representing sound as a continuous wave of amplitude over time. The waveform captures the complete audio signal, including all its nuances, making it a detailed but complex format to work with. In generative music, waveform data is often used when high fidelity and detailed sound manipulation are required. For example, synthesizers may use waveform representations to generate or modify sounds in real-time, allowing for the creation of rich and dynamic audio textures.

Another common typology is the spectrogram, which provides a time-frequency representation of sound. Unlike the waveform, which shows amplitude over time, a spectrogram visualizes how the frequency content of a signal changes over time. This is achieved by applying a Fourier transform to the waveform, breaking it down into its constituent frequencies. Spectrograms are particularly useful in generative music for analysing and synthesizing the spectral characteristics of sound, such as timbre. They enable systems to identify and replicate specific instruments or vocal qualities. An example of this is in style transfer applications, where a piece of music can be transformed to mimic the style of another by manipulating its spectral features.

The third typology is audio embedding, a more abstract representation where audio data is transformed into a dense, low-dimensional vector space. This process captures the essential characteristics of the sound while discarding extraneous details. Audio embeddings are commonly used in machine learning models for generative music, as they provide a compact and efficient way to represent and manipulate audio. For instance, models like OpenAI's *Jukebox*[30] use audio embeddings to learn and generate music that captures the stylistic and structural elements of different genres and artists.

Each of these representations serves a unique purpose in the generative music workflow. Waveforms provide detailed sound information suitable for high-fidelity audio generation, spectrograms offer insights into the frequency content and evolution of sound, and audio embeddings enable efficient and flexible manipulation of musical characteristics. Together, these typologies form the backbone of modern generative music systems, allowing for the creation of innovative and diverse musical outputs.

[30] Jukebox by OpenAI is a powerful neural network designed for generating music, including singing, in a wide range of genres and styles. It uses a sophisticated deep learning architecture to produce high-fidelity audio by modelling music at multiple levels of abstraction. Jukebox can generate music with coherent lyrics, instrumentation, and vocal styles by training on a vast dataset of songs.

As also stated for symbolic representation, this particular form of representation presents its pros and cons. As for the advantages, in comparison to the symbolic one, audio representation surely captures all performance and production nuances and information, resulting in a rich and complex model that automatically generates an audio output – without the need for audio rendering. The disadvantages, however, are many: this model results in large dimension and size output, which is very difficult to manipulate since the representation is not clear. Moreover, there is no compositional information and these models capture with difficulty long-term dependencies – this is a bit deterring considering that for generative models it is an important quality to manifest its predictive ability.

In conclusion, the field of generative music relies heavily on robust music representations to enable machines to understand and create music. The distinctions between symbolic and audio representations highlight the different approaches in generative music, and both can be seen or represented as yin and yang. Symbolic representations, like MIDI, focus on notational elements such as notes, rhythm, and dynamics, providing a structured and discrete view of music. In contrast, audio representations capture the full spectrum of sound, including its timbral and textural nuances, offering a more holistic and continuous perspective. Each approach has its strengths and applications, with symbolic representations excelling in precise compositional tasks and audio representations providing richer, more expressive outputs. Together, they form a complementary framework that drives the evolution of generative music, pushing the boundaries of creativity and technological innovation.

1.5. Generative music techniques

After discussing about the output, it is now time to take some steps back to understand what typologies of generation can be used by AI music models and, in the following paragraph, an overview on the history of algorithmic music generation will be analysed, giving examples and study cases that will help in setting the discussion in a legal perspective for the following chapters.

First and foremost, generative techniques in the context of musical composition and production triggers two questions – or objectives – that need to be answered: what kind of technique does better fit the task I propose myself to achieve? What kind of task I am trying to realize in order for the technique to fulfil my needs?

At this point, it is clear that the two objectives – to choose a technique and to choose a task – are not only interchangeable, but also deeply connected; so, one question cannot be answered without the other. Now it is time to describe – generally and briefly as much as possible – the taxonomy of the generative techniques existent up to date. Initially, a distinction should be made between traditional methods – based exclusively on symbolic representation – such as symbolic AI, optimization, complex systems, statistical methods, and a more innovative and modern approach – that is why it is usually considered as a "cutting edge" – such as Deep Learning, that concentrates both symbolic and audio representations in its output.

Symbolic AI is the most ancient and old-fashioned technique, consisting in giving "reason" to symbols that can be manipulated by the algorithm. If an author wants to use this method, prior deep knowledge of musical theory is needed, since it is necessary to encode it in the rules of the system. The act of feeding the knowledge can be done manually – which is extremely time consuming – or through a corpus, using ML. An example of this technique can be *CHORAL*[31], an expert system with the task of generating four-part chorales in the style of Johann Sebastian Bach.

Optimization represents a fitness (quality) function that tells the operator how "good" the melody generated is, also suggesting compared solutions that every time are subsequently optimized. It consists in an iterative process where each time the result is optimised until the function is completely satisfied. The particularity of this kind of technique is that it can imitate the target style while, at the same time, being "population-based" – meaning that a population of solutions interact with each other to generate the final result. However, it does not come without disadvantages, since it is difficult to program such a function. An example of this technique could be *GenJam*[32], an interactive algorithm that generates real-time jazz solos: it consists of a population of melodies that can be adjusted through immediate human feedback by the user, in order to breed new melodies.

Next in line is complex systems, which generate simple algorithms in which no knowledge of musical theory is needed. However, if such possibility were to happen, the output most likely would not sound good. This technique has proved to be useful only for authors or composers that need a simple inspiration thanks to the generation of raw music material. A few examples can be *CAMUS*[33] and *Conway's Game of Life*[34],

[31] The system contains about 350 rules, written in a form of first-order predicate calculus. The rules represent musical knowledge from multiple viewpoints of the chorale, such as the chord skeleton, the melodic lines of the individual parts, and Schenkerian voice leading within the descant and bass. The program harmonizes chorale melodies using a generate-and-test method with intelligent backtracking.

[32] GenJam, short for Genetic Jammer, is an interactive music system that uses genetic algorithms to generate jazz improvisations. Developed by Al Biles in the mid-1990s, GenJam functions as both a soloist and an accompanist, creating improvisational jazz music that evolves over time through a process inspired by natural selection.

[33] *See* Ebcioglu, K., "An Expert System for Harmonizing Chorales in the Style of J. S. Bach." (1990), The Journal of Logic Programming (8), pp. 145-185. CAMUS (Computer Assisted Music System) is an early generative music system developed in the late 1970s by the composer and researcher Kemal Ebcioglu. It represents one of the pioneering efforts in using computer algorithms to aid in the composition of music, specifically focusing on the automation of counterpoint and harmonization tasks.

[34] Conway's Game of Life is a cellular automaton devised by mathematician John Horton Conway in 1970. It is a zero-player game, meaning its evolution is determined entirely by its initial state, without further input from players. The game takes place on a grid of cells, where each cell can be in one of two states: alive or dead. The state of each cell evolves over discrete time steps according to a set of simple rules based on the states of its neighbouring cells.

both complex systems based on cellular automata that can be alive or dead, depending on the interactions between them. Cellular automata are one of the manifestations of complex systems and in music generation they are utilized to produce sequences of notes or rhythms by mapping the states of the cells to musical elements. For example, a cell's state might correspond to a particular note, rhythm, or dynamic level. As the automaton evolves, the changing states of the cells generate a dynamic and evolving musical output. This approach allows for the creation of complex, intricate musical patterns that might be challenging to achieve through traditional compositional methods.

The last of the traditional techniques is statistical methods, which is considered to be a very primordial attempt at ML, due to the fact that it can extract patterns from a corpus of data. However, this advancement does not keep up with today's ML, because the technique struggles on large scale content or long-term dependencies, and it does not prove to be helpful when the objective is to create an innovative musical piece. A practical example of this would be the *Continuator*: made in the early 2000s and as the name already suggests, this music generator focuses on an interactive performance composition thanks to the upload of pre-existing performances, in order for the AI to extract patterns and generate a probabilistic output. The expression of this technique consists of Markov chains[35], and is based on mathematical rules applied in a creative context. In conclusion, statistical methods play a crucial role in generative music by leveraging probabilistic models to analyse and replicate the patterns found in existing music. These methods enable the creation of new musical pieces that are both stylistically coherent and varied, making them a foundational approach in the field of computational creativity.

Finally, Deep Learning is the first innovative technique in generating music that allows the production of both symbolic and audio outputs. At the base of this technique there are ANNs that are able to target the desired style for the output. However, and as previously discussed, DL is very computationally demanding, meaning that it need vast quantities of data; this is balanced thanks to the fact that it does not need manual input or rules in order to understand its task. The most innovative quality of the application of DL in the music field is the introduction of text-to-music generation, allowing authors and composers to ask with natural language the output they want to receive. This, on one hand, greatly facilitates the users' work – who does not need an in-depth musical knowledge to operate with such systems – but, on the other hand, due to relying on little to no musical knowledge, this model lacks musical coherence and it is difficult to

[35] Markov chains are mathematical models that describe systems or processes where the future state depends only on the current state and not on the sequence of events that preceded it. This property is known as the "Markov property" or "memorylessness." Markov chains are widely used in various fields, including statistics, economics, computer science, and, notably, generative music and language modelling.

steer in a desired direction since they prove to be "wild" and reactive. However, in a study[36] conducted in 2022, it was proposed to create hybrid systems, like neuro-symbolic integration – that combines DL and symbolic AI – to orientate the wilderness of DL models with reasoning on musical knowledge.

In conclusion, there are various techniques available for AI music generation that have evolved over time, reflecting the needs of composers and mirroring the progress in both programming and computational power. As the dissertation follows, it is now necessary to understand the state-of-the-art of generative music, by giving examples and cases that underline the evolution of the systems and AI services offered since its birth.

1.6. Generative Music Eras and examples

The rise of interest and research of AI application in the music field dates back to a time when the concept of artificial intelligence wasn't even at its primordial stage. Generative music has evolved through several distinct eras, each characterized by technological, theoretical, and artistic advancements of the time. These eras reflect the progression from simple algorithmic compositions to complex, AI-driven music generation systems and are generally categorised in five different stages: pre-computer (1700-1956), academic (1957-2009), first startup wave (2010-2016), big tech experiments (2017-2022) and, finally, Music AI music hype – which actually is not defined since it is thought to have started very recently – (2023-present).

The idea of utilizing formal instructions and processes to create music dates back in musical history as far as the ancient Greeks. Pythagoras believed in a direct relation between the laws of nature and the harmony of sounds as expressed in music:

[36] The research was conducted by Yann LeCun and shared through three lectures during the Les Houches Summer School on Statistical Physics and Machine Learning in June 2022. *See* LeCun, Y. "A Path Towards Autonomous Machine Intelligence Version 0.9.2 (2022), Lectures given at the Les Houches Summer School on Statistical Physics and Machine Learning, pp. 1-62, the author proposes an architecture and training paradigms with which to construct autonomous intelligent agents that combines concepts such as configurable predictive world model, behaviour driven through intrinsic motivation, and hierarchical joint embedding architectures trained with self-supervised learning.

> *"The word music had a much wider meaning to the Greeks than it has to us. In the teachings of Pythagoras and his followers, music was inseperable from numbers, which were thought to be the key to the whole spiritual and physical universe. So, the system of musical sounds and rhythms being ordered by numbers, exemplified the harmony of the cosmos and corresponded to it"* (Grout, 1996).

Thus, theoretical applications of numbers (i.e. "data," in a sense) and various mathematical properties derived from nature were the formalisms, or "algorithms," upon which the ancient Greek musicians had constructed their musical systems. Ptolemy and Plato, also, were two others who wrote about this practice. Ptolemy, the "most systematic of the ancient theorists of music," was also a leading astronomer of his time; he believed that mathematical laws "underlie the systems both of musical intervals and of the heavenly bodies," and that certain modes and even certain notes "correspond with particular planets, their distances from each other, and their movements". This idea was also given poetic form by Plato in the myth of the "music of the spheres," the unheard music "produced by the revolutions of the planets", and the notion was later invoked by writers on music throughout the Middle Ages, including Shakespeare and Milton.

These ancient Greek "formalisms", however, are rooted mostly in theory, and their strict application to musical performance itself is probably questionable since Greek music was almost entirely improvised. Thus, while Greek mathematical conjectures certainly created the musical system of intervals and modes with which the musician operated and probably also guided and influenced the performance practice in some ways, the musician was by no means entirely removed from the decision-making process. Ancient Greek music was not "algorithmic composition" in any pure sense, therefore, but it is undoubtably important historically in music for its tendency towards formal extra-human processes.

An extra layer of abstraction would later be achieved with the birth of "canonic" composition in the late 15^{th} century:

"The prevailing method was to write out a single voice part and to give instructions to the singers to derive the additional voices from it. The instruction or rule by which these further parts were derived was called a canon, which means 'rule' or 'law.' For example, the second voice might be instructed to sing the same melody starting a certain number of beats or measures after the original; the second voice might be an inversion of the first or it might be a retrograde [etc.]" (Grout, 1996).

These "rules" of imitation and manipulation are indeed the "algorithm" by which performers unfolded the music. In this case, then, as opposed to the previous one of the ancient Greeks, we can see a clear removal of the composer from a large portion of the compositional process: the composer himself only invents a kernel of music—a single melody or section—from which an entire composition is automatically constructed.

1.6.1. Pre-computer (or pre-digital) era: *The Mozart Dice Game and Mode des valeurs et des intensités*

In this primordial stage of music generation through algorithms, musicians and composers already used algorithms to generate music, way before the birth of actual computer machines. This period starts from the 18th century until the first half of the 20th century, occupying a large amount of time and seeing its conclusion just after the birth of electronic computer machines. During this era, the use of manual algorithms was a key factor for composers, that actually were in the lead and controlled the process fully: these algorithms could follow different techniques such as randomness, re-combination or extraction of musical parameters such as notes and rhythms.

An important example of pre-computer era comes from the great composer and musician himself, Wolfgang Amadeus Mozart. In 1787, Mozart published a musical composition entitled *Musikalisches Würfelspiel*[37] – also known as Mozart Dice Game or Musical Dice Game – which allows one, as the famous composer puts

[37] It is a musical system that functions through the use of dices that generate music from precomposed segments. Detailed instructions can be found at https://www.playonlinedicegames.com/Resources/free/IMSLP20432-PMLP47543-mozart_-_dice_waltz.pdf

it, “to compose, without the least knowledge of music, many German waltzes or ländler as one pleases, by throwing a certain number with two dice”. Even though this idea had been tried to be realised by many other composers[38] [39] before Mozart, but it was Mozart’s work that became popular, and the innovation and vision were mostly attributed to this creation.

The Musical Dice game consists of 272 short measures of music and a table of rules used to select specific measures given a certain dice roll. The result is a randomly selected 16 bar minuet and 16 bar trio. In some cases – like the online version available in public domain for free – instead of two-dice combination, a random number is selected between 2 to 12, that works in the way as the analog form with dices, like shown in the picture below.

Fig. 4

Fig.4. Homepage of the online free version of Mozart Dice Game (via playonlinedicegames.com/mozart)

A second composition that exemplifies the characteristics of pre-computer era is *Mode des valeurs et intensités* written by the musician Olivier Messaien. It is part of a set of four piano compositions, written between 1949 and 1950, with each performance lasting between fifteen and twenty minutes. The specific composition that is being analysed is the second out of the four, the most innnovative and also the most-

---

[38] *See* Hedges, S. A. “Dice Music in the Eighteenth Century” (1978), Music & Letters 59, no. 2, Oxford University Press, pp.180-187. From 1757 to 1812, at least twenty musical dice games were published in Europe. They were all made with the intention to make it possible for the person ignorant to write minuets, marches, polonaises, contredances, waltez etc. by selecting bits of prefabricated music through the use of chance operations.

[39] *Ibidem*, the first ever example of a mathematical music composer dates back to Kirnberger Johann Philipp’s first publication, “*Der allezeit fertige Polonoisen-und Menuettencomponist*” (1757), that served as a model for many of the succeeding musical dice games. The Minuter Composer allowed the novice to compose either a polonaise, a minuet or trio.

discussed, as the first work by a European composer to apply numerical organisation to pitch, duration, dynamics, and mode of attack (timbre). According to the composer's own description, there are separate modes composed of 36 pitches, 24 durations, 12 attacks, and 7 dynamics. The duration scale is separated into three overlapping scales, called "tempi" by the composer, which correspond to the high, middle, and low registers of the piano, and occur in simultaneous superimposition. According to Messiaen:

> *"The durations, intensities and attacks operate on the same plane as the pitches; the combination of modes reveals colors of durations and intensity; each pitch of the same name has a different duration, attack and intensity for each register in which it appears; the influence of register upon the quantitative, phonetic, and dynamic soundscape, and the division into three temporal regions imbues the passage with the spirit of the sounds that traverse them, creating the potential for new variations of colors."*

Mode des valeurs et des intensités departs from traditional notions of rhythm in a way that is characteristic of Messiaen's broader approach to time. Unlike music that adheres to regular, repeating meters, this piece takes a more fluid and unconventional approach to temporal organization. In this composition Messiaen introduces rhythmic values that are free from the regularity of western notation. Instead of using traditional note values that fit neatly within a time signature, Messiaen created rhythms that are based on unusual and irregular durations, with the emphasis placed on the shifting relationships between the notes rather than a fixed pulse. This creates a sense of freedom in the music, as the rhythmic gestures ebb and flow unpredictably. These irregular rhythms allow for a more fluid sense of time, where the individual notes and phrases seem to be suspended or stretched beyond the constraints of conventional musical time.

1.6.2. Academic era: *ILLIAC Suite* and *Experiments in Musical Intelligence*

This era marks the first approach to algorithmic composition in an academic environment, with the purpose of research. It is the first time in history in which people creatively and actively make music with computers. The period starts in 1957, with *Illiac Suite*, the first computer-generated piece to ever be created – and further discussion on this example will follow.

In this environment, two areas need to be analysed: research activity and musical output. As for the research activity, since it was the very first approach of academic studies on algorithmic generation of musical pieces, it was characterised by lots of experimentation and incremental advancement in the field of – what it would later be called – AI music. However, during this period, the research was rather scattered through a very diverse and eterogenous community, where the attention on the matter was drawn by researchers of different fields of study. In terms of musical output, the majority of the algorithms that were created, do not necessarily focus on the generation of full musical pieces, but rather on parts of it (i.e. melodies, progression generation). Also, there is not a particular focus on audio production quality, because the attention of these algorithms was drawn to work in a symbolic setting, in order to generate a score and not to obtain a performance – so the performance was then required to be executed by humans. Finally, most of the experiments gravitate on classical music – with exeptions especially on EDM – due to the vast availability of musical inspiration.

Therefore, it is time to understand the start of this new era, by explaining *ILLIAC Suite*, the first computer-generated musical piece. The *ILLIAC Suite for String Quartet*, composed by Lejaren Hiller in 1957, is widely regarded as one of the first significant works of music to be created with the assistance of a computer. This composition marks a pivotal moment in the history of both classical music and the burgeoning field of computer-assisted composition. Hiller, a pioneering figure in the development of computer music, used the *ILLIAC* computer[40] at the University of Illinois to generate certain elements of the music's structure, integrating technology with traditional musical craftsmanship. Historically, the *ILLIAC Suite* emerges at a time when composers were increasingly engaging with new technologies to explore and expand the boundaries of musical expression. The 1950s were marked by an interest in integrating computers, serialism, and electronic techniques into the compositional process. While many composers during this period were drawn to the potential of computers for generating complex rhythms, harmonies, and forms, Hiller's work stands out because of its conscious effort to use the machine as a tool for organizing musical materials in a non-traditional way. By using the *ILLIAC* computer, Hiller sought to demonstrate that a machine could create musical structures that adhered to the same formal rigor and complexity as human-composed music, while also pushing the limits of creativity through the introduction of randomness and calculation.

At its core, the *ILLIAC Suite* is divided into four movements, each of which was generated using a combination of mathematical processes and human decision-making. The degree to which the computer

[40] The use of the *ILLIAC* computer in generating rhythmic material was critical to this process. The computer was programmed to produce random sequences of twelve different rhythmic values, which were then imposed onto the twelve-tone series. The result was a set of rhythms that, while systematic and structured, also contained an element of randomness and unpredictability. The rhythms, while derived from the mathematical process of serialism, could not be reduced to a simple, predictable pattern.

was involved in the creation of the music varied from movement to movement, with some sections relying more heavily on the computer's ability to generate pitch sequences or rhythmic patterns, while others included greater input from Hiller himself. The role of the computer in the *ILLIAC Suite* was not that of an autonomous composer, but rather as a tool that provided raw material for Hiller to shape and manipulate. In this sense, the work can be seen as a collaborative effort between human intellect and machine-generated randomness. In particular, the second movement of the *ILLIAC Suite*, titled "Mode II: 12-tone Rhythms," provides a fascinating example of Hiller's use of technology to organize time. This movement focuses on the manipulation of rhythm and meter in an unconventional way, employing a twelve-tone series not just for pitch, but also for rhythmic values. The twelve-tone technique, a hallmark of serialism introduced by Arnold Schoenberg, was adapted by Hiller to control the rhythmic organization of the music. The use of a twelve-tone series for rhythm rather than pitch was revolutionary in that it allowed Hiller to apply a systematized, ordered structure to rhythm, pushing the boundaries of traditional rhythmic norms and creating a piece that had a highly abstract, mathematically driven framework.

