

Professor: Who holds the copyright for AI-assisted and AI-generated contents?

Katerina Zdravkova ¹

¹Ss. Cyril and Methodius University in Skopje, Faculty of Computer Science and Engineering, katerina.zdravkova@finki.ukim.mk

Abstract

Generative artificial intelligence (AI) has had a significant impact on intellectual property protection, mainly because model training is usually based on copyrighted content. This problem will deepen because, on the one hand, corporations with large training resources begin to sell content to AI companies, but also because authors whose copyrighted works are being used without permission react against this illegal theft. Parallel to this dilemma, a bigger one opens up: who is the copyright holder when AI is an assistant in the creation of a work, or when AI has competently created a new work? The discussion of these issues is illustrated with the help of nine case studies, which are elaborated in more detail. Based on this, it becomes evident that works created with the support of AI can be considered the copyrighted works of the person whose creativity is aided by AI. In contrast, fully autonomous works, created without any AI assistance, cannot be granted intellectual property protection.

Keywords: *Authorship, copyright, academic dishonesty, scientific misconduct, lawsuits, infringement, artificial intelligence.*

Resumo

A inteligência artificial generativa (IA) teve um impacto significativo na proteção da propriedade intelectual, principalmente porque o treino dos modelos utiliza geralmente conteúdo protegido por direitos de autor. Esse problema agravar-se-á porque, por um lado, as empresas com grandes recursos para o treino destes modelos começam a vender conteúdo para empresas de IA; por outro lado, os autores, cujas obras protegidas por direitos de autor são utilizadas sem permissão, reagem contra esse roubo ilegal. Paralelamente a este dilema, surge um outro ainda maior: quem é o detentor dos direitos de autor quando a IA é um assistente na criação de uma obra ou quando a IA criou competentemente uma nova obra?

A discussão destas questões é ilustrada recorrendo a nove estudos de caso, elaborados com mais detalhes. Com base nisto, torna-se evidente que as obras criadas com o apoio da IA podem ser consideradas obras protegidas por direitos de autor da pessoa cuja criatividade é auxiliada pela IA. Em contrapartida, obras totalmente autónomas, criadas sem qualquer recurso da IA, não podem receber proteção de propriedade intelectual.

Palavras-chave: *Autoria, direitos de autor, desonestidade académica, má conduta científica, processos judiciais, violação, inteligência artificial.*

1. Introduction

One of the key reasons for human progress is the inherent desire of individuals to become better, faster and more efficient – in essence, superior to others. Many tirelessly strive to achieve these goals, investing enormous efforts in their aspirations. Unfortunately, some people take shortcuts and engage in dishonest practices. Cheating takes many forms: athletes use doping (Nicholls et al., 2020), companies inflate revenues or conceal debts (Schilit & Perler, 2010), counterfeiters forge works of art (Spink, Moyer, Park, & Heinonen, 2013), traders engage in insider trading (Woody, 2020), and politicians manipulate votes (Hernández-Huerta, 2020). Similarly, some students resort to academic dishonesty, such as plagiarism, cheating on exams, or falsifying assignments (Zdravkova, 2023), while researchers may engage in misconduct, including data fabrication, falsification, or unethical authorship practices (Gopalakrishna et al., 2022). Because of this dishonesty, it becomes unclear as to who truly won the race, what the company is actually worth, whether the artwork in the gallery is authentic, and what the true value of the stock is. Counterfeiting, piracy, academic dishonesty and scientific misconduct predominantly affect intellectual property, particularly the copyright of the work.

When an author or a group of authors create an original literary or artistic work, according to the Berne Convention, which was adopted in 1886, the authorship is automatically granted, without the need for registration or other formalities (WIPO, 1979). However, the authorship does not mean that the authors hold the copyright (Jaszi, 2017). Copyright gives legal rights regarding the control of use and distribution of the work. For literary works (poems, novels, research papers, computer codes etc.), copyright is usually transferred by the author to a publisher or a software company who then has the legal rights to control how the work is distributed, reproduced and used (Varian, 2005). In most cases, unless explicitly transferred, the authors of artistic works (paintings, sculptures, engravings, photographs, etc.) retain copyright over their artwork even after selling it to a gallery or an individual buyer (Pila, 2010). In music and film, original creators like songwriters and directors may retain some rights, but they often transfer significant rights to record labels or production companies (Hesmondhalgh, 2018).

The copyright is not eternal. According to the World Intellectual Property Organization (WIPO), the United Nations agency responsible for protecting and promoting the rights of innovators and creators worldwide, it lasts for at least 50 years after the creator's death (WIPO, 1996). In the U.S., the protection lasts 70 years after the creator's death, and for some works, such as the anonymous ones, even '120 years from its

creation' (U.S. Copyright Office, 1976, 2024). After this period, the works are no longer protected by intellectual property laws and become part of the so-called public domain, meaning that everyone can own them (Ochoa, 2002). For the works that belong to the public domain, no one can obtain a copyright or use them in other works that are protected by intellectual property rights (Ochoa, 2002). These works become a public good and belong to all of humanity.

In the past, there have been many cases of recorded copyright infringements that were afterwards revoked, affecting the reputation and standing of those who claimed that these were their original works (Temple, 2003). In addition to this, copyright infringement may have severe legal consequences. For example, the Department of Justice foresees a financial penalty of a quarter of a million U.S. dollars and a maximum sentence of 10 years in prison if the defendant has already been convicted of the same crime (U.S. Department of Justice, 2022).

With the rise of digital technologies, the possibility of copyright infringement has significantly increased (Lunney Jr, 2001). This was primarily made possible by the public display of vast amounts of copyrighted works on the web that are freely shared and pirated in abundance. AI-based technologies have significantly transformed the way literary and artistic works are created, directly impacting areas protected by copyright (Yusuf, Pervin, & Román-González, 2024). These tools enable individuals to create their own literary and artistic contents of satisfactory, or sometimes even high quality (Bagenal, 2024).

