

# EU Text And Data Mining Exception and US Fair Use Doctrine: Regulating Copyright Liability in AI Systems

**Hiroka Watanabe Yasui**

IE University, Segovia, Spain.  
Dual Degree of Bachelors of Law and Bachelors in International Relations.

E-mail: [hwatanabe.ieu2024@student.ie.edu](mailto:hwatanabe.ieu2024@student.ie.edu)

Published January 2026

Editor: Johnny Lin, Georgetown University

## Abstract

Artificial intelligence challenges the existing copyright legal framework. Because its training inherently depends on the reproduction of protected data, it raises the question of whether using such data for such purposes and subsequent model development constitutes a lawful action or an infringement. This question finds itself in a fine line: while the European Union treats text-and-data mining as a lawful exception under specific circumstances, the United States regulates it under the fair use doctrine. This paper aims to compare and analyse both approaches to examine their effectiveness in regulating copyright infringement in AI systems.

Keywords: text and data mining, fair use doctrine, AI systems, copyright law

## 1. Introduction

The rapid development of artificial intelligence (AI) is challenging the current legal framework, especially the laws regulating the use of protected material. Under copyright law, authors are granted exclusive rights, including the right of reproduction, distribution, display, performance, and creation of derivative works for a limited time.<sup>1</sup> However, with the rise of AI, questions such as how and when these rights may be applicable, and how they can be effectively protected, especially in its training stage, have arisen. This

is because, in order for AI systems to perform their tasks, they must be trained to learn, reason, and act, which inherently involves the use and reproduction of vast amounts of material, many of which are protected by copyright. For the purposes of this research, the use of protected material refers to the creative form of data, rather than the data itself, such as the facts, ideas, procedures, etc.<sup>2</sup>

<sup>1</sup> U.S. Copyright Office, 'What Is Copyright?' (U.S. Copyright Office 2025) <<https://www.copyright.gov/what-is-copyright/>>.

<sup>2</sup> Mariia Afanasieva, 'Text and Data Mining in Copyright Law' (2025) 55.

Despite the fact that copyrighted material is vulnerable to unlawful use, it is particularly during the training stage that the limit between a lawful use and a copyright infringement is blurred.<sup>3</sup> Even if this distinction were determined properly, the subsequent question would be: who should be held liable for such infringement?

As this area remains relatively new, the legal framework differs across jurisdictions in terms of regulation and how they address specific challenges. While some countries have codified statutory exceptions, other countries depend on judicial doctrines within existing legislation.<sup>4</sup> This paper aims to examine and compare the approaches of the European Union (EU), which has adopted text and data mining exceptions, and the United States, which depends on its fair use doctrine, to further assess how each jurisdiction effectively addresses copyright in AI training.

## 2. Legal Framework: TDM v Fair Use Doctrine

### 2.1. *Text and Data Mining in the European Union*

Despite the absence of one specific legislation regulating copyright law in the field of artificial intelligence (AI), the Directive 2019/790 on Copyright in the Digital Single Market (DSM Directive) is relevant to AI training

---

<sup>3</sup> Thomas Margoni and Martin Kretschmer, 'A Deeper Look into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology' (2022) 71 GRUR International.

<sup>4</sup> Eleonora Rosati, 'Copyright Exceptions and Fair Use Defences for AI Training Done for "Research" and "Learning," or the Inescapable Licensing Horizon' [2025] European Journal of Risk Regulation 1.

within the European Union (EU). Through this directive, the EU has acknowledged the fact that technological advancement 'transforms the way works and other subject matter are created, produced, distributed and exploited'.<sup>5</sup> It can be argued that, in fact, such an adoption was necessary, since the pre-existing legal copyright framework relied mainly on the Copyright and Information Society Directive 2001/29/EC (InfoSoc Directive), which is uncertain when it comes to regulating the digital and cross-border uses of protected material.