A second major example of this period is *Experiments in Musical Intelligence* (EMI), a seminal work on the development of algorithmic composition created by David Cope in the late 20th century (1981). EMI represents a groundbreaking effort to use computer software to compose music that emulates the styles of famous composers, allowing for the replication and expansion of historical musical practices through the lens of artificial intelligence. The project was initially sparked by Cope's desire to explore the potential of computational techniques in the realm of composition, motivated by both academic curiosity and a practical desire to discover how computers might engage with complex musical ideas. The concept of EMI revolves around a system of computer-based tools and algorithms that analyze existing music, deconstruct it into its essential stylistic and structural elements, and then generate new works that mimic the patterns and behaviors found in those analyzed pieces[41]. This method employs both artificial intelligence and an algorithmic understanding of music, and it was revolutionary in demonstrating how a machine could engage with the creative process of composition without direct human input. Basically, in technical knwoledge, EMI works in three steps: first, the user feeds the algorithm a corpus of reference style and genre that it wants to emulate (analysation); secondly, the machine will then extract the passages that make the author's

[41] One of the key innovations in EMI is its use of a "musical intelligence" that is modelled on the composer's individual style. Cope's algorithms are designed to analyse the intricate relationships between the elements of a piece of music—such as how motifs develop, how harmonies progress, or how rhythmic patterns unfold—then use this information to create new, stylistically accurate music. Unlike traditional methods of musical analysis, where rules are applied to pre-existing works, EMI uses its rules to generate new works based on the deep understanding of the composer's style. This method can be thought of as an example of "machine learning," in which the system learns from existing music and improves its ability to generate new music that is increasingly indistinguishable from the original.

style appear distinguishable and recognizable (signature extraction); third, the algorithm will finally use recombination techniques to create new and original pieces, with the influence of the style analysed through the input.

1.6.3. First startup wave: *Melodrive*

The years between 2010 and 2016 witnessed a significant wave of innovation in generative music, driven by advancements in artificial intelligence, algorithmic composition, and creative programming. This era marked the confluence of a growing tech culture, the proliferation of accessible tools for musicians, and the increasing interest in machine learning and computational creativity. During this period, a range of startups emerged, providing novel platforms, software, and hardware designed to democratize music production while fostering new forms of musical expression. During this period, the focus shifts from academic-oriented research to the creation of startups and companies in the new field of music generation through AI systems. This meant that research purposes were now abandoned despite giving space and importance to the development of new products for profit goals. Moreover, the advent of the 21st century allowed generative music to gain more momentum with the rise of computational tools that allowed for more complex and accessible generative processes. Technologically, this wave was propelled by the increasing availability of artificial intelligence, machine learning and also early deep learning techniques that allowed for more sophisticated generative models. The early 2010s marked the intersection of generative music with the broader startup culture, particularly in Silicon Valley and other global tech hubs. Startups in this period sought to capitalize on the growing interest in both creative coding and AI. They developed software tools and platforms that allowed non-experts to generate music using AI-driven processes, often blending accessibility with complexity in ways that appealed to both musicians and technologists. A notable example is *Amper Music*, founded in 2014, which utilized AI to create music in real time. The platform enabled users with little to no musical background to produce high-quality tracks by selecting moods, instruments, and genres. This level of user-friendliness was a key innovation, as it allowed for the democratization of music production. Similar platforms, such as *Jukedeck* (founded in 2012) and *Aiva Technologies* (founded in 2016), followed suit, offering AI-generated music for various applications including content creation, film scoring, and video game soundtracks.

The example this paragraph will explain in detail discusses about *Melodrive*[42], a startup founded in 2016, which focuses on real-time video game music generation that adapts to emotional context. In a research

[42] Melodrive is the first AI music system that composes an infinite stream of original, emotionally variable music in real time. The system uses cutting-edge AI techniques to compose and produce new, original music from scratch that continuously adapts

paper presented by the founders of this startup in a workshop at Queen Mary University of London[43], it is described that is very common for interactive music composers to use adaptive music – music that changes through vertical layering and/or horizontal re-sequencing – to support different player behaviours. However, it's impossible for a human composer to conceive and produce every musical outcome. This means that even an adaptive soundtrack can quickly feel repetitive and break the player's engagement due to listener fatigue. The founders then stated "*This is why we built Melodrive: an AI music generation engine that responds to very granular emotional cues in the experience and dynamically composes and produces music in real time. We call this deep adaptive music.*"

Before putting into the market this innovative and ground breaking generative music system, the founders conducted a psychological experiment designed to understand whether deep adaptive music increases the level of immersion and engagement of a person exploring an interactive experience. The founders created a simple VR space station scene in which there were two rooms connected by a corridor. Each room had its own emotional feel ('tender' and 'angry'), and there were no interactive objects in the scene. Each of the 46 participants were randomly assigned to one of 3 music conditions for the experience: no music, linear music and Melodrive-generated music. The linear music condition had a fixed, looping soundtrack. The music by Melodrive was generated in real-time and adapted to the emotional feel in each room while the participants explored the space station. Both the linear music and the deep adaptive music had the same sound design (instrumentation, effects etc.). Participants had no prior knowledge of what the experiment was designed to test, and were instructed to explore the scene for as long as they liked. They were timed during the experience as an overall measure of engagement, and afterwards were asked to fill out a questionnaire about immersion and music. The results showed that the time spent in the VR scene with the Melodrive-generated music was 42% more than that for the no-music condition and 27% more than that for linear music. Immersion levels with music generated by Melodrive were 30% higher than those perceived with no music, and 25% higher than linear music. Moreover, 90% of participants thought that the music generated by Melodrive was a very important component that helped them to feel immersed. It was also found that the adaptive music generated by Melodrive fitted the VR scene 49% better than the linear soundtrack. Finally, the research concludes that Melodrive and deep adaptive music have the ability

Footnote 42 continued
to user interaction and the emotional scenario on-the-fly. The initial target market is XR/game developers, players and streamers, where adaptive music - music that responds to the player and game environment - is key for player engagement.

[43] Elmsley, A. et al., "Deep Adaptation: How Generative Music Affects Engagement and Immersion in Interactive Experiences", Digital Music Research Network Workshop at Queen Mary University of London, 19th December 2017.

to increase both the time spent in an interactive experience and to significantly amplify the immersion of participants.

This experience perfectly describes the switch that was happening during that time in the context of generative music, but also in AI generally. One of the defining features of this first startup wave was the democratization of music creation. Historically, composing and producing music was an endeavour requiring substantial training, access to expensive equipment, and often the collaboration of multiple skilled individuals. The generative music startups of the 2010s significantly lowered the barriers to entry, providing users with intuitive platforms that automated much of the music-making process. This shift was particularly impactful for independent musicians, content creators, and small-scale entrepreneurs, allowing them to produce high-quality music without the need for expensive studio time or advanced technical training. Furthermore, it created new opportunities for hybrid creative practices, where musicians could leverage machine-generated content alongside their own compositions, facilitating an ongoing dialogue between human intuition and algorithmic precision.

1.6.4. Big Tech Experiments: *AWS DeepComposer* (Amazon) and *Jukebox* (OpenAI)

As we are getting closer to present day, another era of generative music has taken control over the focus on which AI needs to be led. Following the previous era, and starting from 2017, big tech companies like Amazon, Google and OpenAI, started to notice and invest in this new subject of interest regarding the bigger picture of AI. The main focus shift during this period is – compared to the first startup wave, where smaller companies tried to invest in innovative systems with a full-piece production and use of ML algorithms based on high quality datasets – the investment of such tech giants to use generative music as a sort of playground for AI testing. Moreover, due to the increase of computational power and the improvement in deep learning techniques, it feels like this era slightly goes back to a research environment, not only with the capability not only of using massive datasets and computing capacity, but also because there is little intention for these companies to use the final products with a clear commercial end goal – nevertheless, they are put on the market for the use of users.

During this period, a lot of different music generator systems were brought into light with very different purposes. Due to the development of deep learning, machine learning and natural language processing, tech companies invested in different projects that allowed an even vaster democratization of AI music. In 2017, Google's *Magenta* project, an open-source research initiative, emerged as a prominent example of AI's potential to create music. Magenta utilized deep learning techniques to explore the generation of music, art, and other creative expressions. It was particularly notable for its MusicVAE model, which enabled the creation of novel melodies, harmonic progressions, and even the blending of different musical genres.

Another significant development in this era was *AIVA* (Artificial Intelligence Virtual Artist), a music composition AI that gained recognition for creating original classical music. AIVA, initially trained on a dataset of classical music, demonstrated the ability to generate compositions that sounded indistinguishable from works created by human composers. In 2017, AIVA was officially recognized as a composer[44] by the French music rights organization SACEM, further legitimizing the notion that AI could be considered a creative agent in its own right.

As already mentioned, this era gives the opportunity to explore many examples of music generation through AI tools, and the analysis of this paragraph will focus on two of these examples: *AWS DeepComposer* by Amazon and *Jukebox* by OpenAI.

AWS DeepComposer is an innovative product from Amazon Web Services (AWS) that leverages artificial intelligence to allow users to compose music effortlessly. Launched in 2019, AWS DeepComposer is a part of Amazon's broader initiative to bring AI-driven solutions to creative fields, offering a platform where users can create original music using machine learning models. Designed for both musicians and non-musicians, the platform democratizes music creation by allowing anyone with a basic understanding of music to create professional-quality compositions. It is made up of a system that includes both software and hardware (a piano board) that was sold to developers in order to drag them into the world of AI music and machine learning techniques to let them see firsthand the incredible capabilities of this technique, compared to simple natural language processing. At the heart of AWS DeepComposer are generative AI models, particularly deep learning-based generative adversarial networks (GANs). These models were trained on diverse datasets that cover various music genres such as pop, jazz, classical, rock, and more. By inputting a melody or musical motif, users can choose a genre, and the AI will generate a full music composition that matches the selected style. Moreover, this product was thought not only for music production, but also for educational purposes intended for people who wanted to learn more about the connection between AI and music: AWS offers tutorials and resources on how machine learning works in the context of music, enabling users to understand the underlying technology that powers the platform.

[44] The EU-funded AIVA project has developed an AI composition tool, a 'music engine' known as AIVA, that can be integrated into existing video games, providing bespoke music that complements play action. AIVA is the first 'virtual artist' to be recognised by SACEM – the French professional association representing the royalties and rights of original music creators. It was also the first AI composer to be commissioned for a piece for the National Day celebrations in Luxembourg and an anthem for the city of Dubai, both performed in 2017. More on: www.cordis.europa.eu

On the other hand, *Jukebox* – launched in 2020 by OpenAI – is a groundbreaking music generation model developed by OpenAI. Jukebox uses deep learning techniques to generate raw, high-quality music, including both the audio and lyrics, in a variety of genres and styles. Unlike previous AI systems focused primarily on generating sheet music or MIDI files, Jukebox can generate fully-realized audio tracks, complete with singing, instrumentation, and complex musical arrangements. This capability represents a major advancement in the field of generative music, blurring the lines between artificial intelligence and traditional human creativity. It sets a turning point in the history of AI music, due to its demonstration – for the first time in history – of the full potential of raw-audio generation[45]. It uses advanced deep learning techniques and it is able to generate a full piece with lead vocals and, since the output is generated in waveforms, the algorithm can easily embed performance details that complete the work created. Jukebox's architecture is based on a deep learning model known as VQ-VAE-2 (Vector Quantized Variational Autoencoders). This approach allows Jukebox to model complex audio structures at multiple levels of abstraction, enabling it to generate long sequences of coherent audio with high fidelity, also thanks to autoregressive generation.

In the end, this three year long era has proved to show the amazing and astounding capabilities of artificial intelligence in a creative setting, demonstrating in such a short amount of time how technology has improved its potentialities, even more than in the first eras, where the human input and research were immersed fully in imagining what has proven to be realisable almost fifty years later.

1.6.5. AI Music Hype: *MusicLM* (Google) and *MusicGen* (Meta)

Finally, the year 2023 has marked a new era for generative music and artificial intelligence, where musicians and all of the subjects operating in the music industry have finally realised the great potentialities a tool like this has in music composition and production. The democratization of these tools has reached its peak, and now single users are at the centre of attention for new developments and innovations. The main characteristic of this new era is that – after everything that the music industry has experienced until now – both big tech companies and startups are investing in this sector: while the firsts now explore clearer commercial opportunities of generative AI, the latter ones are experiencing a new wave, where more startups are launched with more specific goals and targeted to the different users' needs. The big explosion that AI experienced through the launch of ChatGPT has determined the birth of a clear discussion about

[45] With raw-audio generation it is intended a training for AI models based directly and solely on waveforms, in order to receive an output that is directly playable or hearable with waveforms that compose a musical piece.

generative AI: after its launch in 2022, many users understood the great capabilities of an AI model like this, based on LLM and diffusing globally the concept of generative artificial intelligence.

As already mentioned, due to the explosion of this new ability of AI in the generative field, and thanks to impressive computational power and capability of analysing extremely massive datasets – which, of course, arise some legal and ethical questions that will be further discussed in Chapter 2 – nowadays most of the models available on the market are based on text prompts, either text-to-image and video generation or text-to-music generation.

The first example that will be analysed is *MusicLM* by Google Research, launched in 2023, with the goal of bridging the gap between natural language processing and music generation, through the creation of high quality music from textual descriptions. MusicLM's core architecture builds upon transformer models, which have been successful in a wide range of natural language processing tasks. Transformers have shown remarkable ability in learning and generating sequences, and in the case of MusicLM, they are applied to sequences of musical data. The model is trained on vast amounts of music data, which allows it to generate long, coherent music compositions based on textual prompts. This approach allows MusicLM to integrate musical elements such as melody, rhythm, harmony, and timbre into a unified, structured output. Moreover, one of the defining features of MusicLM is its ability to generate music directly from textual descriptions. This is a novel advancement compared to previous models, such as OpenAI's Jukebox, which generated music based on symbolic inputs like MIDI or pre-recorded audio samples. MusicLM allows users to input a simple or detailed textual prompt (e.g., "a relaxing piano piece with the sound of raindrops in the background") and have the model generate a full, high-quality audio track that fits the description. This text-to-music capability offers significant creative freedom for users, making it easy to produce personalized music without any prior musical knowledge. It is also based on a hierarchical and multi-scale generation, meaning that on one hand it can handle the creation of long and complex compositions, maintain coherence and consistency through the entirety of the piece and on the other hand, multi-scale represents generation in phases, which allows the addition of detail in a firstly and generally produced overall structure. This model shows the capabilities of such technology to make music creation affordable, easily available and personalized: users do not need any kind of music knowledge to use such a service, as it makes music production easy to understand and targeted to the specific needs of the user.

The second example under analysis is *MusicGen* by Meta, also released in 2023, that presents a lot of similarities to the Google model. It is also based on text-to-music generation and transformative models to produce high quality music pieces. However, it is more targeted for shorter musical pieces, but it demonstrates to show a higher quality in the audio synthesis of the piece that is requested. MusicGen focuses on generating clearer, more polished audio, often focusing on producing compositions that are suited for commercial or media production with high fidelity and attention to instrumental detail. While MusicLM provides the ability to influence aspects of the generated music through detailed text prompts – in which users can specify moods, genres, and instruments to tailor the output – its flexibility is typically directed toward complexity and style rather than granular control over every single element of the generated composition. On the other hand, MusicGen places a strong emphasis on user customization with fine-tuning options. The model is responsive to user input and allows for more direct control over the output, whether it's adjusting instrumental choices, tempo, mood, or other specific musical features.

Finally, also startups have started to focus on making an impact in this new era of generative AI explosion. Since January 2023, a lot of the startups in AI music have seen the light (e.g., *Soundraw, Boomy, Riffusion, Beatoven.ai* etc.) and have demonstrated a renewed interest in investing and producing different products to tackle different users' needs. They, however, are more targeted at musicians with more expertise, since big tech companies have conquered large part of the market for the common users, by giving them the right tools to solve their needs and problems when creating music.

1.7. Conclusions

In this first chapter the analysis is focused on providing a more technical insight into understanding how AI works in the context of the music industry. From the very beginning and the birth of the concept of artificial intelligence through manual algorithms and calculators, a deep dive into the history of algorithmic composition clarifies everything for a more in-depth discussion of copyright issues that may arise – and are already arising – in this particular sector.

Nowadays artificial intelligence surrounds us in every aspect of our lives: the fact that we are constantly on the internet and communicate through technology has certainly helped in the development of such inventions and computational power, allowing us to reach new limits. Moreover, the COVID-19 pandemic has definitely increased the global population's use of online services, especially ChatGPT[46], which helped

[46] *See* Hussain T., et al., "The influence of the COVID-19 pandemic on the adoption and impact of AI ChatGPT: challenges, applications and ethical considerations." (2024), Acta Psychologica, Volume 246, Article 104264. In this study, it was found that many users mentioned virtual support and assistance as a main reason for the use of ChatGPT. This suggests that during the pandemic, there was an increased demand for AI ChatGPT systems to provide remote support and assistance in various domains such as healthcare, education, customer service, and more. The pandemic created an opportunity for AI ChatGPT to be utilized as a virtual assistant, bridging the gap created by physical distancing measures.

in grappling with the challenges posed by the complete shutdown and has also taken on newfound significance in providing support, information, and virtual interactions.

In the music industry, many famous artists are recognising the benefits and massive capabilities of generative AI, believing that in the next future it will be a requirement and an omnipresent aspect of this creative sector.

The globally famous artist David Guetta, in an interview with Rolling Stone Magazine[47] said "*You cannot fight with this*", expressing his thoughts on the role of human music-makers that will come to revolve more around choosing the best of AI output. He compared AI to instruments that have led to musical revolutions in the past. "*Probably there would be no rock 'n' roll if there was no electric guitar. There would be no acid house without the Roland TB-303 [bass synthesiser] or the Roland TR-909 drum machine. There would be no hip-hop without the sampler*" says the legendary DJ and producer after being awarded as Best Producer at the Brits 2023. The DJ used two artificial intelligence sites to create lyrics and a rap in the style of the US star for a live show even though he has stated that he will not release the track commercially. But he said he thinks musicians will use AI as a tool to create new sounds in the future, because "*every new music style comes from a new technology*".

Speaking to BBC music correspondent Mark Savage at the Brit Awards[48], Guetta said: "*I'm sure the future of music is in AI. For sure. There's no doubt. But as a tool*". "*Nothing is going to replace taste*" he said. "*What defines an artist is, you have a certain taste, you have a certain type of emotion you want to express, and you're going to use all the modern instruments to do that*". On the other hand, the artist Nick Savage commented on musical pieces created with his voice thanks to AI by users as "*grotesque mockery*" and "*a travesty*". It is possible to analyse, here, two faces of the same medal. The first, where the same artists and musicians use artificial intelligence as a tool to enhance their talent and music taste, and the second, where the same users create music pieces by borrowing someone else's artistry.

This insight lets readers and music listeners understand the thought of a musician regarding AI: of course there might be some fear towards such a powerful technology, that is now able to replicate and mimic perfectly the sound of an artist's voice, but it can also be seen and enhanced as a formidable aiding tool to reach new limits and possibilities in an everchanging industry such as music.

[47] The podcast interview is available on https://www.rollingstone.com/music/music-features/david-guetta-eminem-ai-1234793570/

[48] His intervention at the Brit Awards 2023 can be consulted at https://www.bbc.com/news/entertainment-arts-64624525

CHAPTER 2

COPYRIGHT AND AI HARMONIZATION: HUMAN RIGHTS SET THE TONE FOR COEXISTENCE

2.1. Introduction

After a very thorough analysis of artificial intelligence, its birth, its functioning, and a detailed insight into generative AI and generative music production, it is now time to discuss what kind of challenges this new invention raises towards the protection of music artists and authors that live and earn thanks to the intellectual property rights related to their works.

It is of utmost importance to discuss the origin and reason for the existence of copyright, to deeply understand the necessary implementations needed in such regulation after the breakthrough of artificial intelligence – and generative in particular – which raises many questions and doubts, especially regarding the concepts of creativity, authorship and ownership and originality.

In this chapter human rights are at the center of discussion, with an emphasis on international treaties and regulations, trying to depict both sides of the same coin. Moreover, a detailed analysis of specific AI and copyright legislation will be made, in the context of artificial intelligence training and its legality regarding the use of copyrighted works. An attempt to provide a global insight will be made, taking into consideration both western and eastern points of view and the legislation already available.

2.2. The Universal Declaration of Human Rights: the rights to Intellectual Property and freedom of expression

The Universal Declaration of Human Rights[49] (UDHR), adopted in 1948, marked a historic attempt to lay down a universal standard for the protection of fundamental rights that transcend national borders and legal systems. In the wake of the atrocities of World War II, the Declaration sought to prevent further violations of human dignity by enshrining rights that support both individual autonomy and collective wellbeing. Among its articles, Article 27, which deals with the right to participate in cultural life and to enjoy the benefits of one's intellectual creations, and Article 19, which guarantees freedom of opinion and expression, are particularly relevant in the context of artificial intelligence's growing influence in the domains of intellectual property and freedom of expression[50].

[49] Universal Declaration of Human Rights, 10 December 1948, General Assembly Resolution 217 A.
[50] Geiger, C., "Elaborating a Human Rights friendly Copyright Framework for Generative AI" (2024), International Review for Intellectual Property and Competition Law 2024, Vol. 55, Issue 7, pp. 1129–1165.

