Singapore Journal of Legal Studies
[2025] 271–291

A COMPETITION POLICY ANALYSIS OF COPYRIGHT PROTECTION IN GENERATIVE AI

Giuseppe **Colangelo***

The rise of artificial intelligence (AI) has sparked significant debate, particularly regarding the relationship between generative AI (GenAI) and copyright. Indeed, GenAI appears to challenge every layer of copyright protection. Our analysis focuses on the tensions surrounding the use of copyrighted works to train AI models. Since AI training relies on vast amounts of data, two conflicting interests emerge. On one hand, copyright can act as a major barrier to entry, potentially stifling the next wave of technological innovation. On the other hand, GenAI systems may pose an existential threat to creative industries by replicating human creativity and producing literary and artistic works faster and at lower costs. Against this backdrop, policymakers worldwide are striving to balance these seemingly opposing interests. While most discussions focus on why and how copyright holders should be compensated, this article examines when compensation is appropriate. To this end, it advocates for a competition-based approach in assessing the application of copyright limitations and exceptions. Specifically, it argues that antitrust tools can help courts and policymakers determine when creators suffer commercial harm and when AI-generated content may be considered a substitute for human creations.

I. Introduction

If striking the right balance between access and incentives has always been the central challenge in copyright law,[1] the emergence of new and potentially disruptive technologies has made this task more complex over time. Balancing the public interest in reaping the benefits of innovation with the need to safeguard private investment incentives against the risk of free riding has indeed become increasingly challenging. As a result, questions and doubts continue to arise about the extent of copyright protection needed to effectively encourage innovation.

---

* Associate Professor of Law and Economics, University of Basilicata; Adjunct Professor of Markets, Regulations and Law, Luiss; TTLF Fellow, Stanford Law School and University of Vienna; <https://orcid.org/0000-0002-0089-3545; giuseppe.colangelo@unibas.it>. The study has been conducted as part of the research activities promoted by the International Center for Law & Economics (ICLE). ICLE has received financial support from numerous companies, foundations, and individuals, including firms with interests both supportive of and in opposition to the ideas expressed in this and other ICLE-supported works. ICLE support is in the form of unrestricted, general support. The ideas expressed here are the author's own and do not necessarily reflect the views of ICLE's advisors, affiliates, or supporters.

1 William M Landes & Richard A Posner, "An Economic Analysis of Copyright Law" (1989) 18 J Leg Stud 325 at 326.

Historically, the introduction of new technologies has often led to the creation of additional exclusive rights, compensation mechanisms, and specific limitations or exceptions. However, the rise of the artificial intelligence ("AI"), and in particular of the generative AI ("GenAI"), era is regarded as a true game changer, necessitating a comprehensive rethinking and overhaul of the entire copyright system rather than quick, incremental adjustments.[2] Indeed, the emergence of GenAI as the next evolution of AI — sometimes referred to as creative AI due to its ability to produce content rather than simply analyse or manipulate existing data for decisions or predictions — poses challenges to every layer of traditional copyright protection. These challenges range from the concept of authorship[3] and the requirement of originality[4] to the idea-expression dichotomy[5] and the scope of both existing and new exceptions.[6]

Unsurprisingly, this has sparked a wave of litigation and academic discussion. Key frictions between GenAI and copyright concern the use of copyrighted works to train AI models, the copyrightability of material generated by AI systems, potential liability for infringing works produced by AI, and the treatment of GenAI outputs that mimic the identity or style of human artists.

Industrial policy concerns and geopolitical implications further intensify the pressure and significance of the ongoing debate. On the one side, the expected impact of AI innovation on economic competitiveness and national security is prompting countries to develop strategies that prioritise efforts securing technological leadership.[7] In this scenario, since the initial learning process of an AI model (also known as a foundation model) relies on vast amounts of data for training and fine-tuning, copyright plays a crucial role, serving as a gatekeeper and a significant barrier to entry in the AI sector, potentially hindering the next wave of technological innovation.[8] At the same time, however, GenAI systems may pose an existential threat to the future of creative industries by mimicking human creativity and delivering literary and artistic outputs more quickly and at lower costs.

---

2 Mark A Lemley, "How Generative AI Turns Copyright Upside Down" (2024) 25 Columbia Science and Technology Law Review 190 [Lemley].

3 See, *eg*, Carys Craig & Ian Kerr, "The Death of the AI Author" (2020) 52 Ottawa L Rev 31; Daniel J. Gervais, "The Machine as Author" (2020) 105 Iowa L Rev 2053.

4 See, *eg*, Christophe Geiger, "When the Robots (Try to) Take Over: Of Artificial Intelligence, Authors, Creativity and Copyright Protection" "… in Florent Thouvenin et al, eds. *Innovation – Creation – Markets* (Berlin: Springer, 2024) at 67.

5 Lemley, *supra* note 2.

6 See, *eg*, Tim W Dornis, "The Training of Generative AI is not Text and Data Mining" (2025) 47 Eur IP Rev 65 [Dornis]; Andres Guadamuz, "The EU's Artificial Intelligence Act and copyright" (forthcoming) The Journal of World Intellectual Property [Guadamuz]; Alexander Peukert, "Copyright in the Artificial Intelligence Act — A primer" (2024) 73 GRUR International 497 [Peukert]; Matthew Sag, "Fairness and Fair Use in Generative AI" (2024) 92 Fordham L Rev 1887 [Sag]; Pamela Samuelson, "Fair Use Defenses in Disruptive Technology Cases" (2024) 71 UCLA L Rev 1484 [Samuelson]; Mark A Lemley & Bryan Casey, "Fair Learning" (2021) 99 Tex L Rev 743 [Lemley & Casey].

7 See, *eg*, as one of the first acts of the new US administration, The White House, "Removing Barriers To American Leadership In Artificial Intelligence" (2025) <https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/> (23 January 2025).

8 Christian Peukert & Margaritha Windisch, "The Economics of Copyright in the Digital Age" (forthcoming) Journal of Economic Surveys.

As correctly observed, without significant empirical research, the balance between the value GenAI creates and the value it displaces remains uncertain.[9] In theory, GenAI may have a meaningful substitution effect on human-generated works (particularly when it produces near-verbatim copies), but not all its uses will reduce creators' revenue. Indeed, some applications may enhance consumer surplus by reducing deadweight loss, enabling consumption that would not have occurred otherwise. However, beyond revenue loss, AI-generated content may also diminish the value of human-created works by offering lower-quality versions, thereby reducing overall demand. In this way, AI would not only take a share of the market but may also shrink the market itself. A more complex concern is the potential impact on long-term scientific and cultural innovation. While experimentation is a key driver of creative progress, it remains uncertain whether AI can replicate the trial-and-error process of human creators. Additionally, as generative models depend on human-generated works for training, the displacement of human creators may also slow the development of GenAI. On the other side, AI has the potential to enhance creativity by making human creators more productive and lowering production costs.

For these reasons, policymakers worldwide (including in the Vatican City State)[10] are striving to strike a new balance between the benefits of incentives to create and the costs of restrictions, ensuring the legal framework supports both the creative industries and the AI sector.[11] Specifically, they aim to reconcile seemingly divergent economic objectives and industry interests — namely, fostering a favourable climate for AI development by ensuring broad and lawful access to high-quality data, while also providing adequate incentives and rewards for investments in human literary and artistic creations.

Against this background, the literature has thus far primarily concentrated on three areas of research. First, it examines whether and to what extent the use of works as training data for GenAI models constitutes copyright infringement or falls

---

[9] Joshua Gans et al, "Identifying the Economic Implications of Artificial Intelligence for Copyright Policy: Context and Direction for Economic Research" (2025) US Copyright Office <https://www.copyright.gov/newsnet/2025/1062.html?loclr=licop> (12 February 2025) at 10–13 [Gans et al, "Intellectual property and creative machines"]. See also Gaétan de Rassenfosse, Adam B Jaffe & Joel Waldfogel, "Intellectual property and creative machines" in B Jones & J Lerner, eds. NBER Entrepreneurship and Innovation Policy and the Economy, Volume 4 (Chicago: The University of Chicago Press, 2025) at 47.