To address this challenge, the DSM Directive modernised, expanded, and harmonised EU copyright law by introducing and adapting rules designed to protect rightholders, while also fostering innovation in the digital sphere. Among its key innovations, a specific copyright exception for text and data mining (TDM) was introduced in Articles 3 and 4.

Before expanding on each article, it is essential to explain what TDM consists of. Article 2(2) of the DSM Directive defines it as 'any automated analytical technique aimed at analysing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations'.<sup>6</sup> This process inherently involves the act of reproducing, extracting, or both, protected materials, a right normally reserved to the authors under InfoSoc Directive.<sup>7</sup>

---

<sup>5</sup> 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 [2019] OJ L130/92, recital 3.

<sup>6</sup> DSM Directive (n 3) art 2(2).

<sup>7</sup> Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects

In the context of AI, TDM is undeniably necessary for its training and development since it relies on processing vast quantities of data and content to enable the system to ‘learn’. In fact, this act entails, among other things, copying protected and unprotected material which may constitute an infringement.<sup>8</sup>

The DSM Directive therefore aims to balance between innovation and safeguarding rightholders through exceptions for which TDM may be permissible under specific circumstances. On the one hand, Article 3 provides an exception for TDM conducted by research organizations and cultural heritage organizations, for the purposes of research subject to lawful access, either by contractual agreement, or by free access to the public. On the other hand, Article 4 broadens the scope of exceptions and limitations for TDM to all, and for all purposes insofar as it is necessary and accessed lawfully. Moreover, Article 4(3) introduces what is commonly referred to as an opt-out mechanism, in which rightholders may express to reserve their rights in ‘an appropriate manner’.<sup>9</sup>

## 2.2. *Fair Use Doctrine in the United States*

Unlike the EU, which has legislated specific exceptions for TDM, the United States (US) has no particular statutory framework regulating copyright neither in the digital environment nor in AI systems training. Instead, it relies on a more flexible approach through a judicial interpretation of the fair use doctrine, codified in Section

---

of copyright and related rights in the information society [2001] OJ L167/10, art 2.; DSM Directive (n 3) recital 8.

<sup>8</sup> Afanasieva n (2).

<sup>9</sup> DSM Directive (3) art 4(3).

107 of the Copyright Act. This implies that, traditionally, whether an act constitutes a copyright infringement is determined by the courts through the application of the fair use doctrine.

In principle, rightholders enjoy exclusive rights granted by the law, particularly the right of reproduction. Yet, Section 107 limits those rights by permitting the use of protected material for purposes including, but not exclusively to, ‘criticism, comment, news reporting, teaching, scholarship, or research’ subject as to its fair use.<sup>10</sup> In practice, the fair use doctrine serves as an affirmative defence.<sup>11</sup> Courts assess and evaluate case by case guided by the four factor test consisting of: (1) purpose and character of use, (2) nature of copyright work, (3) amount and substantiality of portion in relation to the whole work, and (4) its effect on the potential market.<sup>12</sup>

Nevertheless, despite the fact that the section itself is not specific to the digital sphere, and even lesser so to AI systems, courts extended the application of said doctrine. This doctrine, however, ‘fluctuates’ in interpretation.<sup>13</sup> Accordingly, different instances have demonstrated that because such application is broad, it cannot automatically

---

<sup>10</sup> 17 USC § 107.

<sup>11</sup> Celeste Shen, ‘Fair Use, Licensing, and Authors’ Rights in the Age of Generative AI’ (Northwestern Pritzker School of Law Scholarly Commons November 2024) <<https://scholarlycommons.law.northwestern.edu/njtip/vol22/iss1/4/>>.

<sup>12</sup> 17 USC § 107.