The UDHR was adopted in the aftermath of World War II, during a time when international cooperation for peace and human rights was seen as crucial for the rebuilding of a war-torn world. The main purpose of the UDHR was to set forth a common standard of human rights for all people, which could serve as a moral guide and legal foundation for the international community. While the UDHR itself is not legally binding, its influence has been profound. It has shaped the development of a broad array of international treaties, conventions, and national laws, which have made the protection of human rights a central feature of international law. At its core, the UDHR aims to ensure the dignity, equality, and fairness of all people. It enshrines civil, political, economic, social, and cultural rights, recognizing that these freedoms are intertwined and essential for human flourishing.

First and foremost, both articles should be analysed to understand their goal and aim to provide universal protection to what, especially at the time of drafting, were considered to be human and fundamental rights to be respected.

Article 27 UDHR introduces the right to cultural, artistic and scientific life, stating, in its first paragraph, that: "*Everyone has the right freely to participate in the cultural life of the community, to enjoy the arts and to share in scientific advancement and its benefits*." This article firmly incorporates cultural rights as human rights for all. They relate to the pursuit of knowledge and understanding, and to creative responses to a constantly changing world. A prerequisite for implementing Article 27 is ensuring the necessary conditions for everyone to continuously engage in critical thinking, and to have the opportunity to interrogate, investigate and contribute ideas, regardless of frontiers. Article 27 is closely linked to Articles 22 and 29[51] in asserting that economic, social and cultural rights are indispensable for human dignity and the development of the human personality. Taken together, they show the UDHR drafters' determination not just to guarantee basic minimum standards, but to help us all become better people[52].

That everyone has a human right to enjoy the benefits of the progress of science and its applications is fundamental, of course, but not to be underestimated. In fact, this right is pertinent to numerous issues at the intersection of science and society: open access; "dual use" science; access to ownership and dissemination of data, knowledge, methods and the affordances and applications thereof; as well as the role

[51] The Right to Social Security and the Right to Duty to your Community (respectively articles 22 and 29 UDHR) are deeply connected to article 27, as they ensure its correct implementation in society. The linkage with article 22 highlights that cultural participation and the protection of intellectual property are part of the broader framework of social and economic rights. While, the connection with article 29 ensures that the exercise of these rights, including cultural and intellectual rights, is balanced with the collective welfare and respect for others' rights. Together, these articles emphasize the interconnectedness of individual rights, cultural participation, and societal obligations.

[52] Office of The High Commission for Human Rights, "Universal Declaration of Human Rights at 70: 30 articles on 30 articles", Press Release, 6th December 2018.

of international cooperation, human dignity and other human rights in relation to science and its products. As we advance towards superintelligence, quantum computing, drone swarms and life-extension technology, serious policy decisions will be made at the national and international levels. The human right to science provides an ideal tool[53] to do so, backed up as it is by international law, political heft and normative weight.

In this analysis, the concepts of science and culture are symbiotic and, actually, the second broadens the application of the first. Both relate to the freedom to engage in creative and mentally developing activities, and cultural life is an "*inclusive concept encompassing all manifestations of human existence*."[54] Cultural life is therefore larger than science as it includes other aspects of human existence; it is however reasonable to include scientific activity in cultural life. Thus, the right of everyone to take part in cultural life includes the right of every person to take part in scientific progress and in decisions concerning its direction.

Leading the discussion towards Intellectual Property, it is now time to analyse the second section of Article 27 UDHR, stating that: "*Everyone has the right to the protection of the moral and material interests resulting from any scientific, literary or artistic production of which he is the author*." Here the concept of intellectual property rights is clear and explicit: the emphasis on the moral and material interests deriving from the authorship of a certain work set the tone for further and subsequent international and national legislation on the matter. This helped recognise cultural impact of subjects as a right – and specifically human right – meaning that such impact needs the highest and most important of protections, since – and connecting with the first section of the same article – the contribution these subjects make is for the common interest of the whole community and helps to foster cultural development for the enjoyment of all people[55].

This section may arise two different and opposite views[56] on the scope and goal of such norm: on one hand, one may see this article as an Hegelian attempt to qualify the concepts of property and personhood with liberal and utilitarian scopes; on the other, however, the recognition of moral and material rights of authors to their own works can be viewed in a Lockean light[57], reflecting natural rights of property. As a support, Article XV of the Inter-American Juridical Committee's bill, could help in the correct interpretation of the

[53] Plomer, A. "The Right to Science: Then and Now", Cambridge University Press, 2021.
[54] E/C.12/GC/25, General comment No. 25 (2020) on article 15: science and economic, social and cultural rights
[55] Geiger, C., Reconceptualizing the Constitutional Dimension of Intellectual Property – An Update (November 29, 2019). in: P. Torremans (ed.), "Intellectual Property and Human Rights", 4th ed., Austin/ Boston/ Chicago/ New York, The Netherlands, Kluwer Law International, 2020, pp. 117-168, Centre for International Intellectual Property Studies (CEIPI) Research Paper No. 2019-11.
[56] Mahalwar, V., "Copyright and Human Rights: The Quest for a Fair Balance" (2017), in Sinha, M. and Mahalwar, V. (eds) "Copyright Law in the Digital World", Springer, Singapore, pp.151-174.
[57] Plomer, A. "The Right to Science: Then and Now", *supra* note 53.

intentions of such article, which literally states: "*The state has the duty to encourage the development of the arts and sciences, but it must see to it that the laws for the protection of trademarks, patents and copyrights are not used for the establishment of monopolies which might prevent all persons from sharing in the benefits of science. It is the duty of the state to protect the citizen against the use of scientific discoveries in a manner to create fear and unrest among the people*." This outlook clears any doubts on the encouragement of monopolies in creative sectors, exploiting the existence of an economic right on works with a creative purpose and putting at risk the open access to science and culture as provided by the norm. In addition, this article is strongly connected to the same provision in the Berne Convention – which will be discussed in a further and specifically dedicated paragraph – which, on the other hand, presents a different outlook and scope than the one in the UDHR.

Finally, another article should be taken into consideration when trying to depict what applies to the general argument of artificial intelligence and intellectual property (in the case of UDHR): freedom of opinion and expression, declared in Article 19 UDHR, enshrines a fundamental principle of exchange of ideas. This right extends not only to the expression of opinions but also to the access and dissemination of information. In an increasingly interconnected world, the digital platforms and technologies that enable communication and the flow of knowledge are essential for personal autonomy and democratic participation. The right to receive and impart information, as laid out in Article 19, also intersects with the cultural and intellectual rights embodied in Article 27.

Article 19 recites: "*Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers*." At the heart of Article 19 lies a deep commitment to the individual's autonomy and dignity and this freedom is understood as a necessary condition for human flourishing. Philosophers such as John Stuart Mill argued that the free exchange of ideas is essential for the development of truth, the discovery of new knowledge, and the functioning of a just society. In Mill's *On Liberty*[58], he emphasized that freedom of expression enables individuals to challenge prevailing ideas, form their own opinions, and participate fully in social and political life. It is clear, just like for Article 27, that the scope and purpose of the drafting of the UDHR underlines a sense of public good, which allows the granting of such freedoms in a universal and equal way.

The discussion, however, needs to shift into a more practical and detailed insight for the scope of this thesis. The analysis of such articles is fundamental to transpose the philosophical debate into present reality,

[58] Mill, J. S. "On Liberty", Dover Publications, 2002.

especially regarding the conflict between IPRs and AI. Considering both subjects of discussion, it is clear that both collide in both articles, because all sides can be interpreted in favour of one or the other matter. If we consider AI first, there is no doubt that this is protected under the two articles analysed: artificial intelligence, as a technological invention, is fostered and developed in sight of public and common knowledge, entering into the concept of right to culture and science as stated in Article 27 UDHR. Moreover, due to the growing reliance on technologically driven information, Article 19 UDHR implicitly allows this kind of freedom of expression and, consequently, information. Democratic societies rely on the open exchange[59] of ideas to encourage social progress, innovation, and public debate. It is within this philosophical framework that Article 19 asserts the right to seek and receive information, regardless of borders. This extends not only to traditional forms of communication like speech and print media but also to modern channels such as the internet and digital technologies. The notion here is that access to information is essential for individuals to make informed decisions, assert their rights, and participate in shaping the direction of society.

On the other hand, however, intellectual property rights counterbalance the exclusive protection of the articles analysed in the view of the scientific progress of artificial intelligence. Article 27 (2) UDHR explicitly protects the moral and material rights of authors of creative works. While on the first part the emphasis is put on the protection of cultural development for a greater good, which is public knowledge and research, the second paragraph tries to weigh the other plate of the scale towards the preservation of personal achievements and works that help that same cultural public development[60]. In the case of AI, the discussion turns a bit tricky, because the analysis of such equilibrium revolves around the technological invention itself. Here the protection of IPRs in a AI setting requires an analysis of the use of already protected works for the training of such models and machines. This means that the equilibrium that needs to be found relates to the legality[61] of the use of copyrighted works for the purposes of making those models

[59] On the matter, *see* Geiger, C. and Jütte, Bernd J., "Copyright, the Right to Research and Open Science: about time to connect the dots" (June 08, 2024), in Enrico Bonadio & Caterina Sganga (eds), A Research Agenda for EU Copyright Law (Edward Elgar, forthcoming 2025), pp.1-21. The concept of 'open science', also defined by UNESCO in its 2017 Recommendation on Open Science, defines perfectly what is the aim of constitutionalizing the protection of artificial intelligence as a technological innovation. Open science manifests in the act of collaboration and cooperation of scientific researchers to foster multilingual knowledge and the sharing of information for the benefits of science and society.

[60] Geiger, C., "Reconceptualizing the Constitutional Dimension of Intellectual Property – An Update" (November 29, 2019), *supra* note 55.

[61] *See* Carroll, M. W., "Copyright and the Progress of Science: Why Text and Data Mining Is Lawful" (December 1, 2019). 53 UC Davis Law Review 893, American University, WCL Research Paper No. 2020-15, pp.893-964.; Samuelson, P., "Text and data mining of in-copyright works: is it legal?" Commun. ACM 64, 11 (November 2021), 20–22.

better and more efficient. On this idea there are different – and also contrasting – schools of thought, that will be further discussed in this chapter.

2.3. Copyright as a human right: *Berne Convention* and *WIPO Treaty*

Even though Berne Convention[62] was drafted over sixty years prior to the Universal Declaration of Human Rights, the discussion of this Convention is located subsequently due to the more detailed provisions and principles that broaden the discussion of the topic. In fact, although its adoption was made in Berne in 1886, the final drafting was completed ten years later, in 1896 in Paris. Moreover, this Convention was subject to many amendments during the span of another sixty years: the first revision took the longest – from 1908 to 1914 – between Berlin and Berne; the second was made in Rome in 1928; the third revision was made in Bruxelles in 1948; almost twenty years later – in 1967 – the fourth revision was done in Stockholm; and, lastly, the ultimate revision dates back to 1971 in Paris. The importance and great value of this international treaty is the setting of the fundamental, and yet not underestimated, concept of mutual recognition of copyright between different states – in this case, the ones that adhered to the Convention. In fact, if the recognition of copyright as an exclusive right of the author of the work was dated before the Berne Convention, it was not possible to enforce that right outside the home state of the author. The innovative outlook this treaty has brought really helped in the diffusion of culture and literature across the globe and, actually, realised what the UDHR set into stone with its Article 27. Moreover, in its first Article, the Convention establishes a Union of the Members that adhered to the treaty, in order to grant the protection of the rights of the authors of *literary and artistic works*.

The explanation of the concept of *literary and artistic works* is defined under Article 2, which includes – and is relevant for the scope of this discussion – also *musical compositions with or without words*[63].

---

[62] Berne Convention for the Protection of Literary and Artistic Works (as amended on September 28, 1979)

[63] Article 2 Berne Convention:

(1) *The expression "literary and artistic works" shall include every production in the literary, scientific and artistic domain, whatever may be the mode or form of its expression, such as books, pamphlets and other writings; lectures, addresses, sermons and other works of the same nature; dramatic or dramaticomusical works; choreographic works and entertainments in dumb show; musical compositions with or without words; cinematographic works to which are assimilated works expressed by a process analogous to cinematography; works of drawing, painting, architecture, sculpture, engraving and lithography; photographic works to which are assimilated works expressed by a process analogous to photography; works of applied art; illustrations, maps, plans, sketches and three-dimensional works relative to geography, topography, architecture or science.*

(2) *It shall, however, be a matter for legislation in the countries of the Union to prescribe that works in general or any specified categories of works shall not be protected unless they have been fixed in some material form.*

(3) *Translations, adaptations, arrangements of music and other alterations of a literary or artistic work shall be protected as original works without prejudice to the copyright in the original work.*

(4) *It shall be a matter for legislation in the countries of the Union to determine the protection to be granted to official texts of a legislative, administrative and legal nature, and to official translations of such texts.*

Nevertheless, in Article 2(2) the treaty leaves the single nations to decide which works that fall under the definition in (1) should not be considered to be protected under the Convention, leaving any doubt and questions to the decision of the single member states. Moreover, Article 2-bis adds that the decision of excluding some legislative or political speeches from the protection is also reserved to the single states.

After analysing the object of protection defined in Article 2, it is important to state that this Convention sets three very important principles in the protection of creative works: the principle of national treatment, the principle of automatic protection and the principle of independent protection.

The first principle, national treatment[64], is stated in Article 5(1) of the Convention, providing that: "*Authors shall enjoy, in respect of works for which they are protected under this Convention, in countries of the Union other than the country of origin, the rights which their respective laws do now or may hereafter grant to their nationals, as well as the rights specially granted by this Convention*." This principle mandates that member states must treat works created by foreign authors or nationals of other contracting countries equally to the works of their own citizens. This principle guarantees that an author or creator's rights are respected and enforced in other member states just as they would be in their country of origin. This creates a uniform international standard for copyright protection, removing the complexities of having different copyright laws and protections in each country. This is a very important principle in the promotion of global harmonization of copyright regulations across different countries and in encouraging universal equality of intellectual property rights, also avoiding discriminatory measures based on the nationality of the author.

The second and the third principles, automatic and independent protection, can be found – and follow the first principle – in Article 5(2) by stating that: "*The enjoyment of and the exercise of these rights shall not*

Footnote 63 continued
(5) *Collections of literary or artistic works such as encyclopaedias and anthologies which, by reason of the selection and arrangement of their contents, constitute intellectual creations shall be protected as such, without prejudice to the copyright in each of the works forming part of such collections.*
(6) *The works mentioned in this Article shall enjoy protection in all countries of the Union. This protection shall operate for the benefit of the author and his successors in title.*
(7) *Subject to the provisions of Article 7(4) of this Convention, it shall be a matter for legislation in the countries of the Union to determine the extent of the application of their laws to works of applied art and industrial designs and models, as well as the conditions under which such works, designs and models shall be protected. Works protected in the country of origin solely as designs and models shall be entitled in another country of the Union only to such special protection as is granted in that country to designs and models; however, if no such special protection is granted in that country, such works shall be protected as artistic works.*
(8) *The protection of this Convention shall not apply to news of the day or to miscellaneous facts having the character of mere items of press information.*
[64] Kur, A. and Dreier, T. and Luginbühl, S., "European Intellectual Property Law : Text, Cases and Materials". Second edition. Cheltenham, England, Edward Elgar Publishing, 2019

be subject to any formality; such enjoyment and such exercise shall be independent of the existence of protection in the country of origin of the work." However, the extension and/or the different provision of modality of enjoyment of these rights by the authors can be modified or regulated differently by the country of origin (Article 5 (3) Berne Conv.). These principles set fundamental standards for the existence of copyright: the fact that there are no obstacles or modalities of registration of such right not only facilitates authors in protecting their creations, but also emphasises that the work itself – due to its creation – embodies the right that is protected under the Convention. This principle aligns with John Locke's theory of property, particularly his idea that ownership of intellectual property is a natural extension of an individual's labour and effort[65]. According to this view, when an individual creates something new, they inherently have a right to own it, as it represents their personal intellectual effort. The absence of formalities underlines the notion that creativity itself is a fundamental right, and any work of creation automatically becomes the creator's to control, whether or not it is formally documented.

Finally, another important aspect of the Berne Convention is the presentation of minimum standards of protection[66], which can be found in the Part II of the treaty, and that set a baseline for all member states to follow and provide. These standards define the scope of the works that can be protected, duration[67] of the protection, and the specific rights of the authors (i.e., right of translation[68], right of reproduction[69], right of public performance and communication of dramatic and musical works[70] etc.).

Now, the second international treaty – and deeply connected to the Berne Convention – that needs discussion and analysis is the WIPO Treaty[71] (also World Intellectual Organization Copyright Treaty or

---

[65] *Ibidem*
[66] *Ibidem*
[67] Article 7 (1) Berne Convention: "*The term of protection granted by this Convention shall be the life of the author and fifty years after his death*."
[68] Article 8 Berne Convention: "*Authors of literary and artistic works protected by this Convention shall enjoy the exclusive right of making and of authorizing the translation of their works throughout the term of protection of their rights in the original works*."
[69] Article 9 Berne Convention:
"(1*) Authors of literary and artistic works protected by this Convention shall have the exclusive right of authorizing the reproduction of these works, in any manner or form.*
(2) *It shall be a matter for legislation in the countries of the Union to permit the reproduction of such works in certain special cases, provided that such reproduction does not conflict with a normal exploitation of the work and does not unreasonably prejudice the legitimate interests of the author.*
(3) *Any sound or visual recording shall be considered as a reproduction for the purposes of this Convention.*"
[70] Article 11 Berne Convention:
"(1*) Authors of dramatic, dramatico-musical and musical works shall enjoy the exclusive right of authorizing: (i) the public performance of their works, including such public performance by any means or process; (ii) any communication to the public of the performance of their works.*
(2) *Authors of dramatic or dramatico-musical works shall enjoy, during the full term of their rights in the original works, the same rights with respect to translations thereof.*"
[71] WIPO Copyright Treaty (WCT), adopted in Geneva on December 20, 1996

WCT). The WCT is a special agreement under the Berne Convention which deals with the protection of works and the rights of their authors in the digital environment. In addition to the rights recognized by the Berne Convention, they are granted certain economic rights. The Treaty also deals with two subject matters to be protected by copyright: computer programs – whatever the mode or form of their expression – and compilations of data or other material ("databases").

The Preamble to the WIPO Treaty is has a very interesting outlook on the view and purpose of drafting of such an agreement – made in 1996 – in which the member states that agreed to the Berne Convention recognize the impact and importance of the digital environment that, already at that time, had a significant role in many sectors of everyday life[72].

The WIPO Treaty emerged in the 1990s, a decade marked by the rapid proliferation of digital technologies. The internet and digital storage fundamentally transformed how creative works were produced, distributed, and consumed. Traditional copyright frameworks struggled to address issues like unauthorized reproduction and global dissemination of works over networks. In this outlook, WIPO decided to draft treaties that would help address this new uncertain situation characterised by issues such as unauthorized reproduction and global dissemination of works over networks. The internet and digital storage fundamentally transformed how creative works were produced, distributed, and consumed and traditional copyright frameworks struggled to address such issues. This led to the creation of detailed international treaties on the footsteps of the Berne Convention[73], to enlarge and specify difficult themes and situations that arose after 1886.

The WCT presents not only an outlook on the digital environment in the context of copyright, but also presents specific rights that are tailored to authors' specific issues and interests that were born after the diffusion of digital distribution of creative works. Articles 6 through 8 of the WCT established the exclusive rights of authors to authorize reproduction, distribution, and public communication of their works. These

[72] WIPO Treaty Preamble:
"*The Contracting Parties,*
Desiring to develop and maintain the protection of the rights of authors in their literary and artistic works in a manner as effective and uniform as possible,
Recognizing the need to introduce new international rules and clarify the interpretation of certain existing rules in order to provide adequate solutions to the questions raised by new economic, social, cultural and technological developments,
Recognizing the profound impact of the development and convergence of information and communication technologies on the creation and use of literary and artistic works,
Emphasizing the outstanding significance of copyright protection as an incentive for literary and artistic creation,
Recognizing the need to maintain a balance between the rights of authors and the larger public interest, particularly education, research and access to information, as reflected in the Berne Convention
[73] The Berne Convention was concluded at a time where countries were trying to create basic protection against rampant piracy.

provisions aimed to protect creators from the novel threats of digital piracy and unauthorized use, particularly in the global digital marketplace.

In fact, Article 6 introduces the right of reproduction, stating that: "*Authors of literary and artistic works shall enjoy the exclusive right of authorizing the making available to the public of the original and copies of their works through sale or other transfer of ownership.*"

This new provision enlarges the right already provided by Berne Convention in Article 9, adapting it to the context of digital environment, where reproduction of a creative work not only is made easier, but also accentuated by the velocity of communication and sharing in such context. This allows authors to have an exclusive control and decide whether their works can be reproduced or copied. In the context of artificial intelligence this article sets an important principle, due to the difficulty on recognizing the actual originality of an AI-generated work because, generally, the training under which the model is submitted consists of majorly copyrighted materials. This principle should help in the provision of a regulation that does not violate such norm, and provides an understanding framework for both enjoyment of copyright by authors and the diffusion of such an innovation that is artificial intelligence[74].

Analysing Article 8 of the same treaty, it is possible to ascertain the broadening of already existing provisions in Berne Convention regarding the communication to the public. In fact, Article 8 states: "[…] *authors of literary and artistic works shall enjoy the exclusive right of authorizing any communication to the public of their works, by wire or wireless means, including the making available to the public of their works in such a way that members of the public may access these works from a place and at a time individually chosen by them*."