[10] See Pontifical Commission for the Vatican City State, "Guidelines on Artificial Intelligence" (2024) <https://www.vaticanstate.va/images/N.%20DCCII.pdf> (16 December 2024).

[11] See, *eg*, Intellectual Property Office, Department for Science, Innovation and Technology and Department for Culture, Media and Sport, "Copyright and Artificial Intelligence" (2024) UK Government <https://www.gov.uk/government/consultations/copyright-and-artificial-intelligence> (17 December 2024) [Intellectual Property Office]. A useful and comprehensive overview of initiatives and policy solutions worldwide is provided by the Stanford Cyber Policy Center: Florence G'Sell, "Regulating Under Uncertainty: Governance Options for Generative AI" (2024) Standford Cyber Policy Center <https://cyber.fsi.stanford.edu/content/regulating-under-uncertainty-governance-options-generative-ai> (September 2024).

within the limitations and exceptions to copyright protection.[12] Second, it evaluates the feasibility, effectiveness, and implications — including those related to regulatory competition[13] — of the solutions and proposals advanced in various jurisdictions.[14] Third, it explores alternative policy solutions.[15]

Focusing on the relevance of training data for GenAI models and given the broad consensus on the need for legal reforms to address the risk that outdated copyright laws may impede research and innovation, this article contributes to the debate by providing a novel perspective on the scope of free uses of copyrighted works in the AI sector. Notably, this article advocates for a competition analysis in the assessment of the application of limitations and exceptions to copyright protection.

Competition issues have always been central to the analysis of intellectual property rights ("IPRs"), as these rights have traditionally been viewed as exceptions to the principle of a free market. They are granted only to the extent necessary to incentivise innovation and are subject to specific limitations, such as duration, scope, and requirements. This perspective is reflected in the exceptions to copyright established by the three-step test, first introduced in the Berne Convention and later incorporated into the Agreement on Trade-Related Aspects of Intellectual Property

---

[12] See, *eg*, European Copyright Society, "Copyright and Generative AI: Opinion of the European Copyright Society" (2025) <https://europeancopyrightsociety.org/opinions/> (January 2025) [European Copyright Society]; Dornis, *supra* note 6; Eleonora Rosati, "Infringing AI: Liability for AI-Generated Outputs under International, EU, and UK Copyright Law" (forthcoming) European Journal of Risk Regulation; Joshua Gans, "Copyright Policy Options for Generative Artificial Intelligence" (2024) NBER Working Paper No. 32106, <http://www.nber.org/papers/w32106> [Gans, "Copyright Policy Options for Generative Artificial Intelligence"]; Andres Guadamuz, "A Scanner Darkly: Copyright Liability and Exceptions in Artificial Intelligence Inputs and Outputs" (2024) 73 GRUR International 111; Nicola Lucchi, "ChatGPT: A Case Study on Copyright Challenges for Generative Artificial Intelligence Systems" (2024) 15 European Journal of Risk Regulation 602 [Lucchi]; Samuelson, supra note 6; Lemley & Casey, supra note 6; Jessica L Gillotte, "Copyright Infringement in AI-Generated Artworks" (2020) 53 UC Davis L Rev 2655; Benjamin LW Sobel, "Artificial Intelligence's Fair Use Crisis" (2017) 41 Colum J L & Arts 45.

[13] See, *eg*, Martin Senftleben, "Generative AI and Author Remuneration" (2023) 54 IIC 1535 [Senftleben, "Generative AI and Author Remuneration"], noting that the US, Canada, Singapore, South Korea, Japan, Israel, and Taiwan have opted for broader, more flexible, copyright limitations, thus markedly departing from the EU approach. Similarly see Niva Elkin-Koren & Neil Weinstock Netanel, "Transplanting Fair Use across the Globe: A Case Study Testing the Credibility of U.S. Opposition" (2021) 72 Hastings LJ 1121.

[14] See, *eg*, Guadamuz, *supra* note 6; João Pedro Quintais, "Generative AI, copyright and the AI Act" (2025) 56 Computer Law & Security Review 106107 [Quintais]; Kaigeng Li, Hong Wu, & 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; Peukert, *supra* note 6; Matthew Sag, "Copyright Safety for Generative AI" (2023) 61 Hous L Rev 295.

[15] See, *eg*, Martin Senftleben, "Win-win: How to Remove Copyright Obstacles to AI Training While Ensuring Author Remuneration (and Why the European AI Act Fails to Do the Magic)" (forthcoming) Chicago-Kent L Rev [Senftleben, "Win-win: How to Remove Copyright Obstacles to AI Training] While Ensuring Author Remuneration (and Why the European AI Act Fails to Do the Magic)"]; Christophe Geiger & Vincenzo Iaia, "The forgotten creator: towards a statutory remuneration right for machine learning of generative AI" (2024) 52 Computer Law & Security Review 105925 [Geiger & Iaia].

Rights ("TRIPS") and in the World Intellectual Property Organization ("WIPO") Copyright Treaty.[16]

The interpretation of the three-step test criteria has long been contentious, and its application in national laws has further deepened uncertainties since contracting parties have adopted diverse approaches and crafted their own provisions.[17] Notably, for the purpose of our investigation, it is worth reminding that, while the U.S. has relied on the broad and flexible fair use exception,[18] other countries have implemented specific exceptions for text and data mining ("TDM"), which however vary significantly in terms of permitted uses and whether they extend to commercial purposes.[19]

The suitability of these varied approaches for data collection in AI model training lies at the heart of the ongoing debate. Nonetheless, there is a shared perception that it will likely take years to clarify the scope of these provisions, as historically demonstrated by the dispute over the proper interpretation of the three-step test, fair use, and other copyright exceptions. Consequently, there is increasing concern that copyright not only fails to address all the challenges posed by the AI sector but is itself a primary source of these issues.[20]

To this end, insights from competition law could prove extremely valuable. Indeed, despite the mentioned significant variations, the core of the three-step test clearly embodies a non-competition principle according to which the use of copyrighted works must not be in competition with the exploitation rights of the holder. The potential role of antitrust tools in copyright infringement cases, particularly in the interpretation of exceptions to copyright protection, was confirmed by the recent US Supreme Court's decision in *Andy Warhol*.[21]

Further, as part of the longstanding and controversial interplay between intellectual property protection and competition law, some jurisdictions have implemented policies to address potential abuses in the enforcement of IPRs. Notably,

---

16 See Berne Convention for the Protection of Literary and Artistic Works (1886; lastly amended in 1979) at Article 9(2) [Berne Convention]; Agreement on Trade-Related Aspects of Intellectual Property Rights (1994; lastly amended in 2017) at Article 13; World Intellectual Property Organization (WIPO) Copyright Treaty (1996) at Article 10. The test requires members to establish exceptions (i) only in "certain special cases", ensuring that the free use does not (ii) conflict with the "normal exploitation of the work" and does not (iii) "unreasonably prejudice the legitimate interests of the author".

17 See Christophe Geiger, Daniel J Gervais, & Martin Senftleben, "The Three-Step-Test Revisited: How to Use the Test's Flexibility in National Copyright Law" (2014) 29 Am U Intl L Rev 581.

18 17 U.S.C. §107.

19 For an overview, see Sean M Fiil-Flynn et al, "Legal Reform to Enhance Global Text and Data Mining Research: Outdated Copyright Laws Around the World Hinder Research" (2022) 378 Science 951 [Fiil-Flynn et al].

20 See, *eg*, Carys J Craig, "The AI-Copyright Trap" (forthcoming) Chicago-Kent L Rev [Craig]; Micaela Mantegna, "ARTificial: Why Copyright Is Not the Right Policy Tool to Deal with Generative AI" (2024) Yale LJ Forum 1126 [Mantegna]; Bertin Martens, "Economic arguments in favour of reducing copyright protection for generative AI inputs and outputs" (2024) Bruegel Working Paper 09/2024, <https://www.bruegel.org/working-paper/economic-arguments-favour-reducing-copyright-protection-generative-ai-inputs-and> (4 April 2024) [Martens].

