Karimli E.E. ♦

DOI: 10.25108/2304-1730-1749.iolr.2024.74.125-131
UDC: 347.778

Big data and copyright protection of artificial intelligence-generated works

Abstract: In the era of digitalization, the convergence of Big Data and copyright protection, specifically for Artificial Intelligence (AI)-generated works, presents a novel and intricate legal landscape. This article explores the challenges and opportunities at this intersection, focusing on how Big Data develops AI's capability to create potentially copyrightable works. As AI systems become increasingly capable of producing works indistinguishable from those created by humans, the need for legal frameworks to evolve accordingly becomes paramount. This research aims to shed light on the complexities involved in copyrighting AI-generated works, underscoring the urgent necessity for legal scholarship and policymaking to adapt to technological advancements. The objective is to ensure that copyright law remains effective in safeguarding creativity and innovation in the digital age, balancing the interests of creators, users, and the wider public.

Keywords: artificial intelligence; Big Data; intellectual property law; copyright protection.

Introduction

In the digitalized world, the juxtaposition between Big Data and Copyright Protection, specifically concerning Artificial Intelligence ("AI")-generated works, puts forward a novel and complex legal landscape. As we delve deeply into the information age, the major advancements of AI primarily lie in the exponential growth of Big Data. These AI systems, fuelled by vast datasets, are now capable of producing creative, intellectual, and original works. This development challenges the traditional boundaries of copyright law, which was primarily designed to protect human-created works. The core of this issue lies in determining the applicability of copyright protection to AI-generated works.

The article aims to unravel the complexities at the confluence of Big Data and copyright protection for AI-generated works. It seeks to analyze how Big Data contributes to the capability of AI systems to create works of potentially copyrightable nature and delves into the legal, implications of extending copyright protection to such works.

Big Data Phenomenon and AI-Generated Works

Big Data is conceived as large amounts of different types of data produced with high velocity from a high number of various types of sources and its handling requires novel tools and methods, such as powerful processors, software, and algorithms. [1, p. 4]. To better understand the notion of Big Data, one needs to explore five dimensions thereof as the following: 1) volume; 2) veracity; 3) velocity; 4) variety and 5) value [2, p.4].

As a first dimension, Big Data is characterized by its volume, which implies its large size which sets it apart from smaller datasets. Considering that Big Data is often processed in an auto-

♦ **Karimli Elvin Eyvaz oglu –** Master's student at the Department of Intellectual Property Law, Faculty of Law, Baku State University (Azerbaijan). E-mail: karimlielvin01@gmail.com

mated manner, the veracity rests upon the data's quality and trustworthiness. Velocity means the speed rate at which data is produced, gathered, utilized, and processed. Variety denotes the diverse types of data and sources, e.g., internet platforms, social media, and various digital devices. Last but not least, once these dimensions converge, the data corpus brings about substantial value, which is vital for significant insights and applications.

As the artificial intelligence is exponentially growing, Big Data corpus has started to play a significant role in the data-driven industries. As an interrelation of these concepts, one needs to understand the process of gathering, utilization, and processing of Big Data. Firstly, Big Data is massively gathered out of different sources, such as Internet of Things, social media platforms [3, p. 4]. Furthermore, Big Data is processed and analysed, which incorporates Text and Data Mining (TDM) process. As using the artificial intelligence, TDM process allows the AI machine to learn and get better what it does over the time. Accordingly, this process is called as machine learning. Dependent upon the advancement level of machine learning, the artificial intelligences vary from each other [4]. In this regard, on the one hand, the human input is required to assist the AI system to effectuate the machine learning, on the other hand, the human contribution is of secondary character and the AI system itself carries out the machine learning process. Namely, the extent of human involvement delineates the boundary between AI- or computer-generated works and computer-assisted works [5, p. 3].

With respect to the computer-assisted works, the computers or artificial intelligence operate as a mere tool to assist the human authors in the same manner that a pen or pencil can help them to generate the work. [6, p. 2]. In case the works are generated in this fashion, it does not cause any challenge vis-à-vis the copyright protection and the human author itself can be qualified as the copyright-holder.

Nonetheless, the extent of human involvement could be so minimal or lacks its existence in the case of AI-generated works, so that there is no human author or intellectual creative activity at hand. Namely, it might be the case that an algorithm of artificial intelligence can be programmed in such a manner that it can behave unpredictably and independently, which makes their original programmers not understand precisely how they function [7, p. 3]. In this vein, it can pose a great difficulty on whether the AI-generated works can fall under the copyright protection. Taking into consideration this challenge, the following chapter will delve into the analysis on the juxtaposition between the copyright protection and the AI-generated works.