In this new era, the communication to the public through wire – and especially wireless – means is the main way of sharing new creative and artistic works. This sets an important precedent and acknowledgement that the society and the global culture is – actually, already has – changed and shifted from analog to digital communication. This is a necessary affirmation to be made in a very digitalised world, especially in a normative context where there is no space for doubts, and where less regulation means free exploitation and infringement of creativity of other people that live through and thanks to their art.

A fundamental discussion, for the purpose of this thesis, should be done regarding Article 10 WCT. It literally provides and allows exceptions and limitations to the enjoyment of authors' rights – upon the

[74] Dinwoodie, Graeme B., "The WIPO Copyright Treaty: A Transition to the Future of International Copyright Lawmaking?" (2007), Case Western Reserve Law Review, Vol. 57, No. 4, pp. 751-766.

decision of the Contracting Parties through national legislation – "*in certain special cases that do not conflict with a normal exploitation of the work and do not unreasonably prejudice the legitimate interests of the author*."

This provision sets the basis for further discussion, allowing the possibility to regulate exceptions in the context of this discussion, that is artificial intelligence, and envisage provisions that allow the use of copyrighted works without them being considered to infringe such right. Further discussion will be led in the next paragraph.

Finally, the analysis of such international treaties sets fundamental boundaries and principles that, even if drafted in a time where no internet existed, are clearly applicable in present days, with concepts that transcend time and space, providing protection to the one of the most important realisations of modern times, intellectual property.

2.4. International Covenant on Economic, Social and Cultural Rights: the right to access to culture, knowledge, education and research in the digitalised world

It seems unreasonable to take a step back now that both intellectual property, the right to culture and science and copyright have been discussed in the context of human and fundamental rights. However, this discussion is useful for the introduction of the concept of artificial intelligence training, an extremely hot topic that is making jurists busy to find the perfect way for balancing the existence of AI models based on copyrighted works training and the rights of the authors whose works are exploited for such training.

Before diving into the more irksome debate, another agreement should be analysed. The International Covenant on Economic, Social and Cultural Rights[75] (from now on ICESCR) is an international treaty that is part of the International Bill on Human Rights – that includes the Universal Declaration of Human Rights – that was previously analysed – and the International Covenant on Civil and Political Rights, and adopted by the United Nations General Assembly in 1966. This treaty introduces important rights for humanity and society as a whole, such as the right to work, the right to social security, the right to family life, the right to an adequate standard of living, the right to health, the right to education and the right to participation in cultural life.

[75] International Covenant on Economic, Social and Cultural Rights, 16 December 1966, General Assembly resolution 2200 A

For the purpose of this discussion the last two rights will be analysed in detail, to help set forward the debate on AI training. First and foremost the right to education, stated in Article 13 ICESCR[76], is considered to be fundamental for "*the full development of the human personality and the sense of its dignity*", setting an important standard in a post-war society for the growth of communities and nations as a whole. The right to education, nowadays and with the help of new technologies and artificial intelligence, has shifted completely. Starting from a standardized way of sharing and teaching education, the digital revolution has not only democratized even more the possibility of realizing universal access to knowledge, but also it has made everything more personalised and better tailored to everyone's needs. Even better, it has made students – and users in general – able to get access to many more materials and information more than ever in history, simply by asking or giving a short prompt on what they are searching for. Education and knowledge has never been more accessible and free than these times, with an extreme cost and time-efficiency on research, but sacrificing the beauty of research and study. Everything is available in a few seconds, every question can be answered and every problem can be solved. But is this really what Article 13 meant by the full development of human personality? This easy accessibility and democratization is definitely and clearly making people less eager to give space and use their imagination and creativity, especially when now artificial intelligence models are able to generate creative works on their own through a simple phrase that indicates the final goal to reach.

---

[76] Article 13 ICESCR: "1. *The States Parties to the present Covenant recognize the right of everyone to education. They agree that education shall be directed to the full development of the human personality and the sense of its dignity, and shall strengthen the respect for human rights and fundamental freedoms. They further agree that education shall enable all persons to participate effectively in a free society, promote understanding, tolerance and friendship among all nations and all racial, ethnic or religious groups, and further the activities of the United Nations for the maintenance of peace.*
2. *The States Parties to the present Covenant recognize that, with a view to achieving the full realization of this right:*
(a) *Primary education shall be compulsory and available free to all;*
(b) *Secondary education in its different forms, including technical and vocational secondary education, shall be made generally available and accessible to all by every appropriate means, and in particular by the progressive introduction of free education;*
(c) *Higher education shall be made equally accessible to all, on the basis of capacity, by every appropriate means, and in particular by the progressive introduction of free education;*
(d) *Fundamental education shall be encouraged or intensified as far as possible for those persons who have not received or completed the whole period of their primary education;*
(e) *The development of a system of schools at all levels shall be actively pursued, an adequate fellowship system shall be established, and the material conditions of teaching staff shall be continuously improved.*
3. *The States Parties to the present Covenant undertake to have respect for the liberty of parents and, when applicable, legal guardians to choose for their children schools, other than those established by the public authorities, which conform to such minimum educational standards as may be laid down or approved by the State and to ensure the religious and moral education of their children in conformity with their own convictions.*
4. *No part of this article shall be construed so as to interfere with the liberty of individuals and bodies to establish and direct educational institutions, subject always to the observance of the principles set forth in paragraph I of this article and to the requirement that the education given in such institutions shall conform to such minimum standards as may be laid down by the State.*

Moreover, and moving the analysis forward to Article 15[77], here is stated the right to culture and enjoyment of scientific progress. In addition, moral and material interests of authors of any work that contribute to culture and scientific progress are protected under this norm. Without being repetitive on the discussion already tackled in Chapter 2.2., when analysing Article 27 UDHR, here we debate on the scope and actualization of such principles in an already fully digitalised environment where artificial intelligence leads the subsequent path to follow. It is important to note that in paragraphs 2, 3 and 4 of the same article, a set of rules are imposed on member states to incentivize the flourishing and benefits that are a result of scientific and cultural research and diffusion. In paragraph 3, the Article states that "*the State Parties* […] *undertake to respect the freedom indispensable for scientific research and creative activity*." As important as it is as a statement and a principle, because it grants no intervention based off of any kind of discrimination, it does not set a boundary to the freedom of research[78]. What if that freedom goes against and infringes other freedoms and rights protected by the same international agreements? What if that freedom allows scientific and cultural research to surpass ethical and human standards for the sake of progress? Artificial intelligence, even though is a very important and fundamental discovery, has demonstrated also to present discriminatory biases[79] and outputs that cannot be considered ethical in the society humanity has worked so hard on to construct in such way.

---

[77] Article 15 ICESCR: "1. *The States Parties to the present Covenant recognize the right of everyone:*
(a) *To take part in cultural life;*
(b) *To enjoy the benefits of scientific progress and its applications;*
(c) *To benefit from the protection of the moral and material interests resulting from any scientific, literary or artistic production of which he is the author.*
2. *The steps to be taken by the States Parties to the present Covenant to achieve the full realization of this right shall include those necessary for the conservation, the development and the diffusion of science and culture.*
3. *The States Parties to the present Covenant undertake to respect the freedom indispensable for scientific research and creative activity.*
4. *The States Parties to the present Covenant recognize the benefits to be derived from the encouragement and development of international contacts and co-operation in the scientific and cultural fields.*"

[78] *See* Geiger, C., "Elaborating a Human Rights friendly Copyright Framework for Generative AI", supra note 50. An indirect obstacle or "negative" aspect that could help this matter is the protection of material and moral interests of the authors. This constitutes a fundamental and binding principle of copyright law, allowing the limitation to the use or access of protected works to protect the author's rights.

[79] For example, online advertising has shown that in search engine ad algorithms can reinforce job role gender bias. Independent research at Carnegie Mellon University in Pittsburgh revealed that Google's online advertising system displayed high-paying positions to males more often than to women. Moreover, in image generation (Midjourney) it was proved that, when asking to represent people in specialized professions, it showed both younger and older people, but the older people were always men, reinforcing gendered bias of the role of women in the workplace. Finally, artificial intelligence has also been used in legislative prediction, especially in policing tools in the criminal justice sector, proving that – due to reliance on historical arrest data – it reinforced existing patterns of racial profiling and disproportionate targeting of minority communities.

2.5. Regulating artificial intelligence training: a global overview

With these questions posed, it is now time to discuss and debate the legality of artificial intelligence training that, naturally, has not been yet regulated in detail in some areas, while in others specific legislation has been drafted to cope with this new challenge, but many nations recall already existing legislation to adapt to this new innovation. But before analysing different points of view in regulation of AI training, an introduction is necessary.

In overviewing and studying how artificial intelligence models are trained it is extremely easy to detect just the disadvantages and challenges for legal purposes, such as infringements and violations, especially in copyright. Nevertheless, the use of copyrighted works and materials for AI training is fundamental and presents many benefits on different categories: firstly, AI models benefit from high-quality datasets[80], allowing for improved accuracy and performance across applications such as natural language processing, computer vision, and automated content creation – for instance, generative AI relies on extensive datasets to produce human-like text and images, enhancing creative and business productivity. Secondly, AI training fosters innovation by enabling the development of new tools and services. In fields such as education, healthcare, and finance, AI-driven systems assist in automating complex tasks, analysing large volumes of data, and providing personalized recommendations. Such advancements contribute to economic growth and enhance public access to knowledge and technology, so this perfectly follows and actually realizes what UDHR and ICESCR provide as fundamental rights[81]. Third, AI can be leveraged for preservation and cultural analysis[82], particularly in the digitization of historical texts and artistic works. By training AI on a diverse range of copyrighted materials, researchers can create sophisticated models that assist in translation, restoration, and archiving, thereby ensuring the longevity of cultural heritage. So, in this case, cultural knowledge and progression – stated as human right – not only should be considered as the advancement of knowledge in a field, but also the ability to consult and experience past history and heritage that helps to reach the same progress that is protected by various international agreements.

---

[80]*See* Geiger, C. and Iaia, V., "Towards an Independent EU Regulator for Copyright Issues of Generative AI: What Role for the AI Office (But More Importantly: What's Next)?" (August 03, 2024). Auteurs & Media, II, 2024, pp. 185-196. The study underlines the great influence creative inputs have on Generative AI models, making it possible for these to obtain higher quality outputs.

[81] Similarly, also the Constitution of the United States of America focuses on the promotion of the same right in its Article 1, section 8 by providing "*The Congress shall have the power to promote the progress of science and useful arts, by securing, for a limited time, to authors the exclusive right to their respective writings and discoveries*"

[82] Chakrabarti, K., "The Role of AI in Cultural Preservation and Heritage", in www.itmunch.com, 2024. The article describes the possible applications of AI in cultural and artistic preservation, as well as the enhancement of artistic and cultural access. The most notable examples of AI use in historic preservation is digital archiving and restoration of historical artifacts such as ancient manuscripts. In addition, in Japan researchers have applied AI in the preservation of ancient traditions such as pottery and textile weaving.

On the other hand, however, it seems that the challenges outweigh the benefits in the protection and enjoyment of the counterbalanced freedom and right that is intellectual property. The major issue here is that artificial intelligence's training on copyrighted materials means infringement of copyright laws in different states. But still, what constitutes copyright infringement? Wouldn't it be considered fair use – or equivalent provision in other nations – when it's done for the sake of scientific progress? Many AI models are trained on and for the internet, when massive quantities of data are already available for usage for users, without the need for an explicit permission. Moreover – but this debate will be further discussed in Chapter 3 – the training based on copyrighted data automatically means that the output will be affected in some way to copy or resemble the materials it was trained on. As seen in Chapter 1, in AI music generation, most of the examples proposed and analysed are obviously based on massive datasets of copyrighted material, and the output that often comes out really shows the ability of these models to resemble and imitate already existing music, by mixing and matching styles and rhythms of other authors and musicians. Additionally, and consequentially, the training and the output have a massive impact on the market, especially since AI-generated material can compete with the original works of the authors, diminishing the latter's' market value.

In the following sub-paragraphs an in-depth analysis will be made on the existing global frameworks – Europe, America and Asia – regarding artificial intelligence training and the threshold allowed in use of copyrighted materials without being considered as infringement. There are different views and opinions, and it is important to note the different approaches to understand which would be the right or best path to follow.

2.5.1. European Framework: *CDSM Directive* and *AI Act*

The European Union has taken significant steps to regulate AI and copyright through legislative initiatives such as the Artificial Intelligence Act[83] (AI Act) and the Copyright and related rights in the Digital Single Market Directive[84] (CDSM Directive). The AI Act, introduced by the European Commission, aims to establish a comprehensive regulatory framework for AI systems, categorizing them based on risk levels and imposing strict compliance requirements for high-risk applications. While the AI Act does not directly regulate copyright issues, it intersects with intellectual property rights by emphasizing transparency,

---

[83] Regulation EU 2024/1689

[84] Copyright and related rights in the Digital Single Market Directive (DIRECTIVE EU 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC)

accountability, and ethical AI use. The CDSM Directive, on the other hand, focuses on the application of copyright principles in the digital environment, providing rules and exceptions with the goal of balancing the rights listed and analysed in the previous paragraphs.

The analysis will focus on the analysis of both legislations that set principles and harmonised standards for all member states to follow and respect, even with additional norms according to national frameworks.

First, the AI Act, as already stated, does not include provisions that explicitly refer to the impact of artificial intelligence in intellectual property rights, focusing on the definition and regulation of AI models based on a risk system. In fact, the new rules establish obligations for providers and users depending on the level of risk from artificial intelligence. Moreover, the provisions in the regulation are explicit on unacceptable systems and high-risk systems, while the ones who are not included in these divisions are required to follow the obligations and principles that are listed subsequently.

Unacceptable risk AI systems, listed in Article 5[85], are systems considered a threat to people and will be banned. They include:

---

[85] Article 5 (1) AI Act: "1. *The following AI practices shall be prohibited*:

(a)*the placing on the market, the putting into service or the use of an AI system that deploys subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective, or the effect of materially distorting the behaviour of a person or a group of persons by appreciably impairing their ability to make an informed decision, thereby causing them to take a decision that they would not have otherwise taken in a manner that causes or is reasonably likely to cause that person, another person or group of persons significant harm*;

(b)*the placing on the market, the putting into service or the use of an AI system that exploits any of the vulnerabilities of a natural person or a specific group of persons due to their age, disability or a specific social or economic situation, with the objective, or the effect, of materially distorting the behaviour of that person or a person belonging to that group in a manner that causes or is reasonably likely to cause that person or another person significant harm;*

(c)*the placing on the market, the putting into service or the use of AI systems for the evaluation or classification of natural persons or groups of persons over a certain period of time based on their social behaviour or known, inferred or predicted personal or personality characteristics, with the social score leading to either or both of the following*:

(i)detrimental or unfavourable treatment of certain natural persons or groups of persons in social contexts that are unrelated to the contexts in which the data was originally generated or collected;

(ii)detrimental or unfavourable treatment of certain natural persons or groups of persons that is unjustified or disproportionate to their social behaviour or its gravity;

(d)*the placing on the market, the putting into service for this specific purpose, or the use of an AI system for making risk assessments of natural persons in order to assess or predict the risk of a natural person committing a criminal offence, based solely on the profiling of a natural person or on assessing their personality traits and characteristics; this prohibition shall not apply to AI systems used to support the human assessment of the involvement of a person in a criminal activity, which is already based on objective and verifiable facts directly linked to a criminal activity*;

(e)*the placing on the market, the putting into service for this specific purpose, or the use of AI systems that create or expand facial recognition databases through the untargeted scraping of facial images from the internet or CCTV footage*;

(f)*the placing on the market, the putting into service for this specific purpose, or the use of AI systems to infer emotions of a natural person in the areas of workplace and education institutions, except where the use of the AI system is intended to be put in place or into the market for medical or safety reasons*;

- Cognitive behavioural manipulation of people or specific vulnerable groups: for example voice-activated toys that encourage dangerous behaviour in children
- Social scoring: classifying people based on behaviour, socio-economic status or personal characteristics
- Biometric identification and categorisation of people
- Real-time and remote biometric identification systems, such as facial recognition

Some exceptions may be allowed for law enforcement purposes. For example, "real-time" remote biometric identification systems[86] will be allowed in a limited number of serious cases, while "post" remote biometric identification systems, where identification occurs after a significant delay, will be allowed to prosecute serious crimes and only after court approval.

On the other hand, high-risk systems (Article 6) AI systems that negatively affect safety or fundamental rights will be considered high risk and will be divided into two categories: AI systems that are used in products falling under the EU's product safety legislation (this includes toys, aviation, cars, medical devices and lifts); AI systems falling into specific areas that will have to be registered in an EU database, that include or are used in fields such as:

- Management and operation of critical infrastructure
- Education and vocational training
- Employment, worker management and access to self-employment

---

Footnote 85 continued

(g)*the placing on the market, the putting into service for this specific purpose, or the use of biometric categorisation systems that categorise individually natural persons based on their biometric data to deduce or infer their race, political opinions, trade union membership, religious or philosophical beliefs, sex life or sexual orientation; this prohibition does not cover any labelling or filtering of lawfully acquired biometric datasets, such as images, based on biometric data or categorizing of biometric data in the area of law enforcement*;

[86] Article 5 (1)(g):
the use of 'real-time' remote biometric identification systems in publicly accessible spaces for the purposes of law enforcement, unless and in so far as such use is strictly necessary for one of the following objectives:
(i)*the targeted search for specific victims of abduction, trafficking in human beings or sexual exploitation of human beings, as well as the search for missing persons*;
(ii)*the prevention of a specific, substantial and imminent threat to the life or physical safety of natural persons or a genuine and present or genuine and foreseeable threat of a terrorist attack;*
(iii)*the localisation or identification of a person suspected of having committed a criminal offence, for the purpose of conducting a criminal investigation or prosecution or executing a criminal penalty for offences referred to in Annex II and punishable in the Member State concerned by a custodial sentence or a detention order for a maximum period of at least four years*.

- Access to and enjoyment of essential private services and public services and benefits
- Law enforcement
- Migration, asylum and border control management
- Assistance in legal interpretation and application of the law.

All high-risk AI systems will be assessed before being put on the market and also throughout their lifecycle. People will have the right to file complaints about AI systems to designated national authorities. Chapter III of the AI Act lists a series of obligations and requirements for high-risk systems to follow, such as: the establishment of a risk management system[87], data governance on the training of such models that is done in accordance with the purpose of such system, the publication of technical documentation before placing on the market, the automatic recording of events during the lifetime of the system, transparency over the system's output to enable deployers to make the correct interpretation and, finally, human oversight.

Consequently, in Chapter IV Article 50 of the AI Act, the Regulation states transparency obligations for all AI models on the market, even for the ones that do not fall into the categories of unacceptable or high-risk systems. In fact, generative models such as ChatGPT are not considered to be high-risk. However, such systems should comply not only with the transparency requirements that are going to be listed, but also with EU copyright law[88]. The transparency obligations that providers should follow, as provided by Article 50 are:

(1) Correct design of the model in such a way that there is explicit information of interaction – for users – with an AI system, except when it is obvious from the point of view of a natural person who is reasonably well-informed, observant and circumspect, taking into account the circumstances and the context of use. This principle is does not apply to AI systems authorised by law to detect, prevent, investigate or prosecute criminal offences, subject to appropriate safeguards for the rights and freedoms of third parties, unless those systems are available for the public to report a criminal offence.

[87] Article 9 (2) of the AI Act explains and defines a risk management system as a "*continuous iterative process planned and run throughout the entire lifecycle of a high-risk AI system, requiring regular systematic review and updating.*"
[88] *See* Senftleben, M., "AI Act and Author Remuneration - A Model for Other Regions?" (February 24, 2024), Institute for Information Law (IViR), pp.1-29. These detailed transparency requirements by the AIA could lead to a less impactful contribution of GPAI models in the EU market, with the risk of having a marginalized role of the availability of EU productions and EU cultural heritage in AI systems.

(2) Plan in a machine-readable format the AI systems that generate audio, video, image or text content and allow detection and reporting of the output as artificially generated or manipulated. The models that have a role of assistance or simple editing – meaning that they do not alter in a substantive way the input data provided by the deployer – are exempted from this requirement.

(3) Information of natural persons when they are subjected to an emotion recognition system or a biometric categorisation system, in accordance with Regulations (EU) 2016/679 and (EU) 2018/1725 and Directive (EU) 2016/680, as applicable, with the exemption of those permitted by law to detect, prevent or investigate criminal offences (and which follow appropriate safeguards and respect EU law).

(4) Disclosure of labelling of "deep fake" on outputs generated by AI systems that manipulate contents, constituting a deep fake. The exceptions are the same as (3). However, when the content forms part of an evidently artistic, creative, satirical, fictional or analogous work or programme, the transparency obligations set out in this paragraph are limited to disclosure of the existence of such generated or manipulated content in an appropriate manner that does not hamper the display or enjoyment of the work. Moreover, in the case of generative text or content published with the purpose of informing the public should be displayed as artificially generated or manipulated – considering the exemption already cited with the addition of subsequent human review or editorial control ad when natural persons are subject to editorial responsibility for the publication of such content.

Nonetheless, Paragraph (7) states that the AI Office should encourage the drawing up of codes of practice at Union level to facilitate the effective implementation of the obligations regarding the detection and labelling of artificially generated or manipulated content[89].