21 *Andy Warhol Foundation for the Visual Arts, Inc. v Goldsmith* 598 US 508 (2023) [*Andy Warhol*]. See Christopher Jon Sprigman, "Copyright, Meet Antitrust: The Supreme Court's Warhol Decision and the Rise of Competition Analysis in Fair Use" (2025) Yale LJ Forum 298 [Sprigman] sharing this view.

the essential facilities doctrine ("EFD") has gained traction in Europe, deeming it anticompetitive — and thus unlawful — to refuse to grant a licence under exceptional circumstances. Two of these circumstances, outlined by the European Court of Justice ("CJEU") in the *Magill* copyright case, could be key in defining the scope of exceptions in the context of AI training.[22] In particular, under the EFD, to balance the conflicting interests between access and incentives, the exercise of an exclusive right is deemed abusive when a refusal to share would likely eliminate all competition in a secondary market and hinder the emergence of a new product for which there is a potential consumer demand.

To be clear, this article does not advocate for the intervention of antitrust authorities or the application of competition law. Rather, it suggests interpreting the copyright exceptions relevant to AI model training through the lens of antitrust tools. This would assist courts and policymakers in establishing a predictable and consistent legal framework, which is much needed for navigating the new balance of conflicting interests brought about by GenAI.

This article is structured as follows. Section 2 explores the roots of the problem in AI training data and critically evaluates the approaches taken by major countries, either by introducing new copyright exceptions or adapting traditional provisions. Section 3 argues that competition rationale has traditionally guided the interpretation of copyright exceptions, with market harm reflecting the balance between incentives to create and the market effects of exclusive rights. Given the regulatory fragmentation and uncertainty around applying copyright exceptions to AI training data, it suggests that using competition insights to assess market harm could make copyright exception analysis more predictable and aligned with copyright's fundamental purpose. Against this background, the proposal is put forward to rely on the exceptional circumstances established by the CJEU for the application of the EFD to IPRs. Section 4 concludes.

II. The (Impossible?) Balance Between Competition, Innovation, and Authorial Incentives in AI Training Data: Testing the Boundaries of Copyright Limitations and Exceptions

AI is rapidly emerging as a transformative innovation, shaping a new technological paradigm and serving as a powerful catalyst for innovation, productivity gains, and economic growth.[23] This wave of innovation and the race for AI have created a moment of heightened market contestability that will shape the future dynamics of competition, making it imperative for policymakers to foster a pro-competitive environment.

---

[22] CJEU Joined Cases C-241/91 P and 242/91 P *Radio Telefis Eireann (RTE) and Independent Television Publications Ltd (ITP) v Commission* [1995] EU:C:1995:98.

[23] See, eg, Giacomo Damioli et al, "Is artificial intelligence leading to a new technological paradigm?" (2025) 72 Structural Change and Economic Dynamics 347. See also Erik Brynjolfsson, Danielle Li & Lindsey Raymond, "Generative AI at Work" (forthcoming) The Quarterly Journal of Economics, finding that access to GenAI suggestions can increase the productivity of individual workers and improve their experience of work.

In this scenario, alongside computing power, technical expertise, and capital, data is widely recognised as an essential building block of GenAI.[24] Data influences the quality of AI foundation models and consequently drives innovation and competition in the field.[25] Since AI models require training on datasets, access to data is crucial and models trained on larger datasets are generally assumed to be superior. Further, the type of data that is used to train a foundation model determines its mode.

This assumption is tempered by some factors: differences in foundation model types, the growing availability of publicly accessible and open-source datasets, diminishing returns to scale in data usage beyond a certain tipping point, the use of synthetic data, and advances in computer science and analytics that enable companies with smaller datasets to compete effectively with larger ones.[26] Therefore, it is more accurate to argue that, while access to large datasets is relevant but not decisive, access to unique datasets is crucial.[27]

In summary, data is an important — though not decisive — factor in the competitive landscape.[28] Since the development of foundation models also depends on data availability, restricted access to this input could hinder competition and stifle

---

[24] See, *eg*, Autoridade da Concorrência, "Competition and Generative AI: Zooming in on Data" (2024) <https://www.concorrencia.pt/en/articles/adc-warns-competition-risks-regarding-access-and-use-data-generative-ai> (September 2024); Autorité de la concurrence, "Generative artificial intelligence: The Autorité issues its opinion on the competitive functioning of the generative artificial intelligence sector" (2024) <https://www.autoritedelaconcurrence.fr/en/press-release/generative-artificial-intelligence-autorite-issues-its-opinion-competitive> (28 June 2024) [Autorité de la concurrence]; Japanese Fair Trade Commission, "Generative AI and Competition (Discussion Paper)" (2024) <https://www.jftc.go.jp/en/pressreleases/yearly-2024/October/1002.html> (October 2024); UK Competition and Markets Authority, "AI Foundation Models: Update paper" (2024) <https://www.gov.uk/government/publications/ai-foundation-models-update-paper> (11 April 2024); Autoridade da Concorrência, "Competition and Generative Artificial Intelligence" (2023) <https://www.concorrencia.pt/en/articles/adc-warns-competition-risks-generative-artificial-intelligence-sector>; UK Competition and Markets Authority, "AI Foundation Models: Initial Report" (2023) <https://www.gov.uk/government/publications/ai-foundation-models-initial-report> (18 September 2023); US Federal Trade Commission, "Generative AI Raises Competition Concerns" (2023) <https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/06/generative-ai-raises-competition-concerns> (29 June 2023).

[25] See Joshua S Gans, "Market Power in Artificial Intelligence" (2024) NBER Working Paper No. 32270, <http://www.nber.org/papers/w32270>, arguing that, from a competition perspective, a useful — though imperfect — distinction between the roles of training data and input data is that the former, by enabling the creation of AI prediction algorithms, serves as a key driver of market entry, while input data, by ensuring prediction algorithms operate effectively, plays a critical role in determining the cost and/or quality of firms' offerings within a market.

[26] See, eg, Thibault Schrepel & Alex Sandy Pentland, "Competition between AI foundation models: dynamics and policy recommendations" (forthcoming) Industrial and Corporate Change [Schrepel and Pentland]; Autorité de la concurrence, supra note 24; Geoffrey A Manne & Dirk Auer, "From data myths to data reality: what generative AI can tell us about competition policy (and vice versa)" (2024) CPI Antitrust Chronicle 9.

[27] Schrepel & Pentland, ibid. See also Yiheng Liu et al, "Understanding LLMs: A Comprehensive Overview from Training to Inference" (2025) 620 Neurocomputing 129190, arguing that not all data contributes equally to the performance of a model.

[28] See OECD, "Artificial Intelligence, Data and Competition" (2024) <https://www.oecd.org/en/publications/artificial-intelligence-data-and-competition_e7e88884-en.html> (24 May 2024) [OECD], arguing that it seems likely that significant data will continue to be required, although where the focal point of the industry will settle is unclear.

innovation in the sector, disadvantaging new entrants *vis-à-vis* incumbent digital players.

However, the policy goal of fostering competition and innovation conflicts with the seemingly opposing goal of protecting creative industries from the potential erosion of incentives to produce literary and artistic works due to GenAI. This is evidenced by the numerous lawsuits filed by copyright holders seeking compensation and accusing AI companies of accessing copyrighted materials without permission to train their models. After all, since AI systems rely on existing works to learn and improve, it is inevitable that some of these training data include copyright-protected material. This raises the question of whether such collection constitutes infringement or if the techniques used for training (*eg*, TDM and deep learning) fall within lawful exceptions.

In this context, copyright protection could not only restrict firms' ability to access the data needed to train their foundation models but also disproportionately impact smaller and newer players, increasing the reliance on proprietary or exclusively licensed datasets.[29] Furthermore, the uncertainty surrounding the legality of certain data uses creates risks for prospective developers, who may need to invest in understanding the regulatory stance of different jurisdictions.[30] Finally, and more generally, the risk should be avoided that copyright, if seen as a barrier to technological progress, may provoke a public policy backlash that ultimately undermines its objectives. Therefore, the prudent course is to adapt copyright law cautiously to enable AI's responsible evolution, while steadfastly preserving incentives for human creativity.[31]

As a result, the ongoing attempt to reconcile these opposing interests essentially hinges on the scope of both old and new copyright limitations and exceptions. While many legislative proposals are being discussed and consulted on,[32] diverging approaches have already emerged, ranging from concerns over copyright underenforcement to risks of overenforcement.