<sup>13</sup> Jenny Quang, ‘Does Training AI Violate Copyright Law?’ (2021) 36 Berkeley Technology Law Journal <<https://btlj.org/wp-content/uploads/2023/02/0003-36-4Quang.pdf>>.

be applied to AI systems.<sup>14</sup> For instance, the debate of the fair use defense is centered on whether it is transformative or not, meaning whether such use “adds value to the original [...] transformed in the creation of new information, new aesthetics, new insights and understandings”.<sup>15</sup> While in some cases the transformative use is deemed as a fair use of copyrighted material, some argue that the mere focus on this nature risks overlooking the rest of elements to be scrutinised.<sup>16</sup> However, as mentioned earlier, since the US lacks a specific legislation governing these systems, courts continue to rely and apply the fair use doctrine for copyright issues, guided through the four factor test.<sup>17</sup>

### 3. Comparative Analysis

It can be argued that the main difference between both jurisdictions is its starting point. Within the EU, because of exceptions stipulated, if an act does not meet the conditions, it would be deemed as an infringement. In contrast, in the US, not all acts are qualified as infringing, insofar as there is fair use of protected material. However, it must be noted that the mere existence of a particular regulation does not constitute being effective in regulating. In practice, most of the cases arising from copyright issues in training AI systems are in the US; in fact, around

mid-2025, there were over forty pending cases before the courts.<sup>18</sup> Meanwhile, the EU has yet to establish its own interpretation, with a pending ruling of the Court of Justice of the European Union in case C-250/25.<sup>19</sup> Thus, any assessment of the effectiveness in regulating copyright must be examined through judicial interpretation. For the EU, cases remain pending before the CJEU. For example, Hungary has sought judicial interpretation and guidance of the provisions set forth in the directives, such as the TDM exception with regards to training AI chatbots, for the case *Like Company v Google Ireland Ltd.*<sup>20</sup> Nevertheless, there are some Member States that have taken a step further and interpreted these provisions. In the German case of *LAION v. Robert Kneschke*, the Hamburg Regional Court ruled that the exceptions for TDM in Article 3 of the DSM can cover AI training, provided that some training datasets can fall under the scope of scientific research.<sup>21</sup> This marks the first landmark case where TDM exceptions have been clarified, especially in AI systems,

<sup>14</sup> *Ibid.*

<sup>15</sup> Pierre N Leval, ‘Toward a Fair Use Standard’ (1990) 103 Harvard Law Review 1105; Quang n (12).

<sup>16</sup> Quang n (12).

<sup>17</sup> Carlos Rivadulla, ‘Training AI with Copyrighted Material: USA vs Europe. Match Point’ (ECIJA 17 October 2025) <<https://www.ecija.com/en/news-and-insights/training-ai-with-copyrighted-material-usa-vs-europe-match-point/>> accessed 29 October 2025.

<sup>18</sup> Stuart D Levi and Mana Ghaemmaghami, ‘Copyright Office Weighs in on AI Training and Fair Use’ (Skadden 15 May 2025) <<https://www.skadden.com/insights/publications/2025/05/copyright-office-report>>.

<sup>19</sup> Nicola Lucchi, ‘Generative AI and Copyright Training, Creation, Regulation Policy Department for Justice, Civil Liberties and Institutional Affairs Directorate-General for Citizens’ Rights’ (2025) <[https://www.europarl.europa.eu/RegData/etudes/STUD/2025/774095/IUST\\_STU\(2025\)774095\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2025/774095/IUST_STU(2025)774095_EN.pdf)>.

<sup>20</sup> Fredericka Argent, ‘CJEU Receives Questions on Copyright Rules Applying to AI Chatbot’ (Covington 4 June 2025) <<https://www.insideprivacy.com/artificial-intelligence/cjeu-receives-questions-on-copyright-rules-applying-to-ai-chatbot/>> accessed 19 November 2025.