It is clear that, even if copyright laws are not explicitly mentioned in this Regulation, the presence of such precise and extensive transparency obligations of course facilitates the detection of infringement and violations of intellectual property rights. The provision of designing, labelling and informing on artificially generated and manipulated content surely guarantees a more transparent adoption of measures by providers.

---

[89] *See* Geiger, C. and Iaia, V., "Towards an Independent EU Regulator for Copyright Issues of Generative AI: What Role for the AI Office (But More Importantly: What's Next)?", *supra* note 80. According to Article 3(47) AIA, the AI Office is a part of the EU Commission's functions, that is in charge of the correct implementation f the AI Act. Regarding the codes of practice, the AI Office opened a call for expression of interest to participate in the drafting of the first general-purpose AI Code of Practice that should experience its entry into force in April 2025.

Finally, when regulating general-purpose AI (GPAI) models (in Chapter V) the AI Act provides, in Article 53, some obligations that take into consideration the application of European intellectual property laws[90]. In fact, as stated in Article 53 (c) and (d), the "[…] *providers of general-purpose AI models shall*:

(c) *put in place a policy to comply with Union law on copyright and related rights, and in particular to identify and comply with, including through state-of-the-art technologies, a reservation of rights expressed pursuant to Article 4(3) of Directive (EU) 2019/790;*

(d) *draw up and make publicly available a sufficiently detailed summary about the content used for training of the general-purpose AI model, according to a template provided by the AI Office*."

This means that every copyrighted content or work used for the training of such models, allowed according to the explicit provision and reference to Article 4 CDSMD, should be disclosed through public summaries, allowing complete and clear information on the authorship of the materials that are artificially manipulated for the users of these services. This is extremely important in this context, because the Regulation forces AI providers to be transparent about the training process of their models. However, as already stated in Chapter 1, this might become difficult when these models are based on such intricate and complex designs – such as deep learning – where all the materials used for the training are difficult to discern, and becomes complicated to explain the reasoning behind the final output of such models[91].

Moving on to the analysis of the Copyright and related rights in the Digital Single Market Directive, it is interesting to note specific provisions on artificial intelligence training that have a significant impact on copyright and intellectual property rights. The focus will be defined on Articles 3 and 4 that provide a specific exception to Text and Data Mining (TDM), the most commonly used technique for generative artificial intelligence training.

[90] AI Act, Recital 105. Here GPAI are recognised as an important innovation, but also a threat to artists and authors and their creative processes due to their need of vast amount of data for their training. Here, the threat of use of copyrighted works in the context of training is explicitly and solely referred to TDM (regulated in Directive EU 2019/790), with the mention of the rightholders' authorisation (opt-out clause) for the use of their protected works.

[91] *Ibidem*, *supra* note 80. The provision of a public detailed summary on the contents and materials used for training could lead to a difficult relationship between providers and the European market. Since the underlining technological explanation of the methodology of AI systems' training is obscure even for developers, this normative provision could sound as an unfeasible compliance path for AI providers; D Kim, D., "What Is Emerging in Artificial Intelligence Systems?" (17 July 2024), Max Planck Law, pp. 1-7, when talking about emergence in AI systems it states: "The limited predictability, and hence control, of emergents is attributed to incomplete knowledge of the rules governing elements' interactions", underlining the incapacity of the same developers to explain the functioning system that sits at the roots of the AI training model.

Before discussing the legal aspects of the provision, it is necessary to make an introduction on what TDM is and how it works. Text and data mining is a technique that is becoming increasingly popular for conducting research. It entails using automated tools to process large volumes of digital content to identify and select relevant information and discover previously unknown patterns or connections[92]. Text mining extracts information from natural language (textual) sources. Data mining extracts information from structured databases of facts. The extracted information is assembled to reveal new facts or to formulate hypotheses that can be further explored using conventional methods[93]. It consists of different four different stages. First, potentially relevant documents are identified (Stage 1 – "identification"). These documents are then turned into a machine-readable format so that structured data can be extracted (Stage 2 – "normalized documents"). The useful information is extracted (Stage 3 – "derived dataset") and then mined (Stage 4 – "extracted information to discover knowledge") to discover new knowledge, test hypotheses, and identify new relationships. In the context of generative music, TDM analyses large datasets with the objective of extracting patterns in order to generate new compositions. The ability of AI to replicate, transform, and synthesize musical works has raised significant copyright concerns, particularly regarding the extent to which AI-generated outputs can be deemed original works. The European Union's CDSM Directive (Directive 2019/790/EU) addresses TDM in Articles 3 and 4, which carve out exceptions to copyright protection to facilitate research and innovation. Moreover, the definition provided in the Directive, in Article 2, explains text and data mining as an "*automated analytical technique aimed at analysing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations*."

With regard to the exceptions that will now be analysed, it is important to note that in the Preamble to the CDSM Directive, Paragraph (6) states that "*the exceptions and limitations provided for in this Directive seek to achieve a fair balance between the rights and interests of authors and other rightholders, on the one hand, and of users on the other. They can be applied only in certain special cases that do not conflict with the normal exploitation of the works or other subject matter and do not unreasonably prejudice the legitimate interests of the rightholders*." The EU places an important focus on balancing the different

[92] *See* Carroll, Michael W., "Copyright and the Progress of Science: Why Text and Data Mining Is Lawful" (December 1, 2019), *supra* note 61, which underlines the social and scientific value of TDM in research by addressing: (i) the ability to sort information caused by "information overload"; (ii) analysis of large amounts of data to identify patterns and correlations that can either directly or indirectly help to explain causal relations associated with the natural phenomena under investigation; (iii) opening new lines of research.
[93] Definition available at Carnegie Mellon University Libraries, at https://guides.library.cmu.edu

fundamental rights already mentioned, in order to achieve the enjoyment of both in a safe and equitable way.

In Title II, named "*Measures to Adapt Exceptions and Limitations to the Digital and Cross-Border Environment*", there are Articles 3 and 4 that list important exceptions for TDM AI training for scientific research and limitations. Article 3[94] of the CDSM Directive provides a mandatory exception of reproduction, extraction and re-utilisation of databases (which are protected by copyright[95]), artistic works, performances, phonograms, films and broadcasts of TDM conducted by research organizations and cultural heritage institutions for scientific research purposes. This provision allows these entities to reproduce and extract copyrighted works to which they have lawful access. This mandatory exception, however, poses significant limitations to the beneficiaries of this provision. In fact, this norm is referred exclusively to research organizations[96] and cultural heritage institutions[97], which are characterised by non-commercial scopes of their research, contributing to the right to cultural knowledge and scientific research stated in international conventions. Moreover, another limitation is the provision of the lawful access to copyrighted databases, works and material, allowed to such subjects. In the context of AI music, this exception finds little to no application since, not only the "academic era" has been greatly surpassed, but also this kind of industry is extremely commercially driven.

---

[94]Article 3 Directive 2019/790:
"1. *Member States shall provide for an exception to the rights provided for in Article 5(a) and Article 7(1) of Directive 96/9/EC, Article 2 of Directive 2001/29/EC, and Article 15(1) of this Directive for reproductions and extractions made by research organisations and cultural heritage institutions in order to carry out, for the purposes of scientific research, text and data mining of works or other subject matter to which they have lawful access.*
2. *Copies of works or other subject matter made in compliance with paragraph 1 shall be stored with an appropriate level of security and may be retained for the purposes of scientific research, including for the verification of research results.*
3. *Rightholders shall be allowed to apply measures to ensure the security and integrity of the networks and databases where the works or other subject matter are hosted. Such measures shall not go beyond what is necessary to achieve that objective.*
4. *Member States shall encourage rightholders, research organisations and cultural heritage institutions to define commonly agreed best practices concerning the application of the obligation and of the measures referred to in paragraphs 2 and 3 respectively.*"
[95] Article 3 Directive 96/9/EC
[96] Article 1 (1) Directive 2019/790: *"'research organisation' means a university, including its libraries, a research institute or any other entity, the primary goal of which is to conduct scientific research or to carry out educational activities involving also the conduct of scientific research:*
(a) on a not-for-profit basis or by reinvesting all the profits in its scientific research; or
(b) pursuant to a public interest mission recognised by a Member State."

[97] Article 1(3) Directive 2019/790: "*'cultural heritage institution' means a publicly accessible library or museum, an archive or a film or audio heritage institution.*"

However, in Article 4[98] of the same Directive, the TDM exemption established in the previous article is also extended to commercial purposes with the limitations of lawful access and, most importantly, the option for the rightholder of the works used for artificial intelligence training to opt out voluntarily from such use. In other words, the authors and rightholders of the works used for TDM purposes have to explicitly provide, in a machine-readable format[99], the exclusive and reserved use of their works. While this exception surely grants a balance and a limited infringement of copyright and related rights of authors, it also restricts artificial intelligence training not only from achieving a higher and better-quality output, but also from pursuing the same right to scientific and cultural progress.

Placing this debate in the context of generative music makes balancing even more difficult to achieve., especially due to the fact that the music industry focuses mainly on profit-driven goals, relying heavily on licensing[100], royalties and performance rights to ensure that artists receive fair compensation for their work. The growing popularity of AI music, and especially the fact that the quality of the music generated is continuously improving, raises concerns not only about economic displacement but also about the potential erosion of traditional revenue streams for human musicians. Finally, the TDM exceptions provided by Directive 2019/790 surely help create a more flourishing environment for cross-border research between Member States, particularly when those have failed to make national implementations up to par. However – with the opt out mechanism in Article 4 – they limit the ability to grant competitiveness of European companies in the global market. Following the three step test outlined in Article 9 (2) Berne Convention – exceptions should be established only (a) *in certain special cases*; (b) *when reproduction does not conflict*

---

[98] Article 4 Directive 2019/790:
"1. *Member States shall provide for an exception or limitation to the rights provided for in Article 5(a) and Article 7(1) of Directive 96/9/EC, Article 2 of Directive 2001/29/EC, Article 4(1)(a) and (b) of Directive 2009/24/EC and Article 15(1) of this Directive for reproductions and extractions of lawfully accessible works and other subject matter for the purposes of text and data mining.*
2. *Reproductions and extractions made pursuant to paragraph 1 may be retained for as long as is necessary for the purposes of text and data mining.*
3. *The exception or limitation provided for in paragraph 1 shall apply on condition that the use of works and other subject matter referred to in that paragraph has not been expressly reserved by their rightholders in an appropriate manner, such as machine-readable means in the case of content made publicly available online.*
4. *This Article shall not affect the application of Article 3 of this Directive*."

[99] *See* Senftleben, M., "Copyright Data Improvement for AI Licensing – The Role of Content Moderation and Text and Data Mining Rules" (May 4, 2024), Institute for Information Law (IViR), pp.1-19. The importance of providing AI models, during their training phase, with high quality metadata (regarding all of the characteristics that qualify a protected work) not only affects the accuracy level of AI outputs, but it also influences authors' and rightholders' opportunities for licensing agreements.

[100] *Idem*, "Guardians of the UGC Galaxy – Human Rights Obligations of Online Platforms, Copyright Holders, Member States and the European Commission Under the CDSM Directive and the Digital Services Act", JIPITEC Vol. 14 No.3, 2023. A mention to Article 17 CDSMD is fundamental in understanding the scope of the legislation in the balancing of different interest and rights. The provision regulates the direct liability of the online content-sharing service providers for users' uploads. To avoid copyright infringements, the OCSSP usually enter in some kind of agreements or licenses with the rightholders.

with a normal exploitation of the work; (c) *when reproduction does not unreasonably prejudice the legitimate interests of the author* – the TDM exception tries to follow these principles while primarily granting the right to scientific research and cultural knowledge. On the other hand, a provision that enlarges the scope of this exception, perhaps with an "open clause"[101], would be helpful for increasing the competitiveness of the European market in a global setting. Moreover, driven by public interest and the granting of fundamental rights, such provision could help the flourishing of this industry with regards to the promotion of freedom of expression and information.

2.5.2. The United States Framework: *Fair Use* and *Title 17 U.S.C.*

Another important piece of legislation to analyse is, of course, the one provided by the United States Code regarding copyright – regulated in Title 17 – that contains the Copyright Act of 1976. Title 17 introduces a very important principle in the American view of copyright, which is fair use. This analysis will discuss and consider this principle in light of AI training, as done above for EU legislation, to examine the legality of such practices (like TDM) in the context of copyright and related rights of authors and rightholders.

Not considering the definition of a copyrightable work at the moment – that will be later addressed in Chapter 3 – the focus of this discussion concerns mainly the topic of copyright infringement committed by developers when designing an AI model and its training. Since, as already mentioned, the training consists in the processing of massive datasets – that mostly include copyrighted works – it is necessary to see if fair use allows such practices without incurring in copyright violations.

Title 17, Section 107 – titled "*Limitations on Exclusive Rights: Fair use*"[102] – presents the doctrine of exemptions to the enjoyment of exclusive rights granted to the authors of creative works. This principle

---

[101] *See* Geiger, C. and Frosio, G. and Bulayenko, O., Text and Data Mining: Articles 3 and 4 of the Directive 2019/790/EU (October 17, 2019). Concepción Saiz García and Raquel Evangelio Llorca (eds.), "Propiedad intelectual y mercado único digital europeo", Valencia, Tirant lo blanch, 2019, pp. 27-71., Centre for International Intellectual Property Studies (CEIPI) Research Paper No. 2019-08.

[102] Title 17, § 107, U.S.C.: "*Notwithstanding the provisions of sections 106 and 106A, the fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include—*
(1) *the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;*
(2) *the nature of the copyrighted work;*
(3) *the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and*
(4) *the effect of the use upon the potential market for or value of the copyrighted work.*
The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors."

presents a quite particular philosophy, evaluating the legality of using copyrighted works based on four factors to determine fair use:

- the purpose and character of the use
- the nature of the copyrighted work
- the amount and substantiality of the portion used
- the effect of the use on the market

These broad factors consider the use, extraction and reproduction of copyrighted works in a broad sense[103], imposing a case-by-case evaluation by the courts and, for that purpose, a more limited approach to copyright infringement. However, while human authors, when relying on fair use, typically pay for the consultation of copyrighted works or – in cases of no payment – they act under legal constraints. On the other hand, the same cannot be said for artificial intelligence models, which engage in large-scale data scraping that often results in the creation of locally stored copies of millions or billions of copyrighted works, usually resulting in some kind of copyright infringement (when considering the output).

Nonetheless, the use of TDM techniques for training AI models is commonly considered to be part of the fair use doctrine. Academics and researchers consider the first stage of AI models design in the use of copyrighted works as "non-expressive use"[104] of such materials, and so, do not constitute copyright infringements because the training process does not implicate the copyright owner's interest in controlling the communication of their original expression to the public – the copying is simply the first step in an analytical process that typically yields abstract metadata[105] that is then used to create new digital artifacts. In other words, AI models do not explicitly include the data they are trained on: the data used for training influences the model's design and output but, simply put, it does not constitute an integral part of the system and the training data serves only the purpose of instructing the machine. So, Generative AI and its training

---

[103] *See* Sag, M., "God in the Machine: A New Structural Analysis of Copyright's Fair Use Doctrine" (2005). Northwestern Law & Econ Research Paper No. 05-10, pp.381-435, when describing the fair use doctrine as a provision, thought by the Congress, with the goal of making copyright law able to adapt and evolve in response to new challenges without necessitating legal intervention.

[104] *See* Samuelson, P. et al., "Comments in Response to the Copyright Office's Notice of Inquiry on Artificial Intelligence and Copyright", Notice of inquiry ("NOI") and request for comments, Artificial Intelligence and Copyright, Docket No. 2023-6

[105] *See* Lemley, M. A. and Casey, B., "Fair Learning" (2021), Texas Law Review, Volume 99 Issue 4, pp.743-785. In this research a very important note is explained in the discussion on why ML learning should be considered legal, falling under the fair use doctrine. Aside from criticism that comes from the economical threat of the use of copyrighted works, AI training – as already stated – benefits from a verified and high-quality dataset from which it extracts information. The main and final statement that justifies the fairness of AI training is especially that AI models use copyrighted works not to extract the part that are protected by copyright. Indeed, the extraction focuses especially on the non-copyrightable parts of the works, including ideas, facts, functions and methods.

are not created with the purpose of copying that data, but rather designed to use that material on an abstract and uncopyrightable level.

Moreover, the training of such models constitutes, as already mentioned, a "non-expressive use" of copyrighted works and, as such, falls under the provision of fair use. Before addressing the concept of non-expressive use, it is important to note first the definition of "expression" in a copyright context. As R. Brauneis notes[106], there are three kinds of expression to analyse: (a) "constitutive expression", meaning the individual expressive choices made in a work by the author, and protected under copyright law because they relate to historical facts, scientific theories or functional elements; (b) "actionable expression", which constitutes a group of expressive choices that are made by an author and are complex enough to be protected under copyright law, constituting in an intricate and "creative" expression of someone's thoughts and imagination; (c) "felt expression", which refers to as an aesthetic or hedonic reaction in response to the experience of a work of creation.

While the first two types of expression are recognizable and present both in human and artificial authors, the last type is difficult to be embodied by a machine, of course.

Now, it is logical to explain the definitions of "non expressive use"[107]: (a) "non (constitutive) expression" means a work that is not the result or does not provide the expressive use of another work; (b) "non (actionable) expression" refers to a work that does not show in a substantive way the use of a complex structure of expressive choices present in another work; (c) "non (felt) expressive use" relates to a work that does not produce an aesthetic or hedonic reaction to the expressive choices presented in that work.

As a fundamental example, the case of *Authors Guild v. Google*[108] helps in the understanding of this debate. This was a copyright case heard in the federal court for the Southern District of New York, and then the Second Circuit Court of Appeals between 2005 and 2015. It concerned fair use in copyright law and the transformation of printed copyrighted books into an online searchable database through scanning and digitization. It centered on the legality of the Google Book Search (originally named Google Print) Library Partner Project that had been launched in 2003. Though there was general agreement that Google's attempt to digitize books through scanning and computer-aided recognition for searching online was seen as a transformative step for libraries, many authors and publishers had expressed concern that Google had not

[106] *See* Brauneis, R., "Copyright and the Training of Human Authors and Generative Machines" (July 29, 2024). 47 Columbia Journal of Law and the Arts (forthcoming 2025), GWU Legal Studies Research Paper No. 2024-52, GWU Law School Public Law Research Paper No. 2024-52, pp. 1-58.

[107] *Ibidem*

[108] *The Authors Guild Inc., et al. v. Google, Inc.*, 804 F.3d 202.

sought their permission to make scans of the books still under copyright and offer them to users. Two separate lawsuits, including one from three authors represented by the Authors Guild and another by the Association of American Publishers, were filed in 2005 accusing Google of copyright infringement. Nonetheless, Google worked with the litigants in both suits to develop a settlement agreement that would have allowed it to continue the program by paying for works it had previously scanned, creating a revenue model for future books that were part of the search engine, and allowing authors and publishers to opt out. However, in late 2013 a class action status was challenged so, in the end, the District Court granted summary judgment in favour of Google, dismissing the lawsuit and affirming that the Google Books project met all legal requirements for fair use. The Second Circuit Court of Appeals upheld the District Court's summary judgment in October 2015, ruling Google's "*project provides a public service without violating intellectual property law*." The U.S. Supreme Court subsequently denied a petition to hear the case.

In summary, the generative AI training process extracts information from millions or billions of works and, in the process, disassembles or tokenizes their elements to construct a very different representation in the models. Actually, the goal of training is to enable foundation models to discern, among other things, “the structure, syntax, and semantics of language,” including “grammar, sentence construction, and how words and phrases are related to each other” to facilitate the generation of “coherent and contextually appropriate text.” This means that, within the limited context of AI training, the analysis and processing of massive amounts of data produce “metadata”, which translates into data about data, that does not interfere with copyright and related rights of the authors.

The case of *Authors Guild v. Google* demonstrates how the simple analysis and storage of copyrighted works are not considered clear copyright infringement under US copyright legislation and, moreover, fall under the scope of fair use.

Finally, OpenAI is also very clear on the matter, stating on their website “*Training AI models using publicly available internet materials is fair use, as supported by long-standing and widely accepted precedents. We view this principle as fair to creators, necessary for innovators, and critical for US competitiveness.*

The principle that training AI models is permitted as a fair use is supported by a wide range of academics, library associations, civil society groups, startups, leading US companies, creators, authors, and others that recently submitted comments to the US Copyright Office.”

2.5.3. The Asian Framework: *Japan's TDM General Exception* and *China's Generative AI Measures*

Since the discussion is based on a global outlook, it is necessary to have also an eastern standpoint on the matter of artificial intelligence. More importantly, since Asian countries are investing a lot of resources in artificial intelligence research and development, their point of view seems fundamental to analyse. The countries that will be discussed in this paragraph will be Japan and China, as they offer an interesting outlook on their legislative position on artificial intelligence training.

Starting from Japan, it is commonly mentioned in western debates how open its approach is to allowing AI training in the context of copyright. In fact, in its Copyright Act[109], amended in 2018 with the introduction of Article 30-4, Japan's view about AI training on copyrighted works seems to be the most permissive framework while still addressing copyright concerns.