Copyright Protection and AI-generated works

As a form of intellectual property right, the copyright protection denotes the exclusive rights granted to the authors of original works for a limited period. Even though the copyright protection is territorially limited, the foundational framework thereof rests upon the following concepts: a) the idea-expression dichotomy, originality, and authorship of a work [8, p. 7]. Prior to determining whether the AI-generated works meet the requirements of these concepts, there is a need to elucidate them in the following part. For that being, the United States of America, the European Union and Azerbaijan copyright standards will be analysed.

US Copyright standards

As per the 1976 US Copyright Act, the copyright protection subsists *"…in original works of authorship fixed in any tangible medium of expression"* [9]. The US copyright legislation explicitly states the requirement of authorship and originality. Specifically, it codifies that *"…in no case does copyright protection for an original work of authorship extend to any idea, procedure, process, sys-*

tem, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such work" [10].

Regarding the originality requirement, it means that the particular work owes its origin to the author. To illustrate more, the originality requirement implies that the work shall be independently created by the author, and it shall have the minimal degree of creativity [11]. Accordingly, the US copyright protection has centred upon the authorial expression of intellectual creative activity.

In this vein, the US jurisprudence also elucidates the concept of authorship under the copyright legislation. In the case of the *Wrecked and Abandoned Vessel R.M.S. Titanic*, the Court put forward that *"…the author of a work is the person who actually creates the work, that is, the person who translates an idea into a fixed, tangible expression entitled to copyright protection"* [12].

EU Copyright standards

In relation to the copyright protection, the Court of Justice of European Union harmonized the concept of work and originality under the requirement of author's own intellectual creation. The CJEU determines the author's own intellectual creation as an indispensable requirement for the copyright protection. In the decision of *Infopaq International v. Danske Dagblades,* the CJEU noted that copyright protection is applied only *"…in relation to a subject-matter which is original in the sense that it is its author's own intellectual creation"* [13, p. 37].

In a nutshell, two essential criteria must be met to be recognized as a "work" under copyright law: first, the subject matter must be original, reflecting the author's personal intellectual creation; and second, it must embody the author's own intellectual creation, ensuring that the expression of the idea is unique to the author.

Azerbaijan Copyright standards

The copyright protection is regulated under the Law "On Copyright and Related Rights" of the Republic of Azerbaijan. As per the Law, the copyright protection applies regardless of the purpose, value, and content, as well as the form and method of expression, to the results of creative activity that exist in an objective form, both published and unpublished, in the fields of science, literature, and art [14, Article 5.1]. Furthermore, the Law defines the author as the human who generates the work [15, Article 4].

Accordingly, the Azerbaijani legislation determines the requirement of author's intellectual creative activity and links the concept of authorship with the human-centred approach.

Applicability of the copyright standards to the AI-generated works

As above-mentioned, the US, EU and Azerbaijan copyright protection standards generally lie in the originality and author's intellectual creative activity requirements. Moreover, the copyright protection has centred upon the human-based approach.

Accordingly, the conceptual framework of copyright protection is based on the human creativity, so that the AI-generated works are devoid of any copyright protection. This is because these works are akin to ideas, concepts, facts, or matters that arise from technological processes rather than from the creative input of an author. Consequently, they do not exist as things in law and thus, cannot be owned.

Practical cases under the US legislation

A Recent Entrance to Paradise

The evolution of copyright law concerning AI-generated works has been notably influenced by efforts from Ryan Abbott and Stephen Thaler. They initiated several test cases across different jurisdictions aiming to enhance intellectual property law towards acknowledging non-human crea-

tivity. Their initial attempt to register a copyright for *"A Recent Entrance to Paradise"* was denied by the US Copyright Office. The denial highlighted the necessity for works to originate from human intellectual effort and creativity to qualify for copyright protection. The Office's review emphasized that for a work to be copyrighted, it must either be proven to be a product of human creation or challenge and change longstanding copyright practices. [16, p. 3] Their analysis referenced key copyright cases, underscoring the consistent judicial stance that copyright protection is reserved for human-authored works.

Zarya of the Dawn

The case involving the comic book *"Zarya of the Dawn"* presents a unique scenario. Artist and AI-educator Kristina Kashtanova initially didn't reveal her use of AI in creating the work, leading to the Copyright Office moving to cancel her registration upon discovering her claims on social media [17, p. 2]. Kashtanova stated that she authored the entire work using AI merely as a tool, and argued that her contributions, particularly the text and the compilation of text and images, were original and thus copyrightable.