<sup>21</sup> *Kneschke v LAION e.V.* (Hamburg Regional Court, 2024).

even though its limits and implementation are still uncertain.<sup>22</sup>

On the other hand, in the US, despite lacking a specific regulation on AI systems, courts have opted to broaden the scope of fair use. For instance, in *Google LLC v. Oracle America, Inc.*, the Supreme Court reversed the Federal Circuit Appeals Court and ruled that Google's copying code lines of Oracle America was within fair use.<sup>23</sup> With this decision, the Court reaffirmed and broadened its basic purpose of promoting progress, as courts have struck down the balance between protecting rightholders' exclusive rights and enabling innovation.<sup>24</sup>

However, recent cases have demonstrated a possible limit to such flexibility. In *Thomson Reuters Enterprise Centre GmbH v ROSS Intelligence Inc*, the District Court for the District of Delaware rejected the fair use doctrine as a defence mechanism.<sup>25</sup> In this case, where the subject matter was an infringement in the training of an AI system, the court dismissed ROSS's affirmative defense that making copies of Bulk Memos for training was a fair use, especially since ROSS intended to act as a competitor,

and it could have had an effect on the market.<sup>26</sup> Hence, despite the fact that the said doctrine serves as a way to progress, such uses in AI systems are still being tested and remain uncertain.

#### 4. Policy Recommendation

Whether exceptions are laid down explicitly in legislation or not, it can be argued that the existing legal frameworks are unable to cover AI systems entirely.<sup>27</sup> This lies in the fact that there is a difference in nature between AI and traditional, human authorship. The law has not and cannot foresee all of the abilities of AI technologies, many of which surpass human capabilities. This does not justify the inability of the current system to regulate and guide rising copyright issues in AI systems. Rather, as the technologies develop and evolve, so should the legal framework.

The four fact test which serves as a guidance in the US may be insufficient in addressing the complexities of AI systems, especially in more technical stages, such as its training. Despite the fact that current courts issued their judgement based on this doctrine, this does not suggest that, for future cases, it may be sufficient. There is a necessity to redefine the boundaries of a copyright infringement and the limitations of fair use, more

---

<sup>22</sup> Maryna Manteghi, 'Kneschke vs. LAION: Opening Fresh Perspectives on AI Training and TDM Exceptions' [2025] European Law Blog.

<sup>23</sup> *Google LLC v Oracle America Inc* 593 US 1 (2021).

<sup>24</sup> United States Copyright Office, 'Copyright and Artificial Intelligence Part 3: Generative AI Training Pre-Publication Version' (2025) <<https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf>>; Rivadulla n (16).

<sup>25</sup> *Thomson Reuters Enterprise Centre GmbH v ROSS Intelligence Inc* Civil Action No 20-613-LPS (D Del, 21 March 2024).

---

<sup>26</sup>Devika Kornbacher, 'The Boundaries of Playing' (Clifford Chance 2025) <<https://www.cliffordchance.com/insights/resources/blogs/talking-tech/en/articles/2025/03/the-boundaries-of-playing-fair-when-training-ai.html>> accessed 17 November 2025.

<sup>27</sup> Eleonora Rosati, 'Copyright Exceptions and Fair Use Defences for AI Training Done for "Research" and "Learning," or the Inescapable Licensing Horizon' [2025] European Journal of Risk Regulation 1.

specifically, determining where the infringement begins, and what mechanisms the law can provide for a lawful use, such as through authorization and licenses.<sup>28</sup>

Because the US legal system is characterized by being flexible and having a clear balance between the legislative and judicial power, an entirely new statute specific on copyright law for AI systems would not be adequate within the US Code. The nature of the US legal framework is industry-oriented and, therefore, lacks “formal mandatory regulations” of the usage of data in training.<sup>29</sup> Nevertheless, as the presence of AI becomes more dominant in the legal field, instead of continuing to rely on the existing legal guidelines for AI, a federal legislation governing AI systems as a whole might be essential to serve as the basis for all branches of law concerning these systems.

