



СЕКЦІЯ 3

АДМІНІСТРАТИВНЕ ПРАВО І ПРОЦЕС; ФІНАНСОВЕ ПРАВО; ІНФОРМАЦІЙНЕ ПРАВО

УДК [004.8:347.732]:347.447.5

DOI <https://doi.org/10.32999/ksu2307-8049/2025-4-3>

RETHINKING ONLINE INTERMEDIARY LIABILITY IN THE AGE OF GENERATIVE AI: FROM “SAFE HARBOR” TO A MODEL OF ALGORITHMIC ACCOUNTABILITY

Shamov Oleksii Anatoliiovych,
Intelligent systems researcher,
Head of the Human Rights Educational Guild
yursprava@gmail.com
orcid.org/0009-0009-5001-0526

The **purpose** of the article is to prove the conceptual and functional inadequacy of the “safe harbor” doctrine for regulating the liability of generative artificial intelligence (AI) providers for copyright infringement and to substantiate the need for developing a special (*sui generis*) liability model. **Methods.** The research is based on a combination of general scientific and special legal methods. Doctrinal analysis was used to study the legal norms governing the liability of online intermediaries in the US (MCA) and the EU (E-commerce Directive, DSA, CDSM Directive). The comparative legal method was used to contrast the American and European approaches to intermediary immunity. The system-structural method allowed for the deconstruction of the generative AI content creation process into stages (data curation, model design, generation) and to define the provider’s active role in each. The formal-logical method was applied to identify contradictions between the premises of the “safe harbor” doctrine and the technological reality of generative AI. **Results.** It has been established that generative AI providers play an active, defining, and co-creative role in the content creation process, which fundamentally distinguishes them from the passive intermediaries for whose protection the “safe harbor” regimes were created. It is demonstrated that the “notice-and-takedown” mechanism is ineffective for the dynamic nature of AI-generated content. A hypothesis is proposed about the need to replace the binary “liable/exempt from liability” framework with a *sui generis* model based on the principles of algorithmic accountability and tiered duties of care. **Conclusions.** It is concluded that the direct application of existing “safe harbor” doctrines to generative AI is a categorical error. The proposed new liability model shifts the focus from reactive response to proactive risk management, imposing a duty on developers to implement technical and organizational measures to prevent infringements. For Ukraine, in the context of European integration, this creates an opportunity to “leapfrog” over outdated stages of regulation and immediately implement an advanced model. Prospects for further research lie in developing specific technical standards for implementing tiered duties of care and in analyzing the economic consequences of introducing the proposed liability model.

Keywords: generative AI, copyright, intermediary liability, safe harbor, algorithmic accountability.

ПЕРЕОСМИСЛЕННЯ ВІДПОВІДАЛЬНОСТІ ОНЛАЙН-ПОСЕРЕДНИКІВ В ЕПОХУ ГЕНЕРАТИВНОГО ШІ: ВІД «БЕЗПЕЧНОЇ ГАВАНІ» ДО МОДЕЛІ АЛГОРИТМІЧНОЇ ПІДЗВІТНОСТІ

Шамов Олексій Анатолійович,
дослідник інтелектуальних систем,
голова ГО “Просвітня фундація з прав людини”
yursprava@gmail.com
orcid.org/0009-0009-5001-0526

Мета статті полягає у доведенні концептуальної та функціональної непридатності доктрини «безпечної гавані» (*safe harbor*) для регулювання відповідальності провайдерів генеративного штучного інтелекту (ШІ) за порушення авторських прав та в обґрунтуванні необхідності розробки спеціальної (*sui generis*) моделі відповідальності. **Методи.** Дослідження ґрунтується на сукупності загальнонаукових та спеціальних юридичних методів. Доктринальний аналіз застосовано для вивчення правових норм, що регулюють відповідальність онлайн-посередників у США (DMCA) та ЄС (Директива про електронну комерцію, DSA, Директива CDSM). Порівняльно-правовий метод використано для зіставлення американського та європейського підходів до імунітету посе-