On one end of the spectrum, for instance, in Singapore, copyright law was amended to introduce an exception to reproduction and communication rights, permitting computational data analysis.[33] Similarly, Japan's 2018 revision of its copyright law granted broad rights to use copyrighted works for any type of data analysis, including training AI models.[34] Consistent with the application of the three-step test, the Japanese exception does not apply in cases that would unreasonably prejudice

---

29 Martens, *supra* note 20.

30 OECD, *supra* note 28.

31 Kristian Stout, Geoffrey A Manne & Emily Corbeille, "Comments of the International Center for Law & Economics to the U.S. Copyright Office, Notice of Inquiry and Request for Comments" (2023) <https://laweconcenter.org/resources/icle-comments-on-artificial-intelligence-and-copyright/> (30 October 2023).

32 See, *eg*, Intellectual Property Office, *supra* note 11; Government of Canada, "Consultation on Copyright in the Age of Generative Artificial Intelligence" (2023) <https://ised-isde.canada.ca/site/strategic-policy-sector/en/marketplace-framework-policy/consultation-paper-consultation-copyright-age-generative-artificial-intelligence>; US Copyright Office, "Artificial Intelligence and Copyright" (2023) 88 Federal Register 59942, <https://www.copyright.gov/policy/artificial-intelligence/>.

33 Copyright Act 2021 (2020 Rev Ed) at sections 243–244.

34 Copyright Act 2018 at Articles 30–34.

the interests of copyright holders.[35] Furthermore, the exception excludes additional training on foundation models (*ie*, fine-tuning) and intentional overfitting.[36]

On the other end of the spectrum, in its recent AI Act,[37] Europe has adopted an input-based remuneration approach by linking the TDM exceptions introduced in the Directive on copyright in the Digital Single Market ("DSM") to the development of general-purpose AI models.[38] Indeed, a last-minute addition to the AI Act clarifies that providers of general-purpose AI models, including GenAI models, must implement a policy to comply with the rights reservations established in the DSM Directive.[39] Specifically, this refers to the provision allowing a TDM exception for commercial purposes — distinct from scientific or research uses — while making it conditional on respecting any opt-outs specified by copyright holders.[40] Therefore, where the rights to opt out have been expressly reserved in an appropriate manner, providers of general-purpose AI models need to obtain an authorisation from rightsholders if they want to carry out TDM over such works.[41] To this end, the AI Act imposes a transparency obligation regarding training data, requiring providers to compile and publicly disclose a sufficiently detailed summary of the content used to train general-purpose models. This summary should be comprehensive in scope rather than technically detailed, making it easier for copyright holders to exercise and enforce their rights.[42]

While several EU Member States believe that such a legal framework would suffice to encourage licensing agreements between rightsholders and AI companies,[43] the European solution has been criticised as unfeasible and economically unsound.[44] On the one hand, the fulfilment of the transparency obligation could be severely hindered by several factors, including the fragmentation of copyright

---

35 See Japan Copyright Office, "General Understanding on AI and Copyright in Japan: Overview" (2024) <https://www.bunka.go.jp/english/policy/copyright/pdf/94055801_01.pdf> (May 2024), providing the example of reproducing a copyrighted database work for the purpose of data analysis for which licences for data analysis are available in the marketplace.

36 See *ibid*, referring to the collection of works for AI training to generate materials similar to copyrighted works.

37 Regulation (EU) 2024/1689 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, (Artificial Intelligence Act), [2024] OJ L *2024/1689* [AI Act].

38 Directive (EU) 2019/790 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC [Directive (EU) 2019/790] and 2001/29/EC, [2019] OJ L130/92.

39 AI Act, *supra* note 37 at Article 53.

40 Directive (EU) 2019/790, *supra* note 38 at Article 4.

41 AI Act, *supra* note 37 at Recital 105.

42 *Ibid* at Recital 107.

43 Council of the European Union, "Policy questionnaire on the relationship between generative Artificial Intelligence and copyright and related rights – Revised Presidency summary of the Member States contributions" (2024) <https://www.consilium.europa.eu/en/documents-publications/public-register/public-register-search/?AllLanguagesSearch=false&OnlyPublicDocuments=false&DocumentNumber=16710%2F24&DocumentLanguage=EN> (20 December 2024).

44 See, *eg*, Quintais, *supra* note 14; Geiger & Iaia, *supra* note 15; Lucchi, *supra* note 12; Peukert, *supra* note 6; Senftleben, "Generative AI and Author Remuneration", *supra* note 13. On the interpretative uncertainties regarding both the application of the European TDM exception to the training of GenAI models and the content of the obligations under the AI Act, see also European Copyright Society, *supra* note 12.

across jurisdictions, the absence of a mandatory registration system, and the complexity of identifying rightsholders through automated processes and managing payments.[45] On the other hand, even if copyright holders succeed in reserving their rights and preventing the use of their works for AI training, this will not necessarily result in licensing agreements. Instead, it may incentivise companies to relocate to countries with more favourable regulations. Consequently, a scenario contrary to Pareto optimality may emerge, where everyone is worse off: the EU risks losing its appeal as a hub for AI development due to restricted access to copyrighted materials, while authors receive no remuneration.[46]

In its consultation launched last December, the UK Government endorsed an approach similar to the European AI Act. However, a lively debate is still ongoing in Parliament, making it difficult at this stage to predict the final outcome.[47]

Finally, to complete our comparative overview, it is worth noting that, with regard to training data, the US currently lacks formal mandatory regulations,[48] so the issue is addressed through the lens of the general fair use exception.[49] Accordingly, the four-factor test for determining whether a use qualifies as fair — thus exempting it from requiring the copyright holder's permission — consists of (i) the purpose and character of the use, (ii) the nature of the copyrighted work, (iii) the amount and substantiality of the portion used in relation to the copyrighted work, and (iv) the effect of the use on the potential market for or value of the copyrighted work.[50]

The very flexibility that gives the fair use doctrine an advantage has also been considered its downfall. Indeed, on the one side, fair use can adapt to technological advancements without requiring frequent amendments to copyright law.[51] On

---

45 See, *eg*, European Commission, "Study to assess the feasibility of a central registry of Text and Data Mining opt-out expressed by rightsholders" (2025) <https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/tender-details/8726813a-bd9b-4f58-8679-01c80f7a1abf-CN?isExactMatch=true&order=DESC&pageNumber=1&pageSize=50&sortBy=startDate> (23 January 2025), issuing a call for tender for a study to assess the opportunity and feasibility of developing a work-based registry of content identifiers and associated metadata that would support the effective expression of TDM opt-outs for copyright-protected works and other subject matter and facilitate their identification by AI developers.

46 Senftleben, "Generative AI and Author Remuneration", *supra* note 13.

47 Intellectual Property Office, *supra* note 11.

48 In April 2024, Representative Adam Schiff introduced the "Generative AI Copyright Disclosure Act" into the U.S. House of Representatives (H.R. 7913, "Generative AI Copyright Disclosure Act of 2024" (2024) <https://www.congress.gov/bill/118th-congress/house-bill/7913> (9 April 2024)), which mandates disclosure of copyrighted content used to train AI models.

49 17 U.S.C. §107.

50 The first factor, also known as the transformative use factor, examines whether the use alters the original work by adding something new or serving a different purpose, even if the copied material remains largely unchanged. The second one assesses whether the copied content is factual or creative, with factual works receiving less protection. Factor (iii) considers the amount and significance of the copied portion in relation to the purpose of the use. The fourth factor — which, along with the first, receives particular consideration by the courts in the overall analysis — evaluates whether the reproduction competes with the original work or diminishes its value by offering a substitute that consumers might prefer. To qualify for the copyright exception, the use should not harm the market for the original work, as copyright protects the creator's ability to profit.

