ZHANG RUNZHE

GENERATIVE ARTIFICIAL INTELLIGENCE CONTENT: THREE DILEMMAS OF COPYRIGHT PROTECTION AND LEGAL ARGUMENTS OF RESOLUTION

Dissertation to obtain a Master's Degree in Law, in the specialty of Law and Technology

Supervisor:

Dr. Giulia Priora, Professor of the NOVA School of Law

2024 September

Anti-plagiarism statement: I solemnly declare that the thesis submitted is the result of my independent research work under the guidance of my supervisor. Except for the contents that have been cited clearly in the thesis, this thesis does not contain any other works that have been published or written by any individual or group. Individuals and groups who have made important contributions to the research of this paper have been indicated explicitly in the thesis. I am fully aware that I shall bear the legal consequences of this declaration.

Signature: 张润哲

Date: 15 September 2024

Abstract

Generative artificial intelligence (G-AI) provokes concerns regarding the protection of its content under copyright law. A spectrum of definitions of G-AI has been introduced but there is still no universal consensus. Historically, the development of G-AI is progressive and there are four main types of G-AI content. This thesis explores three main legal dilemmas.

The first dilemma concerns the lawfulness of using third-parties' contents for G-AI training. The European Union (EU) sets rules for text data mining but limits its applying scope while the EU AI Act imposes transparency obligations on AI providers. The legislation of China on G-AI training emphasizes the importance of lawful sources and respecting IP rights, but lacks specificity.

The second dilemma concerns the copyrightability of G-AI content. Scholars debate whether G-AI content can meet originality criterion akin to human works. The EU copyright law requires the "author's own intellectual creation" for subject matter to be protected. A Czech court denies the copyrightability of G-AI content due to the lack of human contribution. The copyright law of China defines "work" as original intellectual creations, but does not specify the originality criterion either. However, the Beijing Internet Court recognizes the copyright protection on G-AI content where the human user made substantive creative contribution.

The third dilemma concerns the attribution of rights over G-AI content. There are at least five possibilities: G-AI itself, programmer, user, joint rightsholders or public domain. The EU abandoned the proposal to create an electronic personhood of AI by the Draft Report. Instead, the EU AI Act is introduced as a fine-tuning result. The Beijing Internet Court positions G-AI as an assistant tool and acknowledges the user as copyright holder of G-AI content.

Addressing three dilemmas has the reasonability of encouraging innovation and protecting interests and is an urgent response to factual demands for legal certainty. In conclusion, the possible solution to the first dilemma is to adhere to the ex-ante authorization model with the regime of fair use. The possible solution to the second dilemma is to adopt a fine-tuning originality criterion by assessing the creative contribution of its user. The possible solution to the third dilemma is to recognize the user as rightsholder of G-AI content.

Key Words: Copyright Law, Generative Artificial Intelligence Content, Lawful Use, Copyrightability, Attribution of Rights

Resumo

A inteligência artificial generativa (IA-G) lavanta preocupações quanto à proteção dos seus conteúdos no âmbito do direito de autor. Introduzem-se definições diferentes de G-AI, mas ainda não existe um consenso universal. Historicamente, o desenvolvimento de IA-G é progressivo e há quatro principais tipos dos conteúdos de IA-G. Esta tese explora três dilemas jurídicos principais.

O primeiro dilema refere-se à legalidade do uso de conteúdos de terceiros para o treino de IA-G. A União Europeia (UE) estabelece regras para a prospeção de textos e dados, lmitando o seu alcance enquanto o EU AI Act impõe obrigações de transparência aos prestadores de IA. A legislação chinesa destaca a importância das fontes legais e do respeito pelos direitos de propriedade intelectual, mas carece de especificidade.

O segundo dilema refere-se à copyrightability de conteúdos de IA-G. Académicos debatem se os conteúdos de IA-G podem cumprir o critério de originalidade semelhante ao das obras humanas. A legislação da UE exige uma "criação intelectual do próprio autor" para que o objeto seja protegido e um tribunal checo nega a copyrightability dos conteúdos devido à falta da contribuição humana. A legislação chinesa define "obras" como criações intelectuais originais, mas também não especifica o critério de originalidade. Contudo, o Tribunal da Internet de Pequim reconhece a proteção aos conteúdos de IA-G quando o utilizador humano deu contributo criativo substancial.

O terceiro dilema refere-se à atribuição de direitos sobre os conteúdos de IA-G. Há pelo menos cinco possibilidades: IA-G, programador, utilizador, titulares conjuntos ou o domínio público. A UE abandonou a proposta para criar uma personalidade eletrónica da IA, apresentado em Draft Report, mas introduz-se o AI Act como resultado ajustado. O Tribunal da Internet de Pequimês considera a IA-G como uma ferramenta assistante e recenheuce o utilizador como o titular de direitos.

A resolução destes três dilemas tem a razoabilidade de incentivar a inovação e proteger os interesses e constitui uma resposta urgente às exigências factuais de segurança jurídica. Em conclusão, a solução possível para 1º dilema é aderir ao modelo de autorização ex-ante com o regime de uso justo. A solução possível para 2º dilema é a adoção de um critério de originalidade afinado com a avaliação do contributo criativo do utilizador. A possível solução para 3º dilema é reconhecer o utilizador como titular de direitos de autor.

Palavras-chave: Direito de Autor, Contente de Inteligência Artificial Generativa, Uso Legal, Copyrightability, Atribução de Direitos

Table of Contents

Introduction: Global Popularity of Generative Artificial Intelligence

Marked by the victory of AlphaGo, an artificial intelligence (hereinafter referred to as "AI") program developed by DeepMind, which won the march against the top Go player Lee Sedol in 2016[1], the rapid development of AI has attracted both practical and academic concerns increasingly. Just a few years later, generative artificial intelligence (hereinafter referred to as "G-AI") comes to public view, which can generate coherent and logical content, e.g. text, image, video, and code, in response to prompts and guidance from human users. Those G-AIs, exemplified by ChatGPT, Midjourney and other large language models, demonstrate a powerful capacity for language understanding and content generation, allowing users to create content that is difficult to distinguish from the one purely created by humans. Furthermore, the emergence of G-AI not only stirs a new wave of AI development but also provokes intensive concerns with regard to the protection of G-AI content under the framework of copyright law.

In response to posing the specific challenges to current copyright law, it is necessary to examine the existing legislative framework of corresponding jurisdiction and address specific legal problems thereof. Although the copyright law in different jurisdictions is quite different, there are still some common and inevitable issues.

Specifically, on the one hand, G-AI is based on deep-learning algorithms with a rich variety of pre-training on massive content. However, the content used for training may come from the public domain but contain private interests. Should such training behaviour for G-AI be prohibited, restricted or allowed by copyright law? Should such training behaviour be considered as fair use? On the other hand, regardless of the lawfulness of training behaviour, another contentious question is whether the G-AI content is copyrightable? Furthermore, if the answer is affirmative, the next tough question is to whom the rights over G-AI content

[1] Steven Borowiec, 'AlphaGo seals 4-1 victory over Go grandmaster Lee Sedol' (*The Guardian*, 15 M arch 2016) <https://www.theguardian.com/technology/2016/mar/15/googles-alphago-seals-4-1-victory-over-grandmaster-lee-sedol> accessed 15 September 2024

should be attributed? In other words, should the rights over G-AI content be attributed to its programmer, user or the public domain? Only when this matter is resolved, can disputes be settled ultimately. We already have arrived at a stage where those controversial legal dilemmas brought by G-AI *per se* and the G-AI content must be given positive and effective responses.

In this thesis, the following parts will start by introducing the development of G-AI from a dynamic-historical perspective and the main classifications of G-AI content. Next, this thesis will present three main copyright legal dilemmas in terms of G-AI content: (1) the lawfulness of using third parties' contents for G-AI training purposes; (2) the copyrightability of G-AI content; (3) the attribution of rights over G-AI content. By comparing existing scholars' opinions and diverse legal arguments, a comprehensive comparison will explicate the controversy of proposed resolutions. Finally, this thesis will present its contemplation regarding the possible solutions to three legal dilemmas.

For the purpose of conducting this research, the European Union (hereinafter, referred to as "EU") and China have been chosen as two main jurisdictions with a deliberate analysis of their recent legislative advancements and judicial cases in this thesis. Considering the innovative updates on legislation regarding AI governance and connective judicial cases in both the EU and China, the rapid evolution of the legal framework and the richness of its practice cases provide a solid foundation and resources for research. Through analysis and comparison on strategies of the EU and China in terms of G-AI content and copyright law, their attitudes towards three legal dilemmas discussed in this thesis are valuable materials for concluding possible solutions.

The emergence of G-AI has triggered not only challenges in judicial practice, but also many academic debates on its interaction with copyright law. In particular, three main copyright legal dilemmas are presented, which have received attention from copyright law scholars around the world. They narrate rich rationales for justifying their opinions, which are essential references for completing this thesis. Therefore, many doctrinal legal opinions

are introduced and discussed as significant parts of this thesis.

The research methodology of this thesis includes:

First, literature research method. This thesis analyses the previous research results of scholars and then summarizes and analyses them. Through the sub-study of the current legislative provisions of different jurisdictions, and combining the results of existing academic research, it analyses three main legal dilemmas faced by generative AI content in the context of copyright law and tries to decipher the feasible strategies to deal with the dilemmas for reference.

Second, comparative analysis method. This thesis analyses the latest developments in the legislative system and judicial practice of generative AI in the EU, China and the United States, assesses and compares the views of different jurisdictions, summarizes the commonalities and differences, and analyses the legislative logic and judicial views behind.

Thirdly, case analysis method. This thesis mainly explores whether the content of generative AI constitutes a work protected under copyright law and the attribution of the rights over the generative AI content by analysing the selected cases from the EU, China and the United States, in order to demonstrate practicability in judicial scenario.

Chapter I: Get to Know "G-AI" from a Dynamic-historical Perspective

1.1. Definition of G-AI

Before diving into the discussion of specific dilemmas faced by copyright law in terms of G-AI, the first step is to clarify its definition. The concept of AI was first proposed by John McCarthy in the 1956 Dartmouth summer workshop.[2] Then, a spectrum of definitions of AI has been introduced by scholars but there is still no universal consensus around the world.