редників. Системно-структурний метод дозволив деконструювати процес створення контенту генеративним ШІ на етапи (курація даних, дизайн моделі, генерація) та визначити активну роль провайдера на кожному з них. Формально-логічний метод застосовано для виявлення суперечностей між передумовами доктрини «безпечної гавані» та технологічною реальністю генеративного ШІ. **Результати.** Встановлено, що провайдери генеративного ШІ відіграють активну, визначальну та співтворчу роль у процесі створення контенту, що фундаментально відрізняє їх від пасивних посередників, для захисту яких були створені режими «безпечної гавані». Продемонстровано, що механізм «повідомлення та видалення» є неефективним для динамічної природи П-контенту. Запропоновано гіпотезу про необхідність заміни бінарної рамки «відповідальний/ звільнений від відповідальності» на *sui generis* модель, що ґрунтується на принципах алгоритмічної підзвітності та диференційованих обов'язків щодо обачності. **Висновки.** Зроблено висновок, що пряме застосування існуючих доктрин «безпечної гавані» до генеративного ШІ є категоріальною помилкою. Запропонована нова модель відповідальності зміщує фокус з реактивного реагування на проактивне управління ризиками, покладаючи на розробників обов'язок впроваджувати технічні та організаційні заходи для запобігання порушенням. Для України, в контексті євроінтеграції, це створює можливість «перестрибнути» через застарілі етапи регулювання та одразу впровадити передову модель. Перспективи подальших досліджень полягають у розробці конкретних технічних стандартів для реалізації диференційованих обов'язків обачності та в аналізі економічних наслідків впровадження запропонованої моделі відповідальності.

Ключові слова: генеративний ШІ, авторське право, відповідальність посередників, безпечна гавань, алгоритмічна підзвітність.

Introduction. The emergence of powerful generative artificial intelligence (AI) models capable of creating new content has triggered a doctrinal crisis in copyright law, particularly concerning the liability of online intermediaries. Established “safe harbor” doctrines, designed for the internet era of the late 1990s, have proven conceptually unfit for the new technological paradigm.

The contemporary academic discourse is actively exploring this issue, yet often focuses either on analyzing specific aspects, such as fair use (Murray, 2023), or on a general critique of copyright’s inadequacy as a regulatory tool for AI (Mantegna,

2024), without offering a comprehensive alternative liability model.

The central thesis of this article is that the active, defining, and co-creative role of generative AI providers in the content creation process makes them fundamentally different from the “neutral, purely technical and passive” intermediaries for whose protection the “safe harbor” regimes were created. This distinction renders the direct application of existing doctrines not only legally incorrect but also practically impossible. The attempt to apply a legal shield designed for a passive entity to an active one is a categorical error that leads to the collapse of the entire doctrinal structure.

The purpose of this article is to prove the inadequacy of the “safe harbor” doctrine for regulating the liability of generative AI providers and to substantiate the need for a *sui generis* liability model. The research objectives are: 1) to analyze the immunity architecture within the U.S. Digital Millennium Copyright Act (DMCA) and the European E-commerce Directive; 2) to deconstruct the active role of the generative AI provider in the content creation process; 3) to demonstrate the unsuitability of existing frameworks in the U.S. and the EU; 4) to propose a new liability model based on the principles of algorithmic accountability and tiered duties of care; and 5) to consider the implications of this model for legal reform in Ukraine.

Analysis of recent research and publications.

The academic and legal debate is largely a response to the framework established by foundational legislation, most notably the U.S. Digital Millennium Cop-

yright Act (DMCA, 1998). Section 512 of the DMCA created the original “safe harbor” provisions, establishing the “notice-and-takedown” system for online service providers. This legislation, while not a recent publication, represents the doctrinal starting point for any analysis of intermediary liability and has been the subject of continuous scholarly critique and judicial interpretation, forming the essential backdrop against which the challenges of generative AI are now being measured.

The problem of liability for copyright infringements related to the activities of generative AI is a subject of active scholarly debate. A significant portion of research focuses on the application of the fair use doctrine to the training process of AI models. For instance, M. Murray (Murray, 2023) analyzes the roles of various participants in the process - from dataset creators to end-users-through the lens of the four fair use factors, shifting the focus of liability onto the end-user. Other authors, such as M. Mantegna (Mantegna, 2024), put forward the radical thesis that copyright law is fundamentally an inappropriate tool for solving the problems created by AI and call for the search for alternative methods of compensating authors.

A more moderate approach is offered by Lin, Y., & Guan, T. (2025), who develop the concept of an “AI Harbor”, which is a modernization of the existing DMCA “safe harbor.” Their model envisages differentiated duties for different participants in the AI “supply chain” (data suppliers, model developers, service operators), linking immunity to the fulfillment of specific technical and organizational requirements. Similar ideas about distributed liability are expressed by Chen (Chen, 2025), who proposes applying the principle of no-fault liability to providers of generative AI services. Even before the explosive spread of generative AI, scholars such as N. Elkin-Koren and M. Perel (Perel & Elkin-Koren, 2015) laid the foundations for analyzing liability through the concept of “algorithmic accountability”, arguing for the need for transparency and due process in algorithmic copyright enforcement.