51 For an overview of the fair use doctrine's application to technological advancements over the decades, see, *eg*, Samuelson, *supra* note 6; Jane C. Ginsburg, "Fair Use in the US Redux: Reformed or Still Deformed?" (2024) Sing JLS 52.

the other side, due to the lack of clear guidance, the doctrine is often criticised as unpredictable and inconsistent.[52] Sufficient to earn the distinction of being the most important, as well as the most enigmatic, doctrine in US copyright law as it provides ample room for judicial interpretation of both each factor and the weight assigned to them.[53] Therefore, in the context of GenAI, it would not be surprising if the fair use debate continues for several years until the Supreme Court provides clarification.[54]

In summary, despite noble efforts to reconcile these goals, policymakers seem to be struggling to define a clear stance on the relationship between copyright and GenAI, as well as to implement a practical solution that would streamline data access for AI developers while ensuring fair remuneration for copyright holders.[55]

As the list of countries with their own copyright exceptions relevant to AI training continues to grow — and others appear poised to enact specific provisions — regulatory fragmentation is becoming a significant risk, if not already a reality.[56] In addition, both existing and new provisions are often ambiguous and open to divergent interpretations, further increasing uncertainty and, consequently, the risk of slowing AI development.

For these reasons, while some suggest that a practical and effective solution can only be found outside the scope of copyright law,[57] others point to several examples of 'permitted-but-paid' uses already established within the copyright frameworks of various countries.[58] Notably, since, from an economic perspective, the goal should be to encourage both the creation and use of content, to mitigate the risks of both over- and under-enforcement of copyright, the literature has proposed *ex post* mechanisms or a remunerated copyright limitation for GenAI commercial purposes aimed at addressing the transactions costs problem, namely by minimising costs associated with injunctive legal actions or licence negotiations.[59] In this scenario, balancing the protection of creators' incentives with the promotion of an AI-friendly regulatory environment is achieved by reversing the traditional approach to copyright and introducing a liability rule. Accordingly, copyright holders cannot prevent the use of their content in AI training but are compensated for its use.

---

52 See, *eg*, William M Landes and Richard A Posner, *The economic structure of intellectual property law*, (Cambridge: Harvard University Press, 2003) at 115.

53 Barton Beebe, "An Empirical Study of U.S. Copyright Fair Use Opinions, 1978–2005" (2008) 156 U Pa L Rev 549.

54 See, *eg*, Daniel Gervais et al, "The Heart of the Matter: Copyright, AI Training, and LLMs" (2024) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4963711> (1 November 2024) [Gervais et al]; Lemley & Casey, *supra* note 6; Sag, *supra* note 6; Samuelson, *supra* note 6.

55 See, *eg*, Patrick Vallance, "Pro-innovation Regulation of Technologies Review: Digital Technologies" (2023) <https://www.gov.uk/government/publications/pro-innovation-regulation-of-technologies-review-digital-technologies> (15 March 2023), urging the UK Government to prioritise practical solutions to the barriers faced by AI firms in accessing copyright and database materials. See also Craig, *supra* note 20, noting that, in some cases (*eg*, Canada), governments launched multiple consultations on the issue of copyright and AI, supporting different points of view over the years.

56 For a comparative analysis, see Gervais et al, *supra* note 54; Fiil-Flynn et al, *supra* note 19.

57 See, *eg*, Craig, *supra* note 20; Mantegna, *supra* note 20.

58 Geiger & Iaia, *supra* note 15.

59 For a law and economics analysis of the different policy options about the treatment of developers' past and future use of copyrighted works in AI training, see Gans et al, "Intellectual property and creative machines" supra note 9 at 45–53.

More specifically, Gans proposed an *ex post* fair use-like mechanism in which AI providers can use any content for training their models, but if it is later determined that individual content creators have suffered commercial harm, they can compel the AI provider to compensate them for their lost profits.[60] In a similar vein, Senftleben suggested an output-based remuneration system, under which AI developers can use the full spectrum of human literary and artistic works, but must provide equitable compensation for content that could serve as a substitute for human creations.[61] Pasquale and Sun advanced the idea of a system that combines an opt-out mechanism with a compensation levy.[62] Under this approach, AI providers must remove objectors' works from their databases once copyright infringement is documented, while authors who do not opt out — and, therefore, whose works are used without a licence — receive compensation through the levy. Finally, Geiger and Iaia suggested replacing the opt-out mechanism with a statutory licence that includes a remuneration right for commercial TDM activities.[63]

This article aims to contribute to this debate by offering a novel perspective that integrates and complements these proposals. While most of them address the 'why' and the 'how' of compensation, this article will focus on the 'when.' It argues that insights from competition case law could offer clear guidance on when to apply copyright exceptions by assessing the market effects of GenAI, thus assisting courts and policymakers in determining when creators experience commercial harm and when AI-generated content may be considered a substitute for human creations.

III. Interpreting the Market Effects of Free Use Through the Lens of Competition Law

The previous section outlined policymakers' concerns about the role of copyright in AI development, the challenges of balancing competing interests, and the effectiveness of both existing and new copyright limitations and exceptions. Within this framework, despite their differences, these exceptions share two features and a common goal. First, they all follow the guiding principles of the Berne Convention's three-step test.[64] Second, they all face criticism for failing to provide clear guidance, as their criteria have traditionally been subject to divergent interpretations. Finally, they all are aimed at achieving a fair balance between the conflicting interests of authors and users.

Given the importance of copyright exceptions for GenAI training data, this section explores how antitrust law can provide valuable interpretative guidance, helping to clarify the meaning of market harm as the key criterion for applying these exceptions. The relevance of market harm in assessing copyright exceptions hinges on the strong connection between market effects and incentives to create. Indeed,

---

60 Gans, "Copyright Policy Options for Generative Artificial Intelligence", *supra* note 12.

61 Senftleben, "Win-win: How to Remove Copyright Obstacles to AI Training", *supra* note 15.

62 Frank Pasquale & Haochen Sun, "Consent and Compensation: Resolving Generative AI's Copyright Crisis" (2024) 110 Va L Rev 207.

63 Geiger & Iaia, *supra* note 15.

64 Berne Convention, *supra* note 16 at Article 9(2).

losses incurred by a copyright holder normally conflict with copyright's fundamental objective: providing authors with exclusive rights that encourage creative expression.

Therefore, the three-step test conditions the availability of exceptions on ensuring free use does not conflict with the "normal exploitation" of the work and does not "unreasonably prejudice the legitimate interests of the author".[65] Regarding the notion of normal exploitation, the World Trade Organisation Panel, when assessing the compatibility of certain exemptions in the US Copyright Act with the TRIPS Agreement, stated that an exception conflicts with the normal exploitation of a work if the exempted uses enter into "*economic competition*" with the ways in which rightsholders typically derive economic value from copyright protection, thereby depriving them of significant or tangible commercial gains[66] [emphasis added]. Further, the Panel found that the level of prejudice should be deemed unreasonable when an exception results in, or has the potential to result in, an unreasonable loss of income for the copyright owner.[67]

In the EU, despite the adoption of closed lists of specific exceptions and limitations to copyright, the test has been incorporated in several directives, particularly the Information Society Directive and the DSM Directive.[68] The test limits the discretion of Member States in implementing exceptions or limitations introduced by directives, as they remain subject to the three aforementioned conditions. Nonetheless, CJEU decisions have provided little clarity on the specific meaning of each stage, often merely acknowledging that the three-step test helps balance the interests of copyright holders in protecting their IPRs with the interests and fundamental rights of users and the broader public interest.[69]

---

65 *Ibid.*

66 WTO Panel, *United States – Section 110 (5) of the US Copyright Act*, Report WT/DS160/R (2000), para. 6.183, <https://www.wto.org/english/tratop_e/dispu_e/cases_e/ds160_e.htm> (15 June 2000) [WTO Panel].

67 *Ibid* at [6.229].