The World Intellectual Property Organization (hereinafter, referred to as "WIPO") defines AI as a discipline of computer science that is aimed at developing machines and systems that can carry out tasks considered to require human intelligence and emphasizes that machine learning and deep learning are two subsets of AI.[3] Compared to the board definition given by WIPO, the EU defines the AI system in AI Act as a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.[4] The definition provided by the EU focuses on the structure of "input-output" and emphasizes the influence resulting from outputs generated by AI, which has a closer connection with the feature of generativeness. The legislation of China does not provide any direct definition of AI or G-AI at the national level. Article 2 of Interim Measures for the Management of Generative Artificial Intelligence Services (hereinafter, referred to as "Interim Measures") only stipulates its scope of application as the use of generative AI technologies to provide services

---

[2] John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon, 'A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence' (31 August 1955) <http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf> accessed 15 September 2024

[3] WIPO, Frequently Asked Questions: AI and IP Policy - What is artificial intelligence? <https://www.wipo.int/about-ip/en/artificial_intelligence/faq.html> accessed 15 September 2024

[4] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (2024) OJ L 2024/1689, art 3(1)

to the public in the People's Republic of China for the generation of text, images, audio, video, or other content, mentioning some forms of G-AI content.[5] Nonetheless, some local regulations such as in Shenzhen[6] and Shanghai[7] are enacted effectively, which contain a specific definition of AI and provide a valuable reference to better understand the concept of G-AI (see Table 1).

Table 1: Examples of Definition of AI

	Definition	
WIPO	AI is generally considered to be a discipline of computer science that is aimed at developing machines and systems that can carry out tasks considered to require human intelligence. Machine learning and deep learning are two subsets of AI.	
EU	'AI system' means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.	
China	Interim Measures	No definition
	Shenzhen	Artificial intelligence in this Regulation refers to the simulation, extension or expansion of human intelligence through the use of computers or computer-controlled devices, by methods such as environment perception, knowledge acquisition, and deduction.
	Shanghai	Artificial intelligence in this Regulation refers to theory, methods, technology and application systems that use computers or computer-controlled machines to simulate, extend and expand human intelligence, to perceive the environment, to acquire knowledge and to use that knowledge to obtain the best results.

In my opinion, the debate on the definition of G-AI is not the primary concern of this thesis. Nevertheless, as a factual description, it is more adequate to employ the term of "G-

[5] Interim Measures for the Management of Generative Artificial Intelligence Services (CHN) <https://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm> accessed 15 September 2024

[6] Regulation on Promoting Artificial Intelligence Industry of Shenzhen Special Economic Zone (CHN) <https://opendata.sz.gov.cn/article/article/toArticleDetails/1698878093958434816> accessed 15 September 2024

[7] Regulation on Promoting the Development of Artificial Intelligence Industry of Shanghai (CHN) <https://www.shanghai.gov.cn/hqcyfz2/20230627/3a1fcfeff9234e8e9e6623eb12b49522.html> accessed 15 September 2024

AI content" instead of "G-AI work", because the legal nature of what the G-AI generates is not yet clear.

1.2. From Past to Present: Unveiling the Evolution of G-AI

As with any other technology, the development of AI technology is progressive. Indeed, G-AI is an important branch of the development of AI technology with a long history, dating back to the 1950s. In 1950, Alan Turing introduced a famous theory known as the "Turing Test" in his paper "Computing Machinery and Intelligence", where the possibility of creating machines simulating human intelligence is discussed. According to his theory, a machine can be considered intelligent if it can engage in a dialogue with a human (via a teletype device) without being identified.[8] Passing the Turing Test has traditionally been considered a significant milestone for the development of G-AI, signifying the achievement of human-equivalent abilities in language comprehension and logical reasoning.[9] The Turing Test marked an initial endeavour to assess the intelligence capabilities of machines, serving as an ongoing impetus and direction even for contemporary G-AI research.

The next period is from 1970s to 1980s. G-AI entered the stage of knowledge expression and reasoning. In this period, expert systems and generative reasoning techniques become hot spots for research. Expert systems use the knowledge of human experts to generate reasoning rules for problem-solving. For example, the chatbot PARRY[10] created in 1972 and the chatbot RACTER[11] created in 1983 are two representatives.

From 1990s to the beginning of the 21st century, G-AI achieved breakthroughs in the development of machine learning and neural networks. Machine learning is one of key

---

[8] A.M. Turing, 'Computing Machinery and Intelligence' (1950) Mind 49 433

[9] James Moor, 'The Status and Future of the Turing Test 2001' (2001) Minds and Machines 11(1) 77

[10] Harsh Bhattad, Geeta Atkar, 'Review on Different Types of Chatbots' (2021) International Research Journal of Modernization in Engineering 1347
PARRY was made to simulate a person with schizophrenia, created by psychiatrist Kenneth Colby at Standford University.

[11] Md. Al-Amin, etc., 'History of generative Artificial Intelligence (AI) chatbots: past, present, and future development' (2024) <https://arxiv.org/abs/2402.05122> accessed 15 September 2024
RACTER was created by Chamberlain and Etter in 1983, pioneered the random generation of novel conversational text and prose.

natures of G-AI which enables AI models to simulate human intelligent activities and ultimately to achieve the capacity of intelligence to deal with real-time problems autonomously. Neural network is a generalization of mathematical models of biological nervous systems, i.e., how the human brain works, which consists of three layers: input layer, hidden layer and output layer.[12] The neural network allows AI to be applied for problem-solving in pattern recognition, data analysis, control and clustering.

From 2010 until now, G-AI has made further development in more applicational areas, including but not limited to natural language processing, image generation and processing and video generation and processing. It was also at this time that the relationship between G-AI content and copyright law received increasing concerns from scholars.

1.3. Main Classifications of G-AI content

The G-AI has been widely applied in many areas and entered into public daily life unprecedently. Meanwhile, the expressional forms of expression of G-AI content become more diverse. The following part will introduce four main types of G-AI content, helping in getting a specific perception of G-AI content.

1.3.1. Text

Text-based G-AI content mainly refers to the output of text content such as scriptwriting, marketing copywriting, text translation and codes by G-AI after learning human language and dialogue. In practice, one of the representative examples is Xiaoice, an AI system designed by Microsoft that wrote a collection of poems "The Sunshine Lost Windows" independently in 2017.[13] In order to achieve the skill of writing poetry, Xiaoice has learnt modern poems by 519 poets since the 1920s and has been trained more than 10,000 times.[14]

---

[12] Koushai Kumar, Gour Sundar Mitra Thakur, 'Advanced Applications of Neural Networks and Artifi cial Intelligence: A Review' (8 June 2012) 6 Information Technology and Computer Science

[13] Microsoft Asia News Center, 'Microsoft's Xiaoice, China's newest fashion designer, unveils her first collection for 2019' (12 November 2018) <https://news.microsoft.com/apac/2018/11/12/microsofts-xiaoice-chinas-newest-fashion-designer-unveils-her-first-collection-for-2019/> accessed 15 September 2024

[14] People's Daily Online, 'First AI-authored collection of poems published in China' (31 May 2017) <http://en.people.cn/n3/2017/0531/c90000-9222463.html> accessed 15 September 2024

Notably, it would take a human being about 100 years to read these poems 10,000 times. Furthermore, ChatGPT serves as another illustrative example of its ability to generate text-based responses according to the prompts and guidance from users.

1.3.2. Image

Image-based G-AI content mainly refers to the output of image content after receiving natural language descriptions and matching the pre-branching descriptions by an AI system. Currently, many AI models in the market are equipped with the capacity to transfer texts to images, combining the text-based AI system with the image-based AI system. The text-based AI system can transfer the inputted text into internal expression while the image-based AI system can generate images in accordance with the conditions implied by internal expression. For example, Midjourney has over 15 million users, while Stable Diffusion has more than 10 million users, both being famous AI image generators that have generated billions of images in total.[15]

1.3.3. Audio

The process of G-AI for audio generation relies on the model of transformers, which first converts audio data into recognizable coded text, then trains a model that automatically generates the coded data, and finally transforms the generated coded results into an audio-based G-AI content. The application of audio-based G-AI content is mainly in music composition and arrangement, e.g. Suno AI launched in 2023 and Unio launched in 2024. For example, in 2021, a team of computer scientists and musicians trained an AI model to create the third and fourth movements of Beethoven's Tenth Symphony, which was first performed in Bonn and attracted public concern.[16]

---

[15] Alina Valyaeva, 'AI Has Already Created As Many Images As Photographers Have Taken in 150 Years. Statistics for 2023' (*Everypixel Journal*, 15 August 2023) <https://journal.everypixel.com/ai-image-statistics> accessed 15 September 2024

[16] Anthony K. Brandt, 'Beethoven's Ninth and AI's Tenth: A comparison of human and computation al creativity Journal of Creativity' (December 2023) 33(3) Journal of Creativity

1.3.4. Video

Another important branch of the G-AI application is video generation, which assists users in producing high-quality videos, including detecting and deleting specific clips, tracking clips, generating special effects, compositing videos, etc. For example, in the movie Furious 7, film workers used AI technology to bring the late Paul Walker back to life on the screen, who died on 18 December 2013.[17] Such AI technology placed the virtual avatar "Paul" within real-life scenarios perfectly and provided audiences with an opportunity to watch Paul's performance and say goodbye even after his death.

In the next chapters, this thesis will elaborate on three legal dilemmas of G-AI content in light of copyright protection, by presenting diverse academic arguments and outlining the current legislative frameworks of the EU and China.

[17] Carolyn Giardina, 'How 'Furious 7' Brought the Late Paul Walker Back to Life' (*The Hollywood Reporter*, 11 December 2015) <https://www.hollywoodreporter.com/movies/movie-news/how-furious-7-brought-late-845763/> accessed 15 September 2024

Chapter II: The First Dilemma: Lawfulness of Using Third Parties' Contents for G-AI Training Purposes

2.1. Overview and Leading Academic Opinions

As the core of G-AI technology, machine learning simulates the learning mechanism of human brain to generate outputs, achieving a big leap from pure human creation to machine generation. The process of machine learning can be regarded as a process of AI participation in the creation of works. Importantly, machine learning requires a large amount of contents for training purposes, which inevitably entails contact with works protected by copyright law. According to general rules of copyright law, the utilization of works requires to get authorization from copyright holders and monetary consideration is required in many situations. Otherwise, such utilization without obtaining prior authorization constitutes an illegal act.

However, the scale of used contents normally is exponentially vast and it is almost impossible to strictly get authorization from each copyright holder. Notwithstanding, in some circumstances, the copyright ownership is not clear neither, which makes it extremely hard to identify correct copyright holders. The copyright law emphasizes the doctrine of balance of interests, implying that over-strict limitations to the utilization of training contents may harm the development of G-AI technology, ultimately potentially bringing negative influence to social creativity. To certain extend, it explains why the lawfulness of using third-parties' contents for G-AI training purposes is one of three contentious dilemmas.

