COPYRIGHT PROTECTION IN THE CONTEXT OF THE PROVISION OF AI MODELS FOR GENERAL USE

Alexander Ștefan POPA[*]

Abstract

This paper examines how the new European AI Act (Regulation (EU) 2024/1689) governs the use of copyrighted content in the training and deployment of general-purpose AI models, focusing on the pivotal Article 53. Article 53 is a key provision that imposes explicit obligations on AI model providers to ensure transparency, rigorous documentation, and adherence to copyright law. The analysis highlights the legal and ethical implications of using copyrighted works to train AI systems, noting the tension between the need for vast datasets and the rights of authors. It explores how providers are required to maintain detailed technical documentation and publicly disclose summaries of training data, implement policies to comply with copyright (including honoring rights-holders' opt-outs), and guarantee compliance with EU copyright and related rights. These responsibilities aim to increase accountability and enable oversight while fostering trust in AI outputs. At the same time, the paper discusses how the regulation seeks to balance innovation and copyright protection, guided by principles of proportionality and fairness: general-purpose AI development is permitted but constrained to respect authors' economic and moral rights. The broader EU legal context – including existing copyright exceptions (such as text and data mining allowances), moral rights of creators, and the three-step test – is considered to understand the boundaries of lawful AI training. Finally, the abstract addresses enforcement challenges, such as ensuring transparency without compromising trade secrets, difficulty in tracking protected content in large datasets, and cross-border compliance. The study concludes that Article 53 represents a significant step toward aligning AI innovation with European copyright norms, striking a delicate balance between fostering technological progress and safeguarding intellectual property rights.

Keywords: *intellectual property, general purpose AI systems, european AI law, law, moral rights, AI models.*

1. Introduction

Before taking any steps to analyze Regulation (EU) 2024/1689 laying down harmonized rules on AI (AI) 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 (*AI Regulation*)[1] from a copyright perspective, I consider it necessary to point out some historical milestones.

An early copyright law, considered one of the foundations of modern copyright law, was introduced by the *Statute* of Anne, adopted in 1710 in England. This statute was the first to recognize the rights of authors in their works and to establish a legal framework for the protection of these rights.

The year 1791 is another relevant historical point, as it was in this year that one of the first important theoretical works on copyright protection was written by August Wilhelm Rehberg, entitled „*On the relationship between writer and publisher as a contribution to the philosophy of law*". This work explores the rights of authors over their creations as well as their contractual relations with publishers, arguing that intellectual works are the author's personal property and should be protected against unauthorized exploitation, which at the time was an innovative idea for the protection of intangible rights.

The first office dedicated to copyright protection was established in the US and is known as the US Copyright Office. This office was established following the passage of the Copyright Act of 1790, which was the first US copyright law.

In 1956, Allen Newell, Herbert A. Simon and J.C. Shaw, a programme that can be considered an important step towards the development of AI, and it is considered the first practical attempt to implement AI concepts,

[*] PhD Candidate, Faculty of Law, „Nicolae Titulescu" University of Bucharest (e-mail: alexspopa@gmail.com).

[1] Regulation (EU) 2024/1689 laying down harmonized 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, adopted by the European Parliament and the Council of the European Union, OJ of 12.07.2024.

at least from the perspective of the ability of the logical structure to imitate the human logical process to solve certain tasks.

The accelerated evolution of AI technologies and, in particular, of general-purpose AI models has prompted the European Union to propose a harmonized regulatory framework to address the associated technological, social, economic and legal challenges. In this context, protecting copyright and ensuring a competitive, equitable and innovative environment become key objectives.

The basic pillar of AI regulation at the EU level is Regulation (EU) 2024/1689, and the following regulatory instruments for an AI framework are also relevant – European AI Strategy of April 2018 [COM(2018)237], AI for Europe [SWD(2018)137], White Paper on AI – A European Approach to Excellence and Trust (2020), EC Data Communication 2020 [COM(2020)66], Coordinated Plan on AI of April 2021, Action Plan for Digital Literacy 2021-2027 [COM(2020)0624].

Regulation (EU) 2024/1689, a reference document for the regulation of AI at Union level, introduces specific provisions covering the provision and use of general purpose models. The provisions of art. 53[2] are a key element, defining **the obligations of providers of AI designs towards transparency, respect for copyright and other intellectual property rights, including patent rights**.

The aim of this paper is to analyze and interpret the provisions of art. 53 in relation to copyright protection, highlighting the role of providers of AI models in enforcing these rights, as well as the challenges and opportunities that arise in this context. It also includes aspects of the legal, technical and ethical responsibilities associated with the integration of general-purpose AI models into the European digital ecosystem, analyzed from the perspective of doctoral research in the field of intellectual property.