68 Directive 2001/29 on the harmonisation of certain aspects of copyright and related rights in the information society, [2001] OJ L167/10 at Article 5(5); Directive (EU) 2019/790, *supra* note 38 at Article 7(2). See Martin Senftleben, "From Flexible Balancing Tool to Quasi-Constitutional Straitjacket – How the EU Cultivates the Constraining Function of the Three-Step Test" in Jonathan Griffiths and Tuomas Mylly eds. *Global Intellectual Property Protection and New Constitutionalism: Hedging Exclusive Rights* (Oxford: Oxford University Press, 2021) at 83 and 95, arguing that the EU three-step test increases legal uncertainty by creating a bias against limitations and exceptions, as it allows judges to further restrict use privileges that are already narrowly defined in statutory law. Indeed, even when a use fully complies with a statutory copyright limitation, it may still be deemed infringing due to the test's flexible and indeterminate criteria.

69 See, *eg*, CJEU Case C-516/17 *Spiegel Online v Volcker Beck* [2019] EU:C:2019:625 at [46]; CJEU C-476/17 *Pelham v Hütter* [2019] EU:C:2019:624 at [62]. For a concrete application of the test, see, *eg*, CJEU Case C-527/15 *Stichting Brein v Wullems (Filmspeler)* [2017] EU:C:2017:300 at [70], stating that Member States may not extend the scope of the private copying exception to copies made from unlawful sources, as permitting such uses would conflict with normal exploitation by reducing the volume of lawful transactions and would also cause unreasonable prejudice to copyright holders; CJEU Case C-435/12 *ACI Adam v Stichting de Thuiskopie* [2014] EU:C:2014:254 at [38]–[40], stating on the very same matter and arguing that allowing such uses would conflict with normal exploitation; CJEU Case C-360/13 *Public Relations Consultants v Newspaper Licensing Agency* [2014] EU:C:2014:1195 at [54]–[63], concluding that on-screen and cached copies of works made while viewing a website do not unreasonably prejudice the legitimate interests of right holders, as the website publishers themselves

In the US, the fourth factor of the fair use doctrine also examines the market impact of a challenged use, assessing "the effect of the use upon the potential market for or value of the copyrighted work".[70] This factor, along with the transformative use analysis under the first factor, is widely considered the most decisive in fair use cases.[71] As clarified by US Supreme Court in *Campbell v. Acuff-Rose Music*, while the fourth factor requires consideration of both the harm to the original work and the harm to the (actual and potential) market for derivative works, the only "cognizable harm is *market substitution*, not any harm"[72] [emphasis added]. In *Google v. Oracle America*, the Supreme Court further clarified that, in determining market effects, courts must also consider the "public benefits" the copying is likely to produce, along with the potential risks of creativity-related harms arising from allowing a fair use defence.[73]

Interestingly, in its most recent fair use decision, *Andy Warhol Foundation for the Visual Arts v. Goldsmith*, the Supreme Court also evaluated the first factor through a competitive lens, stating that the transformativeness determination relates to "the problem of *substitution* — copyright's *bête noire*"[74] [emphasis added]. Indeed, the use of an original work for a purpose that is identical or closely similar to that of the original is more likely to serve as a substitute and potentially supplant the original. Similarly, in his concurring opinion, Justice Gorsuch argued that, under the first fair use factor, "the salient point is that the purpose and character of the Foundation's use involved *competition* with Ms. Goldsmith's image"[75] [emphasis added]. Finally, Justice Kagan, in her dissenting opinion, acknowledged the role of competition in the fair use assessment, criticising however, the majority for conducting a market analysis under the first factor, effectively transplanting the fourth factor into it.[76]

As noted by Sprigman, by linking the determination of a work's purpose to the prospect of substitution, the Supreme Court made competition between the parties' works for a particular use central to assessing transformativeness.[77] Therefore, after *Warhol*, courts should apply antitrust tools to address the competition questions that have been identified as essential to the fair use analysis. Notably, according to Sprigman, courts should rely on the same antitrust methodologies used to assess

---

must obtain authorisation from right holders to make those works available to the public; CJEU Case C-162/10 *Phonographic Performance (Ireland)* [2012] EU:C:2012:141 at [75]–[76], stating that Member States that Member States may not exempt hotel operators from the obligation to pay equitable remuneration to performers and phonogram producers when making a communication to the public, as such an exemption would unjustifiably harm the legitimate interests of protected artists and performers.

70 17 U.S.C. §107.

71 See, *eg*, Pamela Samuelson, "Generative AI meets copyright" (2023) 381 Science 158 at 160.

72 *Campbell v Acuff-Rose Music* 510 U.S. 569 at 590–594 (1994) [*Campbell*], also stating that, when the second use is transformative, market substitution is at least less certain, and market harm may not be so readily inferred.

73 *Google v Oracle America* 593 U.S. 1 at 31 and 35 (2021) [*Google*].

74 *Andy Warhol*, *supra* note 21 at 528.

75 *Ibid* at 556.

76 *Ibid* at 578: "That issue is no doubt important in the fair-use inquiry. But it is the stuff of factor 4."

77 Sprigman, *supra* note 21 at 310. For a first application to AI training data, see *Thomson Reuters v Ross Intelligence*, Case 1:20-cv-00613-SB (D. Del. 2025), holding that copying was not sufficiently transformative to qualify for fair use because Ross's service, which used Westlaw headnotes to train an AI legal-research tool, operated squarely within the same market as Westlaw.

the substitutability of products or services, particularly in defining relevant antitrust markets.[78]

In the next section, we propose a different tool that we believe is better suited to addressing the role of copyright in GenAI. Nonetheless, we share the same core message: focusing on competition in assessing substitutability and market effects would make the analysis of copyright limitations and exceptions more predictable and more aligned with copyright's fundamental purpose.[79]

A. *Insights from the European Essential Facility Doctrine*

In the EU, the primary antitrust tool for overseeing IPRs is represented by the EFD.[80] The doctrine falls within the refusal to deal framework and serves as an exception to the general rule that firms, including monopolists, are free to choose their business partners. It is based on the principle that, under certain circumstances, a monopolist may be required to share its facilities with those seeking access.

Admittedly, the EFD has always been controversial and its application to IPRs is even more contentious. While the doctrine constitutes an effective tool for lowering entry barriers and preventing the competitive disadvantages of downstream rivals, it also interferes with property rights and freedom of contract by imposing mandatory access, thereby limiting the owner's core right to exclude others. Therefore, ultimately, the debate revolves around defining appropriate limits as an unrestricted duty to share could weaken investment incentives by reducing the ability to secure returns.

This delicate trade-off has shaped the evolving fate of the EFD in the US. The doctrine traces its origins to the Supreme Court's 1912 decision in *Terminal Railroad*,[81] though the term itself first appeared in the 1980s, notably in *Hecht v. Pro-Football*[82] and more prominently in *MCI Communications Corp. v. AT&T*.[83] Lower courts have primarily developed and applied the EFD, recognising its potential relevance across various scenarios, particularly when bottleneck control extends to both tangible and intangible assets.

In the US, the peak of the EFD's success also marked the beginning of its decline, officially affirmed in *Trinko*.[84] In his majority opinion, Justice Scalia asserted that the Supreme Court has never formally recognised the EFD and emphasised that monopoly power, in itself, is a key feature of the free market system that fosters innovation incentives.[85] Accordingly, mere possession of monopoly power is not

---

78 Sprigman, *supra* note 21 at 312.

79 *Ibid* at 330.

80 A similar approach has recently emerged in China: for a critical analysis, see Giuseppe Colangelo & Roberto Pardolesi, "Intellectual property, standards, and antitrust: a new life for the essential facilities doctrine? Some insights from the Chinese regulation" in Peter Drahos, Gustavo Ghidini, Hanns Ullrich eds. *Kritika: Essays on Intellectual Property* (Cheltenham: Edward Elgar, 2017) at 70 [Colangelo & Pardolesi].

81 *U.S. v Terminal Railroad Association*, 224 U.S. 383 (1912).

82 570 F.2d 982 (D.C. Cir. 1977).