Currently, scholars discuss the first dilemma mainly from two aspects: (1) Defining the behaviour of using third parties' content for training purposes; and (2) Discussing the applicability of the regime of fair use.

Scholars have the following different opinions to define the behaviour of using third parties' content for training purposes.

There is an opinion that such behaviour is not an infringement. For example, Edward

Lee assumes that the use of works by AI is "technological fair use" and this argument is supported by a synthesis of existing case law and constitutional underpinnings of US.[18] Matthew Sag thinks that copying expressive works for non-expressive purposes should not be counted as infringement and must be recognized as fair use.[19]

The opposing opinion argues that the behaviour in question should be differentiated first and then evaluated individually to ascertain whether it constitutes an infringement. For example, Abraham Drassinower asserts that the behaviour of using works for machine learning shall be divided into two categories: one is non-works use while the other is fair use.[20] The former is not a use of works in the context of copyright law, for which reason it is not necessary to claim fair use as a defence and the latter can be considered as fair use.

In my opinion, the use of contents for training purposes *per se* is very controversial among scholars while the practical application is even more complex. For this reason, a unitary definition oversimplifies the intricate issue and fails to conduct a detailed analysis in practice.

Regarding whether the regime of fair use is applicable for using third parties' contents for training purposes, there are three main opinions.

The first opinion advocates for applying the regime of fair use. For example, Amanda Levendowski starts by analysing four factors of fair use, asserting that using copyrighted works as contents for training AI systems is highly transformative but such use does not harm the commercial market for copyrighted works, ultimately concluding that the regime of fair use is applicable for training behaviour.[21] Also, Mark A. Lemley and Bryan Casey argue in their paper that fair machine learning should be allowed to use database for training,

---

[18] Edward Lee, 'Technological Fair Use' (3 September 2010) 83 Southern California Law Review
[19] Matthew Sag, 'The New Legal Landscape for Text Mining and Machine Learning' (27 February 2020) 66 Journal of the Copyright Society of the USA
[20] Abraham Drassinower, *What's Wrong with Copying?* (Published on 9 Apil 2015, Harvard University Press)
[21] Amanda Levendowski, 'How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem' (24 July 2017) Washinton Law Review

whether or not the contents restored are copyrighted, which should be treated as fair use.[22]

The opposite opinion holds that the regime of fair use should not be applicable. For instance, Jenny Quang suggests that the regime of fair use shall not be applied to data mining due to the legal uncertainty resulting from the regime of fair use, but another safe harbour is needed because using contents for training AI purposes is fundamentally not copyright infringement.[23]

The third opinion is neutral. For example, Liu Youhua thinks that at current stage, it is not appropriate to completely exclude machine learning from the regime of fair use, nor can it be completely included, but should be analysed according to the specific circumstances of machine learning and adopt the regime of mandatory license.[24] This opinion believes that the regime of fair use is a kind of stop-gap measure, not an ultimate resolution.

2.2. Current Legislation

2.2.1. EU

From the perspective of legislation, there are specific rules for text data mining (hereinafter, referred to as "TDM")[25] that have reference value for analysing the issue of lawfulness of using third parties' contents for G-AI training purposes. Article 3 of Directive (EU) 2019/790 stipulates the exception of text and data mining for the purposes of scientific research, but the scope of applicable subjects is narrowly limited to research organisations and cultural heritage institutions. Furthermore, Article 4 stipulates Member States shall provide for an exception or limitation to the copyrights for reproductions and extractions of lawfully accessible works and other subject matter for the purposes of TDM. However, this

---

[22] Mark A. Lemley, Bryan Casey, 'Fair Learning' (2021) 99(4) Texas Law Review
[23] Jenny Quang, 'Does Training AI Violate Copyright Law?' (2021) 36 Berkeley Technology Law Journal
[24] Liu Youshan, Wei Yuanshan, 'Copyright Infringement Problem of Machine Learning and Its Solution' (2019) 2(2) Journal of East China University of Political Science and Law
[25] 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 (2019) OJ L130/92
Article 2(2): 'text and data mining' means any 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 correlation.

rule only applies to the works that are lawfully accessible, excluding works that are unlawfully accessed. Hence, in practice, using contents for G-AI training purposes normally cannot fall into the applicable scope of Article 3 due to the limitation of subjects. Article 4 indeed opens the possibility of justifying the training behaviour.

In addition, the AI Act imposes transparency obligations for providers, some of which are closely related to the use of training contents. On the one hand, G-AI providers shall provide copyright holders the right to "opt-out". In other words, copyright holders can expressly reserve their works not to be publicly available, therefore not used for G-AI training purposes.[26] On the other hand, the G-AI providers shall make publicly available a sufficiently detailed summary about the content used for training G-AI.[27] Also, Articles 91 provides that the AI Office may request providers to provide the documentation or any additional information for the purpose of assessing the compliance of the provider.[28] Besides, Article 93 provides that the Commission may request provides to take appropriate measures to comply with relevant obligations set out in Article 53.[29] More importantly, the Commission may impose providers fines not exceeding 3 % of their annual total worldwide turnover in the preceding financial year or EUR 15 000 000, whichever is higher., in the case of failing to comply with Article 91 and Article 93. Those specific requirements reflect the basic EU altitude regarding G-AI training: G-AI training shall respect legitimate rights and interests of others, implying that the training process may utilize copyright law protected works of others.

2.2.2. China

Similarly, the legislation of China also sets rules regarding using contents for G-AI

---

[26] Article 53(1)(c) of AI Act: 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

[27] Article 53(1)(d) of AI Act: Providers of general-purpose AI models shall: (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.

[28] Article 91 of AI Act

[29] Article 93 of AI Act

training purposes. The Interim Measures stipulates in Article 7 that G-AI providers shall carry out pre-training, optimisation training and other activities handling training data in accordance with relevant law and regulation, and ensure that the data and foundational models used have lawful sources and the intellectual property rights of others are not infringed.[30] However, the Interim Measures only emphasizes the importance of lawfulness of sources and respecting intellectual property from a very general perspective, but does not provides more specific guidance for further determining whether such use of contents is lawful or not.

Another important legal document is a national standard on G-AI security, namely, Cybersecurity Technology – Security Specification for Generative Artificial Intelligence Pre-training and Fine-tuning Data, which is still an unacted draft for comments. This national standard mandates G-AI providers to document and retain the sources of data used for training purposes, differentiating among various types of data sources. Specifically, it stipulates that G-AI providers shall record: (1) the uniform resource locator for websites if data is from Internet; (2) contracts, cooperating agreements, license and authorization, if data is sourced from organizations or individuals; (3) the service names, users' identification number and authorization if data is sourced from users.[31] Further, the G-AI providers shall also record the version of the G-AI models or services, the time of collection, etc., in the case that the data is sourced from the content generated by G-AI.[32] From the above rules, it can be assumed that the legislation of China also recognizes the importance of respecting others' legitimate copyrights and interests during the process of G-AI training. Unfortunately, the legislation does not stipulate more details.

---

[30] Interim Measures for the Management of Generative Artificial Intelligence Services (CHN) <https://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm> accessed 15 September 2024

[31] Article 6.1 of Cybersecurity Technology – Security Specification for Generative Artificial Intelligence Pre-training and Fine-tuning Data (CHN) <https://www.tc260.org.cn/file/2024-04-01/94e7e6de-2688-472c-af8b-a6cfe7fc7d29.pdf> accessed 15 September 2024

[32] Article 7.1 of Cybersecurity Technology – Security Specification for Generative Artificial Intelligence Pre-training and Fine-tuning Data (CHN) <https://www.tc260.org.cn/file/2024-04-01/94e7e6de-2688-472c-af8b-a6cfe7fc7d29.pdf> accessed 15 September 2024

Chapter III: The Second Dilemma: Copyrightability of G-AI Content

3.1. Overview and Leading Academic Opinions

The copyrightability of G-AI content is the prerequisite for its protection under the copyright law. In other words, it refers to whether the G-AI content could be considered as "work" under the copyright law and then be protected as an eligible subject matter of copyright law. The elements of "works" as stipulated in the copyright law of different jurisdictions are different, but it is commonly agreed that a work at least should possess "originality". The originality criterion serves as a threshold to implement copyright protection on the G-AI content with two fundamental aspects: (1) the subject matter should not be copied, and (2) it should be "intellectual creation". However, there are somewhat different opinions on the interpretation of "originality" in different contexts, subject to different domestic law.

At international level, only general rules are provided, like the "idea-expression" dichotomy, but does not set any specific requirement in terms of the protection of G-AI content, leaving much space for further interpretation. In the context of the EU copyright law, there is no direct harmonization at the EU level regarding the subject matter of copyright but many EU directives and cases address the copyright protection of original work of authorship. Specifically, the G-AI content could be protected under copyright law, only if the content is considered as "an original work of authorship", even though necessary elements and criteria to determine under domestic copyright laws vary from one country to another.

Unsurprisingly, there are two opposing attitudes regarding the copyrightability of G-AI content. However, their arguments and justification are quite different, which will be introduced separately in the following.

The concept of originality is abstract and lacks a universally accepted definition neither in academia nor in practice. As the key element to determine whether G-AI is copyrightable, the originality is placed at the centre of academic debates. Some scholars attempt to affirm

the originality of G-AI content by explaining the meaning and manifestation of "originality". For instance, Atilla Kasap asserts that the G-AI content is copyrightable because it contains a minimum degree of creativity and courts should accept Al-generated work as copyrightable subject matter only if an Al-generated work cannot be differentiated from a human work and possesses as much aesthetic value as a work of the same standard produced by a human would.[33] Moreover, Deng Wen thinks that whether AI-generated content is a work in the sense of copyright law shall be determined in terms of originality and substantive contribution.[34] The underlying logic espoused by these scholars is to adhere to the established originality criterion within the current legislative framework, instead of establishing new criteria for such determination.

Other scholars holding a rejective attitude towards the copyrightability of G-AI content also provide thought-provoking reasoning. They assert that the G-AI *per se* is not a human being and lacks "consciousness" which impedes it from being a qualified "author" in the sense of copyright law. Some scholars refuse to recognize the copyrightability of G-AI content because they argue that the G-AI lacks consciousness and AI cannot become a legal subject under current copyright law.[35] Besides, Wang Qian insists that so far these contents generated by AI are results of the application of arithmetic, logical rules and format, which cannot embody creators' unique personalities, so they cannot be recognised as works in the sense of copyright law.[36]

In my opinion, whether AI can be a proper subject of copyright law and whether the G-AI content is copyrightable are two separate issues. When we are discussing whether the G-AI content is copyrightable, we should set aside the question of the subject of copyright law

---

[33] Atilla Kasap, 'Copyright and Creative Artificial Intelligence (AI) Systems: A Twenty-First Century Approach to Authorship of AI-Generated Works in the United States' (5 June 2019) 19(4) Wake Forest Intellectual Property Law Journal

[34] Deng Wen, 'On the Copyrightability of Content Generated by AI with ChatGPT as a Representative' (2023) 9 Political Science and Law

[35] Selmer Bringsjord, David Ferrucci, 'Artificial Intelligence and Literary Creativity Inside the Mind of Brutus, A Storytelling Machine' (1st ed. 2000) Psychology Press

[36] Wang Qian, 'Argument on the Characterisation of AI-generated Content in Copyright Law' (2017) 5 Journal of Northwest University of Political Science and Law

and focus the discussion on the determining originality criterion In other words, assuming that the same content can be recognized as original if it is created by human beings, its originality should also be recognized if it is generated by G-AI, without taking into account the identity of the subject from which the result is derived, and *vice versa*.

3.2. Current Legislation

3.2.1. EU

The EU copyright law does not explicitly state that copyright requires a natural person creator. However, "the original works of authorship" is highlighted repeatedly in many directives and cases. For instance, both Directive 96/9/EC[37] and Directive 2006/116/EC[38] mention the "author's own intellectual creation" as the criteria to determine the eligibility for copyright protection. Also, the courts stated that the subject matter concerned must be original in the sense that it is the author's own intellectual creation.[39] Hence, the G-AI content can be a subject matter under the EU copyright law only if it is original in the sense that the content is the author's own intellectual creation. Unfortunately, the connotation of "author's own intellectual creation" is not clarified further, reaming ambiguity for interpretation.

In a recent case, the Czech court held that only a natural person can be the author of a copyrighted work and since an AI is not a natural person, an AI cannot be an author.[40] The approach of the Czech court in this case is consistent with the strict requirements for

---

[37] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases (1996) OJ L 77/20
Article 3(1): In accordance with this Directive, databases which, by reason of the selection or arrangement of their contents, constitute the author's own intellectual creation shall be protected as such by copyright. No other criteria shall be applied to determine their eligibility for that protection.

[38] Directive 2006/116/EC of the European Parliament and of the Council of 12 December 2006 on the term of protection of copyright and certain related rights (2006) OJ L 372/12
Article 6: Photographs which are original in the sense that they are the author's own intellectual creation shall be protected in accordance with Article 1. No other criteria shall be applied to determine their eligibility for protection. Member States may provide for the protection of other photographs.

[39] See Case C-310/17 *Levola Hengelo BV v Smilde Foods BV* [2018] CJ, para 36; Case C-5/08 *Infopaq International A/S v Danske Dagblades Forening* [2009] CJ para 37

[40] Tomáš Ščerba, Jaroslav Fořt, 'The first Czech case on generative AI' (4 April 2024) <https://www.technologyslegaledge.com/2024/04/the-first-czech-case-on-generative-ai/> accessed 15 September 2024

copyright to be obtained through the creative choices of human authors, and courts elsewhere in the EU will likely adopt such an approach, in the condition that adverse EU case law has not yet emerged. However, I think that the highlight embodied in this decision lies in its assessment of the inadequacy of the creator's creative contribution, derived from the sole existence of a prompt that, in and of itself, falls short of being eligible for copyright protection. According to the above discourse, it is reasonable to assume that the G-AI content can also be copyrightable if the human creator proves sufficient creative contribution, i.e. inputting detailed and well-described prompts to G-AI.

3.2.2. China

The current copyright law of China defines "work" as intellectual creations with originality in the literary, artistic or scientific domain, insofar as they can be reproduced in a tangible form [41], which requires originality as an essential element but no further interpretation regarding originality is provided in statutes. However, considering Article 11 of Copyright Law, it is reasonable to conclude that the concept of "author" in the context of the copyright law of China includes natural persons, legal entities or other organizations.[42] Hence, G-AI is not an eligible subject under the copyright law of China. In practice, the Beijing Internet Court on December 2023 ruled that only a natural person, legal entity or organization can be author so AI model itself cannot be an eligible author of the picture in question.[43]

Both Czech Court and Beijing Internet Court hold the same opinion that AI cannot be an eligible author in the context of copyright law. However, the conclusions on

---

[41] Article 2 of Regulations for the Implementation of the Copyright Law of the People's Republic of China (2013 Amendment)

[42] Copyright Law of People's Republic of China (2020 Amendment)
Article 11: Except otherwise provided in this Law, the copyright in a work shall belong to its author.
The author of a work is the natural person who has created the work.
Where a work is created according to the intention and under the supervision and responsibility of a legal entity or another unincorporated organization, such legal entity or unincorporated organization shall be the author of the work.

[43] (2023) Jing 0491 Min Chu No.11279; See also Seagull Song, 'China's First Case on Copyrightability of AI-Generated Picture' (7 December 2023) <https://www.kwm.com/cn/en/insights/latest-thinking/china-s-first-case-on-copyrightability-of-ai-generated-picture.html> accessed 15 September 2024

copyrightability are different because both courts consider the level of creative contribution of human beings. Czech courts denies the copyrightability of the picture in question on the basis of the inadequacy of the creative contribution of human beings, while the Beijing Internet Court recognizes the copyrightability of the picture in question because the plaintiff made substantive and creative contribution to the generation of the picture.

Notably, before the case of Beijing Internet Court, Beijing Intellectual Property Court rejected the copyrightability of graphics generated by software in another case[44], ruling that the graphics in question were derived from the data collected by the plaintiff and completed by using the relevant software, and graphics may appear different, but those differences are derived from the differences in the data collected, instead of creative activities. Hence, the court did not recognize these graphics as copyrightable works. In this case, the graphics in question are generated by software, which is traditionally considered as an assistant tool. For this reason, the key point of contention in this case remains around the originality criterion, which necessitates the creative contribution of human beings for the content to be eligible for copyright protection.

[44] The copyrightability of graphics in question was not recognized in the first instance, see (2018) Jing 0491 Min Chu No.239; The original decision was upheld in the second instance, see (2019) Jing 73 Min Zhong No. 2030

Chapter IV: The Third Dilemma: Attribution of Rights over G-AI Content

4.1. Overview and Leading Academic Opinions

Provided that the G-AI content is copyrightable subject to certain circumstances, the subsequent crucial question concerns the attribution of rights over G-AI content. From the perspective of the economic vantage point, the third dilemma involves the commercial utilization and economic value of G-AI content. On one side, people have invested significantly in the invention, utilization and dissemination of G-AI, and are therefore entitled to obtain corresponding remuneration, in line with the theory of utilitarianism. On the other side, the delineation of attribution rights can provide legal protection and a mechanism for rational allocation among relevant stakeholders, thereby achieving a fair balance among them.

Early in 1982, Timothy L. Butler already introduced the theory of "Fictional Human Author", wherein the court initially presumes human authorship in machine-created works, then, after ascertaining that the work meets the other requirements of federal copyright law, determines which individual, including programmers, users and owners of computers, is most deserving of copyright ownership.[45] However, this theory seems to be a gap-stop measure but it fails to clarify how the specific rights are to be allocated, leaving judges with too much discretionary power.

As of now, there are a range of opinions within existing literatures regarding this matter. In addition to aforementioned disagreements in the copyrightability of G-AI content, there are also numerous arguments in the justification of the specific viewpoints and their rationale. Upon the completion of this thesis's research, there are at least five theoretical possibilities for allocating the rights related to the G-AI content: (1) G-AI itself as rightsholder; (2) G-AI programmer as rightsholder; (3) G-AI user as rightsholder; (4) joint rightsholder; and (5) public domain without granting rights or interests. Next, the thesis will further elaborate on

[45] Timothy L. Butler, 'Can a Computer be an Author - Copyright Aspects of Artificial Intelligence' (1982) 4 UC Law SF Communication and Entertainment Journal

the dilemma of the attribution of rights over G-AI content, by critically examining each of the existing proposed solutions.

4.1.1. G-AI Itself as Rightsholder

This viewpoint is extremely controversial among scholars. Those supporting this viewpoint suggest that the AI systems are creative, unpredictable, independent, autonomous, rational, and similar to humans, then AI should be regarded as an independent legal entity entitled to legal and commercial rights and duties.[46] Indeed, there are common reasons for this conclusion. First, the key attributes of AI systems, including intelligence, rational decision-making, and autonomy, resemble those of human beings, therefore AI should be treated as independent legal entities entitled to rights and duties. Alternatively, AI can be very similar to companies which are non-human legal entities that possess legal rights and liabilities.

Nonetheless, more scholars are in denial as to whether AI itself can be the rightsholder of G-AI content, arguing that AI as a legal entity will inevitably create great challenges and uncertainty in the law. For instance, Celine Melanie A. Dee asserts that 'author' pertains to a human author or an 'actual individual who was responsible for creating the work' and author should have a legal personality who may be held legally responsible before the law, implying that G-AI cannot be an author because of the lack of a legal personality.[47] Additionally, Annemarie Bridy argues that the recognition of AI authorship may be a less profound leap than it may seem, because copyright law already recognizes non-human authors (i.e. corporations).[48]

---

[46] Samir Chopra, Laurence F. White, *A Legal Theory for Autonomous Artificial Agents* (2011) University of Michigan Press

[47] Celine Melanie A. Dee, 'Examining Copyright Protection of AI-Generated Art, Delphi - Interdiscipl inary Review of Emerging Technologies' (2018) 1(1) <https://delphi.lexxion.eu/article/delphi/2018/1/11/display/html#63> assessed 15 September 2024

[48] Annemarie Bridy, 'The Evolution of Authorship: Work Made by Code' (8 September 2016) 39 Columbia Journal of Law & the Arts

4.1.2. Programmer as Rightsholder

Given that there are both practical and academic obstacles to recognising the G-AI as the author in the context of the current copyright law in major jurisdictions, another solution is to attribute the rights over G-AI content to human beings who participate in the process of generation. Among those participants, the programmer[49] of G-AI shall be regarded as rightsholder, based on that the algorithms, programs and training data beyond the G-AI models are provided by programmers and/or designers. In other words, the creative activity of G-AI is based on the intelligence and efforts of human programmers, and they contribute more to the generated content than other participants. For this reason, the programmers shall be rightsholders of G-AI content. For instance, Nina I. Brown asserts that programmers exercise the most creative control in determining the parameters for the creative output and the processes the algorithm will use to create that work.[50] Interestingly, in addition to analysing the level of creative contribution of programmers, Brian Golger concluded constitutionally that the programmer should be the author of G-AI content, by looking to the US Constitution.[51]

The opposing reasons are also compelling to some extent. The programmers have been already entitled to copyright in the code itself, which allows them to control its distribution and usage. Assuming that the G-AI content is copyrightable but users still rely on programmers to create the G-AI, the G-AI systems *per se* are valuable enough.[52]

4.1.3. User as Rightsholder

Similar to the arguments in favour of attributing rights to programmers for the G-AI

---

[49] For the purposes of this article, the terms “programmer”, “designer” and “software developer” are used interchangeably to refer to the individual that builds and creates software and applications, tests for errors, and executes the source code of a software application.

[50] Nina I. Brown, ‘Artificial Authors: A Case For Copyright In Computer-Generated Works’ (2019) 20 (1) Science and Technology Law Review

[51] Brian Golger, ‘Copyright in the Artificial Intelligent Author: A Constitutional Approach Using Philip Bobbitt’s Modalities of Interpretation’ (2020) 22 Journal of Constitutional Law

[52] Nina I. Brown, ‘Artificial Authors: A Case For Copyright In Computer-Generated Works’ (2019) 20 (1) Science and Technology Law Review

content, scholars that support recognising the user as rightsholder also acknowledge the value of human creative contribution in the G-AI content, but generally believe that the majority of "creative contribution" comes from users rather than the programmers. For instance, Zach Naqvi thinks that, in the case of AI consumer products, the copyright should belong to the consumer rather than the AI producer because the end-user has final control over what the AI creates.[53] Indeed, this argument seems to be consistent with the rationale of Czech Court[54] and Beijing Internet Court[55], because G-AI here is regarded as an assistant tool. Users or end-users arrange, select, edit and/or significantly modify the original output, making the generated content become a commodity with commercial value, and further creating original works directly by using the G-AI as a tool. As a result, the users are entitled to rights over G-AI content, which also aligns with the legislative objective of encouraging innovation as pursued by copyright law.

4.1.4. Joint Rightsholders

Basically, this opinion argues that the rights over G-AI content should be attributed to different subjects and different participants become joint rightsholders because of the complexity of the process. For instance, Samantha Fink Hedrick believes that when there is a strong argument that the programmer and user both have made substantial contributions to the work, then a justification for joint authorship could be made between them.[56] Admittedly, the unpredictability of G-AI complicates causal responsibility for its acts, and many participants and stakeholders other than the programmers of the G-AI are involved in the creation, design, development, and production of the G-AI system itself, including but not limited to data vendors, trainers, suppliers, holders of AI systems, system operators,

---

[53] Zack Naqvi, 'Artificial Intelligence, Copyright, and Copyright Infringement' (2020) 24(1) Marquette Intellectual Property Law Review

[54] Tomáš Ščerba, Jaroslav Fořt, 'The first Czech case on generative AI' (4 April 2024) <https://www.technologyslegaledge.com/2024/04/the-first-czech-case-on-generative-ai/> accessed 15 September 2024

[55] (2023) Jing 0491 Min Chu No.11279; See also Seagull Song, 'China's First Case on Copyrightability of AI-Generated Picture' (7 December 2023) <https://www.kwm.com/cn/en/insights/latest-thinking/china-s-first-case-on-copyrightability-of-ai-generated-picture.html> accessed 15 September 2024

[56] Samantha Fink Hedrick, 'I "Think," Therefore I Create: Claiming Copyright in the Outputs of Algorithms' (2019) 8(2) NYU Journal of Intellectual Property & Entertainment Law

employers or investors, the public, and the government.

However, opponents provide two main reasons. First, joint authorship requires an intent by both parties to be co-authors but the programmers cannot know all users because of the statistical impossibility, thereby making it impossible to imply that there is an intent to be co-authors. The second reason to be cautious of a joint authorship framework is that it could potentially trigger a 'fractionalization' of ownership rights, leading to various disentangled parties claiming to authorship simultaneously.[57] Furthermore, it may result in uncertainty of copyright law and inability to fundamentally address the attribution of rights over G-AI content.

4.1.5. Public Domain

The last argument asserts that the G-AI content ought to be public property, enabling anyone to use the content without the imposition of any costs or compensation. One of the rationales for this argument is that the users of the "Creativity Machine" have made no creative efforts to generate the results so users cannot be rightsholders while the "Creativity Machine" itself cannot be a rightsholder either, and the content generated falls into the public domain.[58] I disagree with this argument. First, its denial of the users' contribution to the creative process is questionable. Besides, it does not distinguish between the content purely generated by G-AI and the one generated by G-AI with human contribution.

Another rationale justifies that the used contents come from the public domain so the generated content ought to belong to the public domain. However, I think the second rationale is unconvincing because public contents merely constitute a part of the whole training data. In contrast, those contents protected by copyright law serve as the primary focus of training process.

---

[57] Nina I. Brown, 'Artificial Authors: A Case For Copyright In Computer-Generated Works' (2019) 20 (1) Science and Technology Law Review

[58] Ralph D. Clifford, 'Intellectual Property in the Era of the Creative Computer Program: Will the True Creator Please Stand Up?' (1997) 71 Tulane Law Review

Moreover, this argument undermines the legislative purposes of copyright law because it neglects the necessity to incentivize innovation through the recognition of private rights over G-AI content and the inability to offer commensurate rewards for participants.

4.2. Current Legislation

4.2.1. EU

Back on 31 June 2016, the Committee on Legal Affairs of European Parliament published a Draft Report on Civil Law Rules on Robotics[59], which seeks to create a specific legal status for robots, i.e. electronic personhood, that would enable them to have specific rights, obligations and liability in the event of infringements against third parties, instead of being attributed to the manufacturer, developers, owners or users. The Draft Report mentions the intellectual property rights and the flow of data regarding Robots, whereas calls for an "own intellectual creation" for copyrightable works produced by computers or robots[60], as well as the creations of a specific Agency for robotics and AI in order to provide the technical, ethical and regulatory expertise.[61] This Draft Report received controversial debates once published and some scholars were calling for signing an open letter for European Commission to express their concerns regarding the Draft Report.[62] Additionally, it is criticised that, from an ontological perspective, all advanced technologies are not subjects, but merely objects, and there are no reasons to grant them rights, nor hold them legally responsible.[63] Finally, the European Commission affirmed its favour for a fine-tuning of laws and regulations relating to AI on the basis of existing legal norms, rather than granting electronic personhood to AI and robots. Further legal documents like the AI Act do not

[59] Draft Report with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL))
[60] Draft Report with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), para 10
[61] Draft Report with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), para 8
[62] Open Letter to the European Commission Artificial Intelligence and Robotics, <https://robotics-openletter.eu/> accessed 15 September 2024
[63] Andrea Bertolini, 'Artificial Intelligence and Civil Liability' (2020) <https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU(2020)621926_EN.pdf> accessed 15 September 2024

mention the issue of electronic personhood in the future.

The Draft Report was not turned into effective legislation but it indeed sparked a heated debate on the issue of electronic personhood of AI among people from each social sector. The EU has since made its position clear through the AI Act that it is not possible to confer legal personality on AI because AI is not an eligible subject under copyright law, and the rights over G-AI content cannot be attributed to G-AI itself. In my opinion, the Czech court, while rejecting the copyrightability of the image in question, implied the possibility of attributing copyright to its user, but only if the generated content meets the copyrightability requirements of the EU and domestic copyright law.

4.2.2. China

As mentioned earlier, neither the EU copyright law nor China's copyright law does not grant any rights to G-AI itself, at least in the current stage. Interestingly, both Czech Court and Beijing Internet Court, while ruling out the possibility of granting copyright to G-AI itself, underscored the significance of users' creative contribution. The rationale behind this judgement implies that, although the G-AI itself cannot obtain the copyright, if users have sufficient creative contribution in using the G-AI to generate content, i.e. inputting enough intricate prompts as well as modification commands, thereby guiding G-AI's outputs, the G-AI content may be recognized as an eligible subject of copyright law within certain jurisdictions, with the copyright of G-AI content subsequently attributed to users. From the reasoning of those two cases, it can be concluded that Chinese courts *de facto* adopt a similar attitude to the EU: The rights over G-AI content cannot be attributed to G-AI itself but users may be entitled to the copyright of G-AI content under some conditions.

In my opinion, this conclusion of Beijing Internet Court positions the G-AI in the role of an assistant tool, which seems to be a strategic compromise within the framework of the current copyright legislation in order to avoid apparent contradiction with the provisions of the copyright law, while the court has to respond to the urgent demand of a certain confirmation regarding the attribution of rights over G-AI content.

Chapter V: Consolidating Legal Responses to G-AI Content under Copyright Law

A rational and constructive discussion necessitates a mutually agreed foundation. Although perspectives beyond or not included in the framework are not inherently unreasonable, they do not have much relevance with the matter in question. Therefore, it is necessary to clarify the precise scope of topic. In this thesis, the G-AI content should be limited to the content generated by G-AI with human participation, not including the content merely generated by G-AI without any human participation.

Furthermore, the lawfulness of the training behaviour and the copyrightability of the generated content are two relatively independent questions, implying that the generated content can still be not considered copyrightable even though the training behaviour is lawfulness, and *vice versa*. The first dilemma involves concerns regarding the applicability of the regime of fair use, legitimate interests and rights between copyrights holders and G-AI developers. Meanwhile, the second and third dilemmas are more related to the originality criterion in different jurisdiction and interpretation provided by judicial cases.

Notwithstanding, the process of G-AI generating content is dynamic and inexhaustible, which could result in simple and direct outputs as well as complex and deep outputs. Some simple and direct outputs may not be recognized as works in the context of copyright law, due to the lack of originality. Indeed, there are many factors that may affect the output results, including the richness of the training data, the algorithms of G-AI, prompts and guidance of users, etc. Each factor contributes to the generation of ultimate outputs to some extent and the whole process is dynamic, leading to endless possibility of outputs. Likely, the process of generating content is similar to human beings' creation, while there are also factors may affect outputs of human beings, such as education level, knowledge base and personal experience. Undeniably, not all results that are created by human beings are copyrightable for granted. The originality criterion serves as a threshold of copyright protection, excluding contents failing to meet the minimum requirements of copyright law in a specific jurisdiction, either generated by G-AI with human participation or created by human beings.

5.1.Applying Reasonability: Encouraging Innovation and Protecting Interests

Regardless of whether one believes or not, the human beings play an irreplaceable role in the emergence and evolutional development of the G-AI. From Turing Test to machine learning, numerous human scientists and entrepreneurs have invested in the design, development, training, testing and launch of G-AI. Human participation provides not only technical supports, but also countless contents for training G-AI. Also in some situations, human users input meaningful prompts and guidance for G-AI, indicating and adjusting the generated content. Hence, during the process of generating and modifying results by G-AI, it is still obvious to find evidence of human being partial or whole participation. Since humans indeed make monetary investment and intellectual contribution, it is reasonable that those investors and participants should be appropriately rewarded according to the utilitarianism theory of IP. Therefore, it is necessary to provide legislative clarification and affirmation on those three dilemmas related to the G-AI content.

Furthermore, G-AI has obviously become an economic bonanza in the current era, with billions of G-AI users around every corner in the world. Respectively, the commercial utilization of G-AI is innovative and economically valuable, bringing new growth to cultural and creative industries. Legal clarity and certainty will not only facilitate the resolution of conflicts between different parties regarding the G-AI content, but also ensure their interests thereof and encourage further social innovation at the fundamental institutional level. Specifically, the first dilemma should be answered in order to protect respective legitimate interests of G-AI developers and other copyright-holders. The second dilemma regarding the copyrightability and the third dilemma regarding the attribution of rights over G-AI content are deeply combined with more parties, including but not limited to G-AI developers, users and even the public interests.

5.2.Responding to Urgency: Factual Demands for Legal Certainty

AI is a rapidly growing industry, with forecasts indicating an annual growth rate of 37.3% from now until 2030, which is primarily fuelled by the ongoing advancements and

widespread adoption of G-AI technologies.[64] In the current circumstance, G-AI has already infiltrated various fields such as art, music, literature, film and television. However, due to the rapid development of technology and the latency of legislation, those matters related to the G-AI content do not receive a timely response, even though academic discussion has been ongoing in various forums. As of the completion of this thesis, a few legislations respond clearly to the mentioned three dilemmas. Regarding the first dilemma, the EU AI Act impose transparency obligations on AI providers and meanwhile China's AI Interim Measures and drafted national standard stipulate specific compliance requirements from legislative aspect, but neither of them does not give certain answer directly to define the training behaviour. Other two dilemmas are even more contentious, with a wide range of academic arguments, which have been analysed in the previous parts.

First, either defining the behaviour of using contents for G-AI training purposes, and the recognition of rights to the G-AI content shall be considered as an acknowledgement of legitimate interests of G-AI developers and users. On the one hand, G-AI developers have invested largely in designing and training G-AI models with massive interests of others involved. On the other hand, users may make creative contribution by providing precise prompts and detailed guidance, substantially affecting the process of generating and modifying contents by G-AI. If those dilemmas cannot be resolved adequately, enthusiasm and motivation of developers and users are supposed to be negatively influenced, ultimately resulting in grey area as well as legal uncertainty. Although there are not so many examples of infringement involving the G-AI content at now, further advancement of G-AI and the diversification of types of G-AI content portend an imminent escalation of disputes, particularly if the copyrightability of the content and the attribution of its rights remain ambiguous and undefined.

Second, the urgency is also reflected in the demand to address ethical and moral

[64] Melissa Malec, 'Generative AI Statistics: The 2024 Landscape – Emerging Trends and Developer Insights' (HatchWorksAI, 19 January 2024) <https://hatchworks.com/blog/software-development/generative-ai-statistics/> accessed 15 September 2024

hardships brought by G-AI. Traditional copyright revolves around the human being, but such paradigm seems to be facing a challenge from G-AI, because some people argue that G-AI can be an eligible subject of copyright law. Though the courts insist in treating G-AI as an assistant tool that is not fundamentally different from other software, irreversible trends in commercial applications will only exacerbate those dilemmas. A series of ethical concerns, including the protection of privacy, accountability, and transparency, necessitates immediate and effective exploration for answers and countermeasures.[65] Hence, addressing those three legal dilemmas is not only a legislative resolution to legal uncertainty, but also a matter of ethical consideration, ultimately in order to ensure that the training of G-AI and the generation, use and commercialisation of G-AI content are compliance with ethical and moral principles.

5.3.Possible Solution to the First Dilemma: *Ex-ante* Authorization and Fair Use

Based on the data training feeding to form a more mature large G-AI model has become a necessary path for AI upgrading and iteration, but the behaviour of using contents for G-AI training purposes not only conflicts with the existing legal order, but also greatly affect the original business model, impacting on people's understanding of the works transaction, data feeding and relevant existing concepts. The traditional copyright system mandates copyright holders to be fairly compensated by others who use existing works, which is established on the basis of obtaining *ex-ante* authorization from copyright holders. Such model of *ex-ante* authorization is the fundamental for respecting the intellectual achievements of others and maintaining the operation of the market in the current knowledge-based economy. However, G-AI training requires not only massive number of contents, but also a flexible and effective transaction model, in order to achieve free flow of knowledge. Unfortunately, the traditional model of *ex-ante* authorization alone cannot meet the requirements.

---

[65] Thilo Hagendorf, 'The Ethics of AI Ethics: An Evaluation of Guidelines' (2020) 30 Minds & Machines

First, the model of *ex-ante* authorization can easily lead to redundancy and inefficiency in the data transaction process, because this model necessitates multiple and repeated commercial negotiations between copyright holders and G-AI developers but the outcome of negotiations is not as always favorable as it was expected to be. Second, the model of *ex-ante* authorization can result in excessive transaction costs, which are mainly from identifying the true copyright holders and further commercial negotiations. Third, the model of *ex-ante* authorization cannot achieve an efficient operation of a massive knowledge learning model, and is ineffective in terms of knowledge acquisition and data value mining, because the key to the functioning of the G-AI mechanism is the availability of a large amount of contents to support the upgrading and iteration of G-AI models. That is why it is concluded in this thesis that the traditional model of *ex-ante* authorization alone cannot meet the requirements of obtaining a large amount of contents and establishing a flexible and effective transaction model.

In addition, the acquisition, input and output of machine learning carry a high risk of copyright infringement. For instance, the acquisition of contents for G-AI training purposes require to collective massive data by technical measures on Internet, which is likely to be an infringement of the reproduction right[66] when no *ex-ante* authorization is obtained from copyright holders. Besides of possibility of copyright infringement, the training data include different forms, such as data of users, data of enterprise, public data, etc. and all can be an important source for G-AI training purposes, involving multiple data protection interests as well, such as personal data, property interests and public interests. Therefore, copyright law alone is also insufficient for providing a comprehensive protection and it requires the combination with other sectors of law such as data protection law and competition law.

Given that the major context of this thesis is copyright law, I hereby only propose resolution from the perspective of copyright law. The current failures of regulating the

[66] 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 [2001] OJ L 167, art 2

training behaviour urge us for exploring a workable resolution and finding a new settlement mechanism. In this thesis, it is suggested to accept the model of *ex-ante* authorization for lawful data transaction and adopt the regime of fair use where appropriate as a supplementary measure.

Access to lawful and high-quality data is an essential prerequisite for the development of G-AI, thus the traditional transaction model of *ex-ante* authorization is still considered as an important way for G-AI developers to obtain training data. The model of *ex-ante* authorization refers to that anyone seeking to utilize works that are subject to copyright protection must obtain authorization from respective copyright holders prior to the act of utilization. In certain circumstances, this model of *ex-ante* authorization still has the advantages of ensuring data quality, incentivizing the creative industry, avoiding the risk of copyright infringement, and contributing to a certain level of economic efficiency. Notably, the *ex-ante* authorization can be obtained by lawful purchasing. For example, G-AI developers can collaborate with Internet platforms and lawfully purchase their data to acquire massive copyrightable contents for G-AI training.[67] Also, the *ex-ante* authorization can also be obtained directly from users, particularly in case where G-AI developers operate Internet platforms themselves. They can leverage the copyrightable contents uploaded by users by providing corresponding service and explicitly outlining the risks and liabilities involved in terms of their use.

However, as explained before, the disadvantages of the model of *ex-ante* authorization render it insufficient as the only mechanism for G-AI developers to obtain training data. Therefore, a supplementary measure to address the limitations of the exclusive rights of copyright holders is to adopt the regime of fair use for the training behaviour. It is argued that the training behaviour should be exanimated through the lens of three-step test.

[67] In some industries such as online literature, audiobooks, digital music, etc., individual creators often hand over the exercise of copyright of their works to platforms, and G-AI developers can directly purchase data from the platforms to obtain a huge amount of copyright resources.

Theoretically, the famous three-step test, originating from the Berne Convention[68], sets limits to limitations on exclusive rights of copyright holders and orders that exceptions are permitted only if: (a) in certain special cases; (b) which do not result in a conflict with the normal exploitation of a work and (c) which do not unreasonably prejudice the legitimate interests of the author (or other right-holder).[69] The three-step test is also adopted by the EU copyright law[70] and China's copyright law[71], although the stipulated specific situations are somewhat different. On the one hand, the regime of fair use can serve as a legitimate shield for the lawful use of contents for G-AI training purposes under some circumstances, where it is either impractical or unfeasible to obtain *ex-ante* authorization from each copyright holder. On the other hand, it should be considered prudently to apply the regime of fair use to justify the training behaviour, because the of G-AI training technologies remains predominantly dominated by large companies, and the legitimacy of using copyrighted contents for G-AI training purposes without prior authorization is suspicious. For this reason, the application of the regime of fair use ought to be assessed through the three-step test, implying that the specific G-AI training behaviour should be examined rigorously whether it satisfies the three criteria in the formula of the three-step test. Only by adhering to this approach can the regime of fair use effectively function as supplementary measure for the shortcomings of the model of *ex-ante* authorization, without prejudicing to the original value.

---

[68] Berne Convention for the Protection of Literary and Artistic Works
Article 9(2) Possible exceptions: 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.

[69] Jonathan Griffiths, 'The 'Three-Step Test' in European Copyright Law - Problems and Solutions' (22 September 2009) Queen Mary School of Law Legal Studies Research Paper No. 31/2009, The Intellectual Property Quarterly, Forthcoming

[70] 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 [2001] OJ L 167
Article 5(5): The exceptions and limitations provided for in paragraphs 1, 2, 3 and 4 shall only be applied in certain special cases which do not conflict with a normal exploitation of the work or other subject-matter and do not unreasonably prejudice the legitimate interests of the rightsholder.

[71] Copyright Law of People's Republic of China (2020 Amendment)
Article 24: In the following cases, a work may be exploited without the permission from, and without payment of remuneration to, the copyright owner, provided that the name or designation of the author and the title of the work are mentioned and the normal use of the work, or unreasonably damage the lawful rights and interests of the copyright owner shall not be affected: ……

5.4.Possible Solution to the Second Dilemma: Fine-tuning Originality Criterion

As enshrined in statues and practical cases in most jurisdictions, the originality criterion[72] is widely accepted as a substantive element for a work to be eligible for copyright protection. Given that neither Berne Convention[73] nor WIPO Copyright Treaty does not clarify the meaning of originality at international level, there are differences in the specific understanding of originality among jurisdictions, mainly between the common law system and the civil law system.

In the common law system, the requirement of originality is relatively low, as long as the work is completed independently by the author with a minimum level of creativity, not being a plagiarism or reproduction of another's work, it is eligible for copyright protection. For instance, the US Supreme Court explained that the US Constitution mandates originality as a prerequisite for copyright protection and it requires independent creation plus a modicum of creativity.[74] In contrast, the requirement for originality in civil law system is another scenario. In the countries with civil law system, the equivalent term of copyright is "author's right" and it can be found that the copyright in civil law system is positioned as a personal right reflecting author's minds and thoughts. Hence, the requirement of originality is higher, mandating that the content can be protected only if it contains a certain level of creativity. For instance, the Germany Copyright Act stipulates that only the author's own intellectual creations constitute works within the meaning of this Act.[75] There are two concepts in Germany Copyright Act: (a) Photographic works[76] and (b) photographs and products manufactured in a similar manner to photographs[77], while the former one is recognized as works protected by copyright and the latter is not recognized as works but

[72] In this paper, in order to maintain contextual consistency, I use the "originality criterion" which is equivalent to "creativity criterion" and "intellectual creativity criterion" appearing in some literature in terms of terminology.
[73] Berne Convention for the Protection of Literary and Artistic Works
[74] Feist Publications, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340 (1991)
[75] Act on Copyright and Related Rights (Urheberrechtsgesetz – UrhG), Session 2, Division 2, Part 1
[76] Act on Copyright and Related Rights (Urheberrechtsgesetz – UrhG), Session 2, Division 2, Part 1
[77] Act on Copyright and Related Rights (Urheberrechtsgesetz – UrhG), Session 72, Division 2, Part 2

protected by related rights (neighbouring rights). Distinguishing photographic works and photographs reflects that the application of relatively higher originality criterion, because some photographs lack of intellectual creativity then cannot meet the requirements of originality, failing to be recognized as works in the context of Germany Copyright Act.

In this thesis, it is argued that the second dilemma of whether the G-AI content is copyrightable should be determined subject to the originality criterion First, the pursuance of a unified originality criterion is unpractical. This is evident from the fact that international treaties, such as the Berne Convention and WIPO Copyright Treaty, have not succeeded in defining a singular originality criterion. Similarly, at the level of regional copyright framework, the EU fails to achieve straight-forward harmonization but provides the expression of "an original work of authorship" by case law and directives in this regard.[78] Moreover, the understanding of originality of national copyright laws is even more divergent, which is represented by distinct approaches adopted by the civil law system and common law system.

Second, it is not pursuable to reject the copyrightability of G-AI content merely because the subject generating the content is not human being. There is another famous "Monkey Selfie Case", where a monkey took a picture on itself by using the camara placed by a British photographer David Slater.[79] The US court ruled that the monkey cannot be entitled to the copyright of the picture because animals are not eligible subject of US copyright law.[80] This "Monkey Selfie Case" is also cited by some scholars as a justification for rejecting the copyrightability of G-AI content. Indeed, at least until now, animals cannot be an eligible subject of copyright law in most jurisdictions. However, this does not *ipso facto* imply that content generated by G-AI, which is not an eligible subject of copyright law neither, is not copyrightable. Imagine if the photographer had not disclosed at the outset that the picture in

---

[78] See "3.3.1. Legislative Provisions and Cases in Practice: EU" of this paper

[79] Andres Guadamuz, 'Can the monkey selfie case teach us anything about copyright law?' (2018) WIPO Magazine <https://www.wipo.int/wipo_magazine/en/2018/01/article_0007.html> accessed 15 September 2024

[80] Naruto v. David John Slater & Blurb Inc., No. 16-15469 (9th Cir. 2018)

question was taken by a monkey and instead claimed that the picture was taken by himself, the copyrightability of the picture might be another result. In my opinion, the correct approach is to adhere to the originality criterion in determining whether the G-AI content is copyrightable. At the level of domestic law, different country or region ought to apply their domestic rules and precedents to analyse the originality of G-AI content in question. Assuming that the same content is created by a human being and can be recognised as original, then if the same content is generated by G-AI, its originality should also be recognised without regard to the identity of the subject from which the content is derived.

Whether you believe or not, human beings are distinct from other creatures and computers in a philosophical and legal sense. Human users are more or less involved in the process of generating G-AI content, even if it is just a click of the mouse. Thus, in order to determine whether the G-AI content is copyrightable, human beings' creative contribution must be taken into account. It is not only a concrete fulfillment of the adherence to the originality criterion but also a rightful respect for human intellectual activity. Specifically, considering the unique interaction between G-AI and human users during the process of generating content, it is suggested in this thesis to emphasize the human users' creative contribution by analyzing the proportion of human users' participation.

By analyzing the proportion of human users' participation, it is more accurate to assess the user's creative contribution to the G-AI content. On the one hand, the effects of participation can be reflected by concrete endeavours undertaken by users, such as prompts and guidance that are provided by users to guide the G-AI. The amount, complexity and combination of prompts and guidance can be favorable proof of users' active participation, proving the extent to which they contribute to the generation process. On the other hand, it is also noted that users may also make further deletion, addition and/or modification to the generated content. Those adjustments also represent a significant aspect of users' creative contribution to the G-AI content. Notably, it is imperative to articulate that the assessment of the proportion of human users' participation is not a binary determination of Yes or No. Rather, it is a comprehensive evaluation that determines the overall level of participation,

and then the creative contribution. In other words, this is not a zero-sum game where winner achieves victory but the counterparty loses all, but it acknowledges the existence of the hierarchy of users' creative contribution to the G-AI content.

5.5.Possible Solution to the Third Dilemma: User as Rightsholder

In terms of the attribution of rights over G-AI content, it is suggested in this thesis that the rights over G-AI content ought to be attributed to the user. First, as a direct consequence of the development of new technologies, the paradigm of human-AI interaction and co-creation enables us to achieve many innovations that were unimaginable with traditional methods, improving the efficiency of creation and enriching the forms of creation. This paradigm of human-AI interaction and co-creation is increasingly recognized by our industry and society but the key subject traditionally remains on the person who creates the works, and thus priority must be given to human participants when considering this regard.

Second, attributing rights to other subjects is not reasonable. The G-AI itself has not been recognized as an eligible subject of copyright or civil law in most jurisdictions. Without a legal personhood recognized by legislation, the G-AI itself is unable to bear rights and obligations as an independent entity. The problem of attributing rights to G-AI programmers is: Whether is fair to attribute right to G-AI programmers when the content is mainly generated by G-AI with users' creative contribution? The opinion of joint rightsholders is the difficulty in quantifying the specific contributions of different participants and thus attributing rights, resulting in further confusion in practical implementation and exacerbating legal uncertainty. Attributing right to the public domain is substantially equivalent to granting no rights, which means that anyone can use the G-AI content freely without restraint. The state of belonging to nobody is not what the G-AI industry has expected to see.

Last but not least, G-AI users being rightsholder is the most workable in practice. Again, I must cite the reasoning expressed by Czech court and Beijing Internet Courts as justification. Regardless of the opposing attitudes towards the copyrightability of G-AI content, both courts have spontaneously emphasized the significance of users' creative

contribution, which informs the potential possibility of attributions of rights.

Besides, another noteworthy case is that the US Copyright Office decided to grant partial copyright for the manga "Zarya of the Down" in February 2023 (Hereinafter, refer as to "Zarya Case")[81], whereas the creator Ms. Kashtanova used a G-AI, namely Midjourney, to generate the pictures in manga. The US Copyright Office breaks into three parts to analyze the copyrightability and attribution of rights of different elements of the manga: (a) Text; (2) Selection and arrangement of images and text; and (c) Individual images.[82] The Office agree that the text of mange is protected by copyright and also agrees the selection and arrangement of images and text are protectable as a compilation based on that Ms. Kashtanova "selected, refined, cropped, positioned, framed, and arranged" the images in the Work to create the story and it is the product of human authorship, containing sufficient creativity. But regarding the individual images, the Office thinks that, it was G-AI (Midjourney), rather than the human user (Ms. Kashtanova), that originated the "traditional elements of authorship" in the individual images and excluded the individual images generated by G-AI from copyright protection.[83] Further, the Office explained in the end of analysis that, "to the extent that Ms. Kashtanova made substantive edits to an intermediate image generated by Midjourney, those edits could provide human authorship and would not be excluded from the new registration certificate". [84] This interpretation also conveys another dimension of significance, suggesting that the image generated by G-AI could be copyrightable in the case of that the human user makes substantive edits, or in essence, creative contribution. Obviously, the Office's final decision attributed the copyright to the human users, serving as another strong proof to justify the practicability of attributing rights to G-AI users.

[81] United States Copyright Office, Re: Zarya of the Dawn (Registration#VAu001480196)
[82] United States Copyright Offices, Re: Zarya of the Dawn (Registration#VAu001480196), p 4, p 5
[83] United States Copyright Offices, Re: Zarya of the Dawn (Registration#VAu001480196), p 8, p 12
[84] United States Copyright Offices, Re: Zarya of the Dawn (Registration#VAu001480196), p 12

Conclusion

Law necessarily lags behind social norms and behaviour, that it can only be a force for reaction not progress.[85] Yet our thinking about the law should be forward-looking. The rapid development of G-AI brings new challenges to the existing copyright law, particularly three legal dilemmas regarding the G-AI content in the context of copyright law. Academic opinions and judicial practices have diverse arguments on those dilemmas but no universal accepted conclusion has merged. The first dilemma concerns the lawfulness of using data for G-AI purposes. Its controversy exists revealed between G-AI developers and other rightsholders. In other words, the applicability of the regime of fair use would be key factor to determine its lawfulness. It is suggested in this thesis that the model of "*ex-ante* authorization" for lawful data transaction should be adhered while the regime of fair use serves as a supplementary measure. The second dilemma concerns the copyrightability of G-AI content. There are opposing opinions on the issue of whether G-AI content meet the requirements stipulated in copyright law and can be recognized as protected works. This thesis argues that the determination of the copyrightability of G-AI content should be subject to the originality criterion and human users' creative contribution should be taken into account. The third dilemma concerns the attribution of rights over G-AI content. There are five potential candidates for being considered as rightsholder: G-AI itself, programmer, user, joint rightsholders and public domain. This thesis suggests that rights over G-AI content ought to be attributed to human user because of its reasonability and practicability.

In the end, I would like to state that the points proposed in this thesis are not the most satisfactory resolutions for addressing those dilemmas regarding G-AI. Hopefully, this thesis might provide insights to some extent.

---

[85] Richard L. Abel 'Law as Lag: Inertia as a Social Theory of Law' (1982) 80(4) Michigan Law Review

Bibliography

[1] A.M. Turing, 'Computing Machinery and Intelligence' (1950) Mind 49 433
[2] Abraham Drassinower, *What's Wrong with Copying?* (Published on 9 Apil 2015, Harvard University Press)
[3] Alina Valyaeva, 'AI Has Already Created As Many Images As Photographers Have Taken in 150 Years. Statistics for 2023' (*Everypixel Journal*, 15 August 2023) <https://journal.everypixel.com/ai-image-statistics>
[4] Amanda Levendowski, 'How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem' (24 July 2017) Washinton Law Review
[5] Andrea Bertolini, 'Artificial Intelligence and Civil Liability' (2020) <https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU(2020)621926_EN.pdf>
[6] Andres Guadamuz, 'Can the monkey selfie case teach us anything about copyrig ht law?' (2018) WIPO Magazine <https://www.wipo.int/wipo_magazine/en/2018/01/article_0007.html>
[7] Annemarie Bridy, 'The Evolution of Authorship: Work Made by Code' (8 September 2016) 39 Columbia Journal of Law & the Arts
[8] Anthony K. Brandt, 'Beethoven's Ninth and AI's Tenth: A comparison of human and computational creativity Journal of Creativity' (December 2023) 33(3) Journal of Creativity
[9] Atilla Kasap, 'Copyright and Creative Artificial Intelligence (AI) Systems: A Twenty-First Century Approach to Authorship of AI-Generated Works in the United States' (5 June 2019) 19(4) Wake Forest Intellectual Property Law Journal
[10] Brian Golger, 'Copyright in the Artificial Intelligent Author: A Constitutional Approach Using Philip Bobbitt's Modalities of Interpretation' (2020) 22 Journal of Constitutional Law, 2020(22)
[11] Carolyn Giardina, 'How 'Furious 7' Brought the Late Paul Walker Back to Life' (*The Hollywood Reporter*, 11 December 2015) <https://www.hollywoodreporter.com/movies/movie-news/how-furious-7-brought-late-845763/>
[12] Celine Melanie A. Dee, 'Examining Copyright Protection of AI-Generated Art, Delphi - Interdisciplinary Review of Emerging Technologies' (2018) 1(1) <https://delphi.lexxion.eu/article/delphi/2018/1/11/display/html#63>
[13] Deng Wen, 'On the Copyrightability of Content Generated by AI with ChatGPT as a Representative'(2023) 9 Political Science and Law
[14] Edward Lee, 'Technological Fair Use' (3 September 2010) 83 Southern California Law Review
[15] Harsh Bhattad, Geeta Atkar, 'Review on Different Types of Chatbots' (2021) International Research Journal of Modernization in Engineering 1347
[16] James Moor, 'The Status and Future of the Turing Test 2001' (2001) Minds and Machines 11(1) 77
[17] Jenny Quang, 'Does Training AI Violate Copyright Law?' (2021) 36 Berkeley Technology Law Journal

[18] John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon, 'A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence' (31 August 1955) <http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf>

[19] Jonathan Griffiths, 'The 'Three-Step Test' in European Copyright Law - Problems and Solutions' (22 September 2009) Queen Mary School of Law Legal Studies Research Paper No. 31/2009, The Intellectual Property Quarterly, Forthcoming

[20] Koushai Kumar, Gour Sundar Mitra Thakur, 'Advanced Applications of Neural Networks and Artificial Intelligence: A Review' (8 June 2012) 6 Information Technology and Computer Science

[21] Liu Youshan, Wei Yuanshan, 'Copyright Infringement Problem of Machine Learning and Its Solution' (2019) 2(2) Journal of East China University of Political Science and Law

[22] Mark A. Lemley, Bryan Casey, 'Fair Learning' (2021) 99(4) Texas Law Review

[23] Matthew Sag, 'The New Legal Landscape for Text Mining and Machine Learning' (27 February 2020) 66 Journal of the Copyright Society of the USA

[24] Md. Al-Amin, etc., 'History of generative Artificial Intelligence (AI) chatbots: past, present, and future development' (2024) <https://arxiv.org/abs/2402.05122>

[25] Melissa Malec, 'Generative AI Statistics: The 2024 Landscape – Emerging Tren ds and Developer Insights' (HatchWorksAI, 19 January 2024) <https://hatchworks.com/blog/software-development/generative-ai-statistics/>

[26] Microsoft Asia News Center, 'Microsoft's Xiaoice, China's newest fashion designe r, unveils her first collection for 2019' (12 November 2018) <https://news.microsoft.com/apac/2018/11/12/microsofts-xiaoice-chinas-newest-fashion-designer-unveils-her-first-collection-for-2019/>

[27] Nina I. Brown, 'Artificial Authors: A Case For Copyright In Computer-Generated Works' (2019) 20 (1) Science and Technology Law Review

[28] People's Daily Online, 'First AI-authored collection of poems published in China' (31 May 2017) <http://en.people.cn/n3/2017/0531/c90000-9222463.html>

[29] Ralph D. Clifford, 'Intellectual Property in the Era of the Creative Computer Program: Will the True Creator Please Stand Up?' (1997) 71 Tulane Law Review

[30] Richard L. Abel 'Law as Lag: Inertia as a Social Theory of Law' (1982) 80(4) Michigan Law Review

[31] Samantha Fink Hedrick, 'I "Think," Therefore I Create: Claiming Copyright in the Outputs of Algorithms' (2019) 8(2) NYU Journal of Intellectual Property & Entertainment Law

[32] Samir Chopra, Laurence F. White, *A Legal Theory for Autonomous Artificial Agents* (2011) University of Michigan Press

[33] Seagull Song, 'China's First Case on Copyrightability of AI-Generated Picture' (7 December 2023) <https://www.kwm.com/cn/en/insights/latest-thinking/china-s-first-case-on-copyrightability-of-ai-generated-picture.html>

[34] Selmer Bringsjord, David Ferrucci, 'Artificial Intelligence and Literary Creativity Inside the Mind of Brutus, A Storytelling Machine' (1st ed. 2000) Psychology Press

[35] Steven Borowiec, 'AlphaGo seals 4-1 victory over Go grandmaster Lee Sedol' (*T*

he Guardian, 15 March 2016) <https://www.theguardian.com/technology/2016/mar/15/googles-alphago-seals-4-1-victory-over-grandmaster-lee-sedol>

[36] Thilo Hagendorf, 'The Ethics of AI Ethics: An Evaluation of Guidelines' (2020) 30 Minds & Machines

[37] Timothy L. Butler, 'Can a Computer be an Author - Copyright Aspects of Artificial Intelligence' (1982) 4 UC Law SF Communication and Entertainment Journal

[38] Tomáš Ščerba, Jaroslav Fořt, 'The first Czech case on generative AI' (4 April 2024) <https://www.technologyslegaledge.com/2024/04/the-first-czech-case-on-generative-ai/>

[39] Wang Qian, 'Argument on the Characterisation of AI-generated Content in Copyright Law' (2017) 5 Journal of Northwest University of Political Science and Law

[40] WIPO, Frequently Asked Questions: AI and IP Policy - What is artificial intelligence? <https://www.wipo.int/about-ip/en/artificial_intelligence/faq.html>

[41] Zack Naqvi, 'Artificial Intelligence, Copyright, and Copyright Infringement' (2020) 24(1) Marquette Intellectual Property Law Review

- **Laws & Cases**

[42] Berne Convention for the Protection of Literary and Artistic Works

[43] 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 (2019) OJ L130/92

[44] Directive 2006/116/EC of the European Parliament and of the Council of 12 December 2006 on the term of protection of copyright and certain related rights (2006) OJ L 372/12

[45] 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 [2001] OJ L 167

[46] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (2024) OJ L 2024/1689

[47] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases (1996) OJ L 77/20

[48] Draft Report with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL))

[49] Case C-310/17 *Levola Hengelo BV v Smilde Foods BV* [2018] CJ

[50] Case C-5/08 *Infopaq International A/S v Danske Dagblades Forening* [2009] CJ

[51] Copyright Law of People's Republic of China (2020 Amendment)

[52] Regulations for the Implementation of the Copyright Law of the People's Republic of China (2013 Amendment)

[53] Interim Measures for the Management of Generative Artificial Intelligence Services

[54] Cybersecurity Technology – Security Specification for Generative Artificial Intelligence Pre-training and Fine-tuning Data

[55] Regulation on Promoting Artificial Intelligence Industry of Shenzhen Special Economic

Zone
[56] Regulation on Promoting the Development of Artificial Intelligence Industry of Shanghai
[57] (2018) Jing 0491 Min Chu No.239
[58] (2019) Jing 73 Min Zhong No. 2030
[59] (2023) Jing 0491 Min Chu No.11279
[60] Feist Publications, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340 (1991)
[61] Naruto v. David John Slater & Blurb Inc., No. 16-15469 (9th Cir. 2018)
[62] United States Copyright Office, Re: Zarya of the Dawn (Registration#VAu001480196)
[63] Act on Copyright and Related Rights (Urheberrechtsgesetz – UrhG)

Acknowledgement

It seems to be just yesterday when I first walked on the street of Lisbon, paving above those *calçadas*, the sounds and smell are familiar but so closing. I still remember clearly the excitement and uneasiness I brought to NOVA in the autumn of 2023. As a master student, whose environment becomes less familiar, whose future is enveloped with more unknowing but hope, whose life just starts in another period, I encountered with warm friends, good colleagues and professional teachers, in the NOVA School of Law, in the NOVA University of Lisbon, in Lisbon and in Portugal. Those scenarios that happened in campus, in parks and in every corner, indeed impressed me significantly and will continue to be inseparable memory during the rest time.

I would like to express my great gratitude to those people who ever provide unconditional support when I am down. It is your valuable encouragement that helps me in overcoming tough obstacles and reaching my dreaming island, with hands full of happiness. I would also like to say thank you to my supervisor and teacher Professor Giulia Priora for her professional lectures and responsible guidance in completing this thesis.

In the end, I put a famous Portuguese saying "*Aqui, onde a terra se acaba e o mar começa*" as a marker to signify that: Indeed, my student period ends but the brilliant life continues to shine as it has always been!