[2] Art. 53 Regulation 2024/1689: (1) Providers of general purpose AI systems:

(a) establish and update the technical documentation of the design, including the process for its training and testing and the results of its evaluation, containing at least the information set out in Annex XI in order to provide it to the IA Office and the competent national authorities on request;

(b) Develop, update and make available information and documentation for AI system providers intending to integrate the general-purpose AI model into their AI systems. Without prejudice to the need to respect and protect intellectual property rights and confidential business information or trade secrets in accordance with Union and national law, the information and documentation shall be made available to the Member States, in accordance with Union and national law:

(i) Enable AI system providers to have a good understanding of the capabilities and limitations of the general-purpose AI model and to fulfil their obligations under this Regulation; and

(ii) Contain at least the elements set out in Annex XII;

(c) Implement a policy aiming at compliance with Union law on copyright and related rights and, in particular, the identification and enforcement, including on the basis of the state of the art, of a reservation of rights expressed pursuant to art. 4(3) of Directive (EU) 2019/790;

(d) Develop and make publicly available a sufficiently detailed summary of the content used to train the general-purpose AI model, in accordance with a template provided by the AI Office.

(2) The obligations set out in points (a) and (b) of para. (1) shall not apply to providers of AI models which are released under a free and open licence allowing access, use, modification and distribution of those models and whose parameters, including weights and information on the architecture of the models and their use, are made publicly available. This exception shall not apply to general-purpose AI models with systemic risk.

(3) Providers of general-purpose AI models shall, where appropriate, cooperate with the Commission and the competent national authorities in the exercise of their competences and prerogatives under this Regulation.

(4) Providers of general-purpose AI models may rely on codes of good practice within the meaning of art. 56 to demonstrate compliance with the obligations set out in para. (1) of this Article, pending the publication of a harmonized standard. Compliance with harmonized European harmonized standards shall give suppliers a presumption of conformity in so far as those standards cover those obligations. Providers of general-purpose AI models which do not adhere to an approved code of good practice or do not comply with a harmonized European standard shall demonstrate the existence of appropriate alternative means of compliance for assessment by the Commission.

(5) In order to facilitate compliance with Annex XI, in particular with points (d) and (e) of point 2, the Commission shall be empowered to adopt delegated acts in accordance with art. 97 to further detail the measurement and calculation methodologies so that the documentations are comparable and verifiable.

(6) The Commission shall be empowered to adopt delegated acts in accordance with art. 97(2) to amend Annexes XI and XII in the light of constant technological developments.

(7) Any information or documentation obtained pursuant to this Article, including business secrets, shall be treated in accordance with the obligations of confidentiality laid down in art. 78.

2. Legal framework applicable in the EU

2.1. General notions of European substantive law

European substantive law is the body of legal rules adopted at EU level, which sets out the essential rules guiding relations between natural and legal persons and the Member States. Through these rules, the European Union seeks to create a common legal area in which the law is applied uniformly and barriers between Member States are eliminated. In essence, European substantive law aims to protect fundamental rights, ensure the functioning of the internal market, protect consumers and harmonize rules and standards in key areas such as intellectual property.

An important aspect of this right is that it takes precedence over national law, and when a conflict arises between a national law and a European rule, the latter takes precedence. Also, certain European provisions can have direct effect, which means that they apply automatically and as such in the Member States without the need for transposition into national law, ensuring that the rules are applied quickly and uniformly throughout the Union

The areas covered by European substantive law are diverse and essential to the functioning of the Community bloc, with the protection of fundamental rights being one of the major concerns, guaranteed by the Charter of Fundamental Rights of the European Union. The internal market, in turn, is governed by rules that facilitate the free movement of goods, services, persons and capital. European competition law also prevents anti-competitive practices and ensures an open and fair market between all players. At the same time, intellectual property rights, including copyright, benefit from special protection to support innovation and protect the work of authors, helping them to capitalize on their creative output and avoid infringements.

The legal instruments through which the EU regulates these areas are varied, regulations, such as Regulation 2016/679[3] or Regulation 2024/1689, are directly applicable in all Member States without requiring transposition. The Regulation is binding „in all its elements”[4], *i.e.*, „binding as to the end purpose to be achieved and the form and means by which it is to be realized”[5]. Unlike regulations, directives set clear objectives but leave Member States free to decide how to achieve them, whereas decisions are binding only on those to whom they are directly addressed.

One area that deserves particular attention in this context is intellectual property. Copyright is protected by a number of European rules that have been adapted to the new digital realities. The Directive on the Harmonization of Copyright in the Information Society 2001/29/EC or the Directive on Copyright in the Digital Single Market 2019/790 are just some of them. Regulation (EU) 2024/1689 adds relevant new provisions, in particular with regard to the use of protected works by new technologies, such as general-purpose AI models.

2.2. Regulation (EU) 2024/1689: Objectives and scope

With the new AI regulation, EU aims to strike a balance between stimulating innovation, protecting the interests of authors, consumers and ensuring high standards of transparency and accountability[6].

Providers of general-purpose AI systems are companies or entities that develop AI models capable of performing a wide range of tasks in different domains. These models are trained on large datasets and are flexible enough to be integrated into other applications or systems, and are used for many purposes such as data analysis, process automation or even custom application development. Providers of these models must comply with European rules on safety, ethics and transparency to prevent potential risks or misuse.