This aligns with broader calls for accountability in automated systems, such as those articulated by N. Diakopoulos (Diakopoulos, 2016), who has ex-



plored the need for journalistic and public scrutiny of computational power structures. Together, these works underscore a shift from focusing on specific infringing acts to examining the responsibilities inherent in designing and deploying algorithmic systems.

Meanwhile, European researchers are analyzing the problem through the prism of the interaction of new legislative acts. Quintais, J. P., & Schwemer, S. F. (2022) examine the complex interplay between the Digital Services Act (DSA) and the special provisions of Article 17 of the Directive on Copyright in the Digital Single Market (CDSM), which has already abolished the “safe harbor” for a certain type of active platform.

Despite the significant number of publications, the problem of the fundamental conceptual incompatibility of the very logic of “safe harbor,” based on the “active/passive” intermediary dichotomy, with the technological reality of generative AI remains unresolved. Most proposals attempt to “repair” or adapt the existing framework, whereas this article argues for the necessity of its complete rejection in favor of a fundamentally new model based on proactive risk management rather than reactive responses to infringement notices.

Materials and methods. The methodological basis of the research is a system of philosophical-worldview, general scientific, and special-legal methods that ensure the objectivity and completeness of the analysis. The dialectical method of cognition allowed for the examination of the problem of generative AI provider liability in its development and interconnection with technological, economic, and social factors.

Central to the research is the doctrinal (formal-legal) method, which was applied to analyze and interpret the key provisions of U.S. and EU legislation regulating the liability of online intermediaries. Specifically, Section 512 of the U.S. Digital Millennium Copyright Act (DMCA), Articles 12-14 of Directive 2000/31/EC on electronic commerce, the relevant provisions of the Digital Services Act (DSA), and Article 17 of Directive (EU) 2019/790 on copyright in the Digital Single Market (CDSM) were analyzed.

The comparative-legal method was used to contrast the American “safe harbor” model, based on the fair use doctrine and the “notice-and-takedown” mechanism, with the European model, which has evolved from passive hosting to special liability regimes for active platforms.

The system-structural method allowed for the deconstruction of the generative AI content creation process into three key stages: 1) curation and ingestion of training data; 2) design, training, and fine-tuning of the model; 3) content generation (inference). This approach made it possible to identify and legally qualify the active, controlling role of the provider at each stage, which is key to refuting the analogy with passive hosting.

Based on the synthesis of the obtained results, using modeling and forecasting methods, a hypothesis was formulated about the need to create a sui generis liability model, and its conceptual framework was developed, based on the principles of algorithmic accountability and tiered duties of care.

1. A legislative compromise

The “safe harbor” regimes, introduced in the late 1990s in both the U.S. (DMCA) and the EU (E-commerce Directive), were the result of a legislative compromise. Their goal was to balance the interests

of rights holders and the needs of innovative services for protection against endless lawsuits for the actions of their users. The basis of both regimes is the “notice-and-takedown” mechanism, which allows rights holders to notify a provider of a specific infringement, and the provider to react to preserve its immunity.

Key to our argument is that the very architecture of these regimes presupposes a passive role for the provider. The U.S. DMA requires that the material be stored “at the direction of a person other than the service provider” (17 U.S.C. § 512(c)). The European E-commerce Directive is even more direct, stating that the liability exemption applies only to services that play a “neutral, purely technical and passive role” (Directive 2000/31/EC, Recital 42). This legal architecture was built on the notion of a provider as a neutral conduit or repository for content created and uploaded exclusively by third parties.

Generative AI shatters this fundamental assumption. The provider is not a passive host but an active participant-in fact, a co-creator - in the generation of the final product. To demonstrate this, a legally significant technical analysis of the content creation process is necessary. At the training data curation stage, the provider makes active decisions regarding the collection, selection, and processing of vast data arrays, which often include millions of copyrighted works. This is not passive storage but a fundamental act of building the model’s “knowledge base”, which is the basis for direct copyright infringement lawsuits, such as in the cases of *The New York Times v. OpenAI* and *Getty Images v. Stability AI*.

At the model design and training stage, the provider develops the architecture, chooses the training methodology, and makes decisions that directly affect the model’s behavior, particularly its tendency to “memorize” and verbatim reproduce fragments of training data. These architectural decisions are a form of direct control over the method of creating new content.

Finally, at the content generation (inference) stage, the provider’s system receives a request (prompt) from the user, which is essentially an idea, and executes the entire process of creating the expression. The system uses its trained weights and algorithms fully developed and controlled by the provider to synthesize a new work. The user specifies what (the idea, but the provider’s system controls how (the expression)). It is because of this active role that the U.S. Copyright Office has compared the AI user to a client who gives general instructions to an artist, where the artist is the AI system itself.