83 708 F2d 1081 (7th Cir. 1983).

84 *Verizon Communications v Law Offices of Curtis V. Trinko*, 540 U.S. 398 (2004) [*Trinko*].

85 *Ibid* at 407.

unlawful unless accompanied by anticompetitive conduct. Under *Aspen Skiing* — which is in any case "at or near the outer boundary" of antitrust liability[86] — this standard is met only when the termination of a voluntary agreement indicates a willingness to sacrifice short-term profits for an anticompetitive objective.[87]

Although *Trinko* does not address intellectual property, its principles on refusal to deal are particularly relevant to this type of exclusive rights. Indeed, the Supreme Court cautioned that forcing firms to share their competitive advantage conflicts with antitrust law's objectives, as it may reduce incentives for both monopolists and rivals to invest in economically beneficial facilities.[88]

Trinko marks a major transatlantic divergence between the EU and the US. While the legal system that initially adopted the EFD ultimately rejected it, the doctrine has gained considerable influence in Europe over the past decades, extending its application to IPRs.

In fact, EU courts share the same concerns as the US Supreme Court regarding the negative effects on dynamic competition and innovation that arise from interfering with the freedom to contract and the right to property. As noted by the CJEU, while, in the short term, holding a company liable for abusing its dominant position by refusing to contract with a competitor may encourage competition, in the long term, allowing a company to reserve for its own use the facilities it has developed generally benefits competition and serves consumers' interests.[89] If access to production, purchasing, or distribution facilities were too easily granted, competitors would have little incentive to develop their own alternatives. Moreover, a dominant firm would be less inclined to invest in efficient facilities if it could be compelled, at a competitor's mere request, to share the benefits of its own investments.

Therefore, the justification in terms of competition policy for interfering with two fundamental rights (i.e., freedom to contract and right to property) also requires a careful balancing of interests between the short-term and long-term benefits of free competition.

The CJEU addressed this balance in *Magill* by defining four cumulative "exceptional circumstances" that justify interference with the freedom to contract and property rights.[90] Since the dispute involved a refusal to license an IPR (specifically, copyright on television programme schedules), the fourth circumstance is considered applicable only in cases concerning IPRs. The *Magill* exceptional circumstances are constituted by the fact that the refusal: (i) concerns a product (information on the weekly schedules of certain television channels), the supply of which is "indispensable" for carrying on the business in question (the publishing of a gen-

---

86 *Ibid* at 399.

87 *Aspen Skiing Co. v Aspen Highlands Skiing Corp.*, 472 U.S. 585 (1985).

88 *Trinko*, *supra* note 84 at 407–408.

89 See, *eg*, CJEU Case C-233/23 *Alphabet and others v Autorità Garante della Concorrenza e del Mercato* [2025] EU:C:2025:110 at [41]–[42]; Case C-48/22 P *Google LLC and Alphabet Inc. v European Commission (Google Shopping)* [2024] EU:C:2024:726 at [91]; Case C-165/19 P *Slovak Telekom a.s. v Commission* [2021] EU:C:2021:239 at [47]. Previously, see Opinion of the Advocate General Francis Jacobs, Case C-7/97 *Oscar Bronner GmbH & Co. KG v Mediaprint Zeitungs- und Zeitschriftenverlag GmbH & Co. KG*, *Mediaprint Zeitungsvertriebsgesellschaft mbH & Co. KG and Mediaprint Anzeigengesellschaft mbH & Co. KG* [1998] EU:C:1998:264 at [57].

90 *Magill*, *supra* note 22 at [54]–[56].

eral television guide); (ii) is not justified by objective considerations; (iii) is likely to exclude all "competition in the secondary market" on the part of the requesting undertaking; and (iv) prevents the emergence of a "new product" for which there was a potential consumer demand.[91] If these conditions are met, competition concerns take precedence over safeguarding investment incentives, requiring the owner of a physical or intangible asset to share it with third parties.

The latter two *Magill* requirements are particularly important to prevent unintended consequences, such as weakening incentives to innovate, allowing rivals to free ride, and merely filling potential competitive gaps. They form the core of the balancing test between the private and public interests involved and become even more crucial when IPRs are at stake.

Indeed, the 'secondary market' requirement ensures that IP holders and third-party users do not compete directly, thereby protecting investment incentives by preventing direct market harm to the former. At the same time, it seeks to prevent IPRs from being used beyond their intended legal scope to distort competition in a downstream market. The fourth requirement introduces an additional public interest rationale by conditioning unauthorised access to an IPR on the demonstration that the IPR is essential for developing a 'new product' that the IP holder is not producing. In such cases, refusing to share the IPR would harm both competition and innovation, preventing consumers from benefiting from the new product.

The rationale behind the assessment under the fourth requirement is confirmed and clearly defined in *IMS*, the subsequent European EFD decision, which also involved copyright protection.[92] As stated by the CJEU, the 'new product' condition relates to the consideration that, in the balancing of the interest in protection of the IPR and the economic freedom of its owner against the interest in protection of free competition, the latter can prevail "only where refusal to grant a licence prevents the *development of the secondary market to the detriment of consumers*"[93] [emphasis added]. Therefore, the refusal to allow access to a product protected by an IPR, where that product is indispensable for operating on a secondary market, may be regarded as abusive only where the undertaking which requested the licence "does not intend to limit itself essentially to *duplicating the goods or services already offered on the secondary market* by the owner of the intellectual property right"[94] [emphasis added].

Unfortunately, the balance between conflicting interests was undermined by a later judgment from the General Court in *Microsoft*, where the dispute centred on the copyright of interface information necessary to promote interoperability.[95] In that occasion, by providing an expansive interpretation of *Magill*, the Court argued that the new product criterion should be read to include a restriction of technical development.[96] However, it is clear that this interpretation causes a substantial shift

---

91 *Ibid.*

92 CJEU Case C-418/01 *IMS Health GmbH & Co. OHG v. NDC Health GmbH & Co. GH* [2004] EU:C:2004:25.

93 *Ibid* at [48].

94 *Ibid* at [49].

95 General Court Case T-201/04 *Microsoft v Commission* [2007] EU:T:2007:289.

96 *Ibid* at [647].

in the equilibrium, as, from the perspective of IP incentive protection, a significant distinction exists between a new product and a follow-on innovation.

For these reasons, in presenting our proposal to interpret copyright exceptions for GenAI training data in light of the *Magill* criteria, we suggest relying on their original formulation.

B. *Applying the Magill Framework to GenAI: Clarifications and Caveats*

The *Magill* framework appears to complement the Berne Convention's three-step test effectively. While the latter permits exceptions only in "certain special cases," the former allows copyright limitations on competition grounds only in "exceptional circumstances." Further, while the Berne test requires that free uses do not interfere with the "normal exploitation of the work" and "unreasonably prejudice the legitimate interests of the author," in response, *Magill* conditions unauthorised access to copyright-protected materials on the premise that denying access would likely eliminate all "competition in the secondary market" and prevent the emergence of a "new product" for which there is a potential consumer demand.

As previously mentioned, the WTO Panel referred to the concept of "economic competition" to determine whether an exception conflicts with the normal exploitation of a work under the Berne Convention, thereby depriving rights-holders of significant commercial gains.[97] Similarly, in applying the fair use doctrine, the US Supreme Court held that the only cognisable harm to the original work is "market substitution"[98] and that the market effects analysis must also consider the "public benefits" the copying is likely to produce.[99] In summary, to use a recent expression, the problem of substitution is "copyright's *bête noire*".[100]

In this regard, and for the crucial interpretation of the scope of copyright exceptions in the context of GenAI training data, *Magill* appears to offer a much-needed clarification of the Berne Convention's three-step test. Indeed, to balance competing interests, *Magill* incorporates both a requirement of market harm and a condition linked to potential public benefits. Notably, the "secondary market" requirement ensures that copyright holders and third-party users do not compete directly, safeguarding investment incentives by preventing direct market harm to rightsholders. At the same time, it prevents copyright from being misused beyond its intended legal scope to distort competition in downstream markets. Additionally, the "new product" criterion introduces a public interest rationale by linking the prerogatives of copyright holders to potential harm in both competition and innovation, as it could deprive consumers of the benefits of a new product.