As for the EU, since it has a rights-oriented approach, it has historically been prompt in responding to issues arising from the context in which we are living.<sup>30</sup> It is therefore inevitable that, in theory, the EU’s legal framework is structured, and, most of the time, offers predictability. In practice, even though Articles 3 and 4 of the DSM

---

<sup>28</sup>Mohit Porwal and Krupa Vyas, ‘Copyright Implications in Training Artificial Intelligence (AI) Models | Article | Chambers and Partners’ (Chambers and Partners2025) <<https://chambers.com/articles/copyright-implications-in-training-artificial-intelligence-ai-models>>.

<sup>29</sup> Kaigeng Li, Hong Wu and Yupeng Dong, ‘Copyright Protection during the Training Stage of Generative AI: Industry-Oriented U.S. Law, Rights-Oriented EU Law, and Fair Remuneration Rights for Generative AI Training under the UN’s International Governance Regime for AI’ (2024) 55 Computer Law & Security Review 106056 <<https://www.sciencedirect.com/science/article/abs/pii/S0267364924001225>>.

<sup>30</sup> *Ibid.*

Directive explicitly establish the exceptions for TDM, it does not encompass the whole complexity of AI systems, especially its training. Moreover, the concepts defined in the directive may need to be clarified to be directly applicable to AI systems and have a clear binding result, of which way of application would ultimately be up to the discretion of each Member State.

As for Article 4, it allows for an opt-out mechanism for rightholders as an alternative method to the requirement of consent; yet, it can be incompatible with how AI systems work. Due to the fact that training these systems relies on massive amounts of data, some authors argue that it is no different to the “prior consent” model, as it would still induce rightholders to invoke their right to reserve, and therefore can still create conflict between the use of protected material and their rights, leading to, among other things, inefficiency.<sup>31</sup>

In contrast to the US, the EU does have a particular set of regulations about AI that must be harmonized with other existing ones in order to offer clear guidance on what is being regulated. Offering predictability is always gratifying insofar as it can be applicable to the legal issues presented. Since the AI systems cannot be entirely foreseen, it is crucial that its legal framework continues to adapt and evolve as these systems do.

Finally, since it is inevitable that copyright issues may arise beyond its territorial boundaries, it is of utmost importance to consider a collaborative approach between

---

<sup>31</sup> *Ibid.*

countries. Both the rightholders' and the markets' perspectives must be taken into consideration to ensure effectiveness and equity. While some jurisdictions offer more practical approaches, others may provide predictability. It is thus necessary to coordinate and establish a 'foundational framework' guided by shared principles and the universal objective of copyright law: a balance between the protection of rightholders while incentivizing for innovation and progress.<sup>32</sup> Some may be reluctant for a licensing agreement and a fixed compensation mechanism as it may disrupt the market. Yet, if there cannot be a consensus on how to equally balance, this could be a short-term solution that bridges the gap. It would allow for the diffusion of innovation while fairly rewarding rightholders, especially when it involves larger scales of TDM. Eventually, a harmonized, principle-guided framework could strengthen universal copyright law and still support innovation and current authors.

## 5. Conclusion

AI training not only involves access to the material, but it also involves reproduction, extraction, and mining of the text or data discovered.<sup>33</sup> This raises the question of

whether it may constitute an infringement, as such a right is reserved for rightholders. However, the legal framework differs depending on the jurisdiction: both the EU and the US have their own approaches to regulating copyright law in this context.. As discussed throughout the paper, under copyright law, the EU has a specific directive for the digital single market, which includes AI systems, whereas the US uses its traditional doctrine of fair use.

Moreover, their judicial interpretation may not be the same, but it rhymes: both aim to balance between granting exclusive rights and enhancing innovation. The US approach is that, by granting authors rights, it incentivizes their work to be more accessible to the public for progress. The EU, however, concentrates in balancing the protection of rightholders' rights while also promoting innovation as per the DSM Directive. Nevertheless, whether the objectives posed by the EU are achieved is a matter of how the pending cases will be ruled. Thus, there is a necessity to address properly the current and upcoming challenges the legal sphere will face as AI systems continue to progress.