The capabilities of AI-generated systems have reached such a high level that it is almost impossible to prove that a literary or artistic work was not created by humans (Frank et al., 2024). Some of these works seem completely original. Despite recent advances, generative AI cannot be recognized as a copyright owner because both U.S. and EU legislation explicitly reserve copyright protection for human creators: U.S. law requires a 'human being' (Library of Congress, 2023), while EU law refers to a 'natural person' (EU, 2019).

Whether a human-AI-assisted work is eligible for copyright protection depends on several factors, including the way AI was used, whether the human author disclosed AI assistance, and whether the training data used comply with copyright laws. These factors are discussed in detail in the extended version of the paper.

2. Copyright and the nine case studies

The following section introduces the nine case studies that were selected to cover a broad spectrum of applications where AI plays a role, ranging from academic assignments to creative fields like art, music, and literature.

Each case study explores a distinct aspect of AI's involvement: for instance, AI-assisted student assignments highlight the educational dimension, while AI-generated art and music focus on the creative process. The inclusion of both poetry and novels acknowledges the different forms of literary expression that AI can impact, with poetry often involving more condensed, emotional content, and novels typically requiring more complex narrative structure. This diversity in case studies ensures that a wide range of AI's capabilities and challenges are fully covered.

2.1. Copyright of student assignments

For most educational institutions, the student holds copyright of their work, automatically, without any need to register the work to obtain a copyright (Crews, 2020). Certainly, in order for students to obtain this copyright, their assignments, whether they are text, image, video or code, should be an original contribution. Unfortunately, students often plagiarize, i.e., steal other people's authorship and try to obtain copyright for the stolen content.

According to Laura Young (2024), academic programming plagiarism is on the rise. The impression of the professors at my faculty is that after the appearance of the LLMs there has been an increase in copyright infringement.

Based on my long-time experience as a computer ethics teacher, student cheating has significantly changed after December 2022 (Zdravkova, 2023). Before the appearance of ChatGPT, cheaters either literally copied already published texts or engaged in cross-lingual plagiarism. Regarding verbatim copies in student assignments, which refer to the direct reproduction of text or content without alteration, there are two possible scenarios related to copyright:

- If the stolen content was protected by copyright, then the act of verbatim copying is a theft and is subject to sanctions.
- If the stolen content was copyleft, meaning it is licensed in a way that allows others to freely use, modify, and distribute it under the same conditions (Free Software Foundation, 2022), then their derivative work also becomes copyleft, and therefore they lose their copyright.

Cross-lingual plagiarism, which means content automatically translated with machine translation services (MTS) into another language (usually from English), is also connected with two scenarios:

- Although translation is a derivative work, the copyright still belongs to the original author.
- Translating a work using MTS does not transfer the copyright to that service.

As of December 2022, students began to make heavy use of ChatGPT and other large language models (LLMs). Discovering text generated by LLMs is a very complex challenge. Plagiarism checkers, such as: iThenticate (www.ithenticate.com/) and GPTZero (gptzero.me/) claim to be able to verify whether a document was written by artificial intelligence. However, to avoid plagiarism, students resort to using paraphrasing tools such as QuillBot (quillbot.com/), Wordtune (www.wordtune.com/), Paraphrase Tool (paraphrasingtool.ai/), Rephrase (www.rephrase.info/), and Grammarly (www.grammarly.com), both premium and free. Although such plagiarism cannot be formally proven, as an experienced teacher with extensive experience, which includes teaching the computer ethics course since 2003, I have developed a keen ability to recognize when the writing style deviates from what is typical for computer science students (Zdravkova & Ilijoski, 2025). I usually do not announce the grade to suspected students and invite them to discuss their assignment. If they answer my questions competently, which happens extremely rarely, then they receive a grade based on the quality of the assignment. It is interesting that most of the suspected students do not try to dispute the suspicion at all.

Students are typically granted authorship for their assignments by default, regardless of whether generative AI was used in the creation process. If the AI helped create an assignment in accordance with the university or course rules, then they deserved that authorship. But, if AI wrote the whole assignment based on their prompts, then they will again get the authorship because AI cannot become a copyright holder. In the second case, that authorship is a crime.

Namely, AI-generated media are trained with various data, including copyrighted ones, as evidenced by the New York Times journalist's lawsuit against OpenAI, which OpenAI refutes (OpenAI, 2024). This is probably the most well-known of the large number of lawsuits that appear ever more often and refer to the greatest generative-AI software companies in the world (Panettieri, 2024).

The symbiosis of academic dishonesty and copyright infringement by generative AI opens up space for mass theft of intellectual property. If this issue is not addressed, it may lead to significant and far-reaching consequences.

2.2. AI-assisted and AI-generated art

Vincent van Gogh (1853 - 1890) is my favourite painter. He died more than 100 years ago, so all his paintings are public domain. This means that everyone can replicate his paintings and even sell them legally. Nicolae, the owner of Niko Prints (<https://nickoprints.com/>) creates digital art and sells it. He digitally combined two famous van Gogh's paintings: *The starry night* and *Wheat field with cypresses*. The name of the digital picture is expectedly, *Wheat Field with Cypresses under a Starry Night*.



Figure 1. Two van Gogh's paintings and their digital combination

Nicolae obviously knows that he cannot protect his works based on works from the public domain by copyright. What is even more important is that Nicolae does not appropriate van Gogh's works and gives credit to the original works. Therefore, his commercial activity is not a copyright violation. His picture is copyleft just like the original van Gogh's works, and thanks to that I was able to copy them into this article. However, in contrast to his approach, many AI companies involved in generating images and videos do not demonstrate the same level of consideration. A group of young artists from several countries launched the so-called class-action lawsuit against Stability AI, DeviantArt, Midjourney, and Runway AI for using their works without consent, credit, or compensation (Justia, 2023). Further information about the movement and the affected artists can be found on the website (<https://imagegeneratorlitigation.com/>).

Generative AI companies that do not use public domain works or have purchased copyright for their training data are seriously infringing copyright (Chu, Song, & Yang, 2024). Moreover, anyone who chooses to use their software to create their own work for commercial purposes will face copyright infringement sanctions if the AI-generated

works are found to be similar to some original copyrighted art. This will open a complex and contentious issue, leading to an influx of lawsuits regarding authorship theft.