In the light of the above, AI models are capable of performing a wide range of tasks, including in the form of code-based applications. From the perspective of Law no. 8/1996 on copyright and related rights,

[3] Regulation no. 679/2016 on the protection of individuals with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC (GDPR), adopted by the European Parliament and the Council of the European Union (OJ 119/04.05.2016) and rectified by the Corrigendum to Regulation (EU) 2016 (OJ 127/23.05.2018).

[4] Art. 288 TFEU, adopted by the Council of the European Communities (OJ 83/30.03.2010), as complemented by Council Decision no 199/2011 of the Council of the European Union amending Article 136 of the Treaty on the Functioning of the European Union with regard to a stability mechanism for Member States whose currency is the euro (OJ 91/06.04.2011).

[5] A. Fuerea, *Manualul Uniunii Europene*, 6th ed., revised and added, Universul Juridic Publishing House, Bucharest, 2016, p. 235.

[6] Art. 288 TFEU (ex art. 249 TEC): the Regulation is of general application. It is binding in its entirety and directly applicable in each Member State.

republished[7], the right holder, under certain conditions, can be the legal person, and by the definition provided by Regulation (EU)2024/1689 for providers of such systems, the natural person is omitted. In this case, the rule on the ownership of rights mainly concerns the legal person, even for the rights born in the patrimony of natural persons participating in the process of realizing the source codes underlying the IA models.

Under the provisions of Regulation (EU)2024/1689, the Provider of an AI system (including an AI model) is, in principle, the entity that develops or makes available the system under its own brand name or application. The rationale behind this provision is based on the fact that the EU is proposing a unitary approach for all products and services placed on the market, relying on the concept of „economic operator" rather than software author. In the case of a complex software product or digital service, such as an AI model or the end product in which it is embedded, EU legislation focuses on the entity that actually develops and/or produces the system, commercializes it or offers it for use, and in particular assumes responsibility for compliance with legal requirements (copyright, security, transparency and personal data protection).

By analogy with European regulations for tangible goods, where the manufacturer is the one who makes the product and launches it under its own brand, in the field of AI the „AI Model Provider" plays a similar role, being directly responsible for the compliance and safety of the system.

Within Regulation (EU) 2024/1689, art. 53 focuses on the specific obligations of providers of general-purpose AI models and is a central component of this body of regulation. From this perspective, the provisions of art. 53 cover aspects regulating the activity of AI model providers, including documenting the process of training and testing models and providing information to downstream providers who wish to integrate the model into their own systems.

At the same time, the protection of copyright and even related rights in the content used for training is regulated, including the development of proactive policies to respect intellectual property rights, in particular copyright (according to specific legislation) and the publication of summaries of the content used, with the aim of ensuring transparency towards right holders, interested persons and to facilitate the identification of potential infringements. From this perspective, any interested person who applies to a court in Romania may request that a provider of AI models be ordered to comply with the principle of transparency by invoking the direct effect of Regulation (EU) 2024/1689 in defense of his rights[8].

Obviously, one of the sectors targeted by the implementation of technologies operating AI systems is the public sector. In the administrative apparatus, especially in regulated sectors, the use of such systems raises two major questions: the first one is related to the training of AI models with data and information restrictively intended for the public, and the second one concerns issues of administrative fairness, such as the use of these AI systems for modelling principles of judicial fairness[9] and for elaborating the reasons for the decisions handed down

At national level was adopted Decision no. 832/2024 on the approval of the National Strategy for AI 2024-2027(SN-IA)[10], prior to the approval of Regulation (EU) 2024/1689 which established the objective of contributing to **Romania's strategy for the adoption of digital technologies in the economy and society in conditions of respect for human rights and promoting excellence and confidence in AI.** The development of the SN-IA responds to the need for specific regulation in areas such as data protection, consumer protection, protection of intellectual property, given the social risks generated by the impact that decisions based on AI can have on the quality of life and the need for human responsibility for these decisions.

3. Copyright and the challenges of general-purpose AI models

3.1. AI models: definitions and classifications

Starting from the definition that AI models are complex computational systems that mimic human cognitive processes such as learning, decision making and content generation, we can conclude that they are built on advanced algorithms and are trained using large amounts of data, with the ability to adapt and improve

[7] Law no. 8/1996 on copyright and related rights, republished in the Official Gazette of Romania no. 489/14.06.2018, with subsequent amendments and additions.

[8] A. Fuerea, *op. cit.*, p. 251.

[9] F. Martin-Bariteau, T. Scassa, *Artificial Intelligence and the Law in Canada*, LexisNexis, 2021, p. 11.

[10] GD no. 832/2024 on the approval of the National Strategy in the field of Artificial Intelligence 2024-2027, published in the Official Gazette of Romania no. 730/25.07.2024.

their performance as they gain experience. In the current technical and social context of technological advances and European regulations, these models are defined and categorized according to their purpose, the methods by which they are trained and their scope of application.