The Copyright Office reiterated that copyright protection is reserved for human-created works. It recognized the text written by Kashtanova as original and copyrightable but deemed the AI-generated images as non-original since the AI, not Kashtanova, produced the foundational elements of the artwork [18, p. 8]. However, the Office found Kashtanova's role in selecting and arranging the text and images copyrightable as a compilation, presenting that her creative process in compiling the work reflected her personal touch and distinguished her contributions from the AI's input.

Conclusion

As we navigate the complexities at the intersection of Big Data, AI, and copyright protection, it becomes clear that our current legal frameworks are ill-equipped to address the novel challenges presented by AI-generated works. The exploration of US, EU, Azerbaijan legislation and recent cases such as "A Recent Entrance to Paradise" and "Zarya of the Dawn" underscores the pressing necessity for copyright law to evolve in response to technological advancements. These examples highlight the limitations of traditional concepts of authorship and originality, which are rooted in human creativity, in encompassing works produced by AI.

The findings of this article argue for a nuanced approach to copyright protection that acknowledges the unique nature of AI-generated works. Such an approach should consider the extent of human involvement in the creative process, the role of AI as a tool or collaborator, and the implications for creativity and innovation in the digital age. Furthermore, it is evident that a one-size-fits-all solution is unlikely to be effective, instead, a flexible and adaptive legal framework is required to accommodate the diverse and evolving nature of AI-generated content.

References

1. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Towards a thriving data-driven economy, SWD(2014) 214 final, pp. 1-12
2. Daniel J. Gervais, Exploring the Interfaces Between Big Data and Intellectual Property Law, 10 Journal of Intellectual Property, Information Technology and Electronic Commerce Law3, 2019 pp. 1-19

3. Daniel J. Gervais, Exploring the Interfaces Between Big Data and Intellectual Property Law, 10 Journal of Intellectual Property, Information Technology and Electronic Commerce Law 3, 2019 pp. 1-19

4. Cassie Kozyrkov. Are you using the term 'AI' incorrectly? Available at: https://hackernoon.com/are-you-using-the-term-ai-incorrectly-911ac23ab4f5

5. Matt Blaszczyk, Impossibility of Emergent Works' Protection in U.S. and EU Copyright Law,25 North Carolina Journal of Law & Technology 1, 2023, pp.1-55

6. Ryan Abbott, Artificial Intelligence, Big Data and Intellectual Property: Protecting Computer-Generated Works in the United Kingdom, pp. 1-16. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3064213

7. Ryan Abbott, Artificial Intelligence, Big Data and Intellectual Property: Protecting Computer-Generated Works in the United Kingdom, pp. 1-16. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3064213

8. Matt Blaszczyk, Impossibility of Emergent Works' Protection in U.S. and EU Copyright Law, 25 North Carolina Journal of Law & Technology 1, 2023 pp.1-55

9. 17 United States of America Copyright Code § 102(a).

10. 17 United States of America Copyright Code § 102(b).

11. U.S Supreme Court, Feist Pubs., Inc. v. Rural Tel. Svc. Co., Inc., 499 U.S. 340 (1991) Available at: https://supreme.justia.com/cases/federal/us/499/340/

12. Lindsay v. Wrecked & Abandoned Vessel R.M.S. Titanic, No. 97 Civ. 9248 (HB), 1999 WL 816163 (S.D.N.Y. Oct. 13, 1999).

13. Case C-5/08, Infopaq Int'l v. Danske Dagblades, 2009 E.C.R. I-06569.

14. The Law "On Copyright and Related Rights" of the Republic of Azerbaijan Available at: https://e-qanun.az/framework/4167 (in Azerbaijani).

15. Letter from Shira Perlmutter, U.S. Copyright Off., to Ryan Abbott, 14 February 2022, Regarding Second Request for Reconsideration for Refusal to Register a Recent Entrance to Paradise, pp. 1-15 Available at: https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2022cv1564-24

16. Letter from Robert J. Kasunic, U.S. Copyright Off., to Van Lindberg, Taylor English Duma LLP (Feb. 21, 2023) (on file with Library of Congress) (regarding Zarya of the Dawn, registration # VAu001480196)

17. Letter from Robert J. Kasunic, U.S. Copyright Off., to Van Lindberg, Taylor English Duma LLP (Feb. 21, 2023) (on file with Library of Congress) (regarding Zarya of the Dawn, registration # VAu001480196)