Therefore, due to the expected impact of AI and the role of data in training AI models, *Magill* adds significant value to efforts aimed at balancing the public interest in reaping the benefits of innovation with the need to protect private investment

---

97 WTO Panel, *supra* note 66.
98 *Campbell, supra* note 72.
99 *Google, supra* note 73.
100 *Andy Warhol, supra* note 21.

incentives against the risk of free riding. Indeed, the *Magill* interpretation of the three-step test conveys a clear message: the use of works as training data for GenAI models should fall within the exceptions to copyright protection, provided the output is not a mere duplication of goods and services already offered on the secondary market by copyright holders.

In this scenario, some caveats are necessary.

First, this article does not support the involvement of antitrust authorities or the enforcement of competition law.[101] Instead, given the non-competition principle embedded in the three-step test, it proposes interpreting copyright exceptions relevant to AI model training through the lens of antitrust tools.

The interface between antitrust and intellectual property has long been a topic of intense debate, often complex and contentious. Despite frequent statements affirming their shared goal of promoting innovation, the relationship between antitrust and intellectual property is marked more by conflict than complementarity.[102] Rather than exercising caution when addressing issues related to interference with IPRs, antitrust authorities have often fuelled this conflict through an activist approach, challenging practices that are not necessarily beyond the scope of intellectual property protection and, in doing so, questioning its very rationale. Against this background, antitrust intervention to assess the scope of copyright in GenAI training data would only add another layer of complexity and controversy.[103]

Second, since we suggest relying on the *Magill* framework solely to interpret the conditions of the three-step test, the EFD is applied outside its original antitrust context. This necessitates certain adjustments. In particular, the EFD was developed as a tool to assess the anticompetitive nature of refusals to deal involving indispensable inputs. When applied to IPRs, it results in the imposition of a compulsory licence, which is not intended to be free but rather granted in exchange for fair remuneration. In our analysis, instead, the EFD is used to clarify the meaning of the three-step test. Therefore, if the requirements are met, it would establish that third parties can rely on the copyright exception and freely use the materials.

Finally, for this very reason, our adjustments also emphasise the third and fourth *Magill* criteria rather than the entire EFD framework. Indeed, as shown, these criteria are particularly useful for interpreting the scope of copyright exceptions. In any case, meeting the other two *Magill* requirements would not be particularly difficult. Data is indeed an essential input for training AI models. Recent findings show that synthetic data cannot fully replace original data, as training on samples from other generative models can cause distribution shifts, eventually leading to AI model

---

[101] Conversely, see Andrei Hagiu & Julian Wright, "Artificial intelligence and competition policy" (forthcoming) International Journal of Industrial Organisation, arguing that competition authorities should be involved to ensure the IP law adjustments balance the rights of content creators and publishers against the potential effects on competition in the provision of innovative AI models and services.

[102] Colangelo & Pardolesi, *supra* note 80.

[103] Daryl Lim & Peter K Yu, "The Antitrust–Copyright Interface in the Age of Generative Artificial Intelligence" (forthcoming) Emory LJ. See also Pamela Samuelson, Christopher Jon Sprigman, & Matthew Sag, "Comments in Response to the Copyright Office's Notice of Inquiry on Artificial Intelligence and Copyright" (2023) <https://www.regulations.gov/comment/COLC-2023-0006-8854> (30 October 2023).

collapse.[104] Further, regarding the second EFD exceptional circumstance, it is not easy to identify objective justifications for denying access to data for AI training.

C. *From Magill to Andy Warhol: Promoting Consistency Between Civil and Common Law Countries*

As previously illustrated, the international landscape is already fragmented, with each country adopting its own approach to copyright protection in GenAI. Significant differences arise not only between civil and common law systems but also within the same legal tradition. Some countries have introduced specific — albeit diverging — TDM exceptions (*eg*, the EU, Japan, and Singapore) or are considering doing so (*eg*, the UK and Canada), while others currently rely on a general fair use defence (*eg*, the US and South Korea). Meanwhile, countries like Australia and India provide neither a specific TDM exception nor a general fair use clause.

In this context, the proposed framework for determining when data training qualifies as a free use of copyrighted works could offer a workable solution to enhance consistency and reduce regulatory fragmentation. Indeed, the concepts of competition and market substitution form a common thread linking the European *Magill* decision to both the Berne Convention's three-step test and the fair use four-factor test, as recently interpreted by the US Supreme Court in *Andy Warhol*. The conflict between copyright owners and AI developers arises when free use leads to economic losses due to market substitution. Conversely, if no such substitution occurs, creators' incentives remain intact, and free use can serve the public interest. Therefore, rather than proposing new bespoke copyright exceptions, our approach may foster harmonisation between civil and common law systems by offering interpretative tools to identify non-competition scenarios between copyright holders and AI developers.

To this end, the link between the *Magill* EFD framework and the competition-oriented analysis adopted by the US Supreme Court in *Andy Warhol* is particularly relevant. Indeed, beyond the US, other common law countries have incorporated a general fair use clause into their copyright laws. For example, Singapore's Copyright Act includes both a US-style fair use four-factor test and a recent exception for computational data analysis.[105] Similarly, Israel[106] and South Korea[107]'s fair use provisions are modelled on the US Copyright Act. Moreover, the suggested connection between *Magill* and *Andy Warhol* could offer valuable insights for policymakers in common law countries that are following a 'wait and see' approach to AI and copyright, such as Australia and India, or are considering the introduction of specific TDM exceptions, like the UK, Canada, and South Korea.

---

[104] See Ilia Shumailov et al, "AI models collapse when trained on recursively generated data" (2024) 631 Nature 755, finding that access to the original data distribution is crucial for learning tasks, making data collected from genuine human interactions with systems increasingly valuable.

[105] Copyright Act 2021 (2020 Rev Ed) at section 191. For a more in-depth analysis, see David Tan, "The price of Generative AI learning: exceptions and limitations under the new Singapore Copyright Act" (2023) 45 Eur IP Rev 400.

[106] Copyright Act 2007 at section 19.

[107] Copyright Act 2011 at Article 35-3.

IV. Concluding Remarks

Since AI models require training on datasets, innovation and competition in the sector depend on access to data, which inevitably includes copyright-protected material. At the same time, by relying on leveraging human works, GenAI systems raise concerns about the lack of incentives for future authors and creators if they are not adequately compensated.

Reconciling these interests largely relies on copyright limitations and exceptions, though their interpretation has traditionally been contentious.

Despite different national approaches, the guiding principles of copyright exceptions are rooted in the Berne Convention's three-step test, which establishes a non-competition principle ensuring that the use of copyrighted works does not interfere with the rightsholder's exploitation resulting in a market harm. From this perspective, economic losses suffered by copyright holders due to market substitution effects arising from the free use of their work typically conflict with copyright's fundamental objective of encouraging creative expression. On the other hand, if there is no market substitution, private incentives to create would remain unaffected, and promoting the free use of works would serve to enhance the public interest.

Against this backdrop of uncertainty and regulatory fragmentation, rather than advocating for new bespoke copyright exceptions or intervention by antitrust authorities, this article proposes a framework to assist courts and policymakers in determining when data training should qualify as a free use of copyrighted works. Specifically, it argues for incorporating competition analysis into the assessment of copyright limitations and exceptions. Notably, since market harm is the key criterion for applying these exceptions, this article contends that antitrust tools can be highly effective in evaluating when creators suffer commercial harm and when AI-generated content may be considered a substitute for human creations.

To this end, it is suggested that useful insights can be drawn from the European EFD as originally established in the *Magill* decision. The *Magill* framework would effectively complement the three-step test by clarifying the meaning of each stage. As a result, it would strike a balance between the competing interests at stake, ensuring that the analysis of limitations and exceptions remains aligned with the fundamental purpose of copyright. Further, this approach would protect copyright from itself, safeguarding it from the risk of being viewed as a barrier to innovation and progress, rather than one of its key drivers.