2.3. AI-generated music

In July 2024, primary school participants at the European Junior Olympiad in Informatics (<https://ejoi2024.gov.md/>), held in Chisinau, Moldova generated their own anthem (EJOI MK, 2024). Bisera, Vladimir, Marko and Nikola defined the theme of the anthem, and music creator Suno AI (<https://suno.com/>) created and sang it. When I inquired from an experienced colleague about who holds the copyright of the anthem, the response was immediate: 'Children.' This position is entirely accurate, as children, being human beings, fulfil the main prerequisite for obtaining a copyright. Nevertheless, a lawsuit was filed against Suno alleging widespread infringement of copyrighted sound recordings because Suno was trained on copyrighted music (Resnikoff, 2024).

Training the system using copyrighted works is a flagrant example of intellectual property theft. The children's anthem, which by the way, won the desired awards, is their authored work, especially the text that they more or less wrote themselves. They are the copyright holders of that anthem. It is interesting that the melody does not resemble any existing song, so the anthem is not only pleasant to listen to, but also original. However, AI-based systems that compose and create text based on well-known songs violate the copyright of the creators of the works that are part of the system's training base. If musicians use them to enrich their repertoire, they risk being sued whenever another musician recognizes their song. The music industry is powerful and has well-designed methods to protect the rights of creators and performers. Time will tell if composing systems like Suno will survive or if the music machinery will destroy them.

2.4. Deepfakes

Deepfakes are synthetic media created by AI capable of editing or generating images, audio and video. There are many interesting examples of deepfakes, such as, for example, the video of Tom Cruise in the role of Iron Man. If it is not known that Iron Man is portrayed by Robert John Downey Jr, the viewer may not recognize the photo manipulation, as the video appears quite convincing. Deepfakes can be very harmful because they: create confusion and scepticism; initiate the spread of misinformation, pose a threat to privacy and security; convincingly impersonate anyone; orchestrate phishing scams; and identity theft operations with alarming precision (Stanford University IT, 2024).

The connection between deepfakes and copyright lies in the implications they have for intellectual property rights. US Digital Millennium Copyright Act (DMCA) states that: "Deep fakes that take advantage of copyrighted materials could face civil actions for copyright infringement." (U.S. Copyright Office, 2020). According to the recent Artificial Intelligence Act of European Parliament and Council (2024) "artificial or manipulated images, audio or video content ('deepfakes') need to be clearly labelled as such". The act also adds: "These include drawing up technical documentation, complying with EU copyright law and disseminating detailed summaries about the content used for training".

Under these laws, deepfakes are legitimate media as long as they are clearly labelled as such, and do not violate existing copyright laws. However, it is clear that vocal or audio imitations cannot be made without importing the original works. If they are copyrighted, and they almost always are, then the stakes would be high, especially in the US, where copyright infringement is protected by a federal law and the statutory damages can reach \$150,000 per infringed work.

2.5. AI-generated welcome speech

I recently wrote a welcome speech in a language that I spoke bilingually during my youth, but had not used in years. I prepared the welcome speech in my native language and asked the retired GPT-4 to translate it and to improve it. ChatGPT not only corrected it perfectly, but also embellished the sentences.

Although it was no longer my original text, I was pleased with the final result. The backbone of the text was still mine, but the writing style and part of the content was AI-generated. To better illustrate what I am discussing, here is a hypothetical example. First, I wrote a description of the speech and gave a couple of explanations of exactly what I wanted. Before ChatGPT offered me its response, I prepared my own welcome speech and asked ChatGPT to improve it. ChatGPT offered me two responses in both cases, due to space constraints, only the shorter one is shown. Here is how mutual communication took place:

My prompt to ChatGPT:

“Please write me a short welcome speech on the occasion of the 25th edition of the conference organized by my faculty. The conference began at the dawn of the 21st century as a team building event of the Institute of computing and survived the COVID-19. Thank all the participants, representatives of the business community, renowned lecturers, students, members of the program and organizational committee, and at the end declare the conference open.”

The AI-generated response based on my prompts:

“Ladies and gentlemen, esteemed guests, and participants, welcome to the 25th edition of our conference, a remarkable milestone for our faculty and the Institute of Computing. What began as a team-building event at the dawn of the 21st century has evolved into a significant platform for collaboration, innovation, and learning, even persevering through the challenges posed by COVID-19. I want to extend my heartfelt gratitude to all of you: our dedicated participants, representatives from the business community, our esteemed lecturers, enthusiastic students, and the hardworking members of our program and organizational committee. Your commitment and passion are what make this event a success. With that, I am delighted to declare the 25th conference officially open. Let’s make this a memorable and inspiring event together! Thank you.”

This is the draft of my welcome speech that I wrote before asking ChatGPT for help:

“Dear colleagues, dear guests, it is a great honour and pleasure for me to open the 25th anniversary edition of the conference, which began at the dawn of the 21st century as a team building meeting of the Institute of Computing. The conference went through various crises, including COVID-19, and over time turned into a renowned scientific gathering of researchers from all over the world. I am happy that this year the conference was graced by excellent invited lecturers,

a large number of experienced researchers and young students who are taking their first research steps in science. It is a special pleasure that the business community honoured us with their presence. I thank them all from the bottom of my heart. I would like to thank the members of the program committee who wrote valuable reviews, and the members of the organizing committee who were diligent helpers and worked hard. Finally, I sincerely thank all participants who sent valuable works. I declare the conference open and wish you to enjoy all the lectures and exhibitions in the upcoming 3 days and to see you again next year.”