General-purpose models of AI, such as those used in applications like natural language processing or automatic content generation, are particularly relevant. They have the ability to perform tasks ranging from writing text and translations to analyzing images or creating music. These models are the most widely used, the best known being the one developed by Open AI. Unlike specialized models, which are designed to solve precise problems, general-purpose models can operate on many types of data and are used in a wide range of domains, including research, advertising and entertainment. Because of their versatility, they raise complex copyright issues because they frequently re-use existing content to generate new results.

These AI models can be trained in several ways. Some are trained on labelled data, allowing them to learn to recognize patterns and produce accurate outputs, such as identifying objects in an image. Others learn in an unsupervised way, by discovering hidden patterns in the incoming data or through rewards and penalties, adapting in-process to make optimal decisions. Beyond training techniques, AI models are present in fields ranging from natural language analysis and text processing to computer vision and economic forecasting. Natural language models, for example, are essential in the development of chatbots and machine translation systems, but their use requires access to large volumes of text, much of it copyrighted.

The current paradox is models specialized on training other models. This is the case of so-called machines that learn other machines without major human intervention[11].

The major difference between semi-autonomous and fully autonomous models is the degree of independence. Semi-autonomous models require human intervention in certain critical decisions, while fully autonomous models can operate without assistance and are used in complex processes, such as operating autonomous vehicles or performing certain tasks where human intervention would not be possible, such as certain emergency systems or automated safety systems. However, it is general-purpose AI models that are currently causing the most legal wrangling, as their ability to generate new content from existing material raises serious copyright issues. This ability to reuse existing works to create something new is exactly why they are at the centre of attention in current European regulations such as Regulation (EU) 2024/1689.

3.2. Types of protected works and their use by AI models

Copyrighted works are original creations of the human mind, and the law protects them to reward creativity and give authors control over their use. Law no. 8/1996, protects these works in Romania, laying down clear rules on their exploitation and respect for the recognised rights of the owners.

Under EU law, works protected by copyright include, similar to Romanian law, a wide range of works: literary texts, articles, musical works, films, photographs, visual works of art, names that can take the form of trademarks, software and many others. Any original material, be it a poem or a computer programme, can benefit from protection if it meets the requirement of originality. In the digital age, however, these works are increasingly processed and integrated into AI systems, raising complex legal and moral issues.

In my opinion, in the case of AI models, it is useful to distinguish between two categories of protected works: computer programs (including source and machine code or parts thereof) that are used to develop these models and generic copyrighted works, such as literary works, texts, studies and even computer programs that are used to train AI models.

AI models make use of these works at several stages in their development and operation. First, protected works are commonly used in the training phase of models. For example, a natural language processing model, such as the one used for automatic text generation, can learn from the millions of articles, books or other written materials available online. The ultimate goal is for the model to recognize complex grammatical structures, direct and hidden meanings (or even the spirit of the law) and writing styles, so that it can create new texts that look like human-written texts. However, many of these materials used are protected by copyright, and the central question is whether their use for training AI models constitutes an infringement of these rights.

[11] A. Koshiyama, E. Kazim, Ph. Treleaven, P. Rai, L. Szpruch, G. Pavey, G. Ahamat, F. Leutner *et al.*, *Towards algorithm auditing: managing legal, ethical and technological risks of AI*, *ML and associated algorithms*, 2024, at *https://royalsocietypublishing.org/doi/10.1098/rsos.230859*.

An interesting aspect is related to the implementation of tools using AI engines at the European Union level. This is the case of EUIPO, which is deploying an AI based system that is capable of performing semantic searches of goods and services, thus applying the rules laid down in Regulation (EU) 207/2009 on the Community trademark.[12]

Another example is models that generate images or visual works. These are trained using huge databases of images, some of which are copyrighted. For example, an AI model that can create realistic or abstract images can use famous paintings or photographs as a learning source. While the end result may be a completely new creation, it often bears stylistic traces or visual elements inspired by the original works, which opens up discussions about the derivative nature of the new creation and the rights of the owners of the original works.

Musical works are another sensitive category. AI models can analyze and reproduce rhythms, melodies or musical styles based on thousands of existing pieces. For example, AI models are able to generate new musical compositions that resemble the works of famous composers. Of course, these novel musical compositions are in most cases reworking of musical sequences originally used for training. In this case the question arises whether such compositions can be considered new original works, are derivative works, or whether they infringe copyright by intentionally or unintentionally reinterpreting existing works.

Since music applications based on AI have simplified both compositional and interpretive styles into predictable mathematical algorithms, in some cases these styles have been considered sufficient for copyright protection. Although the „style" as such is not fixed in a human-perceptible medium, being an attribute of the author, the seemingly opaque digital files that allow applications to display a digital image are protected by copyright. From this perspective, there is no obvious reason why source code capable of predictably generating a work-like result in a particular artistic style should not also be protected[13]. From this perspective, a new question arises as to who owns the copyright, in the sense of whether it belongs to the creator of the source code, the human operator who trained the AI model, or the author of the original work, if the result is regarded as a derivative work?