Каримли Э.Э.♦

DOI: 10.25108/2304-1730-1749.iolr.2024.74.125-131
УДК: 347.778

Большие данные и защита авторских прав на произведения, созданные искусственным интеллектом

Аннотация: В эпоху цифровизации сближение Больших Данных и защиты авторских прав, в частности для произведений, созданных искусственным интеллектом (ИИ), представляет собой новый и сложный правовой ландшафт. В данной статье исследуются вызовы и возможности на этом перекрестке, с фокусом на то, как Большие Данные развивают способность ИИ создавать потенциально авторские произведения. По мере того, как системы ИИ становятся всё более способными производить работы, неотличимые от созданных человеком, необходимость в соответствующем эволюционировании правовых рамок становится очевидной. Это исследование направлено на освещение сложностей, связанных с авторским правом на произведения, созданные ИИ, подчеркивая срочную необходимость в адаптации правовой науки и политики к технологическим достижениям. Целью является обеспечение того, чтобы закон об авторском праве оставался эффективным в защите творчества и инноваций в цифровую эпоху, сбалансировав интересы создателей, пользователей и широкой общественности.

Ключевые слова: искусственный интеллект; Большие Данные; интеллектуальное право; защита авторских прав.

Библиография

1. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Towards a thriving data-driven economy, SWD(2014) 214 final, pp. 1-12
2. Daniel J. Gervais, Exploring the Interfaces Between Big Data and Intellectual Property Law, 10 Journal of Intellectual Property, Information Technology and Electronic Commerce Law3, 2019 pp. 1-19
3. Daniel J. Gervais, Exploring the Interfaces Between Big Data and Intellectual Property Law, 10 Journal of Intellectual Property, Information Technology and Electronic Commerce Law 3, 2019 pp. 1-19
4. Cassie Kozyrkov. Are you using the term 'AI' incorrectly? Available at: https://hackernoon.com/are-you-using-the-term-ai-incorrectly-911ac23ab4f5
5. Matt Blaszczyk. Impossibility of Emergent Works' Protection in U.S. and EU Copyright Law,25 North Carolina Journal of Law & Technology 1, 2023 pp.1-55
6. Ryan Abbott, Artificial Intelligence, Big Data and Intellectual Property: Protecting Computer-Generated Works in the United Kingdom, pp. 1-16. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3064213

♦ **Каримли Эльвин Эйваз оглы - м**агистрант кафедры интеллектуального права, юридический факультет, Бакинский государственный университет (Азербайджан). E-mail: karimlielvin01@gmail.com

7. Ryan Abbott, Artificial Intelligence, Big Data and Intellectual Property: Protecting Computer-Generated Works in the United Kingdom, pp. 1-16. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3064213

8. Matt Blaszczyk, Impossibility of Emergent Works' Protection in U.S. and EU Copyright Law, 25 North Carolina Journal of Law & Technology 1, 2023, pp.1-55

9. 17 United States of America Copyright Code § 102(a).

10. 17 United States of America Copyright Code § 102(b).

11. U.S Supreme Court, Feist Pubs., Inc. v. Rural Tel. Svc. Co., Inc., 499 U.S. 340 (1991) Available at: https://supreme.justia.com/cases/federal/us/499/340/

12. Lindsay v. Wrecked & Abandoned Vessel R.M.S. Titanic, No. 97 Civ. 9248 (HB), 1999 WL 816163 (S.D.N.Y. Oct. 13, 1999).

13. Case C-5/08, Infopaq Int'l v. Danske Dagblades, 2009 E.C.R. I-06569.

14. The Law "On Copyright and Related Rights" of the Republic of Azerbaijan Available at: https://e-qanun.az/framework/4167 (in Azerbaijani).

15. Letter from Shira Perlmutter, U.S. Copyright Off., to Ryan Abbott, 14 February 2022, Regarding Second Request for Reconsideration for Refusal to Register a Recent Entrance to Paradise, pp. 1-15 Available at: https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2022cv 1564-24

16. Letter from Robert J. Kasunic. U.S. Copyright Off., to Van Lindberg, Taylor English Duma LLP (Feb. 21, 2023) (on file with Library of Congress) (regarding Zarya of the Dawn, registration # VAu001480196)

17. Letter from Robert J. Kasunic. U.S. Copyright Off., to Van Lindberg, Taylor English Duma LLP (Feb. 21, 2023) (on file with Library of Congress) (regarding Zarya of the Dawn, registration # VAu001480196)