Based on my writing, ChatGPT suggested the final welcome speech:

“Ladies and gentlemen, esteemed guests, and dear participants, it is both an honour and a pleasure to welcome you to the 25th edition of our faculty’s conference. As we gather here today, we celebrate not only a significant milestone but also the resilience and commitment that have brought us to this moment. This conference began at the dawn of the 21st century as a team-building event for the Institute of Computing, and over the years, it has evolved into a vibrant platform for collaboration, innovation, and knowledge sharing. The journey has not been without challenges, particularly as we navigated the unprecedented times brought on by the COVID-19 pandemic. Yet, here we are, stronger than ever, ready to engage, learn, and connect. I would like to extend my heartfelt gratitude to all the participants, representatives of the business community, our distinguished lecturers, dedicated students, and the tireless members of the program and organizational committee. Your contributions and support are invaluable, and it is your passion that fuels the spirit of this conference. As we embark on this exciting journey over the coming days, let us embrace the opportunities for learning and collaboration that lie ahead. Together, we can continue to push the boundaries of knowledge and innovation in our field.”

I find both AI-generated speeches extremely valuable, the first based on my detailed explanation, and the second based on my speech proposal. Both were derived because of my author’s works, so whichever answer I choose, my author’s share lies in the basis of the response. If I decide to use one of them, I must determine whether I have the right to hold the copyright. I cannot definitively answer that. I was happy to say the welcome speech that was the inspiration for this example, especially because ChatGPT used almost all my phrases, so the contribution of AI was primarily translation and polishing of the language, which any professional proof-reader would do without taking over my authorship. Moreover, it was not recorded, and therefore, no copyright is claimed. However, if that option is not disabled (OpenAI, 2024), ChatGPT will insert my query and my welcome speech into its training database, so that they will become the basis of some other similar responses that ChatGPT will generate in the future.

2.6. AI-generated research papers

In the first case study presented in this paper, it was mentioned that students are willing to use generative AI as an assistant or ghost writer for assignment preparation. A brief analysis of the title structure reveals a noticeable trend. Comparing titles from papers published ten years ago, or earlier, with those of more recent publications, it becomes apparent that the use of colons in titles has significantly increased. Specifically, when requesting title suggestions from ChatGPT, it is highly probable that at least five out of ten proposed titles will feature two parts separated by a colon.

Another noteworthy observation can be made. By limiting a Google Scholar search to the past two years, it becomes evident that certain terms, such as *Unveiling' Revolutionizing*, and *Bridging*, as well as the phrase *From Barriers to Bridges* have been used frequently from 2023 to the present, as well as between 2010 and 2022. These are not merely buzzwords, but terms favoured by ChatGPT. This can be easily verified through experimentation.

Recently, it has been observed that some research papers exhibit a writing style that significantly differs from previous works by the same authors. It is likely that these papers contain text generated by a LLM, rather than being solely authored by the individuals listed as the creators. Two studies explicitly confirm this impression (Gray, 2024; Liang et al., 2024). They were using approaches similar to those I presented about the welcome speech and the titles. In the first study, by examining LLMs with prompts containing selected keywords or titles, the proportion of generative AI usage was found to be less than 1% (Gray, 2024). In the second, by searching papers with a set of 12 adjectives, adverbs and controls, the Stanford team estimates that generative AI is used in 6.3 to 17.5 percent of papers, depending on the field of research (Liang et al., 2024). While the frequencies of most words moderately increased, the frequency of adjectives commendable, meticulous and intricate, and particularly the frequency of the adverb meticulously increased more than 100%.

If the apparent inclusion of generative AI is not just a harmless aesthetic intervention, and by all accounts it is not, then we are faced with a worrying misconduct in science. With the obvious presence of falsification and fabrications in these papers resulting from algorithmic bias (Ferrara, 2023), lack of quality control, and generative AI misinformation (Wach et al., 2023), another question arises: Do these researchers have the right to hold copyright? Definitely not, primarily because, as was already illustrated in the introduction of this paper, LLMs violate copyright, so everything they generate is plagiarism. Further, what is more than clear is that stolen intellectual property cannot be protected as one's own (Chesterman, 2025). In the paper, Chesterman addresses the theft of copyright by generative AI during the training phase, as well as the ownership, and consequently the copyright, of the results generated by the AI.

2.7. AI-generated paraphrasing

As mentioned in the case study on student assignments, paraphrasing is a technique used to mask plagiarism. On the other hand, it is directly related to copyright infringement. I will prove this claim by using the applications ChatGPT and Quillbot over the same paragraph twice sequentially. The original paragraph is an explanation of the objectives of the COST action LITHME (<https://lithme.eu/>). It consists of these five sentences:

“How will pervasive human-machine language interfaces affect language in areas such as international law, translation, and other forms of language work? What will this mean for how people identify with specific languages? Could increasing reliance on real-time language technologies actually change the structure of language? Longer term, could developments in brain-machine interfaces serve to complement or even supersede language altogether? Linguistics will be far stronger for robust technological foresight, while developers will bene-

fit from better understanding potential linguistic and societal consequences of their creations.”.

First, let us conduct a thorough examination of the paragraph for potential plagiarism. When the first sentence is searched using quotation marks, the result directs us to LITHME. A search without quotation marks leads to the COST page of LITHME (<https://www.cost.eu/actions/CA19102/>). In both cases, it converged to LITHME. Googling all sentences from the source paragraph behaved similarly and revealed no plagiarism. This is great because it allows the comparison of an original text that LITHME has two copyrights, for the COST action and the action website. Two experiments were done step-by-step. First, the source paragraph was paraphrased twice with ChatGPT, and then with Quilbot. These are the results of the paraphrasing:

ChatGPT paraphrase 1:

“In what ways will widespread use of human-machine language interfaces impact language related fields like international law, translation, and various other linguistic tasks? What implications might this have for people’s connection to particular languages? Could a growing dependence on real-time language technologies lead to changes in the structure of language itself? In the long run, could advancements in brain-machine interfaces either enhance or replace language entirely? A solid grasp of technological trends will strengthen linguistics, while developers will gain from understanding the possible linguistic and societal impacts of their innovations.”

ChatGPT paraphrase 2:

“What impact will widespread human-machine language interfaces have on language in fields like international law, translation, and other linguistic work? What effects might this have on people’s relationships with specific languages? Could an increased reliance on real-time language technologies alter the very structure of language? Looking ahead, might innovations in brain-machine interfaces either improve or even replace language as we know it? A clear understanding of technological developments will enhance the field of linguistics, while creators will benefit from recognizing the potential linguistic and societal effects of their work.”