The main issue in all these cases is related to the way in which AI makes use of the protected works: **is this use merely a passive retrieval of the information necessary for learning, or is it an actual reproduction that infringes copyright.**

Current regulations, including those in Regulation (EU) 2024/1689, attempt to draw clear boundaries between what is allowed and what constitutes infringement. On the one hand, proponents of technology development argue that AI models need access to protected works in order to develop, and this should be considered a lawful use, similar to research or education. Copyright holders, on the other hand, fear that such uses could lead to financial losses and diminish the value of their works.

However, in the case of the provisions of Law no. 8/1996, the use of rights-bearing works generates certain obligations towards the right holder, and the exceptions provided by the law *„must be interpreted restrictively, and it is not possible to add new cases through case-law*".[14]

Certainly, the use of AI-based systems optimises professional modalities. By applying innovative methods, „predictions, recommendations or cheaper and more accurate decisions, AI promises to generate productivity gains, improve well-being and help address complex challenges. Exploiting AI requires complementary investments in data, skills and digitized workflows, as well as changes to organizational processes. As a result, adoption varies across businesses and sectors".[15]

In conclusion, the varied typologies of protected works – from text and music to images and software – are essential for the development of AI models, but their use entails legal risks and requires strict regulation. This debate is becoming one of the main topics of European legislation, as the balancing act between protecting copyright and encouraging technological innovation is more important than ever.

[12] G. Irimescu, *Inteligența Artificială și Protecția Mărcilor*, in Universul Juridic no. 12/2024.

[13] R. Abbot, *Research handbook on Intellectual property and Artificial intelligence*, Edward Elgar Publishing Limited, 2022, p. 76.

[14] V. Roș, D. Bogdan, O. Spineanu-Matei, *Dreptul de autor și drepturile conexe. Tratat*, All Beck Publishing House, Bucharest, 2005, p. 311.

[15] Organisation for Economic Cooperation and Development, Artificial Intelligence in Society, OECD Publishing, Paris, 2019, at *https://doi.org/10.1787/eedfee77-en*. See also online Organisation for Economic Cooperation and Development, Recommendation of the Council on Artificial Intelligence (21.05.2019), OECD/LEGAL/0449, s. I, OECD Legal Instruments, at *https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449*.

3.3. Key legal issues on the protection of works

The protection of works of intellectual creation in the EU is governed by a well-established legal framework, which aims to give authors control over the use and distribution of their works[16]. At the heart of this system is the principle that authors have exclusive rights in their works and that any use of their works by third parties requires the prior consent of the right holders.

However, simulated simultaneously with technological developments, particularly in the field of AI, these rights are being challenged, as AI models may use protected works without this process being transparent, legitimate or easily controllable.

One of the most important legal issues related to the protection of works is the **reproduction right**, which gives authors the exclusive right to authorize or prohibit the copying of all or part of their work. The protection offered by the legal provisions does not distinguish *„between partial and complete reproduction of the work. In either case, reproduction by a third party is lawful only with the consent of the author*"[17]. Of course, in some cases, the material or work used in the training process is reproduced in a different way from the original, but even in this case the author's consent is required.

In the case of AI models this right becomes problematic because, in the process of training, AI systems may copy or store fragments of protected works in order to learn whole structures and patterns, such as a language model that is trained using copyrighted books or articles, in which case the question arises whether the mere use of these texts for training purposes constitutes an illegal reproduction or whether it can be considered a permissible exception. The answer can only be that **this type of exploitation is not an exception to the need for the right holder's consent.**

Another key issue relates to the **right of communication to the public**, which applies when a protected work is made available to the public through digital technologies. AI models that generate content, such as text, images or music, can create outputs that are then widely distributed, sometimes without compensating the copyright holders of the original material.

Copyright exceptions and limitations are another important dimension of the protection of works in this context. In European legislation[18], there are certain situations in which the use of protected works is allowed without the consent of the author, such as use for educational or research purposes. Providers of AI models often try to fall within these exceptions, arguing that training their models is a necessary act for innovation and technological progress.

3.4. Key legal aspects of moral rights

Another critical point is related to the **moral rights of authors**, which give them the right to be recognized as the creators of the work and to object to any modification or use that might damage the reputation or integrity of the work. In the case of AI models, there is a risk that protected works may be modified or combined in ways that may distort the author's original intention. Such an approach is in line with the Berne Convention for the Protection of Literary and Artistic Works[19], which provides in art. 6 bis para. (1): *„Independently of economic rights of authorship and even after the transfer of such rights, the author shall retain the right to claim authorship of the work and to object to any distortion, mutilation or other modification of the work or any other prejudice to its honour or reputation*".