Article 30-4, named "*Exploitation without the Purpose of Enjoying the Thoughts or Sentiments Expressed in a Work*", while listing copyright exceptions, states that: "*It is permissible to exploit a work, in any way and to the extent considered necessary, in any of the following cases, or in any other case in which it is not a person's purpose to personally enjoy or cause another person to enjoy the thoughts or sentiments expressed in that work; provided, however, that this does not apply if the action would unreasonably prejudice the interests of the copyright owner in light of the nature or purpose of the work or the circumstances of its exploitation:*

(i) *if it is done for use in testing to develop or put into practical use technology that is connected with the recording of sounds or visuals of a work or other such exploitation;*

(ii) *if it is done for use in data analysis (meaning the extraction, comparison, classification, or other statistical analysis of the constituent language, sounds, images, or other elemental data from a large number of works or a large volume of other such data; the same applies in Article 47-5, paragraph (1), item (ii));*

(iii) *if it is exploited in the course of computer data processing or otherwise exploited in a way that does not involve what is expressed in the work being perceived by the human senses (for works of computer programming, such exploitation excludes the execution of the work on a computer), beyond as set forth in the preceding two items*."

[109] 著作権法 (昭和四十五年法律第四十八号), Copyright Act (Act No. 48 of 1970)

This provision is then better explained by the overview on General Understanding of AI and Copyright in Japan[110] that was published by the Legal Subcommittee under the Copyright Subdivision of the Cultural Council in 2024. This provision allows for the reproduction of copyrighted material without explicit permission, provided the use is limited to the minimum necessary and does not unreasonably harm the interests of the copyright holders. Moreover, the overview underlines the difference of exploitation of copyrighted works in two different stages of AI use: the first, involves the AI training/development stage, where Article 30-4 is fully applicable since there is "non enjoyment"[111] purpose in such methods; the second, concerns the generation and utilization phase where artificial output and content are generated and/or made public. In the second stage there are two possibilities that can materialise: on one hand, the content that is generated can lead to copyright infringement due to violation of the right of original works that are exploited and reproduced; on the other hand, AI-generated works have the possibility to be considered as copyrightable under Japanese law.

The first stage – training and development – allows the use of copyrighted works without the explicit consent of the original authors, when such use does not have a purpose of enjoyment. For "enjoyment" the Japanese Copyright Office gives the definition of "*the act of obtaining the benefit of having the viewer's intellectual and emotional needs satisfied through using the copyrighted work*"[112]. For example, in the case of literary works the enjoyment materialises in the reading process, for computer programmes it is the act of executing the program and for musical and cinematographic works it is the appreciation of such artistry. In the view of Japanese lawmakers, the act of enjoyment directly influences the financial rewards given to the authors of creative works: this means that, without enjoyment, there is no financial prejudice for the authors. Moreover, in case of AI training for non-commercial or research purposes, but realising the enjoyment factor, can be allowed by the copyright holder with an explicit permission. Finally, such exemption does not apply in case the use of copyrighted works unreasonably prejudices the interests of the copyright owner or in case of use of already copyright-infringing works and materials – with a consequent liability[113].

---

[110] "General Understanding on AI and Copyright in Japan", Japan Copyright Office (JCO), Copyright Division, Agency for Cultural Affairs, Japan, May 2024.
[111] Ueno, T., "The Flexible Copyright Exception for 'Non-Enjoyment' Purposes – Recent Amendment in Japan and Its Implication". 70(2) GRUR International, pp. 145-152, 2021.
[112]*Ibidem*. The difference between 享受 (kyoju – enjoyment) and 悲享受 (hikyoju – non enjoyment) is stated by the examples regulated in (i)-(iii) of Article 30-4, with the addition of a basket clause that includes exploitation by stipulation and 'any other cases'.
[113] Ivi, Article 30-4.

As for China, the primary regulatory framework referring to AI is governed by the Cyberspace Administration of China (CAC) that issued – in 2023 – the Interim Measures for the Management of Generative Artificial Intelligence Services[114], providing clear guidelines for generative artificial intelligence training in relation to copyright. Before analysing the specific and detailed provision on AI, it is important to note that, in Chinese copyright law[115], limitations and exceptions[116] of the enjoyment of copyright by authors are permitted in cases of personal interest and research[117] or justified by public interest

---

[114] 生成式人工智能服务管理暂行办法

[115] 中华人民共和国著作权法

[116] Ivi, Article 24:

"*In the following cases, a work may be used without permission of, and without payment of remuneration to the copyright owner, provided that the name or appellation of the author and the title of the work are indicated, the normal use of the work is not affected and the legitimate rights and interests enjoyed by the copyright owner are not unreasonably prejudiced:*

(1) *use of a published work of another for purposes of personal study, research or appreciation;*

(2) *appropriate quotation from a published work of another in one's own work for the purpose of introducing or commenting a certain work, or illustrating a point;*

(3) *unavoidable reproduction or quotation from a published work in newspapers, periodicals, radio stations, television stations or other media for the purpose of reporting news;*

(4) *publication or broadcasting by newspapers, periodicals, radio stations, television stations or other media of current event articles on issues of politics, economy and religion, which have been published by other newspapers or periodicals, or broadcast by other radio stations or television stations, except where the copyright owner declares that such publication or broadcasting is not permitted;*

(5) *publication or broadcasting by newspapers, periodicals, radio stations, television stations or other media of a speech delivered at a public gathering, except where the author declares that such publication or broadcasting is not permitted;*

(6) *translation, adaptation, compilation, broadcasting, or reproduction in a small quantity of copies, of a published work by teachers or scientific researchers for use in classroom teaching or scientific research, provided that such a work shall not be published or distributed;*

(7) *use of a published work by a State organ to a reasonable scope for the purpose of fulfilling its official duties;*

(8) *reproduction of a work in its collections by a library, archive, memorial hall, museum, art gallery, cultural center or similar institution for the purpose of display, or preservation of a copy of the work;*

(9) *free performance of a published work for non-profit purposes, for which the public does not pay any fees and no remuneration is made to the performers;*

(10) *copying, drawing, photographing or video-recording of a work of art put up or displayed in public places;*

(11) *translation of a published work of a Chinese citizen, legal person or unincorporated organization from the standard spoken and written Chinese language into minority nationality languages for publication and distribution in the country;*

(12) *provision of published works to dyslexics in a barrier-free way through which they can perceive; and*

(13) *other circumstances as provided by laws and administrative regulations.*

The provisions of the preceding paragraph shall apply to the copyright-related rights.

[117] Here, doctrine has confronted the provision of such exceptions as the US' fair use. *See* Zhang, C., "Introducing the Open Clause to Improve Copyright Flexibility in Cyberspace? Analysis and Commentary on the Proposed 'Two-Step Test' in the Third Amendment to the Copyright Law of the Prc, in Comparison With the EU and the US" (2017). Computer Law and Security

and public knowledge, unless the authors specifically declare that the reproduction or use of their work is not permitted. Meanwhile, the specific norm for Generative AI allows for a broader and more detailed regulation of this particular action. In fact, in Article 4 (2) of the Interim Measures, the Chinese Government states, when talking about the necessary requirements that generative AI products should present: "*In processes such as algorithm design, selecting training data, model generation and model optimization, service provision, etc., adopt measures to prevent the emergence of discrimination on the basis of race, ethnicity, religious belief, nationality, region, sex, age, or profession*." Moreover, at paragraph (3), a specific mention to IPRs is done, providing that the requirements should "*respect intellectual property rights and commercial ethics; advantages in algorithms, data, platforms, etc., may not be used to engage in unfair competition*."

In addition, Article 7 of the Generative AI measures dedicates to the pre-training and designing phase, requiring that: "*Providers shall bear responsibility for the legality of the sources of generative AI product pre-training data and optimization training data. Data used for generative AI product pre-training and optimization training shall satisfy the following requirements:*

1. *Conforming to the requirements of the Cybersecurity Law of the People's Republic of China and other such laws and regulations;*
2. *Not containing content infringing intellectual property rights;*
3. *Where data includes personal information, the consent of the personal information subject shall be obtained, or other procedures conforming with the provisions of laws and administrative regulations followed;*
4. *Be able to ensure the data's veracity, accuracy, objectivity, and diversity;*
5. *Other supervision requirements of the state cybersecurity and informatization department concerning generative AI functions and services.*"

It is important to note how this provision defines into detail principles that AI developers should follow when designing a generative AI model. It not only focuses on the principles of a socialist state, but also tries to accurately provide a safe framework for the enjoyment of other fundamental rights not simply

Footnote 117 continued
Review, Vol. 33, 2017, pp.73-86. Computer Law and Security Review, Vol. 33, 2017. However, translating the Chinese term "he li shi yong" as fair use is not accurate, since its meaning refers more to a "fair dealing" which, in practice, is a more restricted clause and provision compared to the first one – mainly notable by the explicitly provided examples in the law.

related to copyright: in fact, intellectual property rights are highly regarded and more frequently mentioned throughout the legislation, but also the focus is on the user's privacy and non-discriminatory treatment while using such models. However, as comprehensive as it seems, the provisions seem a bit too general on whether or not techniques such as TDM are allowed and considered not infringing of copyright. It seems that China, with the strict establishment of these principles, puts the authors in a higher position, where they have the power to allow the use of their own works for AI training purposes, emphasizing the protection of rightholders[118]. Moreover, China has been proactive in implementing content moderation and compliance mechanisms to prevent copyright violations. The Administrative Provisions on Algorithm Recommendation Services, which came into effect in 2022, require AI service providers to register their algorithms and disclose how they process data. This regulation underscores China's commitment to maintaining control over AI development while ensuring intellectual property protection.

2.6. A two-way copyright world?

The analysis done in this Chapter, first by listing fundamental principles that justify the existence of artificial intelligence models and the need to counterbalance this with copyright, and secondly by enunciating different global frameworks, makes the picture seem clearer but blurred at the same time. This affirmation is certainly contradictory, but it has its own reason. The existence of different and heterogenous frameworks, as natural as it seems, makes the reader understand that different approaches, of course, lead to different results. It may seem still obvious, but these results have a great impact on the approach and regulation of such a delicate and complex subject that is artificial intelligence. The different outlook leads to the future and subsequent impact of nations on the investment and development of this technology that is rapidly growing. And, since this progress has now reached a level of democratization – even if it might seem too soon to say it – soon the market will be oversaturated, imposing a new perception of creativity and originality in the human sector. The right to culture and scientific research justifies the existence and progress of AI but, at the same time, contemporary approaches are definitely focused on a profit goal, with big tech companies leading the market and allowing little to no space for small startups and research. The focus of this fundamental right should be central in the regulation of this technology, as it easily can lead to misuse and infringement of more and different rights, between which copyright embodies a pivotal role. Between more strict provisions (such as Europe and China) and broader ones (such as US and Japan) the difference in the approach of course influences not only the progress of the innovation, but also the quality of the services and models put into the market. It seems also absurd that, even if distant geographically,

---

[118] *See* Migliorini, S., "China's Interim Measures on generative AI: origin, content and significance", in Computer Law & Security Review, vol. 53, Article 105985, 2024.

both eastern and western countries, in pairs, present some similarities in their approaches. The question now is, since there are two opposite outlooks and views on exempting AI training from copyright infringement, will there be a two-way copyright world?

To better explain this question, let's take the example of GDPR and its impact on global privacy regulations and amendments[119]. The four important principles set into one of the most important European Regulations (users' rights, explicit consent, accountability, data breach reporting), enforced in May 2018, have influenced and transformed the landscape of data privacy and protection legislation all around the world, setting global standards that now are applied worldwide. Moreover, not only has it influenced different legislators – US, India, South Korea, Canada, Brazil etc. – but also it has influenced companies' approach to privacy and data protection when operating in the market. This allowed the transformation in companies' internal rules to follow the maximum standards available on worldwide legislation, providing efficient and maximum protection even when operating in countries that didn't present the principles set forth by the European Union.

The consideration to be made is, in this case, will a similar experience happen with artificial intelligence – where one legislation sets the principles for the other nations to follow and cover jointly and cooperatively – or will there be a sharp distinction between stricter and broader approaches when regulating AI training?

[119] *See* "The Global Impact of GDPR: How it's influenced privacy laws worldwide", available at https://globalailaw.com

CHAPTER 3

DISCUSSING THE COPYRIGHTABILITY OF AI-GENERATED WORKS

3.1. Introduction

The discussion held until now, even if it required knowledge that will be analysed in detail in this last Chapter, was mandatory to allow the understanding of the main topic in this section. It might have seemed that the logical order of the debate was wrong, talking before about copyright exemptions and after about copyright principles, but this distortion of analysis was fundamental to understand the central discussion about the possibility of imagining that artificial intelligence generated works might fall under the scope and principles of copyright protection. It seems reasonable now that, before arriving at the discussion of the output, it was extremely important to discuss the design and preparatory actions for the creation of AI models, since it is something that happens and is done chronologically before the generation of an output. Here, the discussion will continue the analysis of such debate in the context of generative music, explaining the accepted copyright principles in present days. Then, the central focus of the discussion will revolve around the debate of AI-assisted and AI-generated works, and how different global approaches – and mentioning the already analysed legislations in Chapter 2 – lead judicial cases regarding the recognition of copyright protection to AI-generated works (not necessarily related to the music industry). Subsequently, the first ever lawsuit regarding AI music generators will be analysed, even though at its primordial stages – but still important since its impact could and certainly will revolutionize the current music industry – and the side and concern of music labels will serve as a counterbalance in the discussion. Finally, some amendment and modifications proposals will be done in order to keep copyright legislation up to par with the fast-pacing progress of the innovation that artificial intelligence represents, concluding with a future outlook on approaches that could be undertaken to keep the balancing between the different agents fair and equitable.

3.2. Copyright's anthropocentric approach principles

Introducing the last chapter, the most valuable discussion to be made at this point is the principles that characterise and distinguish copyright as a fundamental right to be enjoyed by natural persons. The philosophy behind the existence of copyright resides, starting from the first ever example that is the Statute of Anne (1709), and can be recognized in three principles (in consideration of the nations' legal frameworks analysed in Chapter 2): the Lockean theory of natural rights – comparing physical labour to creative labour, associating both with property of the fruits of a person's work; the utilitarian or incentive-based theory – regarding the consideration of copyright as a legal tool to promote creativity and innovation by offering financial and legal incentives; and societal and cultural justifications – which refer to the public interest of cultural knowledge and scientific research that is also promoted by the creation of new intellectual works. The philosophy residing behind copyright and its corollaries helped in the formation of principles that are valid and recognised all over the world – with the nations' own twist and interpretation of it – and set the standards for a common approach on such a delicate subject.

In the examination of copyright's principles, it is important to cite the international agreements that set the standards for national jurisdiction in the subject of intellectual property rights.

As already mentioned, Article 2(1) Berne Convention describes the kind of works that are protected, literary and artistic works, giving a complete and detailed list. The principle that is derived from this norm is the principle of originality, meaning that every literal and artistic work is considered per se an original work of art for the purpose of copyright protection. However, computer programs are not included in the protection under the Berne Convention: in this case, the TRIPS agreement compensates for this lack of protection, stating in its Article 10: "*Computer programs, whether in source or object code, shall be protected as literary works under the Berne Convention*." In the enunciation of this corollary, many judges have explained the significance of such principle. The EU, in the *Infopaq*[120] decision, defines originality as the author's "own intellectual creation"; in the US, this is explained by referring to "a modicum of creativity"[121]; China gives a list of protectable works under its Article 3, while Japan does the same in Article 2 – both from their copyright legislations – while giving, in addition, the exact definition of a "work"[122].

The creativity threshold for copyright protection is pretty low, with the necessary and sufficient minimal decree of creativity, and carries no requirement of artistic merit, civic virtue or commercial value. Moreover, the human authorship is implicit, requiring that the work must be the materialisation of the fruits of intellectual labour. Copyright protection does not apply to works that result from natural or feral forces[123]

---

[120] *Infopaq International A/S v Danske Dagblades Forening*, C-5/08

[121] *Feist Publications, Incorporated v. Rural Telephone Service Company, Incorporated*, 499 U.S. 340

[122] Article 2, Copyright Law of Japan: "(i) "work" means a creatively produced expression of thoughts or sentiments that falls within the literary, academic, artistic, or musical domain"

[123] *Naruto v. Slater*, 888 F.3d 418. "[A monkey] is not an 'author' within the meaning of the Copyright Act".

but when a human author manipulates these kinds of manifestations, they can be eligible for copyright protection.

Another principle highlighted in the copyright philosophy is the mode of fixation of the work. Most legislations, especially common law ones, require that the work must be fixed in a tangible medium of expression[124] – for example, the US jurisdiction mentions "writings" as a requirement for fixation. However, also in this case, the requirement is largely open-ended, meaning that any form that can be experienced through the senses applies for copyright protection. This requirement is necessary to demonstrate that the work exists in some kind of form and, consequently, can be automatically considered a protected work. In this case, it is important to note that ideas of the mind are not protected by copyright: in *Baker v. Selden*[125], the Court states that the idea itself cannot be protected by copyright, while only its expression is eligible for protection.

Copyright protection is also applicable in the case of derivative works, meaning those works that take inspiration from or are based on pre-existing works, but still demonstrate an original creativity added that makes them materialise into a new creative work.

Copyright protection comes with the enjoyment of moral and economic rights. While the former are related to the personality of the one that makes the work – hence, the recognition of paternity of the work and the right to prevent alterations that may be deleterious to the honour or reputation of the author – the latter ones are related to all of the rights that manifest a somewhat economic income from the use and distribution of the work on the market – hence, reproduction, adaptation, communication to the public, public performance, distribution etc.

After a brief examination of the basic and foundational principles that stand behind copyright protection, it is now time to discuss the actual query that stands behind this thesis: can AI works be considered protectable under copyright laws?

---

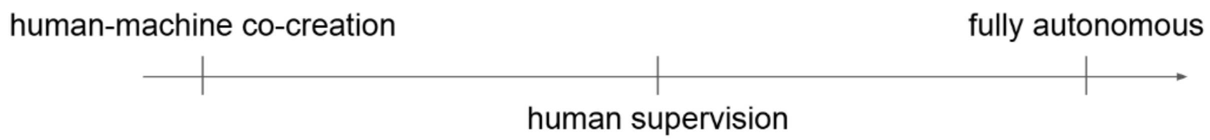
[124] *See* Lemley, M. A., "How Generative AI Turns Copyright Upside Down" (July 21, 2023), Science and Technology Law Review, 25(2), pp. 190-212. On Generative AI, the dichotomy of idea-expression highly discussed in copyright doctrine finds another obstacle. In fact, while it is difficult to grant copyright protection to AI-generated outputs (see Chapter 3.3), the author suggests to recognise the creativity requirement in the prompts made by the user of the AI model. In fact, the most plausible example in which an AI output can be protected by copyright, is when there is a collaboration between the human and the machine. The allocation of copyright protection in the prompts, rather than in the outputs fulfils the criteria of human authorship and creativity, since the prompts are humanly structured. To support this theory, the author describes the latest trend of tech companies that hire "prompt engineers" with the skill of asking the right questions.

[125] *Baker v. Selden*, 101 U.S. 99

3.3. When can an AI work be protected by copyright?: the debate between AI-assisted and AI-generated works in the principle of the US Copyright Office

The fundamental debate to face in this paragraph stands between the recognition of AI-generated works as protectable under copyright law or not. In the case of algorithmic music generation, the line to draw between human creation, intervention, assistance and full computer generation is quite difficult to do. The first questions to pose when talking about Generative music systems are: What is the goal of the system (meaning what type of generative techniques and musical domain it wants to manage)? Who is the user of the system? How autonomous is the system?

Since the more technical discussion has been analysed in detail in Chapter 1, here the examination will revolve around the threshold of autonomy of GM systems. Let's imagine a line where we have three types of music systems:



This scheme clears one's mind on where to pose the threshold for the consideration and distinction between full AI generation and AI assistance in a musical piece. For example, until human supervision a work can be considered to be protectable under copyright laws since there is actual human intervention in the creation of the work, hence there is the expression of one's intellectual effort. However, some examples are needed to make everything clearer: for instance, Melodrive and Aiva are both GM systems that are targeted for professionals (the first for game developers and the second for composers), generating full-piece music and both manifesting a human-machine type of collaboration. The results and outputs of both systems, in this case, could be protected under copyright law – more on the distinction between assistance and full generation later.

Moreover, different techniques produce different outputs and goals between the systems: there is, of course, text-to-music generation which requires just a textual input to generate a full piece of music. In this case, it is a technique which proves to be extremely useful for the users that have little to no knowledge of music theory but still want to have the possibility to produce a musical piece of their own. In addition, there is singing voice cloning, a technique that became very popular in the latest period, which is an algorithm that can replicate the voice of a targeted artist. In this case this technique is targeted at producers that desire to feature and collaborate with a particular artist in their song. Also in this case there is quite a human-machine collaboration but it has proved to be highly criticised from the victims of this method due to possible hijacking of the reputation of the artist that is replicated. As we have already discussed, there is also automatic instrumental accompaniment, which is targeted at singers that have the lead vocals but don't possess the qualities to produce the base of their song. Lastly, as an ultimate example, there is sound synthesis, which follows the goal of creating new sounds that cannot be reproduced by typical music instruments. All of these examples show, in a different percentage, that human and machine collaboration can have different manifestations, and it is quite difficult to distinguish whether the outputs can be protected by copyright or not.

Moving on to the first ever judicial discussion and explanation of AI-generated works and their possibility of falling under copyright protection, the US Copyright Office set a precedent in the history of western law. In a decision dated 11 December 2023, the Copyright Review Board of the United States Copyright Office affirmed the Office's refusal to register an AI-generated artwork submitted by Ankit Sahni. Notably, the starting point of the artwork – entitled "SURYAST" – was an original photograph taken by Sahni. Sahni submitted his photograph into "RAGHAV", an artificial intelligence painting application, then submitted a copy of Vincent Van Gogh's *"Starry Night*" as a "style" to apply to the photograph, then selected the "amount" of style to transfer to the photograph.