---

<sup>32</sup>Mohit Porwal and Krupa Vyas, 'Copyright Implications in Training Artificial Intelligence (AI) Models | Article | Chambers and Partners' (Chambers and Partners2025) <<https://chambers.com/articles/copyright-implications-in-training-artificial-intelligence-ai-models>>.

<sup>33</sup> Kacper Szkalej, 'The Paradox of Lawful Text and Data Mining? Some Experiences from the Research Sector and Where We (Should) Go from Here' (2025) 74 GRUR International <<http://dx.doi.org/10.1093/grurint/ikaf029>> accessed 12 November 2025.

## Bibliography

- Argent F, 'CJEU Receives Questions on Copyright Rules Applying to AI Chatbot' (*Covington* 4 June 2025) <<https://www.insideprivacy.com/artificial-intelligence/cjeu-receives-questions-on-copyright-rules-applying-to-ai-chatbot/>> accessed 19 November 2025
- Coraggio G, 'AI & Copyright - US Fair Use vs EU TDM after Delaware Decision' (*GamingTechLaw* 18 March 2025) <<https://www.gamingtechlaw.com/2025/03/ai-copyright-us-fair-use-vs-eu-tdm-after-the-delaware-decision/>> accessed 29 October 2025
- Cornell Law School, '17 U.S. Code § 107 - Limitations on Exclusive Rights: Fair Use' (*LII / Legal Information Institute* 2019) <<https://www.law.cornell.edu/uscode/text/17/107>>
- Geiger C, Frosio G and Bulayenko O, 'The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market - Legal Aspects' [2018] SSRN Electronic Journal
- Guadamuz A, 'The EU's Artificial Intelligence Act and Copyright' (2024) 28 *The Journal of World Intellectual Property*
- Katarina Hallkvist, 'US vs EU: Diverging Paths on AI Copyright Law - Eris Law Advokatbyrå' (*Eris Law Advokatbyrå* 2025) <<https://erislaw.se/artiklar/us-vs-eu-diverging-paths-on-ai-copyright-law/>> accessed 29 October 2025
- Kaufman R, 'Copyright's Big Win in the First Decided US Artificial Intelligence Case - the Scholarly Kitchen' (*The Scholarly Kitchen* 20 February 2025) <<https://scholarlykitchen.sspnet.org/2025/02/20/copyrights-big-win-in-the-first-decided-us-artificial-intelligence-case/>>
- Kornbacher D, 'The Boundaries of Playing' (*Clifford Chance* 2025) <<https://www.cliffordchance.com/insights/resources/blogs/talking-tech/en/articles/2025/03/the-boundaries-of-playing-fair-when-training-ai.html>> accessed 17 November 2025
- Leval PN, 'Toward a Fair Use Standard' (1990) 103 *Harvard Law Review* 1105
- Levi SD and Ghaemmaghami M, 'Copyright Office Weighs in on AI Training and Fair Use' (*Skadden* 15 May 2025) <<https://www.skadden.com/insights/publications/2025/05/copyright-office-report>>
- Li K, Wu H and Dong Y, 'Copyright Protection during the Training Stage of Generative AI: Industry-Oriented U.S. Law, Rights-Oriented EU Law, and Fair Remuneration Rights for Generative AI Training under the UN's International Governance Regime for AI' (2024) 55 *Computer Law & Security Review* 106056