I undertook my research using all the sentences, but the illustration of the plagiarism check will again be done with the first sentence only. All the words in the sentence, “In what ways will widespread use of human-machine language interfaces impact language-related fields like international law, translation, and various other linguistic tasks?” exist in the blogpost (Schoening, 2023), which is published on the Phrase localization platform (<https://phrase.com/>), which is obviously copyrighted. It can be considered a potluck plagiarism (Myers, 2018) with one major difference: the sentence was a compilation of several sentences from the same and not from different articles. It is possible that the article was used to train the paraphrasing tool, but the plagiarism checker did not detect this.

The first sentence from the second paraphrase: “What impact will widespread human-machine language interfaces have on language in fields like international law, translation, and other linguistic work?” resembles the paper by Vieira, O’Hagan, and O’Sullivan (2021), which is copyrighted by Taylor & Francis. This discloses the copyright controversy of this publishing house (Informa PLC, 2024). The exact focus of the matter

is unclear. Taylor & Francis are owned by Informa PLC. In May 2024, Informa entered into a data access agreement with Microsoft with a term from 2024 to 2027, whereby Microsoft can use their work to train its AI models. The partnership was entered into without the consent of the authors who have already published their research in Taylor & Francis and paid a fee in order for their works to receive open access status. As a result, the publisher faced criticism from authors who had waived their copyright by paying a fee, only to have it sold without consultation or compensation. This situation is, understandably, debatable.

The same paraphrasing experiment was replicated using the professional paraphrasing tool Quillbot. The next two paragraphs present the results.

Quillbot paraphrase 1:

“What effects would ubiquitous human-machine language interfaces have on language use in fields like translation, international law, and other language related tasks? What impact will this have on how individuals relate to particular languages? Could language structure alter as a result of a greater reliance on real-time language technologies? In the long run, may advancements in brain-machine interfaces help to supplement or even replace language? While developers will gain from a better awareness of the possible linguistic and societal repercussions of their creations, linguistics will be significantly stronger for robust technological foresight.”

Quillbot paraphrase 2:

“How might the widespread use of human-machine language interfaces affect language use in domains such as international law, translation, and other language-related tasks? What effect will this have on people’s perceptions of specific languages? Could an increased dependence on real-time language technologies lead to changes in language structure? Could developments in brain-machine interfaces eventually serve to augment or even replace language? Linguistics will be much stronger for solid technical foresight, while developers will benefit from a greater understanding of the potential linguistic and societal effects of their creations.”

I searched each sentence extensively from both paraphrases as a whole and as a combination of keywords. I have not found a single reference in which more than four to five words match. After that, I searched the web looking for any reactions against Quillbot regarding copyright infringement. This search did not reveal any kind of copyright infringement. Commendations are due to Quillbot, which rewrites and paraphrases texts based on AI algorithms without endangering other people’s intellectual property.

2.8. AI-assisted and AI-generated poetry

Using BERT, our young doctoral student managed to convincingly imitate the poetry of the bard of Macedonian literature and the founder of the standard Macedonian language, Blaže Koneski (1921 - 1993). Koneski died recently, so his works are still not part of the public domain. Apparently, no one is concerned about his copyright, neither the publishers who printed his works, nor his daughter, who is the only legal heir to his authorship. Moreover, it turns out that I am the only one who warned the PhD student about the possible copyright implications of his GenAI-generated works. He defends himself that he does not harm anyone, because he does not copyright or publish his

AI-generated poetry with copyright and does not use it for commercial purposes. Is he committing copyright infringement by using Koneski's copyrighted poetry for training his AI model? In my opinion, yes, although some believe that this only promotes Koneski's work.

This example prompted me to find out if there are any examples of imitating someone's poetry using an AI model. I have come across several scientific papers that present similar experiments. For example, the paper by Köbis and Mossink (2021) has already received over 300 citations, which proves the quality of their research. Their system was trained on the works of the famous poet Maya Angelou (1928 - 2014), whose site <https://www.mayaangelou.com/> is copyright protected, so it is very likely that the poetry is also protected by the same type of intellectual property protection. The paper does not give credit to its authorship, but only to the authors of the paper.

An interesting work is related to haiku poetry, which generative artificial intelligence, according to the research participants, was able to imitate (Hitsuwari, Ueda, Yun, & Nomura, 2023). This paper does not mention the authorship of the songs that were used during the training of the system. If it is ancient art, then these works are in the public domain, so there is no copyright infringement. Nevertheless, if the haiku poetry belongs to contemporary authors, then the problem is nothing but obvious.

I will conclude this case study with Hutson & Schnellmann's (2023) suggestions that a compromise must be found between human authorship and technology, primarily for the advancement of artificial intelligence or robotics. I personally do not agree with them at all, and especially with their position that one must always give priority to practical aspects over authorship (Hutson & Schnellmann, 2023).

2.9. AI-assisted and AI-generated novels

By searching Amazon books, I noticed that Amazon offers over 500 books related to AI-assisted content. The amount of novels generated by AI seems to be ten times higher. They mostly include works assisted by AI, where the role of the AI was to support the creation, not to write the novels. Amazon Kindle allows authors and publishers to upload eBooks to their platform, regardless of whether the content was generated by a human author or by AI (Amazon, 2024). Amazon requires authors to notify of AI-generated content when they publish a new book or edit and republish an existing book through Kindle Direct Publishing. Amazon transfers the responsibility of legal violation and violation of property rights to authors, publishers and sales partners who must guarantee that the work does not contradict the laws for the protection of intellectual property. Any additional use of protected works must be guaranteed with the written permission of the copyright holder.

The copyright for books created by generative AI is not mentioned by Amazon. On the other hand, there are many different AI applications that allow for the writing of novels. One example is the reporter at the Verge Adi Robertson (2023) (<https://www.theverge.com/authors/adi-robertson>), who went to great lengths to explain to Sudowrite ([copyshark.ai](https://www.sudowrite.com)) how to write the novel *The Electric Sea* for her (Tumblr, 2023). The Sudowrite website no longer exists and the Copyright Tools section was transferred towards WordPress. Therefore, I am unable to verify what the position is regarding the copyright of the AI-generated novel.