As AI technologies continue to evolve, another important legal issue is the **due diligence obligations of AI modelling providers**. Under the new European regulations, AI model providers must demonstrate that they have taken adequate measures to prevent unauthorised use of protected works in their processes, and this may

[16] Directive 2001/29/EC on the harmonisation of certain aspects of copyright and related rights in the information society, as consolidated from 10.02.2025, is based on the publication in the OJ 167/22.06.2001.

[17] V. Roș, D. Bogdan, O. Spineanu-Matei, *Dreptul de autor și drepturile conexe. Tratat*, *op. cit.*, p. 255.

[18] Directive 2001/29/EC of the European Parliament and of the Council of 22.05.2001 on the harmonization of certain aspects of copyright and related rights in the information society, art. 6.

[19] Berne Convention for the Protection of Literary and Artistic Works of 09.09.1886, supplemented at Paris on 04.05.1896, revised at Berlin on 13.11.1908, supplemented at Berne on 20.03.1914, revised at Rome on 02.06.1928, revised at Brussels on 26.06.1948, revised at Stockholm on 14.07.1967 and at Paris on 24.07.1971 and amended on 28.09.1979, in force since 31.07.1969, consolidated since 10.02.2025, based on the republication (r1) in OJ 156/17.04.1998.

include implementing automated mechanisms to verify licences or limit access to protected data or using databases exclusively to identify possible infringements.

3.5. Technical documentation and transparency in the development of AI models

The provisions of art. 53(1)(a) of Regulation 2024/1689 introduce an essential requirement for providers of general-purpose AI technologies[20], requiring them to develop and keep up-to-date **detailed technical documentation on the process of training, testing and evaluating their models**. This obligation is a cornerstone in the regulation and monitoring of the use of these technologies, providing both safeguards for copyright holders and an accountability framework for providers of AI models.

On the one hand, technical documentation plays a central role in ensuring the traceability of the processes underpinning the development and operation of AI models. Through it, competent authorities and third parties, including copyright holders, can verify that providers comply with legal requirements and rules on data protection and intellectual property rights. For example, this documentation details the data sources used, the data processing methods, as well as the specific algorithms involved in training the model. This makes it possible to identify possible irregularities, such as the unauthorised inclusion of copyrighted works or the misuse of datasets. Therefore, traceability not only facilitates compliance, but also creates a solid basis for holding providers accountable to the authorities or in litigation.

On the other hand, this technical requirement also has an important preventive function. Through the transparency that the documentation requires, potential copyright infringements can be identified at an early stage. In the details provided on how datasets are collected and processed, right holders or authorities can analyse whether the necessary consent has been obtained for the use of protected works or whether there are exceptions allowing their use in a lawful manner. In the absence of such documentation, AI models could function as opaque and unverifiable 'black boxes', increasing the risk of abuse. Technical documentation therefore becomes an essential tool in preventing legal conflicts by providing clarity on the technical and legal processes involved.

4. Copyright protection in the AI context: Interpretation of art. 53

4.1. Use of protected works and permitted exceptions

The use of copyrighted works in the development of AI is one of the most discussed topics in the EU legal sphere because AI models rely on access to large volumes of data to learn and improve their performance, and this data often includes protected works such as books, articles, images, music or films.

The key question that arises is whether such uses, made without the consent of the authors, can be considered lawful and, if so, under what conditions. In order to provide clarity in this respect, EU legislation has established several exceptions, and art. 53 and 78 respectively of Regulation (EU) 2024/1689 seek to draw precise boundaries between permitted use and copyright infringement.

A major exception applies when works are used for scientific or educational research purposes. In this case it is recognised that access to diverse information is essential for technological progress. An AI model that analyses texts to learn how to write articles can use literary works, provided that this use is limited to research needs. However, this exception is not unlimited, and if the model uses too large a volume of works, or if the research is merely a front for commercial purposes, the exception is no longer applicable and the use is illegal.

Another important exception concerns the temporary reproduction of works as part of technological processes necessary for the functioning of the AI system. EU law allows such acts of temporary reproduction when they are indispensable to the technical process, such as the extraction of grammatical structures or statistical data, without preserving the content of the work in its original form. Such use is considered lawful if it is limited to technical purposes.

[20] Providers of general-purpose AI technologies are legal entities or natural persons that develop, own, make available or commercialize AI models capable of performing a wide range of functions, without being restricted to a specific purpose. These versatile technologies are used for diverse applications such as natural language processing, image recognition, content generation, data analysis and autonomous decision-making. Vendors are responsible for ensuring that their designs comply with the applicable legal framework, including copyright regulations, data protection and the ethics of AI usage.

Central to the exceptions is the protection of the public interest. EU law permits the use of protected works in situations which contribute to technological development for the general benefit of society, but only on condition that such use does not adversely affect the commercial exploitation of the work and does not prejudice the legitimate interests of the authors. The principle of balance between the public interest and the protection of authors is essential.

While the exceptions provide some freedom in the use of protected works, they do not exempt providers of AI models from the obligation to respect copyright, as they have to demonstrate that the use of the works is within the limits imposed by law.