Fig.5 The original photograph took by Ankit Sanhi (via https://www.iplawwatch.com)

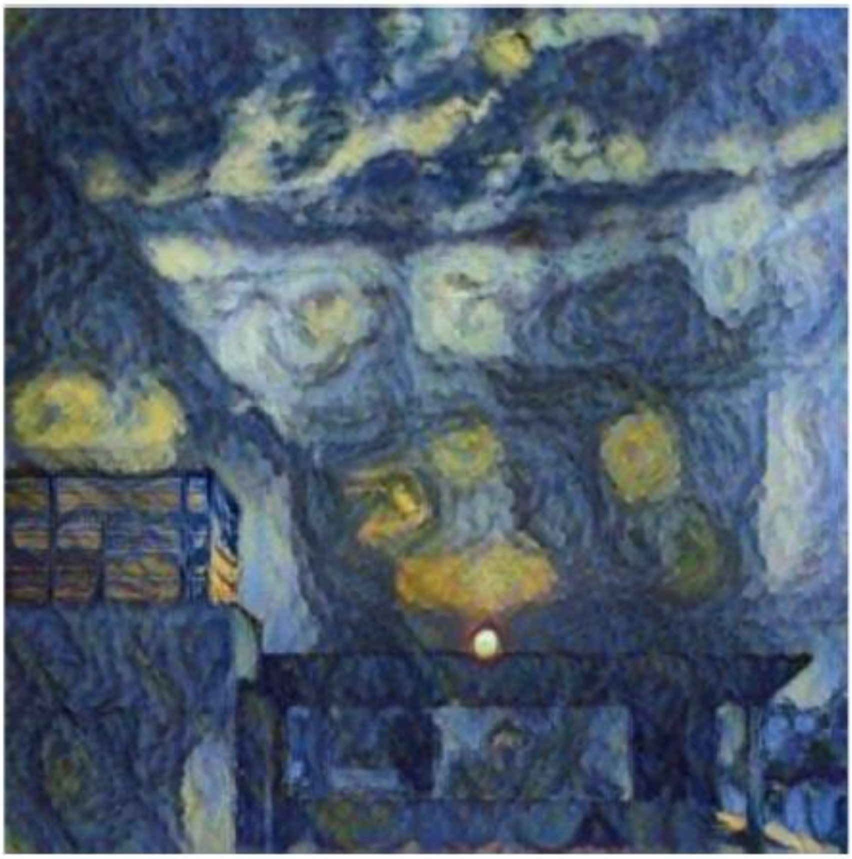


Fig.6 *SURYAST* (via https://www.iplawwatch.com)

On December 1, 2021, Sahni filed an application to register SURYAST with the US Copyright Office. The application was rejected by the Office: initially, on 29 June 2022, on the basis that SURYAST lacked the requisite human authorship; and subsequently, on reconsideration, on the basis that it was a derivative work, being a digital adaptation of a photograph, and that the new aspects of the work were generated by RAGHAV and therefore not the result of human creativity or authorship.

On July 10, 2023, Sahni then requested that the Office reconsider a second time its refusal to register the work. Sahni argued that he was responsible for key creative decisions, including selecting the original photograph, selecting *Starry Night* as the style input and selecting the variable value for the amount of style transfer and that this amounted to sufficient human authorship. Sahni also argued that the work was not a derivative work because the original photograph was "an early stage of what would ultimately become the work." However, The Copyright Review Board affirmed the Office's decision, refusing Sahni's application to register the work. The Board found that Sahni's original photograph was a separate work of authorship because it was fixed separately from the work. The Board went on to find that the work was "not the product of human authorship because the expressive elements of pictorial authorship were not provided by Mr. Sahni." The Board found that RAGHAV, and not Sahni, was responsible for the presence and arrangement of particular elements in the artwork and the colours applied to them. Sahni's contribution – selecting the base image, the style image and the level of style transfer – constituted the unprotectable idea behind the work: an 'altered version of his photograph in the style of *The Starry Night*.

Affirming the Office's refusal, the U.S. District Court for the District Court of Columbia was the first to pronounce itself in the subject of copyrightability of AI-generated works[126] stated that "*copyright law protects only works of human creation*" and that "h*uman authorship is a bedrock requirement of copyright*." It found that "*copyright has never stretched so far [as] . . . to protect works generated by new forms of technology operating absent any guiding human hand*." Because, by his own representation, the "*plaintiff played no role in using the AI to generate the work*" the court held that it did not meet the human authorship requirement. Also, the European Patent Office (EPO) followed the U.S. line of thought and denied patentability[127] of the result of an AI-generated work, for the justification that the author listed and mentioned was not human[128] but, as stated, an AI model named DABUS.

[126] *Thaler v. Vidal*, No. 21-2347

[127] Moreover, the Board of Appeal of the EPO dismissed Thaler's appeal from the EPO's primary decision. The Board of Appeal confirmed that "*under the EPC the designated inventor has to be a person with legal capacity. This is not merely an assumption on which the EPC was drafted. It is the ordinary meaning of the term inventor*."

[128] The European Parliament Resolution of 2017 "Civil Law Rules on Robotics: European Parliament resolution with recommendations to the Commission on Civil Law Rules on Robotics", 2015/2103(INL), 16 February 2017, suggests the

On the other hand, but in a similar situation, China has taken a very different approach compared to the one taken by U.S. Copyright Office. In the case of *Li v. Liu*[129], the Beijing Internet Court (BIC), on November 27, 2023, ruled in an infringement lawsuit that an AI-generated image is copyrightable and that a person who prompted the AI-generated image is entitled to the right of authorship under Chinese Copyright Law. The case started because the plaintiff generated an image of a woman by using Stable Diffusion, an open-source generative AI model that creates images from textual prompts. After publishing the disputed image on a Chinese social media platform (Xiaohongshu), the Plaintiff discovered that the Defendant had used the same image to illustrate an article on a different website without permission. The Plaintiff then sued the Defendant in the BIC.

Specifically, BIC made the following ruling: *"the disputed AI-created image constitutes a "work" pursuant to the Copyright Law of the People's Republic of China."*

In particular, the Court put particular emphasis on the concept that the image was a result of "intellectual achievement": that is because, and as for the Court's ruling, "intellectual achievement" refers to the result of a human being's intellectual activities. In the image, the author's intellectual activities are evinced from the conception to the final creation of the disputed image. Through Stable Diffusion, he selected over 150 prompts, arranged their order and set specific parameters. He continued to adjust and modify those prompts and parameters until the final image aligned with his conception. These steps sufficiently demonstrate that the disputed image was created as a result of the author's intellectual inputs.

As discussed above, it is clear that there are different conceptions and approaches on this matter, since it is extremely difficult, but at the same time very important, to address this new debate that is rising in the field of artificial intelligence. It clearly denotes a hard discussion on which kind of balancing it is necessary to impose to grant for author's rights in the context of copyright and creative works.

Footnote 128 continued
creation of a specific legal status for robots, even though it does not want to recognise legal personality. *See* Frosio, G., "Four Theories in Search of an A(I)uthor", January 8, 2022. in Ryan Abbott (ed), Handbook of Artificial Intelligence and Intellectual Property (Edward Elgar 2022), pp. 155-177.

[129] *Li v. Liu*, (2023) Jing 0491 Min Chu No.11279

3.4. First lawsuits begin: the rise of awareness between music companies and the threats of AI music generators

In June 2024, major record labels, including Sony Music Entertainment, Universal Music Group, and Warner Records, filed lawsuits against AI music generation companies Suno[130] and Udio[131]. The suits, filed in federal courts in Boston and New York, allege that these companies used copyrighted songs to train their AI models without permission or compensation to the rights holders. The plaintiffs seek declarations that Suno and Udio infringed their copyrights, injunctions to prevent future infringement, and damages of up to $150,000 per infringed song.

The core of the dispute centres on the application of the "fair use" doctrine. Suno and Udio argue that their use of copyrighted music to train AI models constitutes fair use, asserting that their AI-generated outputs are transformative and do not replicate existing content. They claim their technology is designed to generate new musical outputs rather than simply replicating existing content.

Conversely, the record labels contend that training AI models to generate music that closely imitates existing songs does not qualify as fair use. They argue that such practices aim to create substitutes for original works without transforming them, thereby infringing on the rights of the original creators. The companies copied music without permission to teach their systems to create music that will "directly compete with, cheapen, and ultimately drown out" human artists' work, according to federal lawsuits filed against Udio in New York and Suno in Massachusetts.

The lawsuits also underscore the lack of licensing deals and transparency around the data used to train many AI models. According to the complaints, Suno and Udio have not disclosed what recordings their models trained on and they have demonstrated that Suno and Udio users have been able to recreate elements of songs including The Temptations' "My Girl," Mariah Carey's "All I Want for Christmas Is You" and James Brown's "I Got You (I Feel Good)," and could generate vocals that are "indistinguishable" from musicians such as Michael Jackson, Bruce Springsteen and ABBA.

---

[130] UMG Recordings, Inc., Capitol Records, Llc, Sony Music Entertainment, Atlantic Recording Corporation, Atlantic Records Group Llc, Rhino Entertainment Llc, The All Blacks U.S.A., Inc., Warner Music International Services Limited, and Warner Records Inc. v. Suno, Inc. And John Does

[131] UMG Recordings, Inc., Capitolrecords, Llc, Sony Musicentertainment, Arista Music,Arista Records Llc, Atlanticrecording Corporation, Rhinoentertainment Company, Warnermusic Inc., Warner Musicinternational Services Limited,Warner Records Inc., Warnerrecords Llc, And Warnerrecords/Sire Ventures Llc. v. Uncharted Labs, Inc., d/b/a Udio.com, And John Does, no. 24-04777.

"[The] *motive is brazenly commercial and threatens to displace the genuine human artistry that is at the heart of copyright protection*" the record labels said in the lawsuits. They said there was nothing about AI that excused the firms from "playing by the rules" and warned that the "wholesale theft" of the recordings threatened "the entire music ecosystem".

The lawsuits come just months after roughly 200 artists including Billie Eilish and Nicki Minaj signed a letter calling for the "predatory" use of artificial intelligence in the music industry to be stopped.

Since these complaints were filed very recently, the federal Courts of Boston and New York have not had the chance to rule on these cases yet. However, both complaints pose serious difficult questions and doubts on the particularity of the claims of the music labels, and it is yet to see what the judges are going to rule on this particular matter.

3.5. Balancing copyright: the solutions of a limitation-based remuneration right

The advent of artificial intelligence in music composition has raised complex legal questions regarding intellectual property rights, authorship, and remuneration. One of the most recent legal debates has centered around the concept of limitation-based remuneration rights, which seeks to balance the interests of human creators and AI-generated music. In this last paragraph this novel approach will be analysed by comparing it with existing legal remedies and demonstrating why it represents a superior solution to the challenges posed by AI music generation.

The awareness regarding the need to address AI in the context of intellectual property, as already discussed in Chapter 2, has been – of course – the hot topic for quite some time. The European Union, before finally drafting the specific regulation on artificial intelligence, brought up to the discussion table the debate on IPRs and the approach needed with regards to this new invention. In February 2019, the European Parliament passed a Resolution[132] in which it called the Commission to "monitor the relevance and efficiency of rules on intellectual property rights to govern the development of AI". A year later, in October 2020, another and more specific Resolution[133] on intellectual property for the development of AI saw the light, suggesting protection for content generated through AI in order to foster and encourage investment in this new technology and '[…] improve legal certainty for citizens, businesses and, since they are among the main users of AI technologies for the time being, inventors'. However, it recognizes that granting

[132] EP, 'Resolution on a comprehensive European industrial policy on artificial intelligence and robotics', (2018/2088 (INI)), 12 February 2019.

[133] EP, 'Resolution on intellectual property rights for the development of artificial intelligence technologies' (2020/2015(INI)), 20 October 2020.

copyright protection to autonomous productions by AI models would be extremely difficult so it suggests the Commission to call for a technologically neutral approach on a common and uniform copyright provision for AI-generated content. Appreciating the effort the EU has made in discussing such a complex topic, however, these debates have not yet materialized an explicit and clear approach on how to deal with copyright protection in the context of AI-generated outputs[134]. Nonetheless, the legal doctrine has suggested the solution of regulating statutory licences or limitation-based remuneration rights to ensure the right protection to rightholders.

First and foremost, limitation-based remuneration rights propose a legal framework wherein human composers and artists receive fair compensation when their works are used to train AI models that generate new musical compositions. This approach is based on the principle that AI-generated music is fundamentally derivative, even if it does not directly infringe copyright law under traditional legal standards. The limitation-based framework suggests:

- a statutory remuneration scheme that mandates compensation for rightsholders whose works contribute to AI training datasets.
- a threshold mechanism where remuneration is due only if AI-generated outputs meet certain qualitative or quantitative similarity criteria.
- collective rights management where licensing bodies distribute revenue collected from AI-related music generation to affected rightsholders.

These principles, and the concept of limitation-based remuneration rights in general, aims to bridge the legal gap between traditional copyright principles and the realities of AI-driven creative industries.

First, under existing copyright law, protection is granted to human-authored works that exhibit originality and creativity. However, AI-generated music often falls into a legal grey area since it lacks human authorship in the traditional sense. Courts and legislators have struggled to define whether AI outputs qualify for copyright protection, leading to uncertainties in ownership and compensation. Between the copyright granted rights applicable to this case, there is the already discussed derivative works doctrine which protects adaptations or modifications of copyrighted works. However, AI-generated compositions

[134] *See* Bulayenko, O. and Quintais, J.P. and Gervais, D. J. and Poort, J., "AI Music Outputs: Challenges to the Copyright Legal Framework" (2022), SSRN, pp.1-153.

often do not meet the strict criteria for derivative works[135] because they may not directly copy any particular original piece but rather generate music in a similar style. For example, and as already discussed, in the U.S. AI training on copyrighted materials might be considered as fair use, further complicating efforts by composers to claim remuneration.

Another option that is important to mention is licensing agreements, with which some AI companies have voluntarily entered licensing agreements with copyright holders. These agreements, however, are non-mandatory and fragmented, leaving many artists uncompensated. Additionally, smaller artists and independent musicians often lack the bargaining power to negotiate favourable terms due to the immense commercial and financial power AI companies have reached in the market. On the other hand, the idea of providing a statutory remuneration based on licenses – which exist on the basis of an exception to exclusive rights of authors by allowing the use of copyrighted works in the context of derivative rights – could help in offering a solution to granting fair and equal remuneration to authors whose works are exploited, especially in the training phase of artificial intelligence models. It is a "permitted-but-paid" regime[136] was particularly provided by U.S. legislation in order to avoid monopoly in the music sector and to reduce transaction costs for the licensing of sound recordings. This particular provision finds its roots and justification in the already cited and analysed provisions of UDHR and ICESCR in Chapter 2. This allows for the application of such principles and the fostering of technological innovation that is now represented by the growth of artificial intelligence. Moreover, this kind of regime is already present in the European Framework, under the InfoSoc[137] Directive, where in its Article 5.2[138] it provides the possibility for Member

[135] That's why it should be considered more of a creative reuse of protected works for the purpose of training and generation of content by AI models. This justifies the provision of a statutory license in the context of such a reuse. On the matter, *see* Geiger, C., "Statutory Licenses as Enabler of Creative Uses" (December 19, 2015). In: R.M. Hilty and K.-C. Liu (eds.), Remuneration of Copyright Owners, Berlin/Heidelberg, Springer, 2017, pp. 305-327., Max Planck Institute for Innovation & Competition Research Paper No. 15-14.

[136] Geiger, C. and Iaia, V., "The Forgotten Creator: Towards a Statutory Remuneration Right for Machine Learning of Generative AI" (October 6, 2023). Computer Law & Security Review, vol 52, 2024, 1-9.

[137] Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society

[138] Article 5.2 Infosoc Directive:
"*Member States may provide for exceptions or limitations to the reproduction right provided for in Article 2 in the following cases:*
(a) in respect of reproductions on paper or any similar medium, effected by the use of any kind of photographic technique or by some other process having similar effects, with the exception of sheet music, provided that the rightholders receive fair compensation;
(b) in respect of reproductions on any medium made by a natural person for private use and for ends that are neither directly nor indirectly commercial, on condition that the rightholders receive fair compensation which takes account of the application or nonapplication of technological measures referred to in Article 6 to the work or subject-matter concerned"

States to allow exceptions and limitations to the reproduction right[139]. In addition, the same Directive offers *fair compensation*[140] in recitals 35 and 36, in the case of specific exceptions or limitations for the use of the rightholders' protected work (Recital 35) and in case of exceptions and limitations that do not require a compensation (Recital 36).

The limitation-based remuneration rights framework presents several key advantages over existing remedies: first, it ensures equitable compensation that, unlike traditional copyright enforcement – which relies on proving direct infringement – provides a more inclusive and systematic approach to compensation by ensuring legal certainty[141]. It recognizes the indirect contributions of human creators by ensuring that those whose works fuel AI training receive due remuneration. Moreover, it surely will help in reducing litigation complexity and high costs since, under current legal frameworks, artists must often engage in costly and uncertain litigation to prove infringement. Limitation-based remuneration rights simplify this process by automating compensation through statutory mechanisms rather than relying on case-by-case judicial determinations. In addition, this solution can surely grant an encourage a more ethical AI development because, by implementing a transparent compensation system, this framework incentivizes AI developers to engage with the music industry in a fair and ethical manner. This fosters collaboration rather than conflict, promoting a sustainable ecosystem where human artists and AI coexist productively.

Lastly, it helps in addressing the vast, and also extremely problematic, nature of AI training. Unlike licensing agreements, which often cover specific works, limitation-based remuneration rights provide collective solutions that ensure broad and proportional distribution of earnings among all contributors.

The limitation-based remuneration rights approach provides a balanced, pragmatic, and forward-looking solution to the challenges posed by AI-generated music. By addressing the limitations of existing legal frameworks, this system ensures that human creators receive fair compensation without stifling technological innovation. As AI continues to reshape the creative industries, adopting such a framework is

---

[139] For example, the German legislator introduced a levy system in its Copyright Act (1965) where the fair compensation or payment to the author whose works are exploited and reproduced, is integrated in a levy imposed in a physical or digital medium that allows the duplication of copyrighted works. Similarly, but in the engineering sector, Article 99 of the Italian Copyright Law (L. 633/1941) provides a fair compensation for the unauthorized use of a technical project for profit purposes. This provision is justified with the reasoning to cultivate technological development and research.

[140] The term fair compensation, however, requires the existence of some kind of damage to the author of the work that has been exploited, which falls out of the scope of the concept of statutory rights. It is interesting to note that in the German version of the Directive, the more neutral term of "remuneration" [Vergütung] is used in the translation of Recital 35, in Geiger, C., "Statutory Licenses as Enabler of Creative Uses", *supra* note 135.

[141] Statutory licenses would replace the uncertain and broad opt-out provision of Article 4 CDSMD for the use of TDM for commercial purposes, in Geiger, C. and Iaia, V., "The Forgotten Creator: Towards a Statutory Remuneration Right for Machine Learning of Generative AI", *supra* note 136.

essential to safeguarding artistic labour[142] while embracing the possibilities of machine-generated creativity.

In light of this new concept to adapt in the artificial intelligence field, the streaming and video platform YouTube published, in August 2023, a set of rules and principles on AI music, fostering a strong collaboration of the platform with major music labels like Universal Music Group. YouTube CEO, Neal Mohan**,** shared the platform's AI music principles and his vision for how the framework will enhance creative expression while also protecting artists on the platform. The principles include:

- "AI is here, and we will embrace it responsibly together with our music partners. As generative AI unlocks ambitious new forms of creativity, YouTube and our partners across the music industry agree to build on our long collaborative history and responsibly embrace this rapidly advancing field. Our goal is to partner with the music industry to empower creativity in a way that enhances our joint pursuit of responsible innovation."
- "AI is ushering in a new age of creative expression, but it must include appropriate protections and unlock opportunities for music partners who decide to participate. We're continuing our strong track record of protecting the creative work of artists on YouTube. We've made massive investments over the years in the systems that help balance the interests of copyright holders with those of the creative community on YouTube."
- "We've built an industry-leading trust and safety organization and content policies. We will scale those to meet the challenges of AI. We spent years investing in the policies and trust and safety teams that help protect the YouTube community, and we're also applying these safeguards to AI-generated content. Generative AI systems may amplify current challenges like trademark and copyright abuse, misinformation, spam, and more. But AI can also be used to identify this sort of content, and we'll continue to invest in the AI-powered technology that helps us protect our community of viewers, creators, artists and songwriters–from Content ID to policies and detection

[142] *Ibidem*. "*A right to a fair remuneration granted by transparency obligations encourages human beings to still produce new works while securing that the use of their work by AI systems generates a fair return. The human creator remains in this way at the centre of the copyright system*". The system of limitation-based remuneration rights follows and respects the anthropocentric approach that constitutes the root of the copyright legislation globally. This provision would help in supporting the use of AI as a tool for the fostering of human creativity and not the contrary, where human input serves for the scope of AI progress.

> and enforcement systems that keep our platform safe behind the scenes. And we commit to scaling this work even further."[143]

The announcement also introduced YouTube's AI Music Incubator, a program that will bring together some of today's most innovative artists, songwriters, and producers to help inform YouTube's approach to generative AI in music.