I watched *The Electric Sea* on Tumblr (<https://www.tumblr.com/>). I have to admit that this “cyberpunk novella” did not grab my attention. But sooner or later, similar AI tools will learn to write more excitingly, so their novels will become fun to read, especially for younger readers. As these books can only be created through the active cooperation of individuals, the authorship belongs to them. When they become completely autonomous, then copyright will become a serious problem.

3. Conclusions

Artificial intelligence is becoming an indispensable part of our everyday life. Based on human prompts, generative AI is capable of autonomously producing creative outputs and intellectual contributions. If generative AI were a human being or a natural person, those creations would be the author’s works. The question of who holds copyright for AI-assisted and AI-generated content is still unanswered.

The current legal framework, especially in the EU, is cautiously trying to balance innovation in AI, including general purpose AI (GPAI), with respect to copyright (Quintais, 2025). The EU AI Act does not change copyright law, but reinforces the need for transparency and lawful data use. Creators have tools to opt out of data mining and to assert their rights, but these mechanisms are still evolving and face many challenges (Lucchi, 2024). This primarily refers to the protection of the integrity of the original works used for training AI models and ensuring adequate compensation if they are not in the public domain.

What can be said with certainty is that translating a work using machine translation does not transfer the copyright to machine translation systems. The same applies to all LLMs with translation abilities. Similarly, the texts, images, and videos generated by LLMs are not copyrighted by that LLM. They belong to the user.

What can be learnt from this paper is that art generated by AI is prohibited for copyright protection because it does not meet the requirement of human authorship according to current laws. Artificial or manipulated images, audio or video content created by AI can become copyrighted works by people who have asked the AI to create them. The only prerequisite is that the share of artificial intelligence must be clearly marked. Book authors are also responsible for reporting content generated by AI.

The greater the contribution of people, the more they have the right to claim authorship. On the other hand, generative AI is becoming more autonomous every day. This leads every human person, especially the inexperienced, to the temptation to appropriate work in which they have invested almost no share. The tension between integrity and ambition highlights a profound dilemma: the desire for greatness can sometimes lead people down a morally ambiguous path. As people navigate their own aspirations, it becomes crucial to reflect on the means they choose to pursue their goals, balancing ambition with ethical standards.

Many new regulatory frameworks have been created to prevent the misuse of AI-assisted and AI-generated content. Training on copyrighted material is a legally grey area. Some argue it’s fair use or allowed under text and data mining (TDM) exceptions, others see it as infringement. Regulatory frameworks can address the concerns related to copyright infringement. Their effect is to protect the true authorship, while preventing the automated creation of both persuasive and irrelevant content.

For now, creators should actively assert their rights wherever possible and stay informed as laws continue to develop. Meanwhile, AI developers must prioritize transparency, fair use, and ethical data practices to ensure compliance and public trust. In the long term, harmonization between AI regulations and copyright laws will be essential. They should explicitly distinguish AI-assisted from AI-generated works.

Laws will soon become stricter to bring order to the sudden influx of new technologies. They will establish licensing frameworks for training data. Moreover, they might suggest easier global mechanisms for creators to control the use of their works in AI systems. But, we must be aware that Walter Mizner was right when he said, "If you steal from one author, it's plagiarism; if you steal from many it's research" (Johnston, 1953, p. 66). Generative AI has thus become a high-calibre researcher. This means that the real challenge is to prevent fraudsters who will be tempted to take the results of generative AI as their own. The obligation of researchers is to prevent them from doing so, at least by avoiding becoming consumers of infringing works (Lucchi, 2024).

Otherwise, an avalanche of stolen works will prevail over the real results of human creativity, mental creations, literary and artistic works. Then the creation of "organic" intellectual property will become extremely rare, which will demotivate most human authors to create new works. I am an optimist that over time "organic" intellectual property will become extremely respected and that its price will become high due to the fact that without new original works even artificial intelligence will not be able to advance a single step forward. This will result in a great regression of humanity. Human civilization has successfully overcome greater trials, so it will surely overcome this too.

References

- Amazon. (2024). *Content Guidelines*. Retrieved 2025-06-26, from https://kdp.amazon.com/en_US/help/topic/G200672390
- Bagenal, J. (2024). Generative artificial intelligence and scientific publishing: Urgent questions, difficult answers. *The Lancet*, 403(10432), 1118–1120. Retrieved from <https://linkinghub.elsevier.com/retrieve/pii/S0140673624004161> doi: 10.1016/S0140-6736(24)00416-1
- Chesterman, S. (2025). Good models borrow, great models steal: Intellectual property rights and generative AI. *Policy and Society*, 44(1), 23–37. Retrieved 2025-12-05, from <https://academic.oup.com/policyandsociety/article/44/1/23/7606572> doi: 10.1093/polsoc/puae006
- Chu, T., Song, Z., & Yang, C. (2024, March). How to protect copyright data in optimization of large language models? In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 38, pp. 17871–17879). Retrieved 2025-12-05, from <https://arxiv.org/abs/2308.12247> (Version Number: 1) doi: 10.48550/ARXIV.2308.12247
- Crews, K. D. (2020). *Copyright Law for Librarians and Educators: Creative Strategies and Practical Solutions*. Chicago: American Library Association.
- EJOI MK. (2024). *Odime po medalji*. Retrieved 2025-06-26, from <https://motika.mk/wp-content/uploads/2024/07/05/ejoi.mp4>
- EU. (2019). *Directive (EU) 2019/790 of the European parliament and of the Council*. Retrieved 2025-06-26, from <https://eur-lex.europa.eu/legalcontent/EN/TXT/HTML/?uri=CELEX:32019L0790>