4.2. Rights and obligations of AI model providers

In the EU, providers of AI models play a key role in the development of emerging technologies, but this freedom to innovate comes with a number of clearly defined legal responsibilities. Art. 53 of Regulation (EU) 2024/1689 sets out a legal framework that allows providers to access copyrighted works, but only under well-regulated conditions. The main objective is to strike a balance between access to data, necessary for training AI models, and the protection of creators' rights.

One right enjoyed by providers of AI models is the possibility to use protected works during the processes of training and testing the models, of course under certain conditions. This is essential as AI models learn from massive amounts of data in order to be able to generate relevant results, from literary texts and academic articles to images and musical works, all of which can be sources of learning for them. However, the use of these materials is not free and unconditional. European law provides that this right must be exercised only within the strict limits imposed by law, respecting the principle of proportionality – *i.e.*, using only as much as is necessary for the correct functioning of the model.

A major obligation imposed on providers of AI models is to verify that the works used are lawfully accessed. Before training their models, providers must obtain the consent of the rights holders or ensure that their use falls within the exceptions allowed by law. For example, if the works are used for scientific research or educational purposes, this may be considered lawful under Directive 2001/29/EC. But if the real purpose is commercial, providers risk breaking the law, which may lead to damages or other penalties.

Moreover, providers are also responsible for implementing technical measures to prevent unauthorized reproduction of protected works. In practice, this means installing filters or automated systems to monitor processed data and prevent the use of unlicensed material. To exemplify, if an AI model is capable of generating written content, it must be configured in such a way that it does not reproduce entire passages of protected works. The implementation of these measures is essential, as many AI models can generate derivative content that, without safeguards, risks copyright infringement.

Transparency is another crucial obligation for providers. They must be open about how their models use protected works and provide clear information about the data sources used, processing methods and safeguards in place. Transparency is not only a legal requirement but also a means to build public trust in these technologies.

„In this context, the transparency and explain ability of algorithms are considered essential to understand how decisions are made and to give citizens, users and operators the possibility to challenge"[21] *their decisions, regarding certain violations or even unanticipated interference.*

Co-operation with rights holders is also essential. If an author complains that his or her work has been used without permission, the AI provider has an obligation to investigate the situation and take steps to remedy the problem. This may include removing protected content from the design database or negotiating a retroactive licence for its use. Thus, the law requires a flexible but responsible approach where providers must be willing to correct any mistakes to avoid legal conflicts.

4.3. Analysing the principles of proportionality and balance

The principles of proportionality and balance are central to the European regulation of the use of copyrighted works in the context of the development of AI. They are not mere theoretical concepts, but legal

[21] Max Craglia, digital, *Artificial Intelligence: A European Perspective*, Luxembourg Publications Office of the European Union, 2018, pp. 63.

mechanisms that are intended to create harmony between the interests of the different parties involved: content creators, AI model providers and the general public. In the absence of these principles, the use of works in processes for training AI models would risk being either overly restrictive, blocking innovation, or too permissive, severely affecting the economic rights of authors.

The principle of proportionality is a key element, as it establishes that the use of protected works must be limited to what is strictly necessary to fulfil the legitimate aim pursued. If an AI model uses literary works to learn complex linguistic structures or to analyse writing styles, proportionality requires that this use be minimal and justified only for technological development, and the application of this principle requires a concrete assessment of how the protected works are used.

The **principle of balance**, on the other hand, is closely linked to proportionality, but goes beyond limiting use. It seeks to protect the economic and moral interests of authors without blocking technological progress. Thus, the balance must be struck between protecting creative rights and ensuring that AI models can be developed and used for the general benefit of society.

The practice of developing AI models has shown that finding the right balance is not always easy, as the interests of the parties are often conflicting, with authors wanting maximum protection and fair compensation, while AI model providers want to have free access to as much data as possible to develop their models.

5. Challenges and critical issues in the implementation of art. 53 – Final issues

The application of the obligations imposed by art. 53(1)(c) of Regulation 2024/1689, in conjunction with national legislation such as art. 10 and 13 of Law no. 8/1996, poses significant challenges for both the providers of AI technologies and the authorities monitoring compliance. One of the biggest difficulties is striking a balance between the requirement for transparency and the protection of commercial secrets or sensitive information.

Providers of AI models are often reluctant to disclose full details of the datasets or methods used to train the models, fearing that doing so could compromise competitive advantages or expose their technologies to risks of unauthorized use or replication. However, without an adequate level of transparency, authorities and stakeholders are unable to assess whether AI models respect copyright, complicating compliance with legal obligations.

Another major challenge is to establish effective and universally accepted methodologies for differentiating between copyrighted content and content in the public domain or available under permitted licenses, such as open-source licenses.

In the context of art. 13 of Law no. 8/1996, which protects authors' economic rights in their works, providers must demonstrate that the works used for training AI models comply with this legal framework. However, the data used to train the models are often extracted from diverse sources and aggregated on a large scale, making it difficult to trace and identify the legal status of each work. In the absence of clear and effective technological tools, the risk of unintentional inclusion of protected works increases considerably, harming both content creators and the reputation of providers.