- <<https://www.sciencedirect.com/science/article/abs/pii/S0267364924001225>>
- Lucchi N, ‘Generative AI and Copyright Training, Creation, Regulation Policy Department for Justice, Civil Liberties and Institutional Affairs Directorate-General for Citizens’ Rights’ (2025) <[https://www.europarl.europa.eu/RegData/etudes/STUD/2025/774095/IUST\\_STU\(2025\)774095\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2025/774095/IUST_STU(2025)774095_EN.pdf)>
- Manteghi M, ‘Kneschke vs. LAION: Opening Fresh Perspectives on AI Training and TDM Exceptions’ [2025] European Law Blog
- Mărginean M, ‘Copyright Infringement & AI: A Case Study of Authors Guild v. OpenAI and Microsoft’ (2024) <[https://www.4ipcouncil.com/application/files/7517/2189/4919/Copyright\\_Infringement\\_and\\_AI\\_A\\_Case\\_Study\\_of\\_Authors\\_Guild\\_v\\_OpenAI\\_and\\_Microsoft.pdf](https://www.4ipcouncil.com/application/files/7517/2189/4919/Copyright_Infringement_and_AI_A_Case_Study_of_Authors_Guild_v_OpenAI_and_Microsoft.pdf)>
- Margoni T and Kretschmer M, ‘A Deeper Look into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology’ (2022) 71 GRUR International
- Porwal M and Vyas K, ‘Copyright Implications in Training Artificial Intelligence (AI) Models | Article | Chambers and Partners’ (*Chambers and Partners*2025)
- <<https://chambers.com/articles/copyright-implications-in-training-artificial-intelligence-ai-models>>
- Quang J, ‘Does Training AI Violate Copyright Law?’ (2021) 36 Berkeley Technology Law Journal <<https://btlj.org/wp-content/uploads/2023/02/0003-36-4Quang.pdf>>
- Rivadulla C, ‘Training AI with Copyrighted Material: USA vs Europe. Match Point’ (*ECIJA* 17 October 2025) <<https://www.ecija.com/en/news-and-insights/training-ai-with-copyrighted-material-usa-vs-europe-match-point/>> accessed 29 October 2025
- Rosati E, ‘Copyright Exceptions and Fair Use Defences for AI Training Done for “Research” and “Learning,” or the Inescapable Licensing Horizon’ [2025] European Journal of Risk Regulation 1
- , ‘TDM Exceptions (Not Just the Three-Step Test) Don’t Allow All Unlicensed AI Development’ (*The IPKat*13 October 2025) <<https://ipkitten.blogspot.com/2025/10/tdm-exceptions-not-just-three-step-test.html>>
- Senftleben M, ‘Compliance of National TDM Rules with International Copyright Law: An Overrated Nonissue?’ (2022) 53 International Review of Intellectual Property and Competition Law 1477
- Shen C, ‘Fair Use, Licensing, and Authors’ Rights in the Age of Generative AI’ (*Northwestern Pritzker School of Law Scholarly Commons*November 2024)

<<https://scholarlycommons.law.northwestern.edu/njtip/vol22/iss1/4/>>

(*Berkeley Library* 3 October 2025)

<<https://guides.lib.berkeley.edu/c.php?g=491766&p=7828496>> accessed 29 October 2025

Szkalej K, 'The Paradox of Lawful Text and Data Mining?

Some Experiences from the Research Sector and Where We (Should) Go from Here' (2025) 74 GRUR International

<<http://dx.doi.org/10.1093/grurint/ikaf029>>

accessed 12 November 2025

Zirpoli CT, 'Generative Artificial Intelligence and Copyright Law' (*Congress.gov*2025)

<<https://www.congress.gov/crs-product/LSB10922>>

Talhoff O, 'Uncharted Territory? Opportunities,

Boundaries and Uncertainties of Text and Data Mining' (*Passle*30 August 2023)

<<https://technologyquotient.freshfields.com/post/102imtz/uncharted-territory-opportunities-boundaries-and-uncertainties-of-text-and-data>>

Tyagi K, 'Copyright, Text & Data Mining and the

Innovation Dimension of Generative AI' (2024) 19

Journal of Intellectual Property Law & Practice

U.S. Copyright Office, 'What Is Copyright?' (*U.S.*

*Copyright Office*2025)

<<https://www.copyright.gov/what-is-copyright/>>

United States Copyright Office, 'Copyright and Artificial

Intelligence Part 3: Generative AI Training Pre-Publication Version' (2025)

<<https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf>>

University of California Berkeley Library, 'Library Guides:

Text Mining and AI Research Resources: Home'