Thus, the origin of the content, the provider’s actions, and the degree of its control are radically different from the traditional hosting model. While a hosting provider passively stores a ready-made file created by the user, a generative AI provider actively synthesizes new content in response to a request. This difference makes the application of “safe harbor” standards impossible. The “actual knowledge” standard of infringement, which underlies the “notice-and-takedown” mechanism, loses its meaning, as the provider cannot have prior knowledge of a dynamic, yet-to-be-generated output.

2. The way out of a doctrinal crisis

The evolution of European law itself points to a way out of this doctrinal crisis. Article 17 of the CDSM Directive has already abolished the “safe harbor” for “online content-sharing service providers” (OCSSPs) that play an active role in communicating works to

the public, establishing a form of direct liability for them. Generative AI providers are an even more vivid example of an “active” platform.

Therefore, creating a special liability regime for them is not a radical break with the European legal tradition, but a logical next step on the path defined by Article 17 (Quintais, J. P., & Schwemer, S. F., 2022).

Instead of trying to adapt the outdated binary “safe harbor” model, this article proposes a fundamentally new approach—a graduated liability framework that reflects the degree of the provider’s control. This model is based on two pillars: algorithmic accountability and tiered duties of care. The principle of algorithmic accountability shifts the focus from “knowledge” of specific infringements to the provider’s duty to exercise care in the design, training, and deployment of the AT system (Perel & Elkin-Koren, 2015). Liability arises from the failure to take reasonable measures to mitigate foreseeable risks.

This principle is operationalized through tiered duties of care, which can be structured according to the stages of the AI lifecycle, developing ideas proposed in academic literature (Lin, Y., & Guan, T., 2025). At the data curation stage, this could be a duty to ensure the legality of training data. At the model development stage, a duty to implement technical measures to reduce risks, such as preventing “memorization” and implementing watermarking. At the deployment stage, a duty to implement dynamic output filtering and robust complaint mechanisms.

Compliance with these duties would function as a defense or a mitigating factor for liability, changing the paradigm from “are you liable?” to “what is the extent of your liability given the reasonableness of your risk mitigation efforts?”.

Conclusions. The analysis conducted demonstrates the conceptual failure of

the passive intermediary doctrine when applied to active providers of generative AI. The fundamental assumptions underlying the “safe harbor” regimes in the U.S. and the EU are destroyed by the active role of providers in selecting training data, designing models, and directly synthesizing content. Consequently, the key “notice-and-takedown” mechanism becomes functionally obsolete and ineffective.

Instead of attempting to artificially extend the boundaries of outdated doctrines, a hypothesis is proposed for the creation of a sui generis legal framework based on the principles of algorithmic accountability and tiered duties of care. This model shifts the focus from reactive remediation to proactive risk management, placing responsibility on AI developers and operators to implement reasonable technical and organizational measures to prevent copyright infringements.

For Ukraine, which is on the path of European integration under the EU Association Agreement, this challenge presents a unique opportunity. Current Ukrainian legislation on intermediary liability is based on a model similar to the E-commerce Directive. However, the European legal framework that Ukraine must implement is fragmented and in a state of transformation. This creates a unique opportunity for Ukraine to “leapfrog” over outdated stages of legal development. Instead of blindly copying the existing, fragmented acquis, Ukraine can immediately implement an advanced, legally consistent, and technologically adequate regulatory model. Such a step would not only fulfill harmoniza-

tion obligations but also create a significant competitive advantage, turning the country into an attractive jurisdiction for the development of responsible artificial intelligence

Prospects for further research in this area lie in the development of specific technical standards and best practices for the implementation of each of the tiered duties of care, as well as in conducting an economic analysis of the consequences of implementing the proposed liability model for the innovative development of the AI industry.

BIBLIOGRAPHY:

1. Chen, Y. Liability Principles for Copyright Infringement of Generative Artificial Intelligence. *The Frontiers of Society, Science and Technology*. 2025. Vol.7. Iss.1. URL: <https://doi.org/10.25236/fsst.2025.070111>
2. Diakopoulos, N. Accountability in algorithmic decision making. *Communications of the ACM*. 2016. Vol. 59. Iss.2. pp.56–62. URL: <https://doi.org/10.1145/2844110>
3. U.S. Copyright Office. *The digital millennium copyright act of 1998*. 1998. URL: <https://www.copyright.gov/legislation/dmca.pdf>
4. The European Parliament. Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market (Directive on electronic commerce). *Official Journal of the European Communities*. 2000. Iss.178. pp.1-16. URL: <http://data.europa.eu/eli/dir/2000/31/oj>
5. The European Parliament. Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC. *Official Journal of the European Union*. 2019. L.130, pp.92-125. URL: <http://data.europa.eu/eli/dir/2019/790/oj>
6. Mantegna, M. ARTificial: Why Copyright Is Not the Right Policy Tool to Deal with Generative AI. *Web-page of Yale Law Journal*. 2024. URL: <https://www.yalelawjournal.org/forum/artificial-why-copyright-is-not-the-right-policy-tool-to-deal-with-generative-ai>
7. Quintais, J. P., & Schwemer, S. F. The Interplay between the Digital Services Act and Sector Regulation: How Special Is Copyright? *European Journal of Risk Regulation*. 2022. URL: <https://doi.org/10.1017/err.2022.1>
8. Murray, M. D. Generative AI Art: Copyright Infringement and Fair Use. *SMU Science and Technology Law Review*. 2023. Vol. 26. Iss.2. URL: <https://doi.org/10.25172/smustr.26.2.4>
9. Perel (Filmar), M., & Elkin-Koren, N. Accountability in Algorithmic Enforcement: Lessons from Copyright Enforcement by Online Intermediaries. *SSRN Electronic Journal*. 2015. URL: <https://doi.org/10.2139/ssrn.2607910>
10. The European Parliament. Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act). *Official Journal of the European Union*. 2022. L.277. pp.1-102. URL: <https://eur-lex.europa.eu/eli/reg/2022/2065/oj/eng>
11. Lin, Y., & Guan, T. From safe harbours to AI harbours: reimagining DMCA immunity for the generative AI era. *Journal of Intellectual Property Law & Practice*. 2025. URL: <https://doi.org/10.1093/jiplp/jpaf043>

REFERENCES:

1. Chen, Y. (2025). Liability Principles for Copyright Infringement of Generative Artificial Intelligence. *The Frontiers of Society, Science and Technology*. Volume. 7. Issue. 1. Retrieved from: <https://doi.org/10.25236/fsst.2025.070111>
2. Diakopoulos, N. (2016). Accountability in algorithmic decision making. *Communications of the ACM*. Volume. 59. Issue. 2. P.56–62. Retrieved from: <https://doi.org/10.1145/2844110>
3. U.S. Copyright Office. (1998). *The digital millennium copyright act of 1998*. Retrieved from: <https://www.copyright.gov/legislation/dmca.pdf>



4. The European Parliament. (2000). Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market (Directive on electronic commerce). *Official Journal of the European Communities*. Issue. 178. P.1-16. Retrieved from: <http://data.europa.eu/eli/dir/2000/31/oj>

5. The European Parliament. (2019). Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC. *Official Journal of the European Union*. L.130. P.92-125. Retrieved from: <http://data.europa.eu/eli/dir/2019/790/oj>

6. Mantegna, M. (2024). ARTificial: Why Copyright Is Not the Right Policy Tool to Deal with Generative AI. *Web-page of Yale Law Journal*. Retrieved from: <https://www.yalelawjournal.org/forum/artificial-why-copyright-is-not-the-right-policy-tool-to-deal-with-generative-ai>

7. Quintais, J. P., & Schwemer, S. F. (2022). The Interplay between the Digital Services Act and Sector Regulation: How Special Is Copyright? *European Journal of Risk Regulation*. Retrieved from: <https://doi.org/10.1017/err.2022.1>

8. Murray, M. D. (2023). Generative AI Art: Copyright Infringement and Fair Use. *SMU Science and Technology Law Review*. Volume. 26. Issue. 2. Retrieved from: <https://doi.org/10.25172/smustr.26.2.4>

9. Perel (Filmar), M., & Elkin-Koren, N. (2015). Accountability in Algorithmic Enforcement: Lessons from Copyright Enforcement by Online Intermediaries. *SSRN Electronic Journal*. Retrieved from: <https://doi.org/10.2139/ssrn.2607910>

10. The European Parliament. (2022). Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act). *Official Journal of the European Union*. L.277. P.1-102. Retrieved from: <https://eur-lex.europa.eu/eli/reg/2022/2065/oj/eng>

11. Lin, Y., & Guan, T. (2025). From safe harbours to AI harbours: reimagining DMCA immunity for the generative AI era. *Journal of Intellectual Property Law & Practice*. Retrieved from: <https://doi.org/10.1093/jiplp/jpaf043>

Стаття надійшла до редакції 30.10.2025

Рекомендована до друку 17.11.2025

Опублікована 28.11.2025