- European Parliament and Council. (2024). Artificial Intelligence Act. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June. In *Official Journal of the European Union, L 1689, July 12*. Retrieved 2025-06-26, from <https://artificialintelligenceact.eu/>
- Ferrara, E. (2023). Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies. *Sci*, 6(1), 3. Retrieved from <https://www.mdpi.com/2413-4155/6/1/3> doi: 10.3390/sci6010003
- Frank, J., Herbert, F., Ricker, J., Schönherr, L., Eisenhofer, T., Fischer, A., ... Holz, T. (2024). A representative study on human detection of artificially generated media across countries. In *2024 IEEE Symposium on Security and Privacy (SP)* (pp. 55–73). IEEE. Retrieved from <https://arxiv.org/abs/2312.05976> (Version Number: 1) doi: 10.48550/ARXIV.2312.05976
- Free Software Foundation. (2022). *What is copyleft*. Retrieved 2025-06-26, from <https://www.gnu.org/licenses/copyleft.en.html>
- Gopalakrishna, G., Ter Riet, G., Vink, G., Stoop, I., Wicherts, J. M., & Bouter, L. M. (2022). Prevalence of questionable research practices, research misconduct and their potential explanatory factors: A survey among academic researchers in The Netherlands. *PLOS ONE*, 17(2), e0263023. Retrieved from <https://dx.plos.org/10.1371/journal.pone.0263023> doi: 10.1371/journal.pone.0263023
- Gray, A. (2024). *ChatGPT "contamination": Estimating the prevalence of LLMs in the scholarly literature*. arXiv. Retrieved from <http://arxiv.org/abs/2403.16887> (arXiv:2403.16887 [cs]) doi: 10.48550/arXiv.2403.16887
- Hernández-Huerta, V. A. (2020). Disputed Elections in Presidential Democracies: Contexts of Electoral “Blackmail”. *The Journal of Politics*, 82(1), 89–103. Retrieved from <https://www.journals.uchicago.edu/doi/10.1086/705599> doi: 10.1086/705599
- Hesmondhalgh, D. (2018). *The Cultural Industries* (Fourth edition ed.). Los Angeles: SAGE Publications Ltd.
- Hitsuwari, J., Ueda, Y., Yun, W., & Nomura, M. (2023). Does human–AI collaboration lead to more creative art? Aesthetic evaluation of human-made and AI-generated haiku poetry. *Computers in Human Behavior*, 139. Retrieved from <https://linkinghub.elsevier.com/retrieve/pii/S0747563222003223> doi: 10.1016/j.chb.2022.107502
- Hutson, J., & Schnellmann, A. (2023). The poetry of prompts: the collaborative role of generative artificial intelligence in the creation of poetry and the anxiety of machine influence. *Global Journal of Computer Science and Technology: D*, 23(1).
- Informa PLC. (2024). *Press Release*. Retrieved 2025-06-26, from <https://www.informa.com/globalassets/documents/investorrelations/2024/informa-plc---market-update.pdf>
- Jaszi, P. (2017). Toward a theory of copyright: The metamorphoses of “authorship”. In *Intellectual Property Law and History* (pp. 61–108). Routledge.
- Johnston, A. (Ed.). (1953). *Chapter 4. In: The legendary Mizners* (pp. 66). New York: Farrar, Straus and Giroux.
- Justia. (2023). *Andersen et al v. Stability AI Ltd. et al*. Retrieved 2025-06-26, from <https://docs.justia.com/cases/federal/districtcourts/california/candce/3:2023cv00201/407208/67>
- Köbis, N., & Mossink, L. D. (2021). Artificial intelligence versus Maya Angelou: Ex-

- perimental evidence that people cannot differentiate AI-generated from human-written poetry. *Computers in Human Behavior*, 114, 106553. Retrieved from <https://linkinghub.elsevier.com/retrieve/pii/S0747563220303034> doi: 10.1016/j.chb.2020.106553
- Liang, W., Zhang, Y., Wu, Z., Lepp, H., Ji, W., Zhao, X., ... Zou, J. Y. (2024). *Mapping the increasing use of LLMs in scientific papers*. arXiv. Retrieved 2025-12-05, from <https://arxiv.org/abs/2404.01268> (Version Number: 1) doi: 10.48550/ARXIV.2404.01268
- Library of Congress. (2023). *Copyright registration guidance: Works containing material generated by artificial intelligence*. Retrieved 2025-06-26, from <https://www.federalregister.gov/documents/2023/03/16/2023-05321/copyright-registration-guidance-works-containing-material-generated-by-artificial-intelligence>
- Lucchi, N. (2024). ChatGPT: A Case Study on Copyright Challenges for Generative Artificial Intelligence Systems. *European Journal of Risk Regulation*, 15(3), 602–624. Retrieved from https://www.cambridge.org/core/product/identifier/S1867299X23000594/type/journal_article doi: 10.1017/err.2023.59
- Lunney Jr, G. S. (2001). The death of copyright: Digital technology, private copying, and the digital millennium copyright act. *Virginia Law Review*, 813–920.
- Myers, C. S. (2018). Plagiarism and copyright: Best practices for classroom education. *College & Undergraduate Libraries*, 25(1), 91–99. Retrieved 2025-12-05, from <https://www.tandfonline.com/doi/full/10.1080/10691316.2017.1391028> doi: 10.1080/10691316.2017.1391028
- Nicholls, A. R., Madigan, D. J., Duncan, L., Hallward, L., Lazuras, L., Bingham, K., & Fairs, L. R. W. (2020, September). Cheater, cheater, pumpkin eater: the Dark Triad, attitudes towards doping, and cheating behaviour among athletes. *European Journal of Sport Science*, 20(8), 1124–1130. Retrieved 2025-12-05, from <https://onlinelibrary.wiley.com/doi/10.1080/17461391.2019.1694079> doi: 10.1080/17461391.2019.1694079
- Ochoa, T. T. (2002). Origins and meanings of the public domain. *U. Dayton L. Rev*, 28, 215.
- OpenAI. (2024). *Europe Terms of Use*. Retrieved 2025-06-26, from <https://openai.com/policies/terms-of-use/>
- Panettieri, J. (2024). *Generative AI Lawsuits Timeline: Legal Cases vs. OpenAI, Microsoft, Anthropic, Nvidia, Perplexity, Intel and More*.
- Pila, J. (2010). Copyright and Its Categories of Original Works. *Oxford Journal of Legal Studies*, 30(2), 229–254. Retrieved from <https://academic.oup.com/ojls/article-lookup/doi/10.1093/ojls/gqq009> doi: 10.1093/ojls/gqq009
- Quintais, J. P. (2025). Generative AI, copyright and the AI Act. *Computer Law & Security Review*, 56, 106107. Retrieved 2025-12-05, from <https://linkinghub.elsevier.com/retrieve/pii/S0267364925000020> doi: 10.1016/j.clsr.2025.106107
- Resnikoff, P. (2024). *UMG, WMG, Sony Music File Litigation Against AI Music Services Suno and Udio for Massive Copyright Infringement*. Retrieved 2025-06-26, from <https://www.digitalmusicnews.com/2024/06/24/umg-wmgsony-litigation-ai-music-suno-udio/>
- Robertson, E. (2023). *I tried the AI novel-writing tool everyone hates, and it's better*