The effectiveness of technologies for identifying and marking protected content is another complex issue. While techniques such as hashing, watermarking or content recognition algorithms can help prevent unauthorized use, they are not without limitations. There is a risk of „false positives", where content is wrongly categorized as protected, thereby blocking legitimate access to works or preventing the use of data that is actually in the public domain. Marking technologies may also have difficulties in identifying complex or derivative works, where the boundary between original and adaptation is often hard to draw. These limitations emphasise the need to develop more advanced technological solutions that are both accurate and affordable.

In addition, the highly dynamic AI market complicates the regulatory process. Rapid technological progress may outstrip authorities' capacity to adapt the regulatory framework to new realities. This mismatch between innovation and regulation can lead to situations where legislation becomes outdated, affecting both copyright enforcement and the competitiveness of the sector covered by the provisions of Regulation 2024/1689. Regular updating of the legislation thus becomes a necessity, but this process is often slow and complicated by the diverging interests of the parties involved.

In conclusion, although copyright enforcement regulations in the context of AI are essential to ensure a fair digital ecosystem, their enforcement is not without challenges, and in this context, the main question remains whether „the current liability and security framework provides adequate mechanisms adequate

mechanisms to deal with products and services, and if not what changes are needed"[22] for an adequate and dynamic level of security.

6. Conclusions

This study provides an analysis of art. 53 of Regulation (EU) 2024/1689, which significantly advances European standards for copyright protection in the context of general-purpose AI models. The key outcomes of the research indicate that the Regulation introduces essential obligations for AI model providers, emphasizing transparency, accountability, and proportionality. Specifically, providers must maintain comprehensive technical documentation, ensure lawful use of copyrighted materials, and disclose their data sources transparently to uphold the economic and moral rights of authors.

The expected impact of these research outcomes includes greater legal clarity and enhanced protection for intellectual property rights, facilitating a balanced coexistence between technological innovation and copyright safeguards. These findings may guide policymakers, legal practitioners, and AI developers in implementing and adhering to best practices and regulatory compliance.

For future research, it is suggested to explore further technological mechanisms to improve the identification and management of copyrighted content within large AI training datasets. Additional studies could also focus on the comparative analysis of similar regulatory frameworks outside the European Union, providing insights into global best practices and enhancing international collaboration in AI governance and intellectual property rights protection.

References

- Abbot, R., *Research handbook on Intellectual property and Artificial intelligence*, Edward Elgar Publishing Limited, 2022;
- Berne Convention for the Protection of Literary and Artistic Works of 09.09.1886, supplemented at Paris on 04.05.1896, revised at Berlin on 13.11.1908, supplemented at Berne on 20.03.1914, revised at Rome on 02.06.1928, revised at Brussels on 26.06.1948, revised at Stockholm on 14.07.1967 and at Paris on 24.07.1971 and amended on 28.09.1979, in force since 31.07.1969, consolidated since 10.02.2025, based on the republication (r1) in the OJ 156/17.04.1998;
- Craglia, Max, digital, *Artificial Intelligence: A European Perspective,* Luxembourg Publications Office of the European Union, 2018;
- Directive 2001/29/EC of the European Parliament and of the Council of 22.05.2001 on the harmonization of certain aspects of copyright and related rights in the information society; current consolidated version: 06.06.2019, at *https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02001L0029-20190606*;
- Fuerea, A., *Manualul Uniunii Europene*, 6th ed., revised and added, Universul Juridic Publishing House, Bucharest, 2016;
- Irimescu, G., *Inteligența Artificială și Protecția Mărcilor*, in Universul Juridic no. 12/2024;
- Koshiyama, A., Kazim, E., Treleaven, Ph., Rai, P., Szpruch, L., Pavey, G., Ahamat, G., Leutner, F. *et al.*, *Towards algorithm auditing: managing legal, ethical and technological risks of AI*, *ML and associated algorithms*, 2024, at *https://royalsocietypublishing.org/doi/10.1098/rsos.230859*;
- Law no. 8/1996 on copyright and related rights, republished in the Official Gazette of Romania no. 489/14.06.2018, with subsequent amendments and additions;
- Martin-Bariteau, F., Scassa, T., *Artificial Intelligence and the Law in Canada*, LexisNexis, 2021;
- Regulation no. 1689/2024 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, OJ of 12.07.2024;
- Regulation no. 2016/679 of the European Parliament and of the Council of 27.04.2016 on the protection of individuals with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC (GDPR), at *https://eur-lex.europa.eu/legal-content/RO/TXT/PDF/?uri=CELEX:32016R0679*;
- Roș, V., Bogdan, D., Spineanu-Matei, O., *Dreptul de autor și drepturile conexe. Tratat*, All Beck Publishing House, Bucharest, 2005.

[22] *Ibidem*.