3.6. Conclusions

This last Chapter posed and tried to answer the focal question of the consideration of granting copyrightability to AI-generated works. Through a brief analysis of already settled and existing copyright principles, to the comparative examination of different legislation, it closed the circle with the proposal of a new concept that may help – according to recent legal debate – to ensure a fair and equitable balancing of copyright protection to authors while fostering and promoting the progress and innovation of artificial intelligence in the context of scientific and cultural research, which both are granted as fundamental and human rights. The different outlooks on different jurisdictions and recent lawsuits make a clear statement that this subject still needs to be analysed and regulated in detail in order to avoid unfair competition and degradation of creativity and originality which makes humanity progress forward.

[143] *See* https://blog.youtube/inside-youtube/partnering-with-the-music-industry-on-ai/

CONCLUSIONS

The purpose of this last section is to analyse what has been discussed all over the dissertation, by giving new insights and summarising the principles that led the discussion to find a closing statement. By concentrating on the history, psychology and purpose of artificial intelligence as a scientific invention and focusing on the specific aspect of its application in the music industry and sector, it is clear to the eye that its journey dates far more back that typical digital users think. The demonisation of such achievement does not find any justification since it is mainly done on the basis that AI is felt as something that happened and developed in the span o f a few years. The amount of research, but also – and most importantly – wait of researchers, developers and investors to make progress in the field is unimaginable. What is happening now was already predicted decades before its realisation. The problem at the time was not the lack of a futuristic outlook, but rather the lack of technological and computational ability and machinery to make those predictions possible. Just like the birth of the – now outdated denomination – "world wide web", artificial intelligence poses many questions, doubts and risks that need to be faced in order to grant and protect what is protected as a human right. Scientific and cultural progress is what justifies the growth and research of such inventions that, of course, need to be counterbalanced by the protection of intellectual property. However, these two rights go hand in hand, since with no intellectual property there would be no cultural progress and vice versa. Both are the balancing and the incentive of one another and these two principles cannot be separated or analysed singularly, especially when talking about artificial intelligence. On the other hand, after training, output is also very important in the legal analysis of artificial intelligence's impact on creativity and the rights of authors and musicians. The two different perspectives brought up by the US on one hand, and China on the other, allow the analysis of benefits and risks for copyright identification or not of an artificially generated creative work.

In the context of a research[144] ChatGPT was asked to define who is the owner of the output generated by this model, and its answer was:

"*As an AI language model, I do not own the copyright of the text generated with my help. The ownership of the text belongs to the user who inputs the prompts and generates the output.*"

However, in the context of music, the matter seems to be much more complicated since it concentrates on concepts like style and musical identity, with the particular detail that artists are identifiable also by their

[144] Lucchi, N. "ChatGPT: A Case Study on Copyright Challenges for Generative Artificial Intelligence Systems" (2024), European Journal of Risk Regulation, 15, pp. 1-23.

voice. Most artists are defining AI music as a threat to their career, as seen in the very recent lawsuits filed against Suno and Udio, while others are recognizing the potentialities of such technology in creating new types and genres never heard before. Just like David Guetta said at the Brit Awards, also the singer named Grimes is incentivizing the use of her voice in the creation of artificial intelligence music. The musician is inviting creators to use AI-generated versions of her voice to make new music, saying she could even provide the raw audio files to facilitate it. The pop singer, whose real name is Claire Boucher, also said she would "*split 50% royalties on any successful AI generated song that uses my voice*."

"*Same deal as I would with any artist I collab with. Feel free to use my voice without penalty. I have no label and no legal bindings*" she added.

Courts have not yet weighed in on AI's use in music and it's unclear how any profit would shake out legally.

This unstable situation makes jurists and regulators think about solutions that cope with this new risk emerging in the copyright sector. The though of including a new limitation-based remuneration right seems the most efficient way to permit the evolution and progress of AI through training on copyrighted works, while still granting and protecting the rights of authors whose works are used in the training process.

At the end, artificial intelligence sets the new frontier of the digital revolution, started with the birth of computer machines, continued with the invention of the internet and progressing even more with the creation of artificial intelligence. As most, if not all, inventions, AI brings benefits and risks like many others, but it certainly needs to be stated that the realisation of such technology makes daily life more efficient and truly helps in tasks and obstacles that human face every day. Moreover, due – or thanks – to its incredible mathematical and predictive powers, AI can help humans to reach new frontiers never imagined before.

Think about the Beethoven X AI Project that tried to use the predictive power of artificial intelligence to complete the great composer’s 10th Symphony. When Beethoven died in March 1827 a part of his legacy were 40 sketches for a 10th unfinished symphony. A team of experts in machine learning and musicology used those sketches to create an AI to finish what the master never could. The 10th symphony completed by AI has then premiered with the Beethoven Orchestra Bonn in front of a live audience. The Beethoven X AI Project has documented the work that led to the completion of the musical piece in a film divided in three segments: the history of L. V. Beethoven, the creation and curation of the 10th Symphony – from the genesis of the 40 sketches to the completed AI opus – and the “human touch”, meaning the live representation of the piece by the orchestra.

This example concludes perfectly the intentions of the thesis. While AI is a new, unpredictable and risky inventions, that arises questions, doubts and worries, on the other hand, developers, researchers and overall users need to understand that this is a tool, a great tool, that helps the human race in reaching new limits that the society has never attempted to reach before. It’s incredible capacity of prediction, by fast mathematical calculations, it undoubtedly helpful when humans cannot reach that fast pace calculations. The trick here is to understand to what extent this new invention can still be considered useful and effective for the sake of progress, and when it becomes a risk that cannot turn back.

BIBLIOGRAPHY

Agostinelli, A. et al., “MusicLM: Generating Music From Text” (2023), arXiv, pp. 1-15.

Anderson, J.R., “The Architecture of Cognition.”, 1983.

Bellier, L. et al. “Music can be reconstructed from human auditory cortex activity using nonlinear decoding models.” (2023), PLOS Biology 21(8), pp. 1-27.

Bender, E. M. et al., “On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?” (2021), Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, pp. 610-623.

Bixio, M.L. and La Rosa, A., “Le tutele telematiche della proprietà intellettuale e industriale” (2020), Giuffré Francis Lefebvre.

Brauneis, R., “Copyright and the Training of Human Authors and Generative Machines” (July 29, 2024). 47 Columbia Journal of Law and the Arts (forthcoming 2025), GWU Legal Studies Research Paper No. 2024-52, GWU Law School Public Law Research Paper No. 2024-52, pp. 1-58.

Bulayenko, O. and Quintais, J.P. and Gervais, D. J. and Poort, J., “AI Music Outputs: Challenges to the Copyright Legal Framework” (2022), SSRN, pp.1-153.

Carroll, M. W., “Copyright and the Progress of Science: Why Text and Data Mining Is Lawful” (December 1, 2019). 53 UC Davis Law Review 893, American University, WCL Research Paper No. 2020-15, pp.893-964.

Chen, Y. et al. “Expected Returns and Large Language Models” (2022), SSRN, pp. 1-62.

Collins, K. “Game Sound: An Introduction to the History, Theory, and Practice of Video Game Music and Sound Design” (2018), MIT Press

Copet, J. et al., “Simple and Controllable Music Generation” (2023), 37th Conference on Neural Information Processing Systems (NEURIPS), pp. 1-17.

Day, B., “In Defense of Copyright: Creativity, Record Labels, and the Future of Music” (2011), Seton Hall Journal of Sports and Entertainment Law, Vol. 21, No. 1, pp. 61-103.

DiCola, P. C. and Sag, M., “An Information-Gathering Approach to Copyright Policy” (October 20, 2012). 34 Cardozo Law Review 173 (2012), pp. 173-245.

Dinwoodie, Graeme B., “The WIPO Copyright Treaty: A Transition to the Future of International Copyright Lawmaking?” (2007), Case Western Reserve Law Review, Vol. 57, No. 4, pp. 751-766.

Dreyfuss, R. and Pila, J. (eds.), “The Oxford Handbook of Intellectual Property Law” (2018), Oxford University Press.

Ebcioglu, K., “An Expert System for Harmonizing Chorales in the Style of J. S. Bach.” (1990), The Journal of Logic Programming (8), pp. 145-185.

Elmsley, A. et al., “Deep Adaptation: How Generative Music Affects Engagement and Immersion in Interactive Experiences”, Digital Music Research Network Workshop at Queen Mary University of London, 19th December 2017.

Eno, B. “A Year of Swollen Appendices”, Faber & Faber, 1996.

Floridi, L. “The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities”, Oxford, 2022.

Floridi, L. & Cowls, J. “A Unified Framework of Five Principles for AI in Society.”, Harvard Data Science Review (1), pp. 1-14.

Frosio, G., “Four Theories in Search of an A(I)uthor”, January 8, 2022. in Ryan Abbott (ed), Handbook of Artificial Intelligence and Intellectual Property (Edward Elgar 2022), pp. 155-177.

Geiger, C., “Elaborating a Human Rights friendly Copyright Framework for Generative AI” (2024), International Review for Intellectual Property and Competition Law 2024, Vol. 55, Issue 7, pp. 1129–1165.

Geiger, C., “Statutory Licenses as Enabler of Creative Uses” (December 19, 2015). In: R.M. Hilty and K.-C. Liu (eds.), Remuneration of Copyright Owners, Berlin/Heidelberg, Springer, 2017, pp. 305-327., Max Planck Institute for Innovation & Competition Research Paper No. 15-14.

Geiger, C. and Jütte, Bernd J., “Copyright, the Right to Research and Open Science: about time to connect the dots” (June 08, 2024), in Enrico Bonadio & Caterina Sganga (eds), A Research Agenda for EU Copyright Law (Edward Elgar, forthcoming 2025), pp.1-21.

Geiger, C., “Reconceptualizing the Constitutional Dimension of Intellectual Property – An Update” (November 29, 2019). in: P. Torremans (ed.), “Intellectual Property and Human Rights”, 4th ed., Austin/ Boston/ Chicago/ New York, The Netherlands, Kluwer Law International, 2020, Centre for International Intellectual Property Studies (CEIPI) Research Paper No. 2019-11, pp. 117-168.

Geiger, C. and Jütte, B. J., “Conceptualizing a ‘Right to Research’ and its Implications for Copyright Law, An International and European Perspective” (2023), 38(1) American University International Law Review, PIJIP/TLS Research Paper Series no. 77, pp. 1-86.

Geiger, C. and Jütte, B. J., “Designing Digital Constitutionalism: Copyright Exceptions and Limitations as a Regulatory Framework for Media Freedom and the Right to Information Online” (August 22, 2023) in Martin Senftleben et al. (eds), Cambridge Handbook of Media Law and Policy in Europe (Cambridge University Press, Forthcoming), pp. 1-18.

Geiger, C. and Iaia, V., “The Forgotten Creator: Towards a Statutory Remuneration Right for Machine Learning of Generative AI” (October 6, 2023). Computer Law & Security Review, vol 52, 2024, pp. 1-9.

Geiger, C. and Iaia, V., “Towards an Independent EU Regulator for Copyright Issues of Generative AI: What Role for the AI Office (But More Importantly: What's Next)?” (August 03, 2024). Auteurs & Media, II, 2024, pp. 185-196.

Geiger, C. and Frosio, G. and Bulayenko, O., “Text and Data Mining: Articles 3 and 4 of the Directive 2019/790/EU” (October 17, 2019). Concepción Saiz García and Raquel Evangelio Llorca (eds.), "Propiedad intelectual y mercado único digital europeo", Valencia, Tirant lo blanch, 2019, Centre for International Intellectual Property Studies (CEIPI) Research Paper No. 2019-08, pp. 27-71.

“General Understanding on AI and Copyright in Japan”, Japan Copyright Office (JCO), Copyright Division, Agency of Cultural Affairs, Government of Japan, May 2024.

Grout, D. J. & Claude V. P., “A History of Western Music”, 5th ed. W. W. Norton & Company.

Hedges, S. A. “Dice Music in the Eighteenth Century” (1978), *Music & Letters* 59, no. 2, Oxford University Press, pp.180-187.

Heaton, J. “Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning” (2018), Genet Program Evolvable Mach 19, pp. 305–307.

Hiller, L., & Isaacson, L. “Experimental Music: Composition with an Electronic Computer” (1959), McGraw-Hill.

Huang, C. et al. “Music Transformer: Generating Music with Long-Term Structure” (2018), International Conference on Learning Representations, pp. 1-14.

Hussain T., et al., “The influence of the COVID-19 pandemic on the adoption and impact of AI ChatGPT: challenges, applications and ethical considerations.” (2024), Acta Psychologica, Volume 246, Article 104264.

Japan’s Approach to AI Regulation and Its Impact on the 2023 G7 Presidency, Habuka H., CSIS, 2023.

Kim, D., “What Is Emerging in Artificial Intelligence Systems?” (17 July 2024), Max Planck Law, pp. 1-7.

Kur, A. and Dreier, T. and Luginbühl, S., “European Intellectual Property Law: Text, Cases and Materials” (2019), Second edition, Cheltenham, England, Edward Elgar Publishing.

LeCun, Y. “A Path Towards Autonomous Machine Intelligence Version 0.9.2 (2022), Lectures given at the Les Houches Summer School on Statistical Physics and Machine Learning, pp. 1-62.

Lemley, M. A., “How Generative AI Turns Copyright Upside Down” (July 21, 2023), Science and Technology Law Review, 25(2), pp. 190-212.

Lemley, M. A. and Casey, B., “Fair Learning” (2021), Texas Law Review, Volume 99 Issue 4, pp.743-785.

Lucchi, N. “ChatGPT: A Case Study on Copyright Challenges for Generative Artificial Intelligence Systems” (2024), European Journal of Risk Regulation, 15, pp. 1-23.

Mahalwar, V., “Copyright and Human Rights: The Quest for a Fair Balance” (2017), in Sinha, M. and Mahalwar, V. (eds) “Copyright Law in the Digital World”, Springer, Singapore, pp.151-174.

Majumdar, A., “Facing the Music: The Future of Copyright Law and Artificial Intelligence in Music Industry” (2023), University College of London, Faculty of Laws, pp.1-18.

McCulloch, W. & Pitts, W., “A Logical Calculus of Ideas Immanent in Nervous Activity” (1943), Bulletin of Mathematical Biophysics vol. 5, pp. 115–133.

Migliorini, S., “China’s Interim Measures on generative AI: origin, content and significance” (2024), in Computer Law & Security Review, vol. 53, Article 105985.

Müller, M. “Fundamentals of Music Processing” (2015), Springer.

Newell, A. et al. “Elements of a theory of human problem solving.” (1958), Psychological Review 65, pp. 151-166.

O'Connor, Seán M., "AI Replication of Musical Styles Points the Way to An Exclusive Rights Regime" (February 15, 2022), Research Handbook on Intellectual Property and Artificial Intelligence, Ryan Abbott ed. (Edward Elgar 2022), pp.65-83.

Pasquier, P. et al. "An Introduction to Musical Metacreation" (2017), Computer Entertainment, Volume 14, Issue 2, pp.1-14.

Plomer, A. "The Right to Science: Then and Now" (2021), Cambridge University Press.

Ramalho, A., Will robots rule the (artistic) world?" (2017), Maastricht University, Journal of Internet Law, pp.12-25.

Roberts, H., Cowls, J., Morley, J. et al., "The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation" (2021), AI & Soc 36, pp. 59–77.

Sag, M., "God in the Machine: A New Structural Analysis of Copyright's Fair Use Doctrine" (2005). Northwestern Law & Econ Research Paper No. 05-10, pp.381-435.

Saiz Garcia C., "Las obras creadas por sistemas de Inteligencia Artifical y su protección por el derecho de autor", InDret (1)2019 (Revista para el análisis del derecho), Barcelona, pp. 1-45.

Samuelson, P. et al., "Comments in Response to the Copyright Office's Notice of Inquiry on Artificial Intelligence and Copyright", Notice of inquiry ("NOI") and request for comments, Artificial Intelligence and Copyright, Docket No. 2023-6

Samuelson, P., "Text and data mining of in-copyright works: is it legal?" Commun. ACM 64, 11 (November 2021), pp. 20–22.

Sarker, I.H. "Machine Learning: Algorithms, Real-World Applications and Research Directions" (2021), SN COMPUT. SCI. 2, 160, pp. 1-21.

Senftleben, M., "AI Act and Author Remuneration - A Model for Other Regions?" (February 24, 2024), Institute for Information Law (IViR), pp.1-29.

Senftleben, M., "Copyright Data Improvement for AI Licensing – The Role of Content Moderation and Text and Data Mining Rules" (May 4, 2024), Institute for Information Law (IViR), pp.1-19.

Senftleben, M., “Guardians of the UGC Galaxy – Human Rights Obligations of Online Platforms, Copyright Holders, Member States and the European Commission Under the CDSM Directive and the Digital Services Act” (September 1, 2023). Journal of Intellectual Property, Information Technology and E-Commerce Law, Vol. 14, No. 3, pp. 435-452.

Simon H. A. & Newell A., “Human Problem Solving: The State of Theory in 1970” (1971), American Psychologist 26(2), pp. 145-149.

Spina Alì, G., “Intellectual Property and Human Rights: A Taxonomy of Their Interactions” (2020), IIC 51, pp. 411–445.

Sturm, B.L.T. and Iglesias, M. et al., “Artificial Intelligence and Music: Open Questions of Copyright Law and Engineering Praxis”, Arts. 2019; 8(3):115, pp. 1-15.

Turing, A. M., “Computing Machinery and Intelligence” (1950), Mind 49, pp. 433-460.

Turing, A. M., “On computable numbers, with an application to the Entscheidungsproblem", Proceedings of the London Mathematical Society, 1936.

Ueno, T., “The Flexible Copyright Exception for ‘Non-Enjoyment’ Purposes – Recent Amendment in Japan and Its Implication” (2021), 70(2) GRUR International, pp. 145-152.

WIPO Intellectual Property Handbook: Policy, Law and Use, 2nd ed., Geneva, 2004.

Xenakis, I. “Formalized Music: Thought and Mathematics in Composition”, Pendragon Press, 1992.

Yu, P. K., “Intellectual Property, Economic Development, and the China Puzzle” (2007). Intellectual Property, Trade and Development: Strategies to Optimize Economic Development in a TRIPS-Plus Era, Daniel J. Gervais, ed., 1st edn, Oxford University Press, Michigan State University Legal Studies Research Paper No. 04-23, pp. 173-220.

Zhang, C., “Introducing the Open Clause to Improve Copyright Flexibility in Cyberspace? Analysis and Commentary on the Proposed ‘Two-Step Test’ in the Third Amendment to the Copyright Law of the Prc, in Comparison With the EU and the US” (2017). Computer Law and Security Review, Vol. 33, 2017, pp.73-86.

CITED CASE LAW

Authors Guild Inc., et al. v. Google, Inc.

Baker v. Selden, 101 U.S. 99

Infopaq International A/S v Danske Dagblades Forening

Naruto v. Slater, 888 F.3d 418

Thaler v. Vidal, No. 21-2347

Li v. Liu, (2023) Jing 0491 Min Chu No.11279

UMG Recordings, Inc., Capitol Records, Llc, Sony Music Entertainment, Atlantic Recording Corporation, Atlantic Records Group Llc, Rhino Entertainment Llc, The All Blacks U.S.A., Inc., Warner Music International Services Limited, and Warner Records Inc. v. Suno, Inc. And John Does

UMG Recordings, Inc., Capitolrecords, Llc, Sony Musicentertainment, Arista Music,Arista Records Llc, Atlanticrecording Corporation, Rhinoentertainment Company, Warnermusic Inc., Warner Musicinternational Services Limited,Warner Records Inc., Warnerrecords Llc, And Warnerrecords/Sire Ventures Llc. v. Uncharted Labs, Inc., d/b/a Udio.com, And John Does, no. 24-0477

LEGISLATION CITED

International legislation

Universal Declaration of Human Rights

Berne Convention for the Protection of Literary and Artistic Works

World Intellectual Property Organization Copyright Treaty

International Covenant on Economic, Social and Cultural Rights

TRIPS Agreement

European legislation

Database Directive (DIRECTIVE EU 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases)

InfoSoc Directive (DIRECTIVE EU 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society)

Copyright and related rights in the Digital Single Market Directive (DIRECTIVE EU 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC)

Artificial Intelligence Act (REGULATION EU 2024/1689)

U.S. Legislation

Constitution of the United States of America

United States Code

Japan Legislation

Copyright Law of Japan (著作権法 (昭和四十五年法律第四十八号) Act no.48 of 1970)

China Legislation

Copyright Law of the People's Republic of China (中华人民共和国著作权法)

Measures for the Management of Generative Artificial Intelligence Services (生成式人工智能服务管理暂行办法)

SITOGRAPHY

https://www.ssrn.com

https://arxiv.org

https://www.semanticscholar.org

https://scholar.google.com

https://www.bbc.com

https://research.gold.ac.uk

https://www.beethovenx-ai.com

https://www.copyright.gov

https://www.senate.gov

https://www.wipo.int

https://thesoundofai.com

https://www.playonlinedicegames.com/mozart

https://musiclm.com

https://musicgen.com

https://eur-lex.europa.eu

https://www.japaneselawtranslation.go.jp

https://www.cac.gov.cn

https://www.sciencedirect.com

https://www.ohchr.org

https://www.icj-cij.org

https://www.aiva.ai

https://www.ibm.com

https://www.europarl.europa.eu

https://artificialintelligenceact.eu

https://copyrightblog.kluweriplaw.com

https://law.mpg.de

https://hdsr.mitpress.mit.edu