- than I expected. Retrieved 2025-06-26, from <https://www.theverge.com/2023/5/24/23732252/sudowritestory-engine-ai-generated-cyberpunk-novella>
- Schilit, H. M., & Perler, J. (2010). *Financial Shenanigans* (3rd ed.). McGraw-Hill.
- Schoening, S. (2023). *Machine Translation Explained: Types, Use Cases, and Best Practices*. Retrieved 2025-06-26, from <https://phrase.com/blog/posts/machine-translation/>,
- Spink, J., Moyer, D. C., Park, H., & Heinonen, J. A. (2013). Defining the types of counterfeiters, counterfeiting, and offender organizations. *Crime Science*, 2(1), 8. Retrieved from <https://crimesciencejournal.biomedcentral.com/articles/10.1186/2193-7680-2-8> doi: 10.1186/2193-7680-2-8
- Stanford University IT. (2024). *Dangers of Deepfake: What to Watch For*. Retrieved 2025-06-26, from <https://uit.stanford.edu/news/dangersdeepfake-what-watch>
- Temple, K. (2003). *Scandal Nation: Law and Authorship in Britain, 1750–1832*. Ithaca, NY: Cornell University Press.
- Tumblr. (2023). *The AI Writers Workshop*. Retrieved 2025-06-26, from <https://robotlit.com/post/717956168948023296/the-electric-sea-an-ai-generated-experiment>
- U.S. Copyright Office. (1976). *The Lifecycle of Copyright*. Retrieved 2025-06-26, from <https://www.copyright.gov/history/copyrightexhibit/lifecycle/>
- U.S. Copyright Office. (2020). *The Digital Millennium Copyright*. Retrieved 2025-06-26, from <https://www.copyright.gov/dmca/>
- U.S. Copyright Office. (2024). *Copyright Law of the United States (Title 17)*. Retrieved 2025-06-26, from <https://www.copyright.gov/title17/>
- U.S. Department of Justice. (2022). *CRM 1500-1999*. Retrieved 2025-06-26, from <https://www.justice.gov/archives/jm/criminal-resource-manual1852-copyright-infringement-penalties-17-usc-506a-and-18-usc-2319>
- Varian, H. R. (2005). Copying and Copyright. *Journal of Economic Perspectives*, 19(2), 121–138. Retrieved 2025-12-05, from <https://pubs.aeaweb.org/doi/10.1257/0895330054048768> doi: 10.1257/0895330054048768
- Vieira, L. N., O'Hagan, M., & O'Sullivan, C. (2021). Understanding the societal impacts of machine translation: a critical review of the literature on medical and legal use cases. *Information, Communication & Society*, 24(11), 1515–1532. Retrieved from <https://www.tandfonline.com/doi/full/10.1080/1369118X.2020.1776370> doi: 10.1080/1369118X.2020.1776370
- Wach, K., Duong, C. D., Ejdys, J., Kazlauskaitė, R., Korzynski, P., Mazurek, G., ... Ziemia, E. (2023). The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT. *Entrepreneurial Business and Economics Review*, 11(2), 7–30. Retrieved 2025-12-05, from <https://eber.uek.krakow.pl/index.php/eber/article/view/2113> doi: 10.15678/EBER.2023.110201
- WIPO. (1979). *Berne Convention for the Protection of Literary and Artistic Works*. Retrieved 2025-06-26, from <https://www.wipo.int/wipolex/en/treaties/textdetails/12214>
- WIPO. (1996). *Copyright*. Retrieved 2025-06-26, from <https://www.wipo.int/copyright/en/>
- Woody, K. E. (2020). The new insider trading. *Ariz. St. LJ*, 52.
- Young, L. (2024). *What are the new and emerging trends in academic misconduct?* Retrieved 2025-06-26, from <https://www.turnitin.com/blog/what-are-the-new-andemerging-plagiarism-trends>

- Yusuf, A., Pervin, N., & Román-González, M. (2024). Generative AI and the future of higher education: a threat to academic integrity or reformation? Evidence from multicultural perspectives. *International Journal of Educational Technology in Higher Education*, 21(1), 21. Retrieved from <https://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-024-00453-6> doi: 10.1186/s41239-024-00453-6
- Zdravkova, K. (2023). Evolution of academic dishonesty in computer science courses. In *9th International Conference on Higher Education Advances (HEAd'23)* (pp. 421–428). Universitat Politècnica de València. Retrieved from <http://ocs.editorial.upv.es/index.php/HEAD/HEAd23/paper/view/16081> doi: 10.4995/HEAd23.2023.16081
- Zdravkova, K., & Ilijoski, B. (2025). Preventing Academic Dishonesty Originating from Large Language Models. In C. Bădică, M. Gušev, A. Iftene, M. Ivanović, Y. Manolopoulos, & S. Xinogalos (Eds.), *Advances in ICT Research in the Balkans* (Vol. 2391, pp. 118–132). Cham: Springer Nature Switzerland. Retrieved from https://link.springer.com/10.1007/978-3-031-84093-7_9 (Series Title: Communications in Computer and Information Science) doi: 10.1007/978-3-031